EMBODIED WORD-FORM LEARNING: THE MOTORIC AND PERCEPTUAL DETERMINANTS OF VERBAL SEQUENCE LEARNING

Amanda Maria Sjöblom

Royal Holloway, University of London

Thesis submitted in fulfilment of the degree of Doctor of Philosophy

July 2019

DECLARATION OF AUTHORSHIP

I, Amanda Maria Sjöblom, 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:

DISSEMINATION OF FINDINGS

Findings from Experiments 1-3 have been disseminated at the following conferences:

Sjöblom, A., & Hughes, R. W. (2018). Hebb sequence learning: A perceptual-motor account. Invited seminar at the School of Psychology, University of Central Lancashire, Preston, UK.

Sjöblom, A. M. & Hughes, R. W. (2018). The motoric and perceptual bases of verbal sequence learning. Poster presented at the Embodied and Situated Language Processing conference, Lancaster University, UK.

Sjöblom, A., & Hughes, R. W. (2017). Verbal sequence learning re-evaluated: The role of perceptual and motor processes. Paper presented at the Meeting of the British Psychological Society (Cognitive section), University of Liverpool, UK.

Sjöblom, A. M. & Hughes, R. W. (2017). Verbal sequence learning re-evaluated: The role of perceptual and motor processes. Paper presented at PsyPAG, Liverpool Hope University, UK.

Sjöblom, A. M. & Hughes, R. W. (2017). Verbal sequence learning re-evaluated: The role of perceptual and motor processes. Paper presented at the Doctoral School Interdisciplinary Conference, Royal Holloway, University of London, UK.

Sjöblom, A. M. & Hughes, R. W. (2016). Verbal sequence learning: The role of perceptual and motor processes. PsyPAG, Poster presented at University of York, UK.

Experiments 1 and 2 are also reported in the following, submitted, paper:

Sjöblom, A. M. & Hughes, R. W. (2019). Verbal sequence learning: A role for motor planning.

ACKNOWLEDGEMENTS

First, huge thanks must go to my supervisor, Dr Rob Hughes, not only for his guidance and for sharing his knowledge and insights, but also for his attention to detail. Many thanks also for the inspiration and the encouragement for further inquiry.

I would also like to thank my advisor, Prof Polly Dalton, for her help and support in both research and the PhD life surrounding it. Many thanks must also go to the rest of the Department of Psychology for the many hours of exciting discussions from the minute details to the grander visions.

Thanks also to Dr John Marsh and Melissa Barker for hosting me at their lab at the University of Gävle, and for borrowing their expertise in fNIRS methodologies to enable my research.

My sincerest thanks also to the research funders and grant providers, Royal Holloway, University of London, British Psychological Society Cognitive Section, and British Psychological Society Study Visit scheme.

Special thanks to Jasmine Virhia, Zsuzsanna Doprontei, Dr Hannah Harvey, Dr Nicky Dawson, Dr Jen Mills, and my favourite conference-buddy Rachel Nesbit for being there for all the highs and all the lows, for the excitement and the hysteria, and for providing much needed help and company. You made the years unforgettable. Rachel, you always knew when tea was required, and undoubtedly contributed to both my alarming consumption of caffeinated drinks and to the survival of my sanity. Zsuzie, PhD life would have been much harder, had we not shared countless of hours at a sweaty gym, thank you for always sparring me to exceed what I thought my limits were. To my lab mates, who were always happy to be a sounding board even for the wildest ideas, and for the support through the successes and the tough times, I could not thank you enough. To the whole PhD community, massive thanks for the great spirit, the encouragement (and the distractions), and the massive over-consumption of coffee we shared. With you, even the smallest victories were always more exciting.

To my family, I have eternal gratitude for your constant support and your ability to retain a sense of humour even when I forgot mine. I knew I could always rely on you when needed.

Anna, Vivi and Oona, I have been especially fortunate to have you, and despite the distance, you have always managed to provide light and laughter when I’ve most needed it. Thank you for being both the sound of reason and the devil on my shoulder.

Last, but perhaps most persistent, Kylo made great contributions to the progress of my writing by sleeping on my feet and keeping me at my desk. Though he will probably appreciate the ready-meal meatballs more than acknowledgments.

SUMMARY

Perceptual and motor processes are often viewed as peripheral systems, subservient to central ‘higher level’ cognitive structures. An alternative approach adopted in the present project characterises the cognitive functioning classically considered to be supported by specialised modules as the product of the embodied processes involved in organising environmental input into candidate-objects for action and producing goal-appropriate behavioural outputs. The present project is the first to test the view that learning novel verbal sequences—attributed classically to the operation of a distinct phonological short-term store—can be reconstrued within this alternative framework. Experiments 1-3 (Chapter 2) used the Hebb sequence learning paradigm—the enhanced serial recall of a repeating sequence amongst otherwise non-repeating sequences—and provided several lines of support for a perceptual-motor account: First, Hebb sequence learning was attenuated when vocal-motor planning of the sequence was restricted by requiring participants to utter an irrelevant verbal sequence (‘articulatory suppression’) or when no recall-response was required. The effect of suppression was smaller with auditory sequences, however, suggesting that passive auditory perceptual organisation processes can independently support auditory Hebb sequence learning. Second, Hebb sequence learning was enhanced for phonologically similar compared to dissimilar sequences when that learning was driven solely by motor planning. Third, disturbing the consistency of the temporal grouping of the repeating sequence abolished learning but only when that grouping was instantiated within a motor-plan. Fourth, demonstrating more direct evidence for a contribution of passive perceptual organisation in learning an auditory-verbal sequence, promoting the perceptual grouping of every-other-item in the repeating sequence by presenting it in alternating male-female voices led to the learning of those non-adjacent-item sub-sequences. Experiments 4-6 provided evidence that motor planning processes also play a role in nonword learning in the paired-associate paradigm, where lists of nonwords (together with known words) are presented and recalled repeatedly. Nonword learning was attenuated when motor-planning fluency was impeded either by articulatory suppression or as the result of phonological similarity within or between the nonwords. The findings are discussed in the context of the debate on modular versus embodied cognition as well as in terms of their implications for word-form learning.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

CHAPTER 1

VERBAL SERIAL SHORT-TERM MEMORY AND LEARNING: DEDICATED MODULE OR EMBODIED PROCESSES?

Since the cognitive revolution of the 1950s, it has been common to view perceptual and motoric processes as peripheral; as merely the input to, and output from, central cognitive structures that carry out fundamental psychological functions such as memory, language, and learning (Hurley, 2001; Wilson & Clark, 2009). A distinctly opposing view is that these functions rely on, and are inseparable from, processes that originally evolved for the processing of perceptual inputs and the preparation and production of motor outputs (Anderson, 2003; Gibbs & van Orden, 2010). The two approaches differ on whether human cognition is built and operates in a modular fashion like a computer, or whether cognition is embodied, inseparable from the particular environmental context and the organism's capacity to act upon it given its particular bodily and motoric constraints. Whether cognition is modular or embodied has, in addition to its theoretical significance regarding philosophy of mind, implications for our understanding of, and empirical approaches to, all cognitive functions, how they have evolved, and how they developed (e.g., Eigsti, 2003; Koziol, Budding, & Chidekel, 2012; Mahon & Caramazza, 2008).

Set within the general distinction between modular and embodied approaches to cognition, interest in the present thesis centres in particular on verbal serial short-term memory and verbal sequence learning. Verbal serial short-term memory refers to the capacity to reproduce, over a matter of a few seconds, a sequence of verbal items in serial order (Baddeley, 1966). Verbal sequence learning is the process of learning, i.e., committing to long-term memory, novel verbal sequences, such as is required for learning new words (Szmalec, Duyck, Vandierendonck, Mata, & Page, 2009). This chapter aims to outline the primary phenomenon of interest—verbal sequence learning and its relation to the capacity to reproduce a novel sequence over the short term—in the context of a modular and embodied approach. The predominant account of verbal serial short-term memory and learning is embedded in the classical modular framework and assumes that these functions are supported by a dedicated phonological short-term store (Baddeley, Gathercole, & Papagno, 1998). An emerging perceptual-motor account, in contrast, posits that these capacities are parasitic on embodied, general-purpose, motor and perceptual processes rather than supported by a dedicated cognitive module. The veracity of these accounts of verbal serial short-term memory and learning will for the first time be contrasted in the context of Hebb sequence learning (e.g., Hebb, 1961; Mosse & Jarrold, 2008; Page & Norris, 2009a, 2009b; Szmalec et al., 2012) and nonword learning (e.g., Martin, Boersma, & Cox, 1965; Noble & McNeely, 1957; Papagno & Vallar, 1992). The general thesis will be that verbal sequence learning is a by-product of the perceptual and motor-planning processes that are engaged in the service of goal-relevant verbal serial short-term behaviours.

1.1 Modular Versus Embodied Mind

In cognitive science, broadly two main approaches exist as to how the mind is constructed. One approach posits that the mind is, to at least some extent, partitioned into specialised modules each of which deals with a specific mental function, ranging from ‘lower-level’ modules for visual object identification, modules for different kinds of remembering (e.g., short- compared to long-term memory), to modules evolved for producing elements of abstract thought or ‘higher cognition’ abilities (Machery, 2008). At the other end of the spectrum, the embodied cognition approach argues for integrated, joined processing, some suggesting that thinking only exists in the context of the perception of, manipulation of, and actions upon, the given environment, to the more conservative embodied theories stating that processes primarily evolved for sensing and perceiving inputs and for producing motoric outputs are an integral part of mental processes at all ‘levels’ of cognition (Wilson, 2002). One of the primary differences between the polar ends of the modular and embodied approaches is the overall purpose of cognition: In the modular view, it is to form internal, symbolic representations of the environment and to manipulate them (e.g., Fodor, 1983). On the embodied view, cognition is for processing environmental information for the purpose of producing the appropriate behaviours to act adaptively in the environment (e.g., Engel, Maye, Kurthen, & König, 2013). In the modular mind, the formation of the abstract-symbolic representations is a goal in its own right, while in the embodied approach, cognition serves adaptive behaviour, and thus cognition must be grounded in the sensorimotor systems.

1.1.1 The Modular Mind Approach

Classically, the cognitive system has been considered to consist of separate mechanisms specialised for distinct functions (Pinker, 1994), such as episodic memory (e.g., Eichenbaum, 2004), short-term memory (e.g., Baddeley, 1989) or language representations, which can be broken down further into independent domains such as grammar (e.g., Näätänen et al., 1997; Pinker, 1994). The different cognitive abilities and functions have been largely considered as separate from each other, and though linked due to functional behaviour simultaneously requiring multiple processes run by many different modules, their processing has been seen as independent (Sternberg, 2011). Due to the functions being conceptualised as separate modules, each of these functions has often been studied in isolation, such as verbal short-term memory and speech production, each generally approached within different fields of study, despite evidence of shared properties of the functions (MacDonald, 2016). These ideas of modularity of mind first became prevalent during the cognitive revolution, around the same time that the computer sciences started growing (Hurley, 2001; Wilson & Clark, 2009). The modular-mind ideas about human information processing were inspired by the workings of a computer, utilising units such as abstract feature lists, and this has led to the extensive computational modelling of many cognitive functions (e.g., Sternberg, 2011). Thus, many models of cognition represent cognitive processes in the working terms of computations, where computations or processes are run by discrete units, requiring a specific type of input and outputting computations in specific formats. These units, though connected to each other, often in a hierarchical manner—in order for each to contribute its function to the larger-scale function—all work in relative isolation (Meunier, Lambiotte, & Bullmore, 2010).

The extent of the modularisation of cognitive functions has long been a topic of controversy. The massive modularity views (e.g., Carruthers, 2003) are characterised by extensive numbers of modules running very specialised, small, functions independently with other modules working as the unification processors by linking the many smaller modules as required for complex task performance. Other models have also proposed modules that are capable of processing larger aspects of information individually, without requiring as large a number of supporting submodules, as well as versions where only some parts of cognition are modular (e.g., Fodor, 1983). Depending on the particular model, there can be specific modules supporting ‘higher-level’ cognitive abilities and functions, dealing with matters ranging from abstract thought across the past, present and future tenses, to orthographic language representations. Alternatively, some models deal with ‘higher level’ cognition by proposing more ‘connectionist’ modules that are able to make more complex inferences, which can be seen as more central functions (Pinker, 2005). The main modules can then be accompanied by sets of peripheral, ‘lower-level’, modules dealing with functions such as generating information from perceptual input derived from different senses, subservient to the higher-level cognitive functions, which are responsible for the manipulations of symbolic representations. Hurley (1998) describes this traditional view of cognition as the “classical sandwich model”, outlined in Figure 1.1, where cognition is like the filling, embraced from either side by a slice of bread: the perception slice (the source of input into cognition) and the motor slice (the means by which the products of cognition can be turned into overt behaviour).

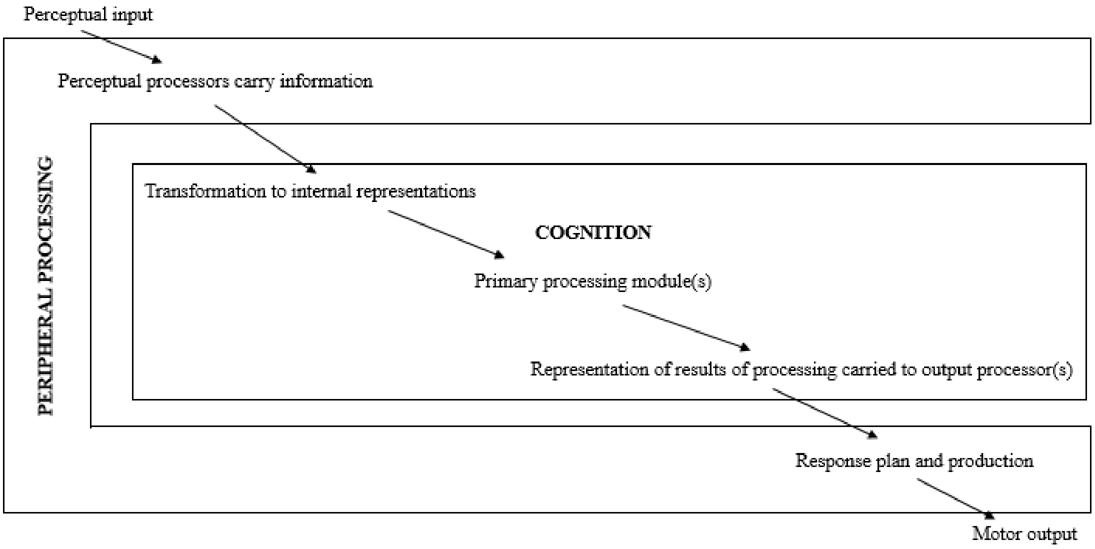


Figure 1.1. A depiction of the modular cognition structure, based on the sandwich model of cognition as described by Hurley (2011).

Regardless of whether a particular model could be described as ‘massively modular’ or is more moderate in its claims of modularity, the actual units of functioning, and the nature of modules themselves as self-contained units of specialised processing, remain broadly similar across variants of the modular approach. The character of modularity in itself, though recently contested (e.g., Carruthers, 2006; Machery, 2008), generally means that each specialised function can only deal with certain types of input and utilises specialist representations of the information relevant to its processing. Thus, the information format contained in each module is domain-specific to the function the module is responsible for performing (e.g., Fodor, 1983). Following the modularisation of functions in these models, knowledge in the semantic memory system, for example, is considered separate from the different modules processing aspects of perception and action or introspection, from which the knowledge would have originated (e.g., Carruthers, 2006). This separation of long-term memory modules and short-term memory modules has been an essential characteristic both in modular models of memory (e.g., Baddeley, Lewis, & Vallar, 1984) and in many general models of cognition (e.g., Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004). These memory modules are a key part of the model architecture, associated with particular cortical regions, where representations of information go through a fixed series of operations in the architecture to produce cognition (Anderson, 1995; Anderson et al., 2004). Knowledge is then stored in an amodal format, separate not only from the initial perceptual signals, but also from other, unrelated cognitive modules, using a unique internal representation format. The specific modules and their nature are quite varied from model to model (e.g., Cosmides & Tooby, 1994; Fodor, 1983; Pinker 2005). The classic models, such as the introduction to modularity in Fodor’s (1983) modular mind–model, are focused largely on specifying the ‘lower level cognition’ modules. In the early models of Fodor, the modularity is located only in the ‘lower-level’ systems that process information in ways that will then be suitable for non-modular higher-level cognitive processes. The Fodorian modules have several defining characteristics: Domain specificity, in which a specific function is performed with information originating from restricted information-types, such as from a specific sensory modality (Cosmides & Tooby, 1994, 1995); limited central accessibility and encapsulation, where the inputs to and the output from modules are limited to specific options; fast and shallow processing, along with mandatoriness, describing the processing as resource-cheap and fast, but also largely out of control; fixed neural structure as fixed neural locations supporting specific functions, with little room for flexibility (Fodor, 1983). Along with the assumption of fixed structure, the modules in the model are considered innate. While newer models and theories have since built on the original ideas of Fodor, regardless of their view of the extent to which more ‘central’ systems are modular, the essential character of modules as independent units of functioning has remained similar (e.g., Carruthers, 2005, 2006). The massive modularity theories posit that the modular nature of processing exists at all levels of cognition, from the Fodorian, lower-level perceptual modules to modules dedicated to functions, or parts thereof, such as problem-solving and various language abilities. Massive modularity, by its requirements for individual modules for complex functions, largely abandoned many of the original ideas proposed by Fodor, but still retain the essentials: the modules are isolatable, function-specific and domain-specific, as well as largely considered to rely on set neural structures (e.g., Carruthers, 2005; Cosmides & Tooby, 1995; Pinker, 1994).

One of the main properties of this type of cognitive model is the assumption of abstract-symbolic representations, ‘above’ perceptual processes—be they, for example, internal representations of visual input, or heard language—on which an organism’s experience of the world relies (O’Regan & Noė, 2001). The format of a particular representation is distinctly related to the properties of the module(s) that form and use it. The putative need for these internal representations, distinct from initial perceptual properties, is linked to the need for ‘offline’ reasoning: the cognitive processes concerning material that is not (or no longer) present in the environment, requiring an internal duplicate available for further processing and manipulation. Each separate domain is, in general, assumed to be functionally specialised, each utilising a unique representational format (Fiddick, Cosmides, & Tooby, 2000). These representations are, due to their domain-specific nature, different from initial perceptual input. This, again, is related to the separation of the modules, for example in the separation of long-term and short-term storage systems, where a key piece of evidence for the separation of these systems is the supposed differential representation format they use (e.g., Atkinson & Shiffrin, 1968; Baddeley, 1966; Broadbent, 1984). However, despite providing tools for characterising a multitude of cognitive functions in computational modelling and other frameworks, there is little empirical evidence that these abstract, module-specific, internal representations are the form in which knowledge resides and is processed (e.g., Barsalou, 2008).

The evidence for massive modularity, or, indeed, modularity at all, has been largely reliant on discussion of how evolutionary adaptation might have favoured modular systems, where new modules evolve to support new behavioural needs (Cosmides & Tooby, 1995), or reliant on some evidence of domain-specificity of learning systems in animals (Carruthers, 2006). Yet modularity does not need to follow from these considerations. Indeed, the preference by adaptation of certain types of changes and learning processes in animals have also largely been accounted for within embodied approaches without the need to specify modules as the focal point of adaptation or development (e.g., Wilson, 2002). However, with the emergence of neuroimaging, the modular models have appeared to gain additional support, with very specific cognitive functions localised to specific regions of the brain (e.g., Buller & Hardcastle, 2000; Carruthers, 2005). The idea of a modular mind is usually accompanied by the assumption that each function is served by a specific region or fixed network of regions in the brain: this discrete-modules–type approach proposes that cognitive functions are served by a specialised region in the brain such as, for instance, the right fusiform gyrus for face processing or the left fusiform gyrus for word-form processing (e.g., Kanwisher, McDermott, & Chun, 1997; McCandliss, Cohen, & Dehaene, 2003). Distinctions have also been proposed between modules and regions that deal with different aspects of a stimulus. For example, it has been suggested that the prefrontal cortex has separate regions that independently deal with what a particular object in the environment is, and another is concerned with where it is (Wilson, O'Scalaidhe, & Goldman-Rakic, 1993). Some disagreement exists on whether these modules are genetically determined evolutionary adaptations or whether they are based on the plasticity of the brain responding to environmental experience during development (e.g., Buller & Hardcastle, 2000; Carruthers, 2005). Along with the neuroimaging evidence, studies of patients with brain lesions have traditionally been taken to support views of localised, informationally encapsulated, modules in the brain, with the (usually tacit) assumption being that, following a lesion, the parts of the brain that were not directly affected continue to function as before (e.g., Brooks, 1975; Geshwind, 1970; Shallice & Warrington, 1970). Amongst these functions that have been considered localised due to lesion data are aspects of language, such as damage to Broca's area related to a language production deficit without affecting comprehension, while damage to Wernicke's area is thought to produce the opposite pattern (e.g., Geshwind, 1970).

Similarly, damage to areas associated with long-term memory has been thought to leave short-term memory relatively intact (Brooks, 1975), while damage in some patients has been observed to result in decreased short-term memory performance with little effect on long-term memory (Warrington & Shallice, 1969). However, others have argued that these data are in fact more supportive of a view of information processing as distributed and interactive rather than localised (e.g., Farah, 1994), though this can be used to support both a distributed-modules approach as well as the embodied standpoint.

1.1.2 The Embodied Cognition View

More recently, ideas of dispersed rather than distinct functions have become more prominent. Many cognitive functions, previously considered discrete and autonomous, have been proposed to be reconceptualisable as the more fluid and interactive operation of more general-function processes. Such views, like the theories falling under the ‘embodied cognition’ approach, propose that perceiving and acting are linked to cognitive processes as more than merely peripheral systems and, in many of these theories, are an integral part of cognitive processing (e.g., Wilson, 2002). Though there are many different views within the embodied approach, they are generally joined in the idea that the traditional focus on abstract-symbolic representations falls short in explaining cognitive functions (Anderson, 2003). Embodied theories are generally united by the coupling of perception and action, and the rejection of modules in favour of cognition being supported by embodied sensory and motor processes (e.g., O’Regan & Noė, 2001; Thomas, 1999). Indeed, many embodied theories are centred on the view that cognition has evolved to produce movement, the only way we are able to affect the environment (e.g., Koziol, Budding, & Chikedel, 2011). In particular, even ‘offline cognition’—

cognitive processing focused on objects that are no longer present in the environment where in-the-moment perception of those objects cannot be used to support processing—is assumed to be embodied. The embodied models generally achieve this via recourse to the idea that long-term memory representations retain perceptual and motor properties of objects or events, rather than being removed from the original perceptual or motor context in which they were first encountered (e.g., Wilson, 2002).

In essence, in embodied or grounded theories, 'perception' is the exploration of the environment and 'motor' is the acting upon knowledge about the environment. In terms of accounting for 'offline' reasoning, some embodied theories have conceptualised this as the simulation of sensing and acting (Clark, 1999). The various specific models of embodied cognition also share the view of behaviour or action as purposeful, goal-driven and generally involving volitional control and planning, and not simply as movement. One of the main differences between the modular and embodied views is that the latter denies that specific modules have evolved to produce specific functions. Instead, complex cognitive functions are parasitic on processes whose adaptive function is the organisation of information from the environment and the planning and outputting of context-appropriate behavioural responses. Rather than having evolved new modules to perform the more complex behaviours required, the earlier-evolved processes are considered to have evolved further in such a way as to be effective in a variety of specific contexts and for a variety of purposes to produce goal-relevant behaviours. In addition, though some research focuses on basing parts of cognition solely on bodily states, the embodied approach also includes views allowing cognition to be built on multiple bases, including sensory, bodily states, but also others, such as simulated and environment-situated actions (Barsalou, Brezeal, & Smith, 2007). This has led to criticisms of the approach as only able to account for sensory or motor functions, and that an embodied or a grounded system would be unable to deal with abstract concepts not directly related to sensory input or motor output (Pylyshyn, 1973). However, these criticisms have generally been successfully met with reference to simulation–type function and internal, introspective sensory information as a way of representing abstract information, where the processing of that information can be relevant for future perception of, and action in, the environment (e.g., Barsalou, 1999).

The shift towards embodiment within experimental cognitive psychology has extended also into cognitive neuroscience where many functions are starting to be considered as being supported by distributed neural networks. These networks are considered to have evolved initially for perceptual and motor processing, concerned with the demands of interacting adaptively with the environment, and which now also provide the basis of varied ‘central’ cognitive functions (Wilson, 2002). This contrasts with the view that new modules would have evolved to support the evolutionarily newer ‘central’ functions (e.g., Carruthers, 2005). This view can better account for large individual differences, with the network connections highly modifiable by experience instead of being shaped into pre-determined modules. Indeed, recently, the modular views have been criticised for discounting overlapping activation areas in favour of separating out dissociated functional areas (Postle, 2006). As an example, some evidence suggests that the visual perception of faces and that of words, traditionally separated into different hemispheres, are in fact supported by overlapping neural circuits, where the activation pattern differs according to the task at hand (Behrmann & Plaut, 2013). A recent emphasis has been on how cognitive processes are irremovably intertwined with the production of actions (e.g., Engel et al., 2013), leading to consideration of motor behaviour not only as output from the cognitive system but as a critical part of the system. Interesting contributions from motor-imagery research, closely linking imagined movement and actual bodily actions both in behavioural data and in overlapping brain activation patterns (e.g., Lacourse, Orr, Cramer, & Cohen, 2005) can also be easily viewed from the embodied approach: Though the neural activations are not identical, as would be expected due to the lack of actual bodily movement in imagery, the overlap suggests that simply imagining moving activates the neural networks involved in the planning and production of those motor actions. The recent data have also shown that even in motor imagery there is somatosensory processing of what producing the movements feels like, and that this activation is sufficient to support the processing of the information about the action, without the need to involve additional abstract-symbolic representations (Hanakawa, Dimyan, & Hallett, 2008). Motor-related processing has also been linked to memory retrieval accuracy: Dijkstra, Kaschak, and Zwaan (2007) showed that a retrieval posture congruent to an encoding posture improves recall.

The view that cognition is based on the perception of the environment and on the actions of the organism is also supported by a growing body of neuroscientific evidence. These findings have been particularly prevalent in the neuroscientific literature focused on memory, such as evidence that the motor properties of an object are encoded as part of the semantic memory for that object, and the recruitment of motor-related neural networks when perceiving (Weisberg, van Turennout, & Martin, 2007) or mentally manipulating such objects (Richter et al., 2000). Within cognitive-experimental psychology, the importance of the role that motor-related processing plays in cognitive functioning has been demonstrated in imaginative thought and language, such as in a sensorimotor grounding of comprehension of metaphors (Gibbs, 2005), or in embodied emotions as the basis for understanding moral concepts (Prinz, 2005). Additionally, an increasing amount of evidence has also linked brain regions that have previously been thought to only be involved in producing movements to ‘higher-level’ processing such as language, timing, and spatial processing (Cona & Semenze, 2017). Indeed, the evidence for both wide network processing and for a multitude of functions being supported by the same anatomical regions suggest that many cognitive functions can be produced adaptively in a domain-general fashion, contrary to the notion of specialised regional modules for specific functions. In short, the neural evidence indicates both that several brain regions can participate in producing a specific function and that particular anatomical regions can play a role in many different functions, implying domain-generality of regions and the utilisation of domain-general processors when they are applicable to the particular task at hand. These findings are easily aligned with the embodied approach but are less commensurate with modular models.

The embodied cognition approach has re-defined the relations between environmental input, perceptual processing, and cognitive processing as action-guiding functions. Bodily sensation and the perceptual feedback relating to the movement in the environment are essential to cognition, rather than perception simply producing the pathway for re-coding environmental sensory data into abstract-symbolic representations (O’Regan & Noe, 2001). The action-guiding purpose of cognition is also evident in the observation that the motor-related neural networks that are active when making a movement are also activated when observing others perform the same movement, which appears to indicate that the perception and comprehension of the actions of others is tied to the processing required to produce the same actions (Rizzolatti & Craighero, 2004). Applying this general idea specifically to speech processing, it has been argued that the development of speech should be viewed as perceptual-motor co-structuring, with both the perception and production of speech developing together, resulting in an intrinsic coupling of the two. The perceptual-motor processing of speech is then defined by articulatory coherence, i.e., its gestural nature, and its perceptual value, which makes speech functional: Vocal gestures are shaped by their auditory properties (Schwartz, Basirat, Ménard, & Sato, 2012).

The plasticity of neural networks also appears more compatible with adaptive, embodied processing. The ability of the neural networks to change with learning, and to re-configure in order to re-learn functions following brain damage (Nudo, Plautz, & Frost, 2001; Thulborn, Carpenter, & Just, 1999) is conceptually closer to an embodied system, where available resources are utilised according to task goals. Additionally, the uptake of the same structural systems for different purposes—such as the involvement of the visual cortex not only in visual perception but in tactile perception in the blind (Sadato et al., 1996)— undermines the notion of function-specific modules intrinsically tied to specific neural regions. Buller and Hardcastle (2000) also noted that developmental plasticity largely undermines the innateness-assumption of many modular models, although not all modular models have assumed the innateness of modules, at least not for all modules (e.g., Carruthers, 2006). In an example of developmental plasticity, an investigation of infant walking development showed that in infants, skills such as standing, crawling and walking do not appear to follow from the maturation of pre-wired systems and modules generating walking behaviour (Smith & Thelen, 2003). Instead, the development is based on the interaction of initial random and spontaneous movements and the changing contexts, where the repeated interactions result in an emergent, organised product. This type of sequence learning is the basis of all complex behaviours (Ellis, 1996; Melton, 1963).

In the present thesis, interest centres on contrasting the embodied and modular approaches to cognition in the context of serial short-term memory and sequence learning, particularly of verbal input. Verbal serial short-term memory refers to the temporary retention and reproduction of a sequence of verbal items (Baddeley, 1966), and verbal sequence learning is the process by which the verbal sequence encountered in the short-term becomes represented in long-term memory (e.g., Burgess & Hitch, 2005). Such sequence learning is, in the verbal domain, critical for spoken language learning but the general process is also applicable to the learning of other complex behaviours (Ellis, 1996). A classical, modular, account of verbal short-term remembering is that it is supported by a discrete, dedicated, system, which in turn is fractionated into multiple sub-components (the phonological loop of the Working Memory model; Baddeley, 1986). However, short-term remembering of a sequence has recently been increasingly considered in terms of the use of embodied and general-purpose motor and perceptual processes that are exploited opportunistically to meet task demands (e.g., Jones, Macken, & Nicholls, 2004). These two accounts of verbal serial short-term memory will now be reviewed in some detail, followed by how they relate to verbal sequence learning.

1.2 A Modular Approach to Verbal Serial Short-Term Memory: The Phonological Loop Account

Historically, the Working Memory model (Baddeley & Hitch, 1974) grew out of the so-called modal model of memory, the name given to numerous early cognitive models of memory that all shared basic assumptions that were first formalised most clearly in the model of Atkinson and Shiffrin (1968) (Baddeley, Hitch, & Allen, 2019; Norman, 1970). In the modal model, short-term and long-term memory were fundamentally separate modules, though ones that interacted. The modal model distinguished between a set of pre-categorical sensory stores with a very fast decay rate, and a unitary short-term memory that had a limited capacity (see also Broadbent, 1958). Attending to the relevant subset of the contents of the sensory store(s) gave it access to the short-term store. The short-term store was served by the pre-categorical sensory stores and rehearsal of the contents of the short-term store resulted in longer-term memories.

From the short-term memory component of the modal model grew the Working Memory model (Baddeley & Hitch, 1974), which has since been one of the most influential models of short-term memory, particularly of verbal short-term memory. While the modal model assumed that short-term memory functioned as a single unit to hold and manipulate information for a brief period, with an automatic but relatively slow transfer of information to long-term memory, the Working Memory model introduced more active components. The Working Memory model advanced the view of short-term memory from a mere temporary store to an active processing system with multiple components in order to account for complex task performance such as learning and comprehension (Baddeley, 1983, 2000), hence the preference for the term ‘working’ as opposed to ‘short-term’ memory. The Working Memory model consists of a short-term phonological store that holds verbal input and a separate short-term memory module for visuo-spatial input (the visuo-spatial sketchpad) as well as a central executive for directing attention, dividing resources and other organisation–type functions (Baddeley, 2012). Acting as an interface between the short-term stores and between these and long-term memory is a fourth component, the episodic buffer, a limited capacity module, controlled by the central executive, capable of utilising a multimodal code (Baddeley, 2000).

Common to the modal model and the Working Memory model is a structural separation between short-term systems and long-term memory. An historically important line of evidence cited in support of such a separation was that remembering over the short term and remembering over the long term appeared to involve different representational codes (e.g., the case from *differential coding*, e.g., Baddeley, 1996). Specifically, recalling a list of words over the short term appeared to involve some kind of speech-based code, as evidenced by the fact that phonologically similar items (e.g., *mad, man, cap…* or *B, G, D…*) were more poorly recalled in order over the short term than were phonologically dissimilar items (e.g., pen, day, sup… , *R, Q, H…*) whereas there was no difference in the short-term recall of semantically similar compared to semantically dissimilar words (Baddeley, 1966; Conrad, 1964). In contrast, recall of a given set of verbal items across several successive trials (‘serial learning’), where recall is assumed to rely increasingly on long-term memory, did indeed show a semantic similarity effect (Baddeley, 1966). Thus, the short-term system was thought to use phonological coding while long-term memory was thought to rely predominantly on semantic coding (Baddeley, 1996).

The short-term store for verbal (or, more generally, verbalisable) input in the Working Memory model was developed into the concept of a *phonological loop*, easily the most fully specified component of the Working Memory model (Baddeley, 1988, 1989, 2012). The main empirical paradigm on which the phonological loop account, and verbal short-term memory research and theorising in general, has been based is verbal serial recall where participants are presented with a sequence of verbal items, and are required to recall them in serial order (e.g., Baddeley, 1986; Crowder, 1971; Jones et al., 2004). A typical verbal serial recall task involves presenting a seven-or-so item sequence of letters, words, or digits, presented one at a time, at a rate of roughly one item per second. After the sequence has been presented, participants are asked to reproduce the sequence either in written, spoken, or typed form or by mouse-clicking the (re-presented) items in the correct order. The phonological loop construct is made up of two subcomponents: its core component is a passive phonological store dedicated to holding verbal material in phonological form for around two seconds before it decays (Baddeley, 2012), which is served by an active articulatory control process mapping onto the phenomenological experience of inner speech. The articulatory control process is used to refresh the phonological contents of the store (subvocal rehearsal), thereby preventing their loss through decay (Baddeley, 2012). However, there is an important difference in how items enter the store according to their modality. Auditory-verbal material has direct, obligatory, access to the phonological store but before visual-verbal input can enter the store it must go through a deliberate grapheme-to-phoneme conversion process for it to be encoded into the format used by the phonological store. This is achieved by the articulatory control process, its other main function in addition to that of refreshing the store’s contents (Baddeley, 2003). Critically, though the entry-routes of verbal material from the different sensory modalities are different, once the items are in the store, the representations share a common status regardless of their original input modality. This is what gives the phonological loop its abstract-phonological character: two different routes of entry for input from different modalities meeting in a common store, with their representational status thus separated from their original, either specifically auditory-perceptual characteristics of spoken input or the specifically articulatory processes required in the grapheme-to-phoneme conversion of visual-verbal material. The notion that auditory-verbal material has direct entry to the phonological store but is represented there in abstract-linguistic form specific to the store is essential to two key tenets of the account: that a (passive) store exists independently of articulatory processes and that the store is phonological in nature (Baddeley, 2003).

As alluded to above, the empirical signature of the passive phonological store component is the *phonological similarity effect:* verbal serial recall is poorer when the items in the sequence sound similar (*B, T, D…*) than when they sound dissimilar (*R, Q, H...),* even when the items are presented visually[1]. This is the central piece of evidence for the notion that short-term memory uses phonological, as opposed to, for example, semantic coding (Baddeley, 1989). According to the phonological loop account, phonologically similar representations are more easily confused with each other during retrieval from the store, leading to decreased accuracy in recalling the items in the correct serial order (Baddeley, 1989; Baddeley & Larsen, 2007). Pivotal to the assumption that there exists a passive phonological store—that is, that appeal to articulatory rehearsal or planning processes alone is not sufficient—is the way in which phonological similarity enters into a three-way interaction with sensory modality (visual vs. auditory list-presentation) and articulatory suppression (where participants are either free to rehearse the items subvocally or are required to utter, whisper or mouth a repeating verbal item or short sequence, e.g., “the the the…” or “8, 9, 10, 8, 9, 10…”). Articulatory suppression is assumed to impede if not block the rehearsal of the items in the store but also, of greater relevance here, to impede if not

---

[1] It is worth noting that the phonological similarity of these letter-names is based on the ‘ee’ sound at the end of each of the ‘similar’ letter-names, rather than on the phonology of how each letter sounds in the phonetic alphabet.

block the grapheme-to-phoneme conversion process, thus preventing the access of visual-verbal material into the store in the first place. Suppression does not, however, impede the access of auditory-verbal material due to the obligatory access of such material to the store. The key three-way interaction in question, then, is that articulatory suppression eliminates the phonological similarity effect with visual presentation but not with auditory presentation (Baddeley, Lewis, & Vallar, 1984). That the signature of the phonological store is still present even when the articulatory control process is impeded by articulatory suppression has been taken to indicate both the obligatory access of auditory-verbal material into the phonological store and that a passive phonological store must, therefore, exist independently of the articulatory control process. As described by Baddeley and Larsen (2007), the phonological loop system “was initially termed the articulatory loop, but 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” (p. 497).

Further evidence cited as support for the existence of a passive phonological store is the finding that verbal serial recall is impaired by the mere presence of task-irrelevant sound (e.g., the *irrelevant sound effect* (ISE); Beaman & Jones, 1997; Colle & Welsh, 1976; Salamé & Baddeley, 1982, 1989). This effect is proposed to be due to the obligatory access of auditorily presented material into the phonological store. While early phonological loop-based accounts posited that the sound interfered directly with phonological item-representations (e.g., Salamé & Baddeley, 1982), this view proved untenable (e.g., Jones & Macken, 1995) and so more contemporary loop-based accounts posit that representing the order of the sounds (cf. Jones & Macken, 1993) interferes with the passive representation of the order of the to-be-remembered items in the phonological store (Norris, Baddeley, & Page, 2004). Similar to the logic relating to the phonological similarity effect, the argument that the sound interferes with the passive store and not, for example, motor rehearsal is based on evidence that that effect is eliminated under articulatory suppression unless the to-be-remembered items are also presented auditorily (as opposed to visually) and hence when both they and the irrelevant material gain obligatory access to the store (Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987).

1.3 An Embodied, Perceptual-Motor, Account of Verbal Serial Short-term Memory

A recent challenge to the modular phonological-store based account of verbal serial recall has led to the development of a more embodied, perceptual-motor, account of serial short-term memory. In particular, it has been argued that the key evidence taken classically to support the existence of a short-term verbal store dealing with abstract-phonological representations separate from perceptual and motor processes instead demonstrates the ‘centrality’ of perceptual and motoric processing to verbal serial short-term performance (e.g., Jones et al., 2006, 2004; Maidment & Macken, 2012). Of particular importance is that the character of the key three-way interaction described earlier between phonological similarity, articulatory suppression, and modality—which forms the main basis of the notion of a phonological store separate from articulatory rehearsal processes—is not as it first appears. Through more detailed analysis of the interaction, it becomes apparent that the phonological similarity effect is indeed eliminated by articulatory suppression even with auditory presentation throughout most of the serial position curve. The survival of the effect under suppression with auditory input is found primarily at *recency* (the end of the sequence), which had previously gone unnoticed due to the omission of a consideration of serial position data. That is, the residual similarity effect under suppression is a product 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). Critically, the modality effect is much more pronounced, and in some studies is only apparent, with phonologically dissimilar sequences compared to similar sequences (Crowder, 1971; Jones et al., 2004; Watkins, Watkins, & Crowder, 1974). Given that the elimination of the phonological similarity effect by articulatory suppression can be observed at the same level of general recall performance at which, in other conditions or experiments, it has been present (e.g., Jones et al., 2006; Maidment & Macken, 2012), the elimination of the effect cannot be attributed to floor/proportional scaling effects (cf. Beaman, Neath, & Surprenant, 2008; Wang, Logie, & Jarrold, 2016) or to the phonological store being abandoned due to excessive task demands (Baddeley & Larsen, 2007).

Key to a reconceptualisation of the interaction between phonological similarity, modality, and articulatory suppression is that auditory recency is considered a decidedly acoustic effect rather than a phonological one. Indeed, proponents of the Working Memory model consider the modality effect to be “peripheral to the working memory system” (Baddeley, 1986, p. 95; see also Hurlstone, Hitch, & Baddeley, 2014). Good evidence now exists that auditory recency arises from exploiting the products of acoustic-based perceptual organisation rather than reflecting the action of a phonological store (Nicholls & Jones, 2002). More specifically, the evidence that the modality effect is larger with ‘phonologically’ dissimilar sequences than similar ones and that it is restricted to the recency portion of the sequence can be explained in acoustic-perceptual terms. It is well established that temporal order perception in auditory sequences is a nonmonotonic function of the acoustic similarity of the items within the sequence. For reliable order perception, the sequence must contain acoustically distinct items that still cohere as a group or single perceptual ‘stream’ (Bregman, 1990; Hughes et al., 2009, 2016). Thus, relatively poor order perception is observed when the items are acoustically indistinct. However, should the items of the sequence be so distinct that they are no longer perceptually grouped together, order perception is poor again (Bregman & Campbell, 1971; Jones, Alford, Bridges, Tremblay, & Macken, 1999; Lackner & Goldstein, 1974; Warren, Obusek, Farmer, & Warren, 1969). This can be seen in auditory-verbal serial recall experiments showing that acoustically distinct items are recalled better that acoustically indistinct items (i.e., the classical ‘phonological’ similarity effect) but recall is markedly reduced again if the acoustic distinctiveness between items is increased still further by presenting them in distinct voices (Hughes et al., 2009, 2011). The conclusion is, then, that the modality effect is larger with ‘phonologically’ dissimilar sequences because ‘phonologically’ similar sequences are too acoustically, not phonologically, similar to yield strong order cues.

The restriction of the modality effect to recency can be viewed as a *perceptual-group boundary effect*. The silence that follows the sequence of auditory items creates a boundary to the sequence, making the items near that boundary, along with their order, particularly salient (Bregman & Rudnicky, 1975), similar to how the boundaries—or edges—of visual objects are salient (Wagemans et al., 2012). The impact of the perceptual-group boundary can be seen in the elimination of auditory recency when a redundant auditory item (a suffix) is added to the end of the to-be-recalled auditory sequence. The suffix now constitutes the boundary, preventing the use of that boundary for accessing the last to-be-remembered items. In turn, this suffix effect can be eliminated, that is, auditory recency can be restored again, by contriving conditions designed so that the suffix is likely to be perceptually grouped separately from the to-be-recalled sequence. This can be done, for example, by having the suffix captured into another perceptual group formed by other redundant stimuli. Thus, despite the presence, position and characteristics of the suffix itself remaining the same—including most notably its phonological characteristics—auditory recency is restored due to the perceptual partitioning of the suffix away from the to-be-remembered sequence (Jones et al., 2006, 2004; Maidment & Macken, 2012; Nicholls & Jones, 2002). Accordingly, the survival of the 'phonological' similarity effect under articulatory suppression with auditory presentation—which is, as noted above, synonymous with the survival of auditory recency (or the modality effect)—is eliminated by a suffix (Jones et al., 2004), but evident again if that suffix is captured into a different perceptual group (Maidment & Macken, 2012).

The fact that, notwithstanding the acoustic-based similarity effect at recency, the 'phonological' similarity effect is eliminated under suppression regardless of modality has two major, inter-related, theoretical implications. First, given that it is the apparent survival of the phonological similarity effect under suppression with auditory presentation that forms the primary basis of the notion of a passive phonological short-term store separate from motor processes, the phonological store construct is now lacking a strong empirical justification. Second, it locates the phonological similarity effect in the motor rehearsal process. This is further supported by findings indicating that the characteristics of the errors in recalling phonologically similar sequences reflect the kind of errors that can be observed occasionally in normal speech production (Acheson & MacDonald, 2009; Ellis, 1980; Page, Madge, Cumming, & Norris, 2007). Speech and verbal short-term memory have thus been shown to use shared processes, with performance in both tasks having a number of parallels across multiple phenomena: For example, phonological similarity has been shown to affect short-term memory and speech production in similar ways, with vowel similarity affecting consonant exchanges, and with the location of the vowels and consonants within the sequence affecting the likely location at which errors occur (Acheson & MacDonald, 2009; Ellis, 1980). Long-term factors such as the lexicality and phonotactic frequency of list items—referring, respectively, to whether the word-status or word-likeness of an item and the frequency with which a phonological segment appears in a given position in words of a particular language—have also been shown to influence both speech production and short-term reproduction in similar ways. All of these findings suggest that verbal serial short-term memory has a motoric basis shared with speech production.

In sum, the 'phonological' similarity effect has now been shown to demonstrate the roles of both motor planning and perceptual organisation in verbal serial short-term memory: When the participant is free to engage in articulatory rehearsal, the effect results from speech-planning errors and hence would be more suitably named the 'articulatory similarity effect'. When the presentation modality is auditory and articulatory planning is impeded, it is an 'acoustic similarity effect'[2]. Thus, the interaction that has been described as "crucial to separating the two components of the articulatory loop, the phonological store and the articulatory control process" (Baddeley, 1989, p. 257) does not adhere to a form that supports a

---

[2] Though 'phonological similarity' is considered a misnomer in this thesis, this term will be generally used throughout in line with decades of literature on verbal short-term memory.

structural separation of an abstract storage system from a motor rehearsal process, or to a system utilising abstract–symbolic representations.

The foregoing findings point instead to an alternative view that explains verbal serial short-term memory (and indeed serial short-term memory more generally) as reliant on motor output-planning functions and in addition, with acoustic input, the perceptual organisation of events into streams. (Hughes et al., 2009, 2016; Jones et al., 2006, 2004). Placing performance in these tasks in "peripheral" processes removes the need for a separate short-term store dealing with abstract-linguistic representations (Jones, Hughes, & Macken, 2007). A particular point of separation between the phonological loop model and the perceptual-motor account is the conceptualisation of the role of articulatory processes. The phonological loop model characterises articulatory processes as subservient to a phonological store, working to counteract a negative effect (item-decay within the store) and converting visual items into a form suitable for the phonological store (Baddeley et al., 1989). In contrast, in the perceptual-motor approach, subvocal articulatory processing, or *motor planning*, functions instead as a 'positive' process, acting as a substrate for short-term sequence retention in itself rather being subservient to a separate entity (e.g., a store) responsible for such retention. That is, the skill of motor planning is co-opted opportunistically due to the fact that it is well-suited—given its inherently sequential, continuous, and rhythmic character—to bind items that are, by design, sequentially unrelated. For example, the paralinguistic act of (covertly) co-articulating verbal items in a serial recall task is thought to support retention as it provides information—not present in the list itself—about which item followed which (Murray & Jones, 2002; Woodward, Macken, & Jones, 2008).

In addition to the evidence relating to phonological similarity, articulatory suppression, modality effects, and their interaction, the perceptual-motor account is supported by a number of other strands of evidence (e.g., Hughes et al., 2009, 2016; Jones et al., 2004; Macken, Taylor, & Jones, 2014; Macken et al., 2016; Woodward et al., 2008). For example, the irrelevant sound effect is better accounted for in terms of perceptual-motor interference than in terms of the disruption of the operation of a passive phonological store (e.g., Hughes & Marsh, 2017; Jones et al., 2004): Sounds that are changing (a key precondition for marked disruption, e.g., Jones, Madden, & Miles, 1992) yield order cues as a by-product of auditory perceptual organisation and this process conflicts with the similar but deliberate process of serially ordering the to-be-remembered items in the form of a motor sequence-plan (regardless of their presentation-modality), thereby impairing serial recall. The key piece of evidence for locating the interference in the phonological store rather than motor rehearsal (Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987) has since been overturned: The survival of the ISE under suppression with auditory (but not visual) to-be-remembered items appears not to have been a true ISE but an artefact of acoustic-level masking of the relevant by the irrelevant. When this artefactual effect is avoided by ensuring that the irrelevant sound sequence is perceptually partitioned from the auditory to-be-remembered items, articulatory suppression eliminates the effect, just as it does with visual to-be-remembered items (Jones et al., 2004; see also Hanley & Hayes, 2012). Thus, motor planning is again implicated as the key process supporting serial recall.

The perceptual-motor account of performance in verbal serial short-term memory tasks is in line with the embodied cognition approach, where short-term ‘memory’ performance is reliant on processes that are not short-term memory–

specific, but are relevant due to specifics of the task, such as the sensory modality through which the information enters. This does not require a simultaneous running of multiple separate modules, some extracting relevant sensory information, some maintaining internal representations and running different manipulations of them, others assigning resources, as supposed by modular models. If the task at hand is the short-term serial recall of auditory-verbal material, the associated processes would be the ones evolved for auditory input organisation and object formation, inseparable from the nature of the incoming information (rather than ones that transform the input into an abstract form), along with the processes responsible for the planning and execution of motor output required to make the appropriate, goal-matching, response in the current environment. The perceptual-motor approach to verbal serial short-term memory is in line with other embodied explanations of short-term memory phenomena outside the verbal context. For example, the notion of short-term memory performance being supported by the application of general-purpose mechanisms according to the requirements of the current goals and environmental conditions has also been successfully applied to non-verbal short-term functions. For example, in the context of the serial recall of hand gestures, effects equivalent to those considered classically to reflect key characteristics of a specialised verbal short-term system—the effects of gesture (cf. phonological) similarity, motor-manual (cf. vocal-articulatory) suppression and gesture- (cf. word-) length—are found despite the non-verbal nature of the task (Wilson & Fox, 2007). The appearance of such effects outside the verbal context indicates that they reflect much more general-purpose processes being utilised for task performance rather than processes uniquely characteristic of a specialised verbal short-term storage system.

1.4 Verbal Sequence Learning: The Function of the Phonological Store?

Since the initial development of the phonological store construct in the context of theorising about verbal short-term memory (e.g., Baddeley, 1986), it has been hypothesised that the structure in fact evolved for language-form learning, or more generally verbal sequence learning, a hypothesis that will be referred to as the *Phonological Store as Language Learning Device* (PS-LLD) hypothesis (cf. Baddeley et al., 1998). This re-specification of the function of the phonological store was motivated by the lack, otherwise, of a clear biological function, as patients with deficits in phonological short-term storage appeared to have little trouble in everyday life (Baddeley, Papagno, & Vallar, 1988). The deficit was, however, accompanied by a difficulty with the long-term retention of new word-forms (Baddeley et al., 1988). However, given the questioning of the empirical basis of the phonological store concept in the domain of verbal short-term memory (e.g., Jones et al., 2006, 2004), a re-evaluation of the phonological-store based account of verbal sequence learning is warranted. As such, the central goal of this thesis is to investigate the mechanisms supporting verbal sequence learning, in particular to examine the extent to which such learning may be a by-product of the same general-purpose perceptual and motor functions that, according to the perceptual-motor account, are exploited to support the reproduction of a verbal sequence over the short term.

Learning skilled behaviours involves the integration of a succession of components of the behaviour, which in themselves are already familiar, into a new order and encoding that new sequence into long-term memory (Melton, 1963). For example, the ability to learn a new word—a basic building-block of language acquisition—involves the unitisation of a novel sequence of already-familiar sounds into a new long-term memory representation (Ellis, 1996). The long-term learning in itself means that the retrieval of a stimulus/event relies more and more on retrieval from long-term memory. The phonological store-based account of this ability posits that the phonological store temporarily retains the unfamiliar phonological form to allow for the creation of a more permanent long-term representation, with the short-term retention of verbal input for short-term recall now considered a by-product rather than the main purpose of the store (Atkins & Baddeley, 1998; Baddeley, 2012; Baddeley et al., 1998). In particular, the quality of the representations in the phonological store determines the accuracy of the new long-term representation. I turn now to discuss the three main strands of evidence for the PS-LLD hypothesis, the first two being of most relevance to the empirical work reported in the present thesis.

1.4.1 Hebb Sequence Learning

In recent years, the most prominent line of evidence for the PS-LLD hypothesis has been based on Hebb sequence learning. An experiment within this paradigm typically involves a serial recall task in which one to-be-recalled sequence (e.g., of letters or words, generally 5-8 items long) is presented repeatedly (e.g., every third sequence) among other, non-repeated ‘filler’ sequences, without informing the participant. Long-term sequence learning is indicated in the performance accuracy pattern, with enhanced recall of the repeated sequence compared recall of the filler sequences (e.g., Hebb, 1961; Melton, 1963; Mosse & Jarrold, 2008; Page & Norris, 2009a, 2009b; St-Louis, Hughes, Saint-Aubin, & Tremblay, 2018; Szmalec et al., 2012; Szmalec, Loncke, Page, & Duyck, 2011; Yanaoka, Nakayama, Jarrold, & Saito, 2018). This Hebb repetition effect has reliably been found to be a good laboratory analogue of word-form learning, both in children learning their first language and for adults learning a new, non-native language (e.g., Norris, Page, & Hall, 2018; Szmalec et al., 2009). The Hebb effect has been shown to correlate with children's performance in more direct word-learning tasks such as the paired-associate learning (Mosse & Jarold, 2008), and has even been shown in very young children (Yanaoka, Nakayama, Jarrold, & Saito, 2018). Performance in the Hebb repetition task produces lexical learning, which has been demonstrated in the slower rejection in an auditory lexical decision task of nonsense syllable sequences that had previously been learned in a Hebb repetition task (Szmalec et al., 2009). Nonwords learned during a Hebb paradigm have also been shown to produce lexical competition with similar real words when tested in a lexical-decision task the following day (Szmalec, Page, & Duyck, 2012). These findings suggest that learning in the Hebb paradigm and that in other verbal learning paradigms relies on the same processes, and that material learned in the Hebb repetition task has a similar lexical representation to that of real words.

Computational models of the phonological loop theory and in particular of the role of the phonological store in verbal sequence learning (and hence word-form learning) have focused on studies of Hebb sequence learning (Burgess & Hitch, 2006; Hitch, Flude, & Burgess, 2009; Page & Norris, 2009a; Szmalec et al., 2012). The model of Burgess and Hitch (2006) explains verbal sequence learning as the strengthening of connections between the item representations and their position in the sequence according to a separate abstract (non-motoric) context/timing signal, which represents the temporal order of the items. Each sequence is recalled through competitive queuing, where the most active candidate item is first selected and its phonemic composition is retrieved, after which that item is inhibited to allow the recall of the next item. The context signal is considered responsible for the repetition-learning of the sequence, with each sequence associated with its own context-set. A second computational model of the phonological loop—the primacy model (Page & Norris, 1998)—has also been applied to Hebb sequence learning (e.g., Page & Norris, 2009b). In the primacy model, serial order is represented by several layers of processing in a primacy gradient. An occurrence layer signals the appearance of a familiar item, followed by a recognition layer for competing units and an order layer for storing the order of the sequence. At the order layer, the primacy gradient works to code serial order, with earlier items receiving higher activation. The primacy gradient representing a particular serial order will build up over learning trials to commit previously uncommitted units to recognising the particular learned sequence (Page & Norris, 2009b). Again, while articulatory processes can refresh item representations, they do not play a direct role in order processing (as in the original phonological loop model; Page & Norris, 2009b).

Importantly, both models posit that Hebb sequence learning, and thus language-form learning, relies on the storage of the individual items in a serial recall task in a domain-specific phonological store and on a domain-general temporal or positional context-signal that represents their order (Burgess & Hitch, 2005; Page & Norris, 2009a; Szmalec et al., 2012). More specifically, the models assume that phonological item-content and item-order are stored separately, and that the order-representation mechanism is unaffected by the phonological properties of the items; the phonological content of an item selected for output at a given position is determined only at a second stage in the recall process. Crucially, while the integrity of item storage is required for the long-term learning of a repeating sequence, it is the order-representation mechanism that drives the learning process. This assumption leads to the prediction that variables that are assumed on these models to impair short-term serial recall specifically via their influence on phonological item-storage (rather than item-order)—such as phonological similarity and articulatory suppression—should not affect long-term verbal sequence learning. And the available evidence does appear to confirm this prediction: Page et al. (2006, Experiment 1) reported that the Hebb effect was not modulated by articulatory suppression, suggesting that motor processing in the form of articulatory response-planning does not play a role in verbal sequence learning. Hitch et al. (2009) also reported that the Hebb effect was not affected by articulatory suppression but also that it was not affected by phonological similarity. Further evidence cited against the notion that motor-related functions play a role in verbal sequence learning comes from studies showing that Hebb sequence learning can occur even when responses to the sequence are not required (Kalm & Norris, 2016; Oberauer & Meyer, 2009, although see Cohen & Johansson, 1967; Cunningham, Healy, & Williams, 1984).

What should indeed affect learning according to the phonological-store based models, then, are factors that are assumed to specifically affect the representation of item-order (Page & Norris, 2009b) or item-position (Burgess & Hitch, 2006). In line with this, Hitch et al. (2009, Experiment 2) found that a repeating list presented with a different temporal grouping on each presentation was learned more poorly than when the temporal grouping was consistent for each presentation (see also Bower & Winzenz, 1969). However, whereas this finding is consistent with the positional context-signal based model of sequence learning (Burgess & Hitch, 2006), it is unclear how the stages of processing in the primacy model can account for the effect of temporal grouping because the order layer exhibits declining activation levels for successive items regardless of how they are grouped (Page & Norris, 2009b).

1.4.2 Nonword Learning in the Paired-Associates Task

A role for the phonological store in language-form acquisition has been studied more directly in the context of paired-associate learning. This paradigm involves the presentation of a list of nonwords (or real words foreign to the participants), each of which is preceded by a known word. After presentation of the list, usually comprising eight pairs, the participants are cued with the known words and required to recall the associate nonword, with a recalled nonword only scored as correct if it is recalled in response to the correct cue. Generally, an increasing number of the nonwords are correctly recalled across repeating trials (or ‘cycles’) with the same list (though the order of the items is randomised on each trial), indicating long-term learning. It is argued that the role of the phonological store in paired associate learning is evident in the observation that such learning is impaired by some of the same factors that impair verbal serial recall (e.g., Papagno et al., 1991). More specifically, articulatory suppression and phonological similarity between the nonwords have both been shown to impair learning in the paired-associate task (Baddeley et al., 1998; Papagno & Vallar, 1995; Papagno et al., 1991). As in serial recall, phonological similarity is suggested to impair paired-associate learning by affecting the phonological store, where the similar nonword representations are more confusable than dissimilar nonwords (Papagno & Vallar, 1992). In particular, phonological similarity impairs the learning of *word-nonword* pairs but not the learning of *word-word* pairs, which is taken to indicate that *word-nonword* associate learning requires the phonological store, while semantic coding can instead be used to support the learning of *word-word* pairs. Articulatory suppression has also been shown to have a greater effect on the learning of *word-nonword* pairs compared to *word-word* pairs (Papagno et al., 1991), again due to the hypothesised use of the phonological store for the learning of the nonwords and the reliance on semantic coding in the case of the *word-word* pairs. The impairment of learning by articulatory suppression is theorised, as it is in the context of serial recall, to be due to articulatory suppression occupying the articulatory control process responsible for giving visual input access to the store and, regardless of input-modality, for counteracting item-decay (Baddeley et al., 1998).

1.4.3 Nonword Repetition and Vocabulary Acquisition

A final line of evidence cited as support for the PS-LLD hypothesis is the positive correlation between measures assumed to reflect the capacity of phonological storage and vocabulary size (Atkins & Baddeley, 1998; Gathercole & Adams, 1993, 1994; Gathercole & Baddeley, 1989; Gathercole, Service, Hitch, Adams, & Martin, 1999; Jarrold, Baddeley, Hewes, Leeke, & Phillips, 2004). In particular, a child’s ability to repeat back a spoken nonword (nonword repetition) correlates positively with vocabulary size (Gathercole, 2006). While the phonological store model recognises that vocabulary size could causally determine nonword repetition ability to some extent (for strong variants of this position, see Bowey, 2001; Melby-Lervåg et al., 2012; Metsala, 1999), it also posits that nonword repetition ability depends on the phonological store and that the capacity of the store in turn causally determines the speed of vocabulary acquisition (Baddeley et al., 1998; Gathercole, 2006). The assumption that nonword repetition ability depends on the phonological store is based on evidence that nonword repetition performance mimics the characteristics of verbal serial recall, such as a U-shaped serial position curve and the fact that longer items are more poorly recalled (e.g., Baddeley et al., 1998; Gupta, 2005).

1.4.4 On the Coherence of the PS-LLD Hypothesis

Each of the three strands of literature on verbal sequence learning would appear to provide good evidence for the PS-LLD hypothesis. However, one can raise questions both about the coherence of some of the lines of evidence in relation to one another and about the coherence of the hypothesis in relation to some of the evidence cited in support of a separate (phonological) short-term store in the first place. The evidence for the PS-LLD hypothesis based on the nonword repetition paradigm and nonword learning in the paired-associates task do converge with one another: The phonological store is implicated in both cases on the basis that both tasks exhibit key signatures of the operation of the store in the context of verbal serial recall (e.g., disruptive effects of articulatory suppression, item-length, phonological similarity). However, this directly contrasts with the evidence for the hypothesis drawn from the Hebb learning effect: Here, factors that are assumed to affect the phonological store are predicted *not* to affect learning because learning is driven by mechanisms that can be linked to the phonological store (when the input happens to be verbal) but which are not an integral part of it (Burgess & Hitch, 2006; Page et al., 2009). The comparison of the line of reasoning based on the paired-associate paradigm and that based on the Hebb repetition paradigm is of particular interest as the contradiction seems particularly sharp here: While the phonological store-based models of Hebb sequence learning predict that articulatory suppression and phonological similarity should not affect sequence learning, in the paired-associate learning paradigm, it is the *presence* of these very same effects that is taken as support for the PS-LLD hypothesis (e.g., Papagno et al., 1991).

In addition, some of the evidence for the PS-LLD hypothesis appears to be at odds with the case from differential coding for the fractionation of a (phonological) short-term store from a long-term memory system (e.g., Baddeley, 1996). Specifically, the finding that nonword *learning*—that is, an increasing reliance on long-term memory (as opposed to short-term memory) to recall a nonword—exhibits a phonological similarity effect undermines such a fractionation because that fractionation is based on the notion that only recall from the phonological short-term store itself should exhibit a phonological similarity effect. The phonological store-based models based on the Hebb effect, however, insofar as they separate the phonological store from the long-term sequence learning mechanism, appear to better fit the basic assumption at the core of the phonological store concept of a separation of long-term and short-term memory structures. But this further brings into relief the apparent inconsistency in the reasoning based on the various strands of evidence for the PS-LLD hypothesis.

1.5 The Present Empirical Work

In light of the challenge to the empirical basis of the phonological store concept in the context of verbal serial short-term memory phenomena (e.g., Jones et al., 2004) and questions about the coherence of the PS-LLD hypothesis, the empirical work reported in this thesis had two interrelated aims: to re-evaluate the PS-LLD hypothesis and to examine the extent to which verbal sequence learning may instead be understood as the by-product of the same general-purpose processes—particularly motor planning but also (auditory) perceptual organisation—that have supplanted phonological-store based constructs in the perceptual-motor account of verbal serial short-term memory (Hughes et al., 2016; Jones et al., 2004). These aims were addressed in the context of both Hebb sequence learning (cf. Chapter 2) and nonword learning in the paired-associate task (cf. Chapter 3). The evidence for the PS-LLD hypothesis based on the nonword repetition/vocabulary paradigm will also be revisited in light of the empirical studies in the General Discussion chapter (Chapter 4).

1.5.1 Preview of Chapter 2 (Experiments 1-3)

Regarding Hebb sequence learning, some of the features of the studies conducted to test phonological-store based models raise questions about the degree to which the results do in fact support the PS-LLD hypothesis. For instance, in some cases, unusually long sequences of 12 items were used (e.g., Hitch et al., 2009, Experiments 1 and 2), which is problematic for an account based on the phonological store concept since it has been claimed repeatedly that the store is considered to be over-burdened, and hence abandoned, when the sequence is longer than around seven items (Baddeley, 2000; Baddeley & Larsen, 2003, 2007; Larsen & Baddeley, 2003). It is also notable that in several cases, the Hebb and filler lists contained at least partially non-overlapping sets of items (e.g., Hitch et al., 2009, Experiment 3; Page et al., 2006, Experiment 4) while the classic paradigm typically involves having all sequences comprise the same closed set of items. It is possible, therefore, that the repetition effect in those experiments reflected in part the learning of which items were a part of the repeating set (i.e., item-set learning) rather than sequence learning per se. It could then be argued that the lack of articulatory suppression effect on learning may not, in fact, have constituted reliable evidence for a separation of a sequence learning mechanism from articulatory motor-planning functions.

On the face of it, the absence of a phonological similarity effect on Hebb sequence learning (Hitch et al., 2009) is also problematic for a perceptual-motor account. This is because, on this account, the phonological similarity effect primarily reflects motor-planning errors (Jones et al., 2004). From this standpoint, one would expect some kind of effect of phonological similarity on learning. For instance, it might be expected that a non-fluent motor-plan (due to phonological similarity between its elements) would stand to gain more from repeated opportunities to correct the errors therein and hence result in a larger learning effect than when the motor-plan is already relatively fluent to begin with (i.e., with phonologically dissimilar items) (e.g., Newell & Roosenbloom, 1981). However, the possibility that Hebb sequence learning exhibits a phonological similarity effect has only been examined in one previous experiment, one in which, again, the attempt to study sequence learning per se may have been contaminated by an item-set learning effect (Hitch et al., 2009, Experiment 3).

In the current Experiment 1, then, I tested the prediction of the perceptual-motor account that there should indeed be effects of articulatory suppression and phonological similarity on Hebb sequence learning using the standard Hebb repetition procedure in which filler and repeated sequences contain the exact same items and using relatively short lists (seven items). Finding such effects would also clearly go against the predictions of the phonological-store based models of Hebb sequence learning.

In Experiment 2, a role for motor planning in Hebb sequence learning was examined further by testing whether a change in the temporal grouping of the items within the repeating sequence across successive instances of that sequence attenuates learning. It was presumed that such inconsistent grouping within the list itself would lead to inconsistently structured motor-plans across repetitions, resulting in poorer learning. While such a grouping effect on learning has been demonstrated previously (Hitch et al., 2009), a motoric locus for the effect has been rejected on the grounds that it is not modulated by articulatory suppression. Again, however, interpretation of the Hitch et al. (2009) experiment is complicated by its use of twelve-item lists.

The notion that the irrelevance of overt recall to the Hebb effect (e.g., Kalm & Norris, 2016; Oberauer & Meyer, 2009) goes against a role for motor processing in general in verbal sequence learning also seems premature. In particular, these experiments do not rule out a role for pre-production covert motor-planning of the sequence. This is because participants were only informed about whether or not they were required to recall a particular sequence after it had already been presented. It is likely that a motor-plan for a recall response would have been formulated during presentation, due to the possibility that such a response would be required. As a result, the condition requiring a response may not have, in effect, differed from the one not requiring a recall response. A need for a re-evaluation of the role of motor processes in the Hebb repetition effect is also indicated by recent findings showing that overt recall enhances the Hebb repetition effect, suggesting that either motor production or planning (or both) do indeed play some role: Poorer learning is observed when the motor sequence to be produced in response to each presentation of the repeating sequence is changed between the end of presentation and the start of the recall test (from forward to backward recall, but no effect was observed when recall was changed from backward to forward; Guerrette, Saint-Aubin, Richard, & Guérard, 2018). The present Experiment 3, then, included a manipulation whereby participants knew before sequence-presentation that they would either have to recall the sequence—and hence, presumably, construct a motor-plan—or would not have to recall (nor therefore plan to produce) the sequence but simply passively monitor it (cf. Glass, Krejci, & Goldman, 1988). It was predicted that Hebb sequence learning would be diminished in the absence of a need to assemble a motor-sequence plan.

In Experiments 1 and 2, the modality of sequence-presentation (visual vs. auditory) was also manipulated. This was done in order to examine whether passive perceptual organisation processes that bypass motor processes (cf. Jones et al., 2004) may, in addition to motor planning processes, support verbal sequence learning when sequence-presentation is auditory. For example, if this is the case, impeding the contribution of motor planning to learning through articulatory suppression or through introducing variable temporal grouping across sequence repetitions may have less effect with auditory compared to visual presentation. A role for auditory perceptual organisation was also examined more directly in Experiment 3 by including a condition in which the (auditory) sequences were presented in alternating female-male voices and hence likely to be partitioned into two by-voice perceptual streams. A role for perceptual organisation would be indicated if it were found that the sub-sequences made up of every other item in a sequence (e.g., **A, C, E, G** and B, D, F from the sequence **A,** B, **C,** D, **E,** F, **G**) could be learned with alternating-voice sequences but not single-voice sequences.

1.5.2 Preview of Chapter 3 (Experiments 4-6)

The case for the PS-LLD hypothesis based on paired-associate *word-nonword* learning has been reliant on the finding that such learning exhibits similar effects to those observed in verbal serial recall, such as a detrimental effect of phonological similarity between the nonwords (e.g., Papagno & Vallar, 1992). However, as noted, an effect of phonological similarity on retrieval from long-term memory contradicts the case from differential coding for a separate *short-term* phonological store (e.g., Baddeley, 1966, 1996). Experiment 4 will therefore examine whether the phonological similarity effect on nonword-word learning is a general cue-based interference effect due to having to match each nonword with the correct cue (i.e., the word associate) rather than affecting learning of the new word-form. If this is the case, this would resolve the contradiction between the PS-LLD hypothesis and the differential coding argument but the phonological similarity effect in paired-associate learning could no longer be taken as support for the PS-LLD hypothesis. This possibility was investigated in Experiment 4 by comparing nonword learning in the classical cue-based word-nonword associates task with nonword learning in free recall task where no known-word associates were presented. If phonological coding is restricted to the short-term store and the phonological similarity effect in paired-associate learning is a specious one (i.e., general interference effect), phonological similarity should impair learning in the paired-associate task, but not in the free recall task. The possibility that the similarity effect in paired-associate learning is a product of its particular methodology gains further credence from the fact that phonological similarity has only ever been manipulated *across* the nonwords in this setting (i.e. the nonwords were similar to each other; Papagno & Vallar, 1992). Thus, Experiment 4 also investigated for the first time whether similarity across the syllables within the nonwords also attenuates nonword learning. Such an effect would more clearly indicate disruption of internal word-form learning rather than list-wide general interference. Finding a phonological similarity effect in nonword learning regardless of the particular recall-task and regardless of the particular manipulation of similarity would be consistent with the phonological-store account but, to the extent that the phonological similarity effect can be reascribed to motor-planning errors (cf. Jones et al., 2004), also consistent with the view that verbal sequence learning is supported in part by motor planning. A secondary interest in Experiment 4 was in how speaking more than one language enhances word-form recall and learning (Papagno & Vallar, 1995) and in particular whether the different manipulations of phonological similarity implemented in this experiment may be informative as to the mechanism underpinning that advantage.

The role of motor planning in nonword learning was investigated more directly in Experiment 5 through a manipulation of articulatory suppression. On the PS-LLD hypothesis, articulatory suppression has been theorised to impair nonword learning by restricting the articulatory rehearsal process, which is responsible for the refreshing of decay-prone representations in the phonological store (Papagno & Vallar, 1992). However, this effect could, on the perceptual-motor account, reflect instead the contribution of constructive motor planning to verbal sequence learning. Experiment 5 also involved crossing an articulatory suppression manipulation with that of modality of presentation (visual vs. auditory) with a view to examining whether passive perceptual organisation or phonological storage support word-form learning with auditory sequences when motor processes are impeded.

Finally, Experiment 6 replicated the design of Experiment 4 but also included a brain-imaging component carried out during a research visit to the University of Gävle, Sweden. In particular, using functional near infrared spectroscopy (fNIRS), I sought to examine possible changes in activation during nonword learning in the supplementary and pre-supplementary motor areas of the brain, in light of their established involvement in language production, gesture-learning and motor sequencing (e.g., Ackermann & Riecker, 2010; Hatakenaka et al., 2007; Hirano et al., 1996; Koziol et al., 2014). Of particular interest was whether activation in these motor areas would differ according to phonological similarity condition, in line with the contention that the phonological similarity effect has a primarily motoric basis (Hertrich, Dietrich, & Ackermann, 2016).

CHAPTER 2

HEBB SEQUENCE LEARNING: A PERCEPTUAL-MOTOR APPROACH

Abstract

The three experiments in this chapter test the idea that Hebb verbal sequence learning—in which short-term serial recall is enhanced for a repeated sequence in amongst otherwise non-repeated sequences—reflects not the action of a phonological short-term store but rather the legacy of motoric and perceptual processing, engaged in producing short-term goal-relevant behaviours. The perceptual-motor account was supported by several convergent findings: First, Hebb sequence learning was either attenuated (Experiment 1) or abolished (Experiment 2) when vocal-motor planning of the to-be-remembered sequence was restricted by articulatory suppression. This was less the case with auditory sequences, however, suggesting that passive auditory perceptual organisation processes that bypass motor planning can independently support auditory Hebb sequence learning (Experiments 1 and 2). Second, Hebb sequence learning was also absent when no serial recall response—and hence no motor planning—was required (Experiment 3). Third, Hebb sequence learning was enhanced for phonologically similar compared to dissimilar items when that learning was driven solely by motor planning (i.e., with visual sequences). That this enhanced learning was eliminated when motor planning was restricted also points to a motoric basis for this ‘phonological’ similarity effect (Experiment 1). Fourth, an inconsistent temporal grouping of items across instances of the repeating sequence also abolished learning but only when that grouping—based on independent evidence from output response-times during serial recall—was instantiated within a motor-plan (Experiment 2). Fifth, demonstrating more direct evidence for a contribution of

passive perceptual organisation in learning an auditory-verbal sequence, promoting the perceptual grouping of every-other-item in the repeating sequence by presenting it in alternating male and female voices led to the learning of those non-adjacent-item sub-sequences. The results support an embodied view in which Hebb sequence learning is a by-product of motoric and perceptual processes evolved for immediate interaction with the environment and challenge the currently dominant phonological-store based account.

2.1 Introduction

Essential to language acquisition, particularly word-form learning, is the capacity to learn familiar verbal elements (phonemes, syllables) encountered in a new order. The currently dominant account of such verbal sequence learning posits that it is supported by, and indeed is the evolved function of, a dedicated phonological short-term store (e.g., Baddeley, 2012). The phonological store construct is predicated primarily on a three-way interaction between phonological similarity, modality and articulatory suppression. Specifically, the phonological similarity effect is found despite the incapacitation of the articulatory control component of the phonological loop, thereby indicating, on this model, that there must be a passive store to which auditory (but not visual) input gains obligatory access and which is separate from articulatory processes (Baddeley et al., 1984). However, further scrutiny of this interaction—particularly relating to serial position—has not supported this theoretical interpretation. It transpires that articulatory rehearsal is indeed a precondition for the phonological similarity effect regardless of modality: It is eliminated by articulatory suppression even with auditory presentation throughout most of the serial position curve (Jones et al., 2004). The survival of the effect under suppression is restricted primarily to recency and hence driven by the modality effect (Jones et al., 2006, 2004; Maidment & Macken, 2012), an effect deemed to be “peripheral to the working memory system” (Baddeley, 1986, p. 95). Instead, such data have suggested that serial short-term memory performance can be conceptualised more parsimoniously as being parasitic on motor planning (regardless of presentation-modality; Wolpert et al., 2001) and processes involved in the perceptual organisation of acoustic input (when material is presented auditorily; Bregman, 1990; Oxenham, 2018; Hughes & Marsh, 2017;

Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016). In this view, the motor-plan functions as a substrate for short-term sequence retention in itself rather being subservient to a separate entity (e.g., a store) responsible for such retention.

Given that the principal evidence for the existence of a phonological store has not held up under closer empirical scrutiny, the question at the centre of this thesis is what, therefore, fulfils the function—verbal sequence learning—for which that store is said to have evolved (Baddeley et al., 1998)? The hypothesis tested in the present thesis is that such learning reflects the legacy of the same general-purpose processes—motor planning and also, when the sequences are presented in spoken form, auditory perceptual organisation—that have supplanted phonological store-based constructs within the perceptual-motor account of verbal serial short-term memory. This chapter focuses on addressing this question in the context of Hebb sequence learning (Hebb, 1961), which has been the primary focus of attention in recent years in the development of the phonological-store based account of verbal sequence learning (e.g., Burgess & Hitch, 1999, 2006; Hitch, Flude, & Burgess, 2009; Page & Norris, 2009a; Szmalec, Page, & Duyck, 2012). The paradigm involves a serial recall task in which one sequence (e.g., 5-8 letters, digits, or words) is repeated several times (e.g., on every third trial) among otherwise novel, ‘filler’, sequences. Long-term verbal sequence learning is indicated by the enhanced short-term serial recall of the repeating sequence compared to filler sequences (e.g., Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015; Hebb, 1961; Stadler, 1993; St-Louis, Hughes, Saint-Aubin, & Tremblay, 2018; Yanaoka, Nakayama, Jarrold, & Saito, 2018).

Phonological-store based models posit that Hebb sequence learning relies on the storage of the items in the phonological store and on a domain-general and non-motoric temporal or positional context-signal that represents their order (Burgess & Hitch, 2005; Page & Norris, 2009a; Szmalec et al., 2012). A major assumption of these models is that the representation of serial order information is separate from a second stage of processing at which the representation of the phonological identities of items residing in the phonological store are linked to their serial or ordinal positions. Crucially, it is the first, order-representation, stage that drives the long-term learning of a repeating sequence. This key assumption leads to the prediction that variables that are, on these models, assumed to impair short-term serial recall via their influence on phonological item-storage/retrieval (rather than item-order)—such as phonological similarity and articulatory suppression—should not affect long-term verbal sequence learning (e.g., Burgess & Hitch, 2006; Page & Norris, 2009). And indeed, the available evidence appears to confirm this prediction (Hitch et al., 2009; Page et al., 2006).

The present experiments examined whether Hebb sequence learning, rather than being supported by a bespoke phonological store, is driven by the same general-purpose motor and perceptual processes that have been implicated in the perceptual-motor account of verbal serial short-term memory performance (Jones et al., 2006, 2004). That is, it is hypothesised that Hebb sequence learning (and verbal sequence learning in general) reflects the by-product of processes engaged during the attempt to meet short-term goals, not separate long-term learning processes linked to a short-term store. In particular, when successful behaviour requires the formation of an articulatory motor plan of a verbal sequence to produce the goal-relevant short-term response, learning of that motor sequence can occur as a consequence of the short-term task goal. In Experiment 1, Hebb sequence learning was examined for the first time in the context of the intricate interplay of factors—phonological similarity, articulatory suppression, and presentation modality—that has been instrumental in the emergence of this alternative construal of serial recall. By examining the impact of these interacting factors on short-term serial recall and, simultaneously, on the enhancement of the serial recall of a repeating Hebb sequence, this experiment aimed to reveal the contributions of motor planning and auditory perceptual organisation to long-term verbal sequence learning. Experiment 2 had the same general aim but involved examining the impact on verbal sequence learning of a more direct manipulation of the temporal organisation of the motoric and auditory-perceptual representation of a sequence. Experiment 3 sought convergent evidence for the role of motor planning by manipulating the likelihood that a motor-plan would need to be assembled at all. Experiment 3 also involved a manipulation of perceptual variability within an auditory sequence to examine more directly the contribution of perceptual organisation to Hebb sequence learning.

2.2 Experiment 1

In this experiment, the role of motor planning in Hebb sequence learning was investigated in two ways. First, if verbal sequence learning reflects in large part the by-product of repeated short-term motor-planning of the same sequence, then restricting motor planning via articulatory suppression should attenuate such learning, at least with visual sequences when any contribution of passive auditory perceptual organisation to Hebb sequence learning can be ruled out. As noted earlier, this strong prediction of the perceptual-motor account seems to have already been disconfirmed (Hitch et al., 2009; Page et al., 2006). However, to anticipate, the data from Experiments 1 and 2 reported here suggest that the conclusion that Hebb sequence learning is immune to articulatory suppression was premature.

The second way in which the role of motor planning was examined was through a manipulation of phonological similarity as well as the interaction of phonological similarity with articulatory suppression and modality. Given the evidence that the phonological similarity effect is, notwithstanding the acoustic-based effect at recency, an articulatory similarity effect that reflects speech-planning errors (Acheson & MacDonald, 2009; Jones et al., 2004; Page et al., 2007), a role for motor planning in verbal sequence learning may also be revealed in the form of a modulation of Hebb sequence learning by phonological similarity. Surprisingly, only one previous experiment has addressed this possibility: Based on their phonological-store based model of verbal short-term memory and sequence learning, Hitch et al. (2009, Experiment 3) tested and seemed to confirm the “counterintuitive prediction that phonemic similarity should not impair sequence learning, despite having its normal effect of disrupting STM [short-term memory] for serial order” (p. 106). However, the design of this study was rather atypical in that there were two Hebb sequences in a single block drawn from two different item-sets, one in which the items were phonologically similar to each other and another in which they were phonologically dissimilar, while the filler sequences comprised items drawn equally from both sets. This means that at least part of the ‘Hebb effect’ observed in this experiment may have been driven by the learning of which items were likely to be presented in a given list (i.e., item-set learning), not (or not only) the learning of a sequence per se. Such item-set learning would not necessarily be expected to be hindered by phonological similarity (e.g., Fallon, Groves, & Tehan, 1999; Nairne & Kelley, 2004). Furthermore, it is noteworthy that the only hypothesis entertained in previous work is that phonological similarity could be expected to *impair* sequence learning (Hitch et al., 2009), presumably on the grounds that it impairs short-term serial recall. But this is not the only possible—or indeed necessarily most plausible—hypothesis. On a motor-planning based account, there are good reasons to expect a greater learning effect for phonologically similar compared to dissimilar sequences; that is, the recall of phonologically similar sequences may benefit more from repeated opportunities to plan that sequence than is the case for a phonologically dissimilar sequence: The less fluent a motor-skill is to begin with, the more that skill stands to benefit from practice (e.g., Heathcote, Brown, & Mewhort, 2000; Newell & Roosenbloom, 1981). In addition, it was predicted that if the greater learning effect with a phonologically similar sequence is indeed located in the motor-planning process, the enhancement should be attenuated or eliminated under articulatory suppression, at least with visual sequences in which passive auditory perceptual organization could not support any learning.

As well as investigating the role of motor planning in verbal sequence learning—which should be evident regardless of presentation modality—this experiment also examined the possible additional contribution of passive auditory perceptual organisation processes to the learning of an auditorily presented sequence. Such a contribution should be evident in differences in the Hebb repetition effect according to the modality of presentation (i.e., auditory as opposed to visual), at least under articulatory suppression when the contribution of motor planning—common to both modalities—would be reduced. Specifically, it was predicted that although learning should still be diminished with auditory sequences when motor planning is restricted by articulatory suppression, this diminution should not be as marked with such sequences (compared to visual) due to the independent contribution to (auditory) sequence learning of passive auditory perceptual organisation processes that by-pass motor-planning processes (Jones et al., 2004).

In sum, then, in this experiment participants were required to serially recall sequences of seven letter-names that were either phonologically similar or dissimilar to one another and to do so while being free to engage in vocal-motor planning or whilst engaging in articulatory suppression. Moreover, the sequences were presented either visually or auditorily.

2.2.1 Methods

2.2.1.1 Participants. The Hebb effect was identified as being, typically, a medium-sized effect (estimated Cohen's *d* ranging between ~ .4 and ~ .7; Bogaerts et al., 2015; Hitch et al., 2009; Page et al., 2006) but given the large number of factors in the current experiments and the interest in their interactions, the appropriate sample size was determined to allow the potential detection of small- to medium sized effects (e.g., Cohen's d ~ .3) with a relatively large amount of power. A sample size of 52 was calculated to allow this with a power of .9, and therefore 52 participants were recruited for both the present experiment and Experiment 3, and 52 participants were recruited for each of the two between-participants groups in Experiment 2. The participants in this experiment consisted of six males and 46 females, all students at Royal Holloway University of London (mean age: 19.17 years, SD = 1.63). They received either course credits or a small honorarium for their participation. Two participants were excluded from data analysis due to not completing all conditions.

2.2.1.2 Apparatus and Materials. The experiment was conducted using E-Prime software (Psychology Software Tools, Pittsburgh, PA) running on a PC. The visual stimuli were presented on a flat monitor and the auditory stimuli via headphones. For the with-suppression condition, a microphone was used for on-line monitoring of each participant’s compliance with the whispered articulatory suppression instruction (see below). The to-be-remembered sequences consisted of a random ordering of either the phonologically similar letters *B, C, D, G, P, T,* and *V,* or the phonologically dissimilar letters *F, H, K, L, Q, R,* and *Y* and these could be presented either visually or auditorily. Regardless of input-modality, the letters were presented for 250 ms with an interstimulus interval of 750 ms. The auditorily presented letters were recorded in a female voice at a pitch corresponding to a fundamental frequency of approximately 210 Hz, sampled with a 16-bit resolution at a rate of 48kHz, and compressed to 250 ms (without altering pitch) with Sonic Forge 5.0 software (Sonic Foundry, Inc., Madison, WI; 2000). The visually-presented letters were presented in a 72-point Times Roman font in the centre of the monitor.

2.2.1.3 Design. The experiment involved five repeated-measures factors in all: Modality (visual, auditory), Articulatory suppression (no-suppression, with-suppression), Phonological similarity (similar, dissimilar), List-type (Hebb, Filler), and Cycle (referring to each successive triplet of trials across a block, each triplet comprising a Hebb sequence and two preceding filler trials). There were eight blocks of serial recall trials in total, each consisting of 36 sequences of seven letters where every third sequence (starting with trial 3) was the same, repeating, Hebb sequence, amounting to 12 instances of the Hebb sequence within a given block. The experiment was divided into two order-counterbalanced blocks according to Modality and these were undertaken on different days. Each modality block/session was itself sub-divided into four 36-trial blocks; two of these blocks comprised all phonologically-similar sequences and two comprised all phonologically-dissimilar sequences. Finally, in one block in each phonological similarity condition (i.e., similar and dissimilar), participants engaged in articulatory suppression (with-suppression blocks) whilst in the other block in each phonological similarity condition they did not (no-suppression blocks). The four blocks [2(Phonological similarity) × 2(Articulatory suppression)] within each modality were presented in a randomised order, with a different order in each modality for each participant. There were four possible block-orders: A1-B2-A2-B1, A2-B1-A1-B2, B1-A2-B2-A1 or B2-A1-B1-A2, where A = similar, B = dissimilar, 1 = no-suppression, and 2 = with-suppression. For each of the eight blocks, the participant undertook one of two possible sets of trials, which had different sequences with different item-orders for both the Hebb sequences and the filler sequences.

2.2.1.4 Procedure. The experiment was divided into two sessions held between 1 and 14 days apart, with a randomly assigned half of the participants completing the four visual blocks in the first session followed by the four auditory blocks in the second and vice versa for the other half of participants. Participants were tested individually and wore the headphones throughout both sessions (even though only in the auditory modality condition was sound presented). At the beginning of the first session, participants gave informed consent and were then given task instructions. These included a description of the immediate serial recall task, the four-block structure of the session and the articulatory suppression that would be required in a sub-set of the blocks. The articulatory suppression involved repeatedly whispering ‘8, 9, 10’ at a rate of approximately three items per s during both the presentation of the letters and during the recall attempt (note that the recall mode was manual; see below). The Experimenter demonstrated the approximate (whispered) form and rate of articulatory suppression and participants then practiced the suppression before any serial recall trials. With the permission of the participant, compliance with these articulatory suppression instructions was monitored ‘live’ by the Experimenter throughout the experimental trials through an audio link. Before each block, the participants were instructed on whether or not they had to undertake articulatory suppression and received two practice trials that corresponded to the nature of the trials in the upcoming block. In the experiment, each presented sequence was followed immediately by a serial recall cue in which the participants clicked the letters from a circular array presented on the monitor in the order they saw/heard them. Importantly, the order of the letters in the circular response array was randomised anew for each trial, including the Hebb trials. This means that learning the repeating sequence could not be based on a repeating spatial sequence of clicks or on the planning or production of a repeating sequence of finger-movements (cf. Fendrich, Healy, & Bourne, 1991; Page et al., 2006). After recall, participants moved to the next sequence by clicking an icon to start the new sequence. Neither the particular phonological similarity condition nor the Hebb repetition manipulation was mentioned to participants at any point until the debriefing following the last block of the second session. Each of the two sessions lasted approximately 1 hr.

2.2.2 Results

2.2.2.1 Serial recall. Serial recall performance per se, as opposed to Hebb sequence learning, was examined first. The data for this analysis were, for each of the eight [2(Phonological similarity) × 2(Modality) × 2(Articulatory suppression)] blocks, those from the 24 filler trials and the first instance of the Hebb sequence (which in effect was equivalent to a filler sequence as it would not have been presented previously at that point). For each sequence, an item was scored as correct only when recalled in the same absolute position as that in which it was presented.

Figure 2.1 shows the percentage of items recalled correctly at each serial position in each of the eight conditions. In the absence of articulatory suppression (left panel), a clear phonological similarity effect is evident for both visual and auditory sequences. Under suppression (right panel), however, the phonological similarity effect is eliminated with visual sequences but remains with auditory sequences. Critically, however, replicating previous studies (Jones et al., 2004; Maidment & Macken, 2012), it can be seen that this survival of the phonological similarity effect with auditory presentation under suppression is located primarily at recency.

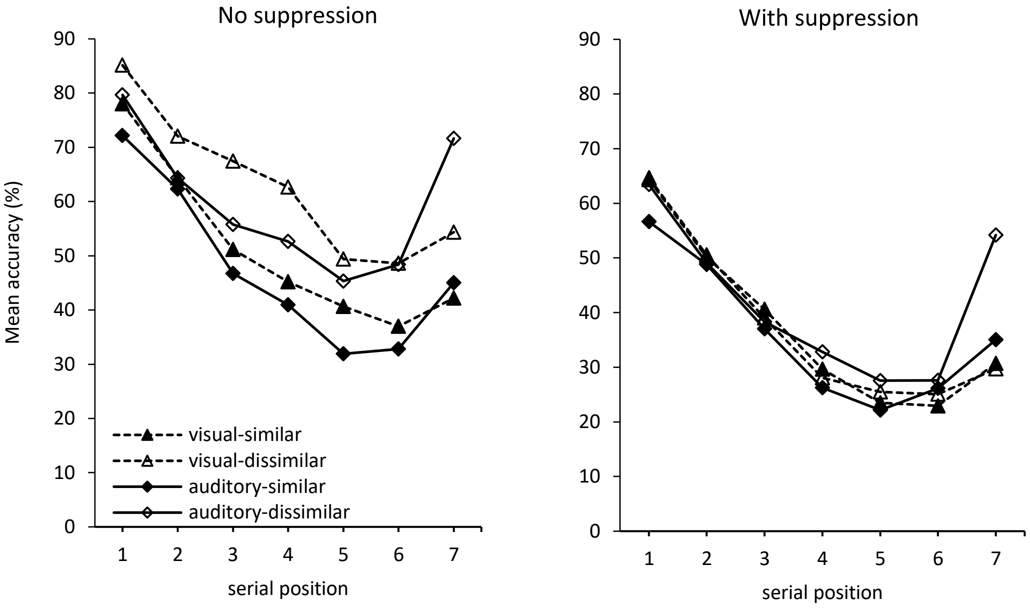


Figure 2.1. Accuracy of serial recall performance for filler sequences in the eight conditions of Experiment 1 according to serial position.

In line with this impression of the data, a 2 (Modality) × 2 (Phonological similarity) × 2 (Articulatory suppression) × 7 (Serial position) repeated measures ANOVA showed main effects of Phonological similarity, $F(1, 49) = 42.9$, $MSE = .09$, $p < .001$, $\eta_p^2 = .47$, Articulatory suppression, $F(1, 49) = 176.12$, $MSE = .12$, $p < .001$, $\eta_p{}^2 = .78$, and Serial position, $F(6, 294) = 231.18$, $MSE = .03$, $p < .001$, $\eta_p{}^2 =$

.83. No main effect of Modality was observed, $F(1, 49) = 1.20$, $MSE = .07$, $p = .28$, $\eta_p^{\ 2} = .02$. Importantly, however, a significant four-way interaction was found, $F(6, 294) = 3.13$, $MSE = .01$, $p = .005$, $\eta_p^{\ 2} = .06$, in line with the observation based on Figure 2.1: While the phonological similarity effect survived suppression only with auditory sequences, this was primarily the case at recency. (For completeness, other significant interactions subsumed within this four-way interaction are included in Supplementary Table 1 in Appendix 1, which provides the full set of results from the analyses of Experiment 1.)

Thus, the pattern of serial recall performance replicates closely that which has formed the empirical basis of the argument that such performance can be explained solely by recourse to auditory perceptual organisation and motor-planning processes (e.g., Jones et al., 2006, 2007; Macken et al., 2016; Maidment & Macken, 2012). As such, the serial recall data provide a strong platform from which to examine the role that these same processes may play in the long-term learning of a verbal sequence.

2.2.2.2 Hebb sequence learning. The analysis of Hebb sequence learning involved, for each of the eight [2(Phonological similarity) × 2(Modality) × 2(Articulatory suppression)] blocks/conditions, the serial recall data from the twelve Hebb sequences and the average recall of each pair of filler sequences that preceded each instance of the Hebb sequence (hereafter: ‘fillers’). For the purpose of this analysis, performance accuracy for each list was collapsed over serial positions. These data, shown in Figure 2.2, were entered into a 2 (List-type: Hebb vs. Filler) by 2 (Modality) by 2 (Phonological similarity) by 2 (Articulatory suppression) by 12 (Cycle) repeated measures ANOVA. First, this analysis revealed several main effects that are not reported here in detail—those of Phonological similarity, Articulatory suppression and Modality—because they simply reflect the same effects as already reported in the previous sub-section on serial recall performance per se (rather than pertaining specifically to sequence *learning*). Turning now to effects that are indeed relevant to the assessment of Hebb sequence learning, the main effect of List-type was significant, $F(1, 49) = 59.25$, $MSE = .19$, $p < .001$, $\eta_p{}^2 = .547$, reflecting the better recall of Hebb sequences compared to fillers (i.e., the classic Hebb effect), as was the List-type by Cycle interaction, $F(11, 539) = 7.64$, $MSE = .04$, $p < .001$, $\eta_p{}^2 = .135$, which likely reflects, primarily, the fact that the benefit of repetition increases as a function of the number of repetitions (i.e., that Hebb sequence learning is progressive, at least across eleven repetitions of the Hebb sequence as was the case in the present experiment). The main effect of Cycle was also significant, $F(7, 389.8) = 9.48$, $MSE = .06$, $p < .001$, $\eta_p{}^2 = .162$, which also likely reflects the increasingly beneficial effect of Hebb repetition across a block.

There were several reliable interactions that, like some of the reliable main effects, reflect patterns in the serial recall data per se that have already been reported and that do not relate to the Hebb effect (the full set of results is, however, reported in Supplementary Table 2 of Appendix 1). However, there were also several significant interactions that do indeed reflect a modulation of Hebb sequence learning by one or more of the other factors: Of particular interest was a reliable interaction between List-type and Articulatory suppression, $F(1,49) = 7.8$, $MSE = .15$, $p = .007$, $\eta_p{}^2 = .14$, whereby the advantage in the recall of the repeating sequence was attenuated under articulatory suppression. In addition, while the List-type by Phonological similarity interaction was not significant, this was because these two factors entered into a reliable three-way interaction with Modality, $F(1, 49) = 5.06$, $MSE = .15$, $p = .029$, $\eta_p{}^2 = .09$, as well as a reliable four-way interaction with Modality and Articulatory suppression, $F(1, 49) = 4.6$, $MSE = .06$, $p = .037$, $\eta_p{}^2 =$

.086. To aid in the interpretation of this complex interaction, Figure 2.2 is supplemented with Table 1, which shows the results of the critical Hebb *vs*. Filler pairwise contrast as a function of modality, similarity, and suppression.

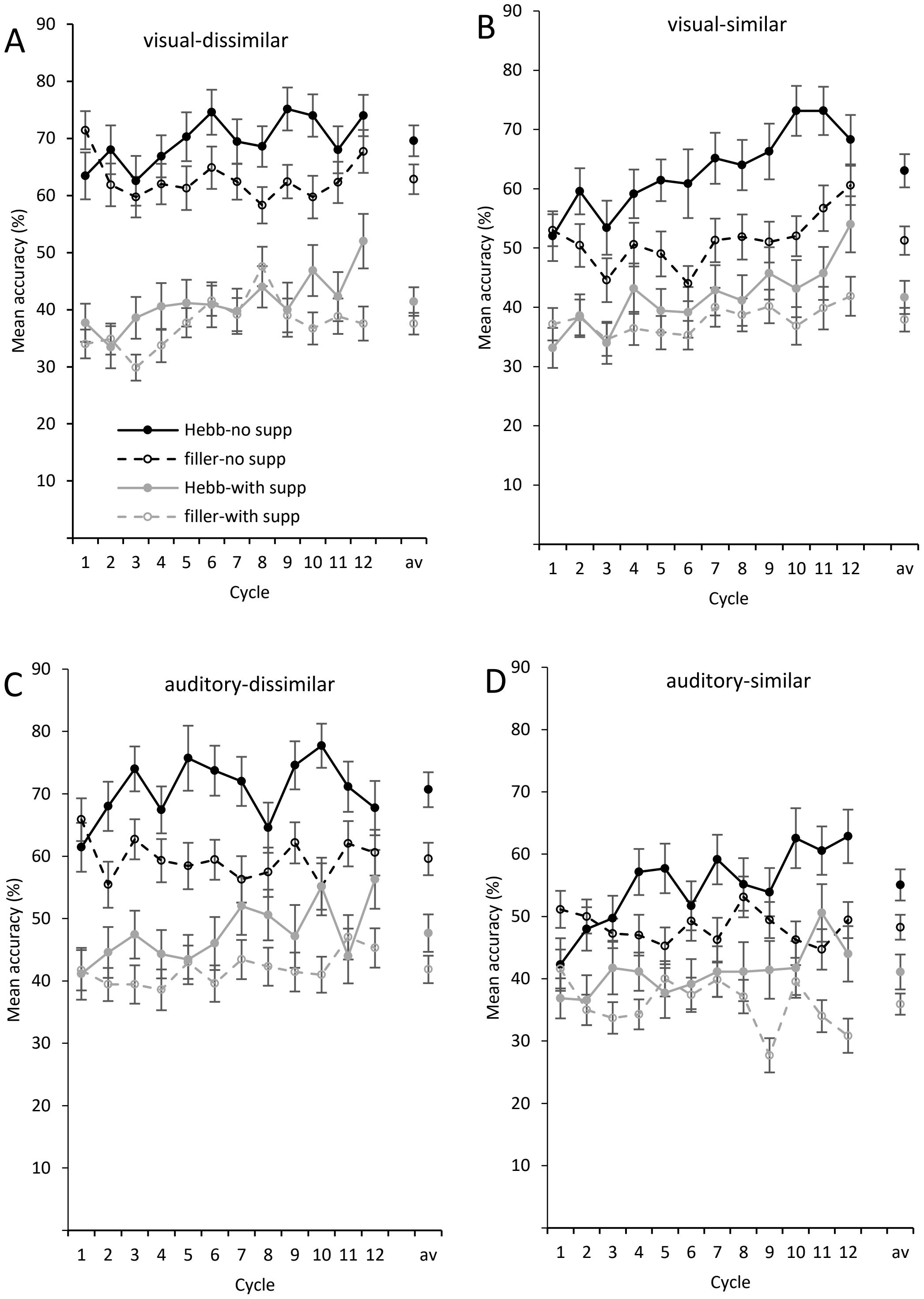


Figure 2.2. Serial recall accuracy (collapsed across serial position) at each Cycle according to List-type, Suppression, Modality, and Phonological similarity in Experiment 1.

Inspection of Figure 2.2 and Table 2.1 suggests that the reliable four-way interaction reflects the following pattern of effects: With visual sequences (cf. Panels A and B of Figure 2), and in the absence of articulatory suppression, there was a Hebb effect with both phonologically dissimilar sequences (Panel A) and similar sequences (Panel B). However, regardless of similarity, the effect was clearly attenuated by articulatory suppression. Also evident with visual presentation was a modulation of the Hebb effect by phonological similarity: Without suppression, the Hebb effect was *larger* for phonologically similar than dissimilar sequences (compare panel B with A), a difference no longer apparent under suppression. The pattern was different in a number of ways with auditorily presented sequences however (cf. Panels C and D of Figure 2.2): In the absence of articulatory suppression there was again a Hebb effect with both phonologically dissimilar and similar sequences but now the magnitude of the effect was greater for phonologically

Modality	Similarity	Suppression	Hebb (%)	Filler (%)	Magnitude of the Hebb effect (Hebb – Filler) (%)	*p*
Visual	Dissimilar	No-supp	69.9	62.8	7.1	.003
		With-supp	41.4	37.5	3.9	.023
	Similar	No-supp	63	51.2	11.8	<.001
		With-supp	41.7	37.9	3.8	.038
Auditory	Dissimilar	No-supp	70.7	59.6	11.1	.019
		With-supp	47.7	41.8	5.9	.024
	Similar	No-supp	55.1	48.3	6.8	.021
		With-supp	41.1	35.9	5.2	.018

Table 2.1. Hebb vs. Filler pairwise comparisons according to Modality, Phonological similarity, and Articulatory suppression (supp).

dissimilar than similar sequences. Moreover, the impact of articulatory suppression on the Hebb effect with auditory sequences was weaker than with visual sequences.[3]

2.2.2.3 Supplementary linear mixed-effect analysis of Hebb learning. A question was raised by an anonymous reviewer of a submitted manuscript containing the first two experiments reported in the present chapter regarding the suitability of ANOVA in the context of the Hebb paradigm, specifically concerning the fact that there is only a single data point for each participant for each instance of the Hebb list. The first point to make here is that this would only be an issue for results (main effects or interactions) that include the Cycle factor; when this is not the case, the data for the Hebb condition are averaged across 12 instances of the Hebb list per participant. Thus, the most important conclusions from the ANOVA are not affected by the 'single-data-point' issue because they do not involve the Cycle factor. Nevertheless, a supplementary linear mixed-effects analysis was conducted, as it has been argued to produce more reliable results in cases where there is only a single data-point per participant for at least some conditions (e.g., Boisgontier & Cheval, 2016; Jaeger, 2008). The data were modified to utilise binary coding for the variables included in the ANOVA. The new data also included factors for block-order, with separate factors for each modality, and a factor for the order in which each participant performed the two modality conditions, as the linear mixed-effects model will be used to account for any random effects block-order may have caused. R (R Core Team, 2012) and lme4 (Bates, Maechler, & Bolker, 2012) were used to perform a linear mixed effects analysis. List-type, Modality, Articulatory suppression,

---

[3] It has been reported that participants tend to make the same recall errors repeatedly in response to the repeating sequence and that this can sometimes obscure a 'true' sequence learning effect (Lafond, Tremblay, & Parmentier, 2010). However, an analysis of response-error learning conducted following the protocol of Lafond et al. (2010) found little evidence of this in the present experiment; a given response error was generally not repeated more than once across a 12-cycle block.

Phonological similarity and Cycle were included as fixed effects. Slopes for order-of-modalities for each participant, block-order for visual blocks for each participant, and block-order for auditory blocks for each participant were included as random effects. Six models were constructed: The null model (Model 1) contained only the fixed effects for List-type and Cycle, and all of the random effects. The main effects model (Model 2) contained all of the fixed effects and all of the random effects, but no interaction effects. Three three-way interaction models (Models 3-5) contained the factors from the main effects model as well as a three-way interaction of either List-type, Similarity and Suppression, or List-type, Similarity and Modality, or List-type, Suppression and Modality. The four-way interaction model (Model 6) contained the effects of interest from the ANOVA (List-type, List-type × Cycle, List-type × Suppression, List-type × Similarity × Suppression and, most importantly, the four-way List-type × Modality × Similarity × Suppression interaction) as well as the random effects. Significance was tested by likelihood ratio tests of the main effects model against the null model, the three-way interaction models against the main effects model and the four-way interaction model against the three-way interaction models in order to determine whether the model of interest, the four-way interaction model, was a better fit to the data than the models with fewer factors. The main effects model was a better fit than the null model, $\chi^2(3) = 1790.9$, $p < .001$. Of the three-way interaction models, the models containing the interactions of list-type, similarity and suppression, and list-type, similarity and modality were a better fit to the data than the main effects model, $\chi^2(4) = 143.66$, $p < .001$ and $\chi^2(6) = 107.52$, $p < .001$, but the third three-way interaction model (List-type × Suppression × Modality) was not, $\chi^2(4) < 0.1$, $p < .9$. Of most relevance, however, was that the four-way interaction model was a better fit to the data than any of the three-way interaction models, $\chi^2(8) = 85.54$, $p < .001$, $\chi^2(6) = 121.68$, $p < .001$ and $\chi^2(8) = 362.17$, $p < .001$. This success of the four-way interaction model (accuracy ~ listtype + listtype × cycle + listtype × suppression + listtype × similarity × suppression + listtype × modality × similarity × suppression + (1 + firstmodality | participant) + (1 + auditoryblockorder | participant) + (1 + visualblockorder | participant) therefore supports the conclusions based on Figure 2.2 and the ANOVA, as the model points to the importance of the 4-way interaction to modelling the data. Estimates of effects, standard errors and *t* values for the four-way interaction model are reported in Table 2.2. The pattern of estimated effects generally suggests an increasing effect on accuracy as more factors are included in the interactions.

Fixed effects	Estimate	SE	*t*
Intercept	0.62	0.02	32.79
List-type	0.01	0.02	0.59
List-type × Cycle	0.01	0.00	6.62
List-type × Suppression	-0.03	0.02	-1.51
List-type × Similarity × Suppression	-0.05	0.03	-2.01
List-type × Modality × Similarity × Suppression	0.09	0.04	2.40

Table 2.2. Estimates of fixed effects in the interaction model, with estimated differences from visual to auditory, from no suppression to with suppression, from dissimilar to similar, and from filler to Hebb.

2.2.3 Discussion

The results of Experiment 1 are in line with a perceptual-motor approach to Hebb sequence learning and at the same time disconfirm basic predictions of the phonological store-based approach (Burgess & Hitch, 2006; Page & Norris, 2009). Critical to the aims in relation to verbal sequence learning was the replication of the intricate pattern of short-term serial recall data (i.e., ignoring the Hebb repetition manipulation) that has served recently to challenge the main empirical basis of the phonological store construct (Jones et al., 2006, 2004; Maidment & Macken, 2012).

Specifically, the survival of the phonological similarity effect under articulatory suppression with auditory presentation—an observation that has been pivotal to the notion of a passive phonological store separable from vocal-motor rehearsal (Baddeley et al., 1984; Hitch et al., 2009)—is located primarily at recency, a portion of the serial recall curve considered to lie outside the explanatory compass of the phonological store (Baddeley, 1986). The present data therefore reinforce the view that this vestige of the phonological similarity effect under suppression is an acoustic similarity effect, reflecting the contribution of passive, acoustic-based, perceptual organisation factors to serial recall performance (Nicholls & Jones, 2002). Thus, as argued earlier, aside from this perceptual-acoustic effect, the apparent empirical signature of the passive phonological store—the 'phonological' similarity effect—is in fact an articulatory similarity effect; a product of motor-planning errors (Jones et al., 2006, 2004).

The analysis of the Hebb effect provided several converging lines of evidence for the hypothesis that verbal sequence learning can also be explained by recourse to motor planning and auditory perceptual organisation processes. First, the prediction that restricting motor planning (through articulatory suppression) should impair Hebb sequence learning was confirmed, especially with visual presentation in which such learning would be expected to be driven mainly by motor planning, that is, where passive auditory organisation processes could not contribute. This attenuation of Hebb sequence learning under articulatory suppression (a result replicated in Experiment 2) contradicts phonological-store based accounts, which predict that articulatory suppression should not affect Hebb sequence learning. This is because, on these accounts, suppression disrupts phonological item-memory but not the stage of processing (order processing) assumed to underpin verbal sequence learning

(Burgess & Hitch, 2006; Page & Norris, 2009). The results of Experiment 1 contradict those of Hitch et al. (2009) who reported no effect of suppression on the Hebb effect with either auditory sequences (Experiments 1 and 2) or visual sequences (Experiment 3). However, closer examination suggests that there may well have been an attenuation of Hebb sequence learning in their Experiment 3 with visual sequences (the condition in which the greatest effect of suppression was found): They reported that there was an interaction between list-type (Filler, Hebb-dissimilar, Hebb-similar) and suppression but attributed this to an attenuation of the similarity effect on serial recall (and hence not to do with sequence learning). However, based on their Figure 6, it seems this interaction may well have been attributable also to an attenuating effect of suppression on the difference between filler and Hebb sequences, that is, an effect of suppression on the Hebb effect. Independently of the foregoing observation, as noted in the Introduction, it can be questioned whether the repetition learning effect in the experiments of Hitch et al. (2009) was, in any case, a pure sequence learning effect insofar as the filler sequences were not made up of the exact same set of items as the repeating Hebb sequence. That is, the repetition learning may, at least in part, have reflected item-set learning and so the relative resistance of this learning effect to articulatory suppression may have obscured an otherwise (larger) attenuating effect of suppression on the Hebb effect proper. The same ‘item-set learning’ issue arises in relation to the only other experiment showing an absence of a suppression effect on (auditory) sequence learning (Page et al., 2006, Experiment 1).

One might question, however, whether the attenuation of the Hebb effect under articulatory suppression in the present experiment was due to a proportional scaling effect (cf. Wang et al., 2016) whereby an effect (here Hebb learning) becomes less likely to be empirically detectable the lower the general level of performance (due in this case to the highly disruptive effect of articulatory suppression on serial recall, cf. Figure 1). However, this argument would be difficult to sustain. First, the Hebb effect under articulatory suppression was still marked with auditory presentation (as predicted by the perceptual-motor account; see below) despite a comparably poor overall level of performance (with recall of filler lists at 38.9%) to that in the visual-with-suppression condition (37.7%). Second, the Hebb effect was sometimes larger at lower overall levels of performance (that found with visually-presented phonologically similar lists) than it was at higher overall levels of performance (that for visually-presented phonologically dissimilar lists). Both these observations suggest that, in the present data at least, there was not a clear association between overall levels of performance and the magnitude of the key effect of interest.

A second key result that supports a perceptual-motor account of Hebb sequence learning and at the same time is problematic for phonological-store based accounts is that learning was modulated by phonological similarity. Phonological-store based accounts predict no such modulation because the order-representation stage that drives sequence learning is ‘blind’ to the phonological identity of the items being ordered (Burgess & Hitch, 2006; Page & Norris, 1998, 2009). The only previous study to have examined the possible effect of phonological similarity on Hebb sequence learning (Hitch et al., 2009, Experiment 3) found no phonological similarity effect using visual sequences (they did not include an auditory condition). But again, this result is difficult to interpret due to the possible contribution of item-set learning to the ‘Hebb effect’ in that experiment. From a motor-planning standpoint, however, the finding that the greater learning of phonologically similar compared to dissimilar sequences with visual presentation—in which the role of motor-planning should be evident in relatively pure form—is readily explicable by recourse to the notion that a relatively disfluent, error-prone, motor activity would stand more to gain from opportunities to re-plan the same sequence (cf. Heathcote et al., 2000). It is worth noting also that this enhanced effect cannot be ascribed simply to greater general task-difficulty or to recall being at a relatively low level before learning commenced (i.e., at Cycle 1) and hence to there being more 'room' for learning to manifest empirically: Performance started at an even lower level under articulatory suppression and yet learning was attenuated, not enhanced, under suppression. Further reinforcing a motor-planning locus for the enhanced learning of (visually-presented) phonologically similar sequences, the enhancement was not observed when motor planning was restricted by articulatory suppression.

Turning to the hypothesised contribution of auditory perceptual organisation to Hebb sequence learning, it is argued that this was manifest in the observation that the Hebb effect remained relatively strong with auditory sequences despite articulatory suppression (compared to the case with visual sequences). It is also evident in the finding that, in contrast to the case with visual sequences, the Hebb effect was weaker with phonologically similar than dissimilar sequences: This may be due to the fact that the strength of the passive processing of order in an auditory sequence is a positive function of the acoustic distinctiveness of its successive elements (so long as that distinctiveness is carried on a common ground such as a common voice; Bregman & Rudnicky, 1975; Hughes et al., 2009; Jones & Macken, 1995). That is, the relatively poor passive processing of order in an auditorily-presented 'phonologically' similar sequence counteracted what is otherwise (e.g., with a visually presented sequence) a larger learning effect with such a sequence.

The results of Experiment 1 support the tenet of the perceptual-motor account that a key part of what underpins Hebb sequence learning is the increasing fluency of the motor-plan generated initially to support the short-term recall of the Hebb sequence. The rationale for the next experiment is based on the notion that a motor-plan embodies not only the sequence-items but also a particular temporal and prosodic organisation of those items. Indeed, as discussed in Chapter 1, it is such paralinguistic features of vocal-motor planning that act as the scaffolding that binds the otherwise unrelated items together (Hughes et al., 2016; Jones et al., 2004; Macken et al., 2016). Experiment 2, therefore, tests the prediction that changes in the temporal grouping within the motor-plan across repetitions of the Hebb sequence—at least when the contribution to learning of auditory perceptual organisation can be ruled out (i.e., with visual sequences)—should attenuate Hebb sequence learning.

2.3 Experiment 2

There is evidence that presenting a sequence of verbal items in two or more temporally-defined sub-groups for serial recall (e.g., *F, H, K, L*----*Q, R, Y*; where the dashed line represents a temporal gap between the *L* and *Q* that is longer than that between any other pair of items) invokes the (qualitatively) equivalent psychological grouping of the sequence. For example, the serial position function with such grouped lists is characterised by two or more (depending on the number of sub-groups) micro serial position curves, suggesting that the sequence is represented, at least to some extent, as separate sub-sequences (e.g., Frankish, 1985, 1989; Hitch, Burgess, Towse, & Culpin, 1996; Ryan, 1969). The finding that this modulation of the serial position function is attenuated under articulatory suppression (Hitch et al., 1996) suggests further that the internal grouping is, at least in part, instantiated within the motor-plan. Furthermore, it has been found that the timing of responses when serially recalling a grouped sequence qualitatively mimics the presented grouping (Maybery, Parmentier, & Jones, 2002). Here, the grouping effect in serial recall is used to provide convergent evidence on the role of motor planning in Hebb sequence learning. It was predicted that presenting the Hebb sequence with different temporal groupings across repetitions should attenuate sequence learning because such variability in presentation will invoke variability in the motor-plan generated in response to each iteration of the Hebb sequence.

It has already been reported that presenting the Hebb sequence with different temporal groupings across repetitions attenuates the Hebb effect (Hitch et al., 2009, Experiment 2; see also Bower and Winzenz, 1969). However, in contrast to the current suggestion, Hitch et al. (2009) argued that temporal grouping-inconsistency affects an abstract representation of the positions of the items that is independent of motor planning processes. In support of this, they reported that restricting motor-planning via articulatory suppression did not affect the modulation of learning by grouping-inconsistency (nor indeed the Hebb effect in general). However, the present Experiment 1 has already shown that their conclusion that the Hebb effect is immune to articulatory suppression seems to have been premature. Here, therefore, the aim is to demonstrate that the effect of temporal grouping-inconsistency on Hebb sequence learning does indeed reflect the role of motor planning in verbal sequence learning. The experiment involved two complementary analyses. First, following Maybery et al. (2002), an assessment of the extent to which different presentation-groupings promote at least qualitatively similar groupings within participants’ temporal organisation of their responses as they output the sequence. To assess the extent to which any such output-grouping reflects the overt execution of a grouped motor-plan, it was also examined for the first time whether or not the match between output- and presentation-timing is diminished under articulatory suppression. It was predicted that when, according to the perceptual-motor account, serial recall is more purely based on motor planning—that is, with visual presentation—articulatory suppression will attenuate markedly the degree to which the output-RTs resemble the presentation-timing. With auditory presentation, in contrast, where the temporal organisation of the presented sequence is likely to be replicated within output-RTs due to passive perceptual grouping processes that proceed regardless of any deliberate motoric grouping (e.g., Jones et al., 2004), articulatory suppression should have less effect. The second analysis will then involve examining the extent to which the evidence for temporal grouping within the motor-plan for serial recall (derived from the first analysis) maps onto the extent to which temporal grouping-inconsistency across repetitions of a sequence attenuates Hebb sequence learning.

In sum, then, in this experiment the timing of serial recall-output was examined as well as Hebb sequence learning for visual- and auditory-verbal sequences with or without articulatory suppression. Of most interest in the present experiment, the temporal grouping of the items was also manipulated; in particular, the temporal grouping of items was either consistent or inconsistent across repetitions of the Hebb sequence.

2.3.1 Methods

2.3.1.1 Participants. One hundred and four students (18 males, 86 females) from Royal Holloway, University of London, aged 18-49 years (mean 20.22 years, SD = 4.05) took part in return either for course credits or a small honorarium.

2.3.1.2 Apparatus and Materials. The apparatus and materials were identical to those of Experiment 1 except that all sequences comprised permutations of the seven letters *F, H, K, L, Q, R,* and *Y* (i.e., the dissimilar set from Experiment

1). The duration of each item was always 400 ms but the items could be presented in a number of different temporal groupings: 2-2-3, 2-5, 3-2-2, 3-4, 4-3, and 5-2, where the numbers represent the number of items in each group and a hyphen representing a between-groups interval. The within-group interstimulus interval was 200 ms while the between-group interval was 1000 ms, resulting in an overall sequence length that varied between 4800 ms and 5600 ms from the onset of the first item to the end of the seventh item.

2.3.1.3 Design. The experiment involved four within-participant factors and one between-participants factor. The first within-participant factor was Grouping-consistency (referred to simply as ‘Grouping’ for the purposes of the analysis of output RTs): In the consistent-grouping condition, all sequences across a block of trials was presented with the same grouping (one of the six possible groupings) while in the inconsistent condition, all six groupings occurred 3 times across the block of trials, once each for each instance of the Hebb sequence and twice each for Filler sequences. The other three within-participant factors were Articulatory suppression (no-suppression, with-suppression), List-type (Hebb, Filler), and Cycle (1-6). The between-participants factor was Modality of presentation, with 52 participants receiving the sequences visually and 52 participants receiving the sequences auditorily. Each Modality group received four blocks of trials, each consisting of 18 sequences of seven letters and in which every third sequence (starting with trial 3) was the same (Hebb) sequence, amounting to 6 instances of the Hebb sequence within a given block. Six instead of 12 cycles were used so that each of the available six groupings could be used once without repetition in the inconsistent-grouping condition. These four blocks corresponded to the 2 × 2 combination of Grouping-consistency (consistent, inconsistent) and Articulatory suppression (with-suppression, no-suppression) and the four blocks were presented in one of four possible orders: A1-B2-A2-B1, A2-B1-A1-B2, B1-A2-B2-A1 or B2-A1-B1-A2, where A represents the no-suppression condition, B represents the with-suppression condition, 1 represents the consistent-grouping condition and 2 the inconsistent-grouping condition. In the inconsistent-grouping condition, where the Hebb sequence, across the block, was presented in all six possible groupings, it was ensured that particular organisations containing the same groups, such as the first group in 2-2-3 and 2-5, were not used for the Hebb sequence in successive cycles. For each block, the participant saw/heard one of two possible counterbalanced sets of sequences with different item-orders for both the Hebb sequences and the filler sequences.

2.3.1.4 Procedure. The procedure was the same as for Experiment 1 except that each participant took part in only one session lasting approximately 1 hr.

2.3.2 Results

2.3.2.1 Output RTs during serial recall. Figure 2.3 shows the extent to which output RTs during mouse-click driven serial recall of the filler sequences aligned with the timings of the items as-presented for each of the six groupings as a function of Articulatory suppression and Modality. In the absence of suppression, with only a couple of exceptions (cf. 5-2 grouping and part of the 3-2-2 grouping), there was a high degree of alignment between presentation and output timings, with RTs, once output was initiated, tending to be longest at group boundaries, indicating a temporal organisation of responses that mimicked how the items were presented. Of particular interest is that with visual sequences, this output-grouping was greatly attenuated under articulatory suppression, consistent with our supposition that with visual presentation (where there can be no passive auditory organization/grouping of

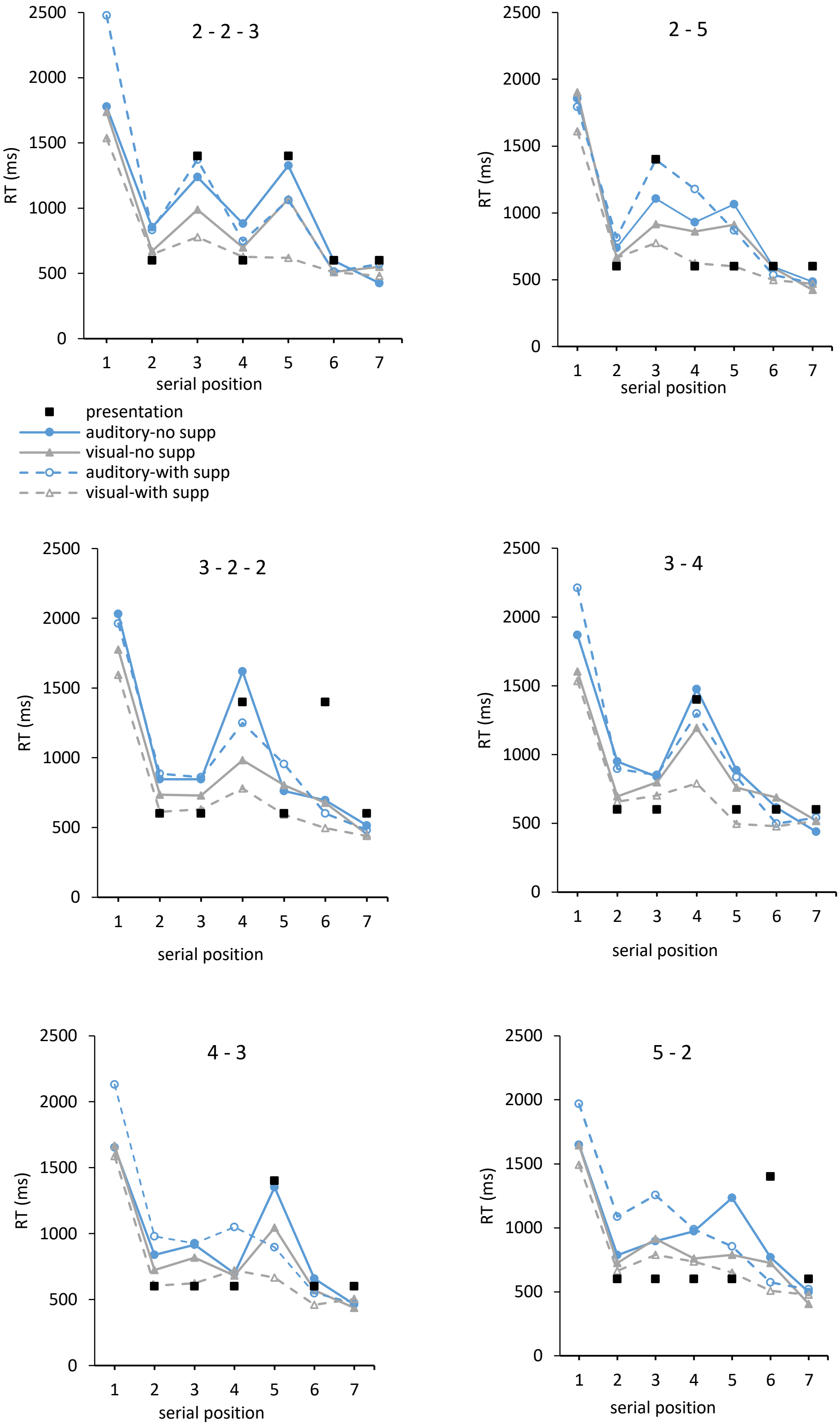


Figure 2.3. Presentation timing and output reaction times during serial recall within each of the six groupings according to modality, suppression and position in Experiment 2.

the items) the output RTs reflect how the items are deliberately assembled into a motor-plan.

Accordingly, with auditory presentation, the alignment of input and output grouping is still very much evident despite articulatory suppression. Thus, the output-grouping effect with auditory, unlike visual, presentation, is not reliant to any large extent on motor planning but rather reflects, I would argue, a direct ‘reading off’ from an episodic record that reflects the passive perceptual organisation of the presented sequence (cf. Jones, Beaman, & Macken, 1996).

These observations were supported by a mixed ANOVA applied to the output-RTs which indicated a significant interaction of Modality, Articulatory suppression, Grouping and Serial position, $F(15.3, 1557.9) = 1.71$, $MSE = 605565.9$, $p = .042$, $\eta_p^2 = .016$ (Greenhouse-Geisser corrected). Other significant effects subsumed within this interaction are reported in Supplementary Table 3 in Appendix 2.

2.3.2.2 Hebb sequence learning. Turning now to Hebb sequence learning, Figure 2.4 shows recall performance accuracy across cycles with visual lists (Panels A and B) and auditory lists (Panels C and D) as a function of Grouping-consistency and Articulatory suppression. It is evident that with visually-presented lists (Panels A and B)—for which the output-RTs analysis suggested a high degree of motor-plan based grouping—learning was considerably weaker with inconsistent grouping of the repeated sequence. Another important feature of the data with visual sequences is that learning was again markedly attenuated under articulatory suppression regardless of grouping-consistency. Indeed, the learning of visual sequences appears to have been abolished by suppression in this experiment as opposed to merely attenuated as in Experiment 1. In contrast, with auditory sequences (Panels C and

D)—for which the output-RTs suggested grouping but grouping driven by passive perceptual organisation rather than motor-planning—there is a strong Hebb effect regardless of grouping-consistency. Thus, only when the grouping is driven solely by motor planning (i.e., with visual sequences) does grouping-inconsistency have a strong disruptive effect on sequence learning. Moreover, while suppression again markedly attenuated learning with auditory sequences, this attenuation was not as emphatic across cycles as was the case with visual sequences (see Panel C).

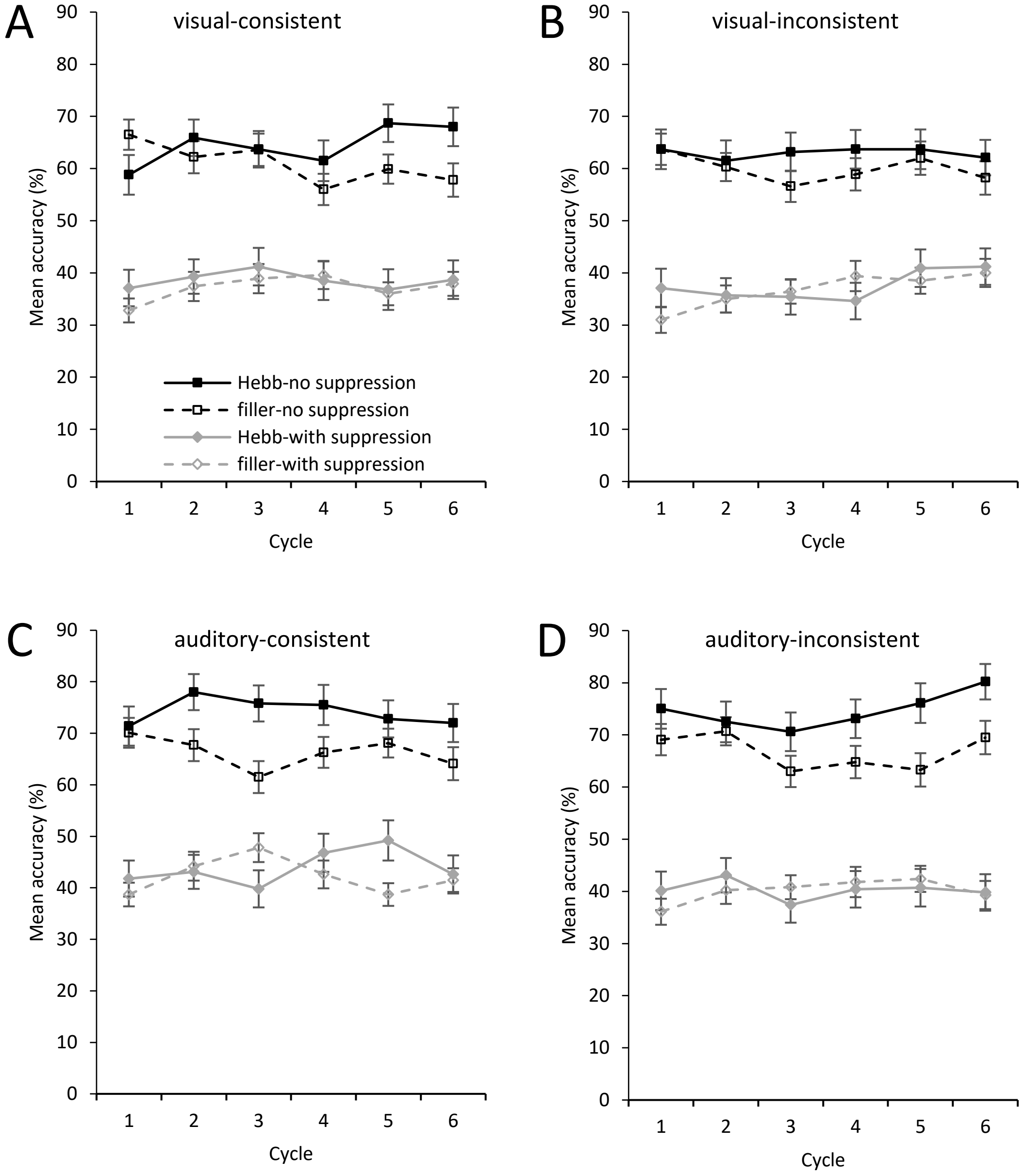


Figure 2.4. Performance accuracy across Cycles by List-type and Suppression conditions according to Modality and Grouping.

A 2 (List-type) × 2 (Modality) × 2 (Articulatory suppression) × 2 (Grouping-consistency) × 6 (Cycle) mixed-design ANOVA showed reliable main effects of List-type, $F(1, 102) = 20.93$, $MSE = .07$, $p < .001$, $\eta_p^2 = .17$, Articulatory suppression, $F(1, 102) = 497.71$, $MSE = .18$, $p < .001$, $\eta_p^2 = .83$, and Modality, $F(1, 102) = 10.93$, $MSE = .45$, $p = .001$, $\eta_p^{2} = .097$. There was also a significant interaction between List-type and Articulatory suppression, $F(1, 102) = 8.71$, $MSE = .07$, $p = .004$, $\eta_p^{2} = .079$, replicating the attenuation of Hebb sequence learning by articulatory suppression observed in Experiment 1. Corroborating the impressions of the pattern evident in Figure 2.4, the five-way interaction was also significant, $\mathrm{F}(1.9, 200.3) = 4.38$, $\mathrm{MSE} = .02$, $\mathrm{p} = .014$, $\eta\mathrm{p}\ 2 = .041$ (for the full set of results from this ANOVA as well as a simple effects analysis of the 5-way interaction, see Appendix 2). Of particular relevance, when suppression was not required, the enhanced recall of the Hebb compared to filler sequences was reliable by cycles 5 and 6 in the visual-consistent condition ($p = .023$ and $p = .017$ respectively) while there was no reliable Hebb effect at any cycle in the visual-inconsistent condition (all ps > .05).For auditory sequences, without suppression, there was a Hebb effect at all cycles except cycles 1 and 5 in the consistent-grouping condition (though only marginal at cycle 6, $p = .066$) but, in contrast to the case with visual sequences, there was clear learning also in the inconsistent-grouping condition, with either a reliable effect or marginally reliable effect observed at all cycles except the first two (see Supplementary Table 5 in Appendix 2). Under suppression, there remained some evidence of learning with auditory sequences, though only at one cycle within the consistent-grouping condition (Cycle 5, $p = .016$).

2.3.2.3 Supplementary linear mixed-effect analysis of Hebb learning.

Following the procedure used in relation to the results of Experiment 1, a supplementary linear mixed-effects analysis was conducted. The data were modified to utilise binary coding as in Experiment 1, and the new data also included a factor for block-order. List-type, Modality, Articulatory suppression, Grouping-consistency and Cycle were included as fixed effects. A slope for block-order for each participant was included as a random effect. Again, six models were constructed: The null model contained only the fixed effects for list-type and cycle, and the random effect. The main effects model contained all of the fixed effects and the random effect, but no interaction effects. The two-way interaction model contained the effects from the main effects model as well as an interaction of List-type and Cycle. The three-way interaction model contained the effects from the main effects model and an interaction of List-type, Cycle and Suppression. The four-way interaction model was similar to the three-way interaction model, but also included the four-way List-type × Cycle × Suppression × Grouping-consistency interaction. Finally, the five-way interaction model contained the main effects and the five-way interaction from the ANOVA as well as the random effects. For the sake of brevity, it is simply reported here that the five-way interaction model was a better fit to the data than any other tested model, $\chi^2(15) = 28.42$, $p = .019$. This interaction model (accuracy ~ list-type × modality × grouping × suppression × cycle + (1 + block-order | participant) again supported the conclusions based on Figure 2.4 and the original ANOVA, as the model points to the importance of the 5-way interaction in modelling the data. Estimates of effects, standard errors and *t* values for the interaction model are reported in Table 2.3. Suppressions appears to be the largest estimated effect on accuracy on its own, but the five-way interaction is also significant in the model.

Most notably, the importance of the five-way interaction in the model supports the findings from the ANOVA.

Fixed effects	Estimate	SE	*t*
Intercept	0.67	0.03	22.0
List-type	0.08	0.04	1.94
Cycle	0.02	0.01	2.31
Suppression	-0.32	0.04	-8.13
Grouping	-0.05	0.04	-1.23
Modality	0.02	0.04	0.47
List-type × Modality × Grouping × Suppression × Cycle	0.06	0.02	1.98

Table 2.3. Estimates of fixed effects in the interaction model, with estimated differences from visual to auditory, from no suppression to with suppression, from consistent to inconsistent grouping, and from filler to Hebb.

2.3.3 Discussion

The results of Experiment 2 provide strong convergent support for an account of Hebb sequence learning in which motor planning plays a central role. First, replicating the finding from Experiment 1, when motor planning is restricted by articulatory suppression, learning is diminished markedly; indeed, for visual sequences, it was abolished in the present experiment. There was again some evidence of the learning of auditory sequences being more resistant to articulatory suppression than visual sequences: While only apparent at one cycle in the consistent-grouping condition, it remains the case that only with auditory sequences was there any evidence of learning surviving the otherwise emphatic impact of articulatory suppression. Turning to the novel aspects of the present experiment, the serial recall data showed first that RT-indexed output-grouping during serial recall is diminished under articulatory suppression but only with visual presentation, where there can be no automatic, auditory-perceptual based, grouping. When such passive auditory grouping can occur (i.e., with auditory sequences) the grouping remained strong under suppression. Thus, the output-grouping in this case appears to reflect a direct motoric translation of the way in which passive perceptual process have organised the auditory input. Second, this pattern in the RT data mapped systematically onto the pattern of verbal sequence learning: Only when the grouping was dependent on motor planning (i.e., with visual sequences) did an inconsistency in the presented-grouping across Hebb repetitions attenuate (indeed eliminate) learning. I contend that the inconsistency in the input-grouping across the repeated sequence produced a corresponding inconsistency in the motor-plan generated for its serial recall, thereby reducing the motor fluency-gain that is otherwise made from repeatedly planning the same sequence.

When the sequence was subject to auditory perceptual organisation as well as motor-planning (i.e., auditory, no-suppression condition), learning was evident regardless of grouping-inconsistency across repetitions. One possibility is that the co-occurrence of motor-planning and auditory-perceptual processes provides sufficiently strong cues to the order of successive items to resist the otherwise disruptive impact on learning of a change in the way the items are organised into sub-groups across repetitions. This may also account for the particularly strong and rapid learning found for the auditory-dissimilar sequences (but not similar sequences, where acoustic-order cues would be weak) in Experiment 1. Further research will be required to examine this tentative account of this particular finding however.

The findings of this experiment are again problematic for phonological-store based accounts. Not only was the attenuating effect of articulatory suppression on Hebb sequence learning replicated, temporal grouping effects also interacted with the presence of articulatory suppression and presentation-modality. Such effects and interactions go against an explicit (null) prediction of phonological-store accounts because these accounts posit that articulatory suppression affects the representation of item information but not the (ordering) stage of processing responsible for long-term sequence learning (Burgess & Hitch, 2006; Hitch et al., 2009; Page et al., 2006). In particular, the survival of the Hebb effect with auditory sequences with inconsistent-grouping replicates a result observed by Hitch et al. (2009, Experiment 2). However, whereas Hitch et al. focused on the fact that the learning was weaker in their inconsistent-grouping condition than their consistent-grouping condition, what they could not have shown—as they only included auditory sequences—is that the effect is eliminated completely by inconsistent grouping with visual sequences[4]. This is problematic for any phonological-store account because so long as items access the phonological store—which, in the absence of suppression, should be the case regardless of presentation-modality—the store does not, by definition, discriminate between representations derived from reading the items from representations derived from hearing them. Thus, any factor that modulates Hebb sequence learning should do so to an equivalent degree regardless of presentation modality, contrary to the current data.

A potential counterargument to the interpretation of the grouping effects in Experiment 2 as having a motor-planning locus could be based on a study by Farrell and Lelièvre (2012), the results of which are sometimes interpreted as demonstrating that grouping at output reflects the structure of memory storage during encoding, not motor planning. Farrell and Lelièvre (2012) asked participants to start serial recall of a list at various serial positions and then wrap back around to the beginning of the list (e.g., to recall items in positions 4-7 followed by those in positions 1-3) in an attempt

---

[4] Bower and Winzenz (1969) reported that the Hebb effect was eliminated with inconsistent grouping even with auditory sequences (again, they did not include visual sequences). However, this result (which is contrary to both our findings and those of Hitch et al., 2009) is difficult to interpret: In the relevant experiments (Experiments 3-4, 6-8), few details about the structure and timing of the groupings used are provided and the power of the experiments was relatively low (n = 10-18 compared with n = 104 in the present experiment).

to disentangle the role of input and output processes in output-grouping effects. They found peaks in both recall accuracy and RTs at group boundaries that were similar regardless of recall start-point, suggesting that temporal grouping affects the input encoding of memory traces in short-term storage, rather than the production of the output. However, as Farrell and Lelièvre (2012) acknowledge, the results of their Experiments 1 and 2 only bring into question the idea that output-grouping effects reflect the action of a late-stage motor-output buffer and that they are still compatible with such effects reflecting the temporal structure of a motor-plan generated during encoding. Indeed, their final experiment (Experiment 3) was specifically designed to try to also rule out a motor-planning account. The results of that experiment are ambiguous however: While impeding motor planning through articulatory suppression was found to have little effect on output-grouping as evident from RTs, it did attenuate grouping as evident in recall accuracy (see also Hitch et al., 1996). Doubts can be raised also about the effectiveness of their articulatory suppression manipulation insofar as they used steady-state suppression ("blah, blah, blah...") which is known to be significantly less effective at impeding motor planning than changing-state suppression ("8, 9, 10…") such as used in the current experiments (Macken & Jones, 1995). The rate of suppression was also rather slow in Farrell and Lelièvre (2012; approximately two items/s) compared to the more typical rate of approximately three items/s as used in the present experiments). Their findings are not, therefore, as troubling for a motor-planning account of output-grouping effects as often thought. Furthermore, the current interaction between modality and articulatory suppression in relation to the output-RTs, where grouping was diminished under suppression only for visual sequences, is particularly relevant: This interaction is precisely predicted by an account in which there are two sources of grouping, one motor-planning based and the other, with auditory sequences, perceptual organisation-based. Such an interaction is not, however, predicted by an account where grouping has a single, input-mnemonic, basis.

2.4 Experiment 3

The argument presented thus far that Hebb sequence learning is driven in large part by motor planning may appear to be at odds with studies indicating that the production of the repeating sequence during serial recall is not necessary for such learning (Kalm & Norris, 2016; Oberauer & Meyer, 2009; but see Cohen & Johanson, 1967; Cunningham, Healy, & Williams, 1984). However, whether or not such *overt production* of the sequence is necessary for the Hebb effect does not speak directly to the role of the *covert planning* of the sequence. This is because participants in such studies are typically only informed after the presentation of the sequence whether or not overt recall is required (Kalm & Norris, 2016; Oberauer & Meyer, 2009). A motor-plan for the recall response is therefore likely to be assembled during sequence-presentation regardless of the identity of the subsequent cue due to the potential need for that plan. In terms of the contribution of motor planning, this may have resulted in no substantive difference between the condition that did require recall of the repeating sequence, and that which did not. Indeed, the results of studies in which there was no requirement or reason to assemble a motor-plan for the repeating sequence at all are entirely consistent with the present argument: No Hebb effect is observed under such conditions. Cunningham, Healy, and Williams (1984) presented participants with sequences containing two separated segments, only one of which participants were required to recall, and they observed that learning was only evident for the items that were recalled, with no improved accuracy demonstrated for the segment that had only been observed and not recalled, when that recall accuracy was tested later in block. However, their design is open to the potential problem that the not-to-be-recalled sub-sequence was inhibited and hence this was the reason that no learning of this sub-sequence was evident. The recall of only one part of the sequence could have led to intentional forgetting of the other, no-to-be-recalled part, as it became unnecessary for meeting the goals when the recall cue was presented. Participants have been shown to be proficient at actively inhibiting or intentionally forgetting parts of a serial recall list while retaining other parts that are relevant for the task-goals (e.g., Harnishfeger & Pope, 1996). The only other experiment in which participants knew before the sequence that recall of the repeating Hebb sequence was not required did not suffer from this complication: Glass, Krejci, and Goldman (1989) presented a long auditory-verbal sequence in which a repeating sequence was intermittently embedded. Participants who were required only to monitor the sequence for discrepancies against a written transcript of the sequence or to shadow it by pronouncing each digit as they heard it—that is, they were not required to recall, or therefore generate a motor-plan for, the repeating sequence—did not show a Hebb effect when the repeated sequence had to be serially recalled later.

The current experiment adopts a similar approach to Glass et al. (1989) to seek further convergent evidence for the role of motor planning in Hebb sequence learning. However, the different task goals (recall vs. passive processing) were implemented in a design more closely matched to the present Experiments 1 and 2 (and hence the standard Hebb paradigm). This contrasts with the design of Glass et al. (1989), which involved continuous presentation which would prevent the partitioning of each sequence into a clear auditory object, or a simultaneous active task of comparison or production, with only short, 2 s breaks between the sequences, that could promote item-level rather than sequence-level processing. In the present experiment, the sequences were presented individually and only required a response once the entire sequence had been heard, so that any sequence in a given block had to either be reproduced after presentation or only had to be monitored. This also allowed the direct comparison of the learning of Hebb sequences that were either recalled (and thus required motor planning) or monitored through the block amongst filler sequences requiring either monitoring or recalling. Thus, within a given block of trials, some sequences had to merely be monitored for the presence or not of a target item (the spoken letter "A") while other sequences had to be recalled (as in the standard paradigm). In one condition, the Hebb sequence was always associated with a recall cue while in another the Hebb sequence was always monitored during the learning phase (see Phase 1 demonstrated in Table 2.4) and only recalled during a test phase to measure any learning after the initial learning phase. While performance accuracy in the monitoring task in itself is not of particular interest, it was recorded to ensure participants continued to be engaged in the task regardless of the response-requirement. There was no target in the Hebb-monitor sequences to ensure that participants attended to the entire sequence, as they could be expected to discontinue actively monitoring the remainder of a sequence if a target was in a position other than the final one. If learning is reliant on the motor planning of the repeating sequence then learning should be reduced when motor planning is not required to meet the short-term task-goal (i.e., in the to-be-monitored Hebb sequences condition). This was tested by having a 'test phase' whereby, regardless of condition (monitor-Hebb vs. recall-Hebb), the seventh instance of the Hebb sequence (and the preceding filler) had to be recalled. This response-requirement manipulation is illustrated in Table 2.4 (please ignore, for the time being, the columns labelled ‘presentation format’).

			Presentation format				Response-requirement	
			Phonologically dissimilar		Phonologically similar			
	Trial	List-type	Single voice	Alt voice	Single voice	Alt voice	Recall Hebb	Monitor Hebb
Phase 1	1	F	*LFHYKRQ*	**H**L**Q**R**Y**F**K**	*PTDBGVC*	**T**D**G**V**C**P**B**	Recall	Recall
	2	F	*QKRFAHY*	**F**Y**A**K**L**R**Y**	*CTGBVPD*	**G**B**T**D**P**C**V**	Monitor	Monitor
	3	**H**	*KLFQHYR*	**R**F**Y**K**Q**L**H**	*BPDTCVG*	**P**G**T**B**V**D**C**	**Recall**	**Monitor**
	4	F	*FQRYHLK*	**Y**R**K**L**H**F**Q**	*DCTPGBA*	**B**C**T**D**V**A**G**	Monitor	Monitor
	5	F	*YFHQKRL*	**Q**H**F**Y**R**K**L**	*VBGCTDP*	**C**P**T**B**D**G**V**	Recall	Recall
	6	**H**	*KLFQHYR*	**R**F**Y**K**Q**L**H**	*BPDTCVG*	**P**G**T**B**V**D**C**	**Recall**	**Monitor**
	7	F	…				Recall	Recall
	8	F					Monitor	Monitor
	9	**H**					**Recall**	**Monitor**
	10	F					Monitor	Monitor
	11	F					Recall	Recall
	12	**H**					**Recall**	**Monitor**
	13	F					Recall	Recall
	14	F					Monitor	Monitor
	15	**H**					**Recall**	**Monitor**
	16	F					Recall	Recall
	17	F					Monitor	Monitor
	18	**H**					**Recall**	**Monitor**
Phase 2	19	F					Monitor	Monitor
	20	F test	*FKRYHLQ*	**K**L**Y**Q**F**R**H**	*VDTGCPB*	**C**T**G**B**D**P**V**	Recall	Recall
	21	**H** test	*KLFQHYR*	**R**F**Y**K**Q**L**H**	*BPDTCVG*	**P**G**T**B**V**D**C**	**Recall**	**Recall**
	22	F					Monitor	Monitor
	23	F transfer	*HLKYQFR*	**YKRL***HFQ*	*PDTBGVC*	**DGTV***BPC*	Recall	Recall
	24	**H** transfer	*KFHRLQY*	**RYQH***FKL*	*BDCGPTV*	**PTVC***GBD*	**Recall**	**Recall**

Table 2.4. A schematic illustration of the block-structure in, and design of, Experiment 3, showing the manipulations of Voice-presentation, Phonological similarity, and Response-requirement. Example sequences are also shown. List-type ‘H’ stands for Hebb sequence and ‘F’ for filler. The 24 trials of each block comprised 8 cycles, with each cycle containing one instance of the Hebb sequence and two preceding fillers. For alternating (Alt) voices, the underlined and bold items represent items spoken in one voice and the italicised items represent items spoken in the other voice. For the transfer Hebb sequence (trial 24), alternate items in the Hebb sequence presented during the previous seven cycles were now presented successively. A response-requirement cue (‘Recall’ or ‘Monitor’) was always shown before the sequence was presented. In Monitor trials, participants had to listen out for the target letter ‘A’ which was presented on half the Monitor trials (see Method for further details).

A particular benefit of using a response-requirement manipulation in the present experiment stems from the fact that the case for the role of motor planning in Hebb sequence learning in the present chapter has so far relied to a considerable degree on the effects on learning of articulatory suppression. That is, one potential challenge to the conclusions from Experiments 1-2 could come from the theoretical standpoint that articulatory suppression does not necessarily affect performance by restricting motor planning but by producing a general interference effect (Lewandowsky et al., 2008; Oberauer et al., 2015). In this approach, reduced learning under articulatory suppression in Experiments 1-2 might also be construed as being due to general interference. If, in the present experiment, the absence of motor planning (i.e., in the Monitor-Hebb condition) has a similar effect on learning to articulatory suppression, that would leave Experiments 1-2 less open to this potential challenge because the monitor condition does not involve introducing any kind of secondary task (and hence the possibility of general interference) but rather removing the need for the process of interest.

Another goal of this experiment was to provide a convergent test of the role of passive perceptual organisation in auditory-verbal sequence learning. First, interest centred on the way in which the passive organisation of the input can modulate the readiness with which a suitable motor-plan can be assembled. This perceptual-motor mapping has been shown to have an important role in verbal serial short-term memory performance: The introduction of perceptual variability into an auditory sequence such as presenting items in an alternating female-male voice or alternating-ear fashion produces poorer serial recall because, it has been argued, this reduces the degree of alignment between the perceptual organisation of the sequence and the need to assemble the items in serial order in a motor-plan (Hughes et al., 2009, 2016). To elaborate, a key cue for streaming is acoustic similarity and so when the alternate items are more acoustically similar to one another than are successive items, they are grouped into the same stream. As illustrated in Table 2.5., this results in two interleaved by-voice streams within which the order is at odds with the actual temporal order of the items. When the auditory-perceptual object is compatible with the goal of the motor-planning (single-voice condition in Table 2.5)—that is, to reproduce the true temporal order of the items—this should more readily support the active assembly of the motor-plan and hence support both serial recall and sequence learning.

Second, the co-manipulation of response-requirement and presentation-format (single voice vs. alternating voices) provided a way of examining the direct role of passive perceptual organisation in sequence learning: By examining the effect of the voice-format manipulation within the monitor-Hebb condition, it should be possible to witness the role of perceptual organisation in sequence learning uncontaminated by the influence of motor-planning processes. Specifically, it will be possible to examine whether learning occurs in the absence of a motor-plan by contrasting recall of the Hebb-test sequence compared to Filler-test sequence in the Monitor-Hebb condition (cf. Table 2.4). In addition, it can be examined whether the perceptual organisation of items by voice in the alternating-voices condition results in the learning of the by-voice sub-sequences. This was assessed using a transfer cycle (cycle 8) in which the alternate items in the Hebb sequence presented at cycles 1-7 (which would be the by-voice items in the alternating-voices condition) were now presented successively [e.g., Cycle 7: **R***F***Y***K***Q***L***H** to Cycle 8 (transfer): **RYQH***FKL*]. If passive perceptual organisation is sufficient for learning then recall of the transfer Hebb sequence should be better than that of the preceding transfer-Filler sequence (which was matched to the Hebb sequence in terms of its separate voice-by-voice arrangement; see Table 2.4).

Presentation-format	Perceptual organisation
Single voice: KHLQRYF	KHLQRYF
Alternating voices: **K***H***L***Q***R***Y***F**	**K L R F** *H Q Y*

Table 2.5. The perceptual organisation of items presented in a single voice sequence or in an alternating voices sequence, where the bold and underlined letters represent items presented in one of the voices, and the italicised letters in the other.

As in Experiment 1, the phonological similarity of the items was also manipulated in this experiment (also illustrated in Table 2.4). This was done for two reasons. The first was to check whether the larger learning effect for phonologically dissimilar compared to similar auditory sequences (cf. Experiment 1) replicates. Second, it has been shown that, in serial recall, phonological similarity interacts with voice presentation-format such that the effects of phonological similarity and voice alternation cancel one another out to some extent (Hughes et al., 2009, Experiment 4). This is consistent with the view that both variables impair motor planning (Hughes et al., 2009). In the present context, then, it is possible that while voice alternation and phonological similarity may each independently impede sequence learning, their effects on learning, as in serial recall, will not be additive.

In contrast to the view that motor planning and auditory perceptual organisation underpin verbal sequence learning, if verbal sequence learning is supported mostly or entirely by central modular systems, a different pattern should emerge. Rather than verbal sequence learning being affected when perceptual organisation and motor-planning processes are affected, learning should be comparatively free from the influence of particular sensory properties or the availability of motor-planning processes. As learning, on these models, results from the encoding of the sequences, and is not based on the production level (Page et al., 2006), there should be no difference in the learning patterns between Hebb sequences that are recalled throughout a block compared to those that are merely monitored. Additionally, these models would not predict an effect of alternating voices on learning. It appears that the effect of alternating voices in reducing short-term performance accuracy—as that of phonological similarity or articulatory suppression—would be considered to act on the phonological store, rather than on the temporal context signal or primacy gradient, which would be required if voice alternating were to affect learning on these models (Hitch et al., 2009; Page et al., 2006).

2.4.1 Methods

2.4.1.1 Participants. Fifty participants of the intended fifty-two (14 males, 36 females) from Royal Holloway, University of London, aged 18-41 years (mean 22.5, SD = 4.19), took part in the experiment in return for a small honorarium. One participant's data was excluded from data analysis due to their not completing all conditions.

2.4.1.2 Apparatus and Materials. The apparatus and materials were adopted from Experiment 1 with the following changes: Only auditory presentation was used, and the stimulus set included, for some of the to-be-monitored filler sequences, the target letter *A*, which was recorded and edited similarly to the other stimuli. The letters were presented for 500 ms with an interstimulus interval of 300 ms. All letters were recorded in both a female and a male voice and compressed to 500 ms without altering other acoustic features with *Adobe Audition Creative Cloud* software. A slower presentation rate compared to previous alternating-voice experiments (Hughes et al., 2009) was used to increase the likelihood that learning of the Hebb sequence would occur.

2.4.1.3 Design. Each of the eight blocks [2(Voice) × 2(Phonological similarity) × 2(Response-requirement to Hebb)] of the experiment contained two phases, one running uninterrupted into the next: The initial six cycles (Phase 1) and the final two cycles (Phase 2) (see Table 2.4). Phase 1 embodied five factors: Phonological similarity (dissimilar, similar), Voice [single, alternating; in the single voice condition, all items were presented in the same (male or female) voice, and in the alternating-voices condition the female (F) and male (M) voice alternated (i.e., FMFMFMF)], Response-requirement (monitor, recall; referring to the response-requirement for the six Hebb sequences during Phase 1, with the conditions containing a to-be-monitored Hebb sequence ignored for the purpose of the Phase 1 analysis), List-type (Hebb, filler), and Cycle (referring, in Phase 1, to the first 6 cycles only). Phase 2 also manipulated the same five factors, but in Phase 2 Cycle refers to the ‘test’ and ‘transfer’ cycles: At ‘test’, the Hebb sequence and the preceding filler had to be recalled regardless of whether the previous seven instances of the Hebb sequence had been recalled or had been monitored. At ‘transfer’ the items that had been in alternate positions in the Hebb sequence during Phase 1 were presented successively in the transfer Hebb sequence [in the alternating-voices condition, the filler transfer sequence, like the Hebb transfer sequence, was presented in a voice-by-voice (FFFFMMM) fashion]. The eight blocks were presented in counterbalanced order. Additionally, for each block participants heard one of two possible sets of sequences, with different item-orders for both filler and Hebb sequences for each condition. The voice used for the single voice condition (male or female) was counterbalanced across participants. Similarly, the voice beginning each list in the alternating-voice sequences was counterbalanced across participants. In each block, half of the fillers required monitoring and half had to be recalled, regardless of the Hebb sequence response-requirement. The target letter could appear at any of the seven serial positions of a to-be-monitored filler sequence, but was never presented in a Hebb sequence to ensure the Hebb sequence was monitored for its entire duration. The target could appear in each block between one and four times.

2.4.1.4 Procedure. The procedure was identical to that of Experiment 1 with the following changes: The instruction included descriptions of the two types of tasks to perform (monitor and recall), and participants were told that before each trial, they would be informed on the screen which task they were to do. The response-requirement description appeared on the screen together with a ‘Begin Trial’ button. After the participant pressed the button, the description remained on the screen for 200 ms, until the letter presentation began. After the seven letters were presented, depending on the task, either a question appeared on whether the letter ‘A’ was present in the sequence, and participants had to press a button either to indicate ‘yes’ or ‘no’. If the task was to recall, the letters would appear on the screen in a circle, and the participant was required to press the letters on the screen in the order in which they thought they heard them using the mouse. The experimental session lasted 1 hr.

2.4.2 Results

2.4.2.1 Monitor task. Performance accuracy in the monitor task was very high (96.1%, SD = 3.9%), indicating that participants engaged well in the task.

2.4.2.2 Serial recall. First, serial recall performance in the four serial recall conditions with serial recall data (2 [Voice] ×2 [Similarity]) of the experiment was examined, excluding to-be-monitored trials. The data for each participant for each condition included the averaged accuracy at each of the seven serial positions of the to-be-recalled filler trials, with 14 sequences for both single voice and alternating voice conditions. Figure 2.5 shows the classic phonological similarity effect, with more accurate recall for phonologically dissimilar sequences, and the detrimental effect of voice alternation. The data were analysed with a 2 (Voice) × 2 (Similarity) × 7 (Position) within-participant ANOVA. There was a significant main effect of Position, $F(3.7, 177) = 87.65$, $MSE = 635.2$, $p < .001$, $\eta_p^2 = .65$ and a significant main effect of Similarity, $F(1, 48) = 77.72$, $MSE = 1124.1$, $p < .001$, $\eta_p^2 = .62$, replicating the phonological similarity effect. The interaction of Position and Similarity indicated that this effect was particularly strong in the middle

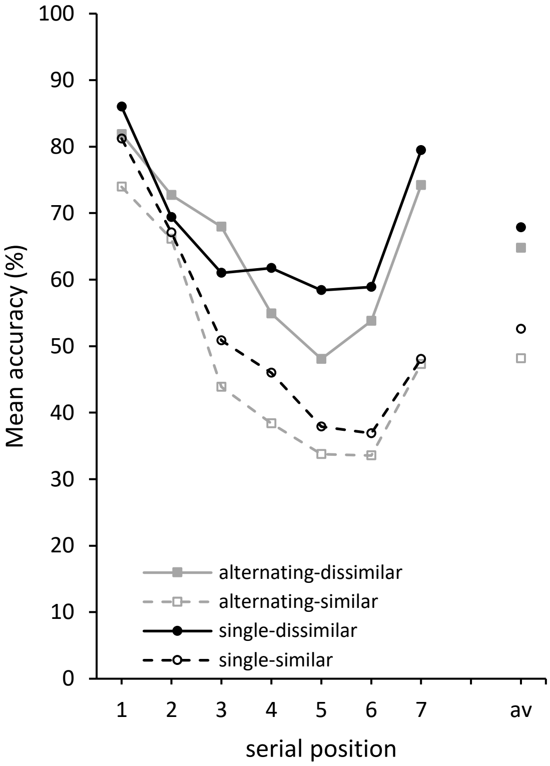


Figure 2.5. Serial recall accuracy for the filler sequences across serial positions and voice conditions according to similarity.

and the end of the sequence, and less apparent at the first two serial positions; $F(4.3, 204.9) = 16.24$, $MSE = 302.6$, $p < .001$, $\eta_p^2 = .25$. There was also a main effect of Voice, $F(1, 48) = 8.39$, $MSE = 573.1$, $p = .006$, $\eta_p^2 = .15$, with alternating voice sequences recalled significantly less accurately than single voice sequences. Despite the implication of Figure 2.5, Voice and Similarity did not significantly interact; $F(1, 48) = .28$, $MSE = 580.2$, $p = .603$, $\eta_p^2 = .01$. There was, however, a significant three-way interaction of Voice, Similarity and Position, $F(3.4, 162.9) = 3.06$, $MSE = 319.3$, $p = .025$, $\eta_p^2 = .06$. Figure 2.5 and the simple effects analysis, reported in Supplementary Table 6 of Appendix 3, of the three-way interaction indicated that the alternating voice effect is particularly prevalent in the first item and the middle of the sequences, with a more pronounced effect for dissimilar compared to similar sequences.

2.4.2.3 Hebb sequence learning in Phase 1. This first analysis of sequence learning was designed to assess learning across Phase 1 (i.e., the first six cycles) in those blocks in which the Hebb sequence had to be recalled throughout that phase (note that this cannot of course be assessed for the to-be-monitored Hebb sequences due to the absence, by design, of responses to those sequences). Inspection of Figure 2.6 suggests that there was a clear sequence learning effect across the first six cycles regardless of voice or similarity condition.

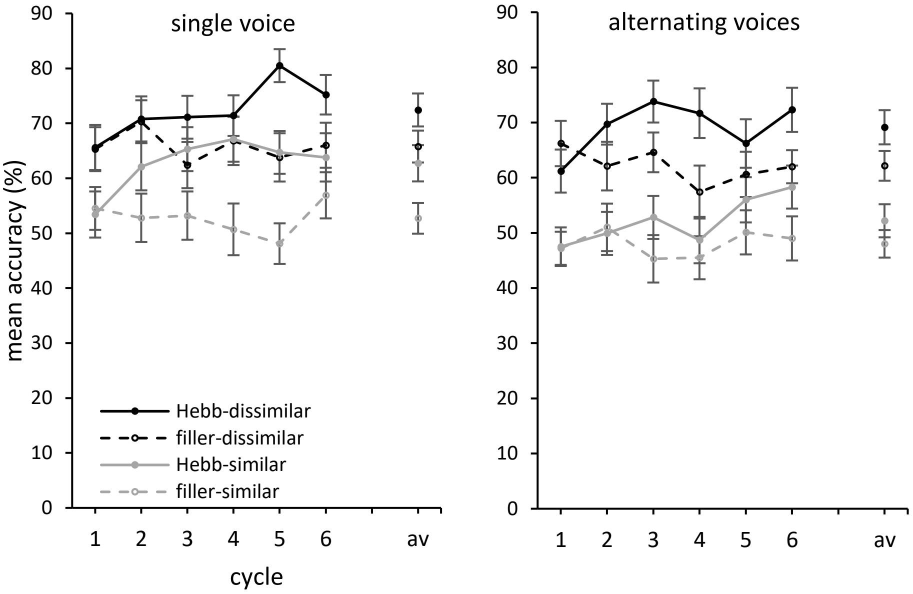


Figure 2.6. Serial recall accuracy according to list-type, cycle, similarity, and voice in Phase 1 of Experiment 3.

The data were analysed with a 2(List-type) × 2(Voice) × 2(Similarity) × 6(Cycle) repeated-measures ANOVA. Hebb sequence learning was confirmed by the significant main effect of List-type, $F(1, 48) = 34.35$, $MSE = .08$, $p < .001$, $\eta_p^2 = .417$, and the interaction of List-type and Cycle, $F(5, 240) = 4.34$, $MSE = .05$, $p = .001$, $\eta_p^2 = .083$. There were also significant main effects of Voice, $F(1, 48) = 15.17$, $MSE = .12$, $p < .001$, $\eta_p^2 = .24$, and of Similarity, $F(1, 48) = 39.27$, $MSE = .27$, $p <$

.001, $\eta_p^2 = .45$—reflecting the known disruptive impact of these variables on serial recall per se (see previous sub-section)—and a main effect of Cycle, $F(5, 240) = 2.6$, $MSE = .05$, $p = .026$, $\eta_p^2 = .051$. There were no reliable interactions; thus, other than Cycle, none of the factors influenced Hebb sequence learning in the Hebb-recall condition during Phase 1 of the experiment. Thus, while it may appear from Figure 2.6 that learning was diminished in the phonologically-similar condition within the alternating-voices condition, this was not borne out by the statistical analysis. The full set of results is reported in Supplementary Table 7 of Appendix 3. Thus, there was no strong evidence in these data for a larger Hebb effect with phonologically dissimilar sequences as found in the auditory condition of Experiment 1. Nor was there evidence of an impairment of learning due to voice alternation. As the larger learning effect for dissimilar compared to similar sequences in auditorily presented conditions appeared gradually, this may not have been replicated here for single voices due to the fewer cycles included in each block. Voice alternation also appeared to not have a significant effect on sequence learning, but given the reduced effect observed in serial recall data compared to earlier findings (Hughes et al., 2009), this could have made an effect on learning undetectable.

2.4.2.4 Hebb sequence learning in Phase 2. The data from Phase 2 were analysed separately by voice condition, as interest centred on different questions for the single voice condition and the alternating-voices condition. To re-cap, interest in the single voice condition focused on the requirement for a sequential motor-plan for a response in supporting sequence learning, and the interest in the alternating-voices condition centred on the role of the passive by-voice organisation of the auditorily stimuli either into two objects according to voice-identity and its effect on learning the sequence according to voice identity compared to the true temporal order.

For each of the four blocks within a given voice condition [i.e., 2(Response-requirement) × 2(Phonological similarity)], the data of interest were those from the four trials of Phase 2: the 7th Hebb sequence (Hebb-test), the preceding filler sequence (filler-test), the re-organised 8th Hebb sequence (Hebb-transfer) and the preceding filler sequence (filler-transfer). Importantly, while the two response requirements are labelled ‘monitor’ and ‘recall’, this of course refers to the response-requirement for the six preceding presentations of the Hebb sequence; both the Hebb-test and Hebb-transfer sequences were, necessarily, recalled. To anticipate, though the usual effect of phonological similarity was found on serial recall per se, as phonological similarity did not appear to affect learning in Phase 1, or enter into any significant interactions of interest, the data presented in the following sections are collapsed over Similarity.

Single voice. The data from the single voice condition are depicted in Figure 2.7 as a function of Cycle (test, transfer), List-type (Hebb, filler) and Response-requirement (Recall-Hebb, Monitor-Hebb). The data were analysed with a 2(List-type) × 2(Response-requirement) × 2 (Cycle) within-subject ANOVA. The Hebb effect was replicated; that is, overall, collapsing across all other factors, the Hebb sequence was recalled better than the Filler sequence, $F(1, 48) = 8.19$, $MSE = .04$, $p = .006$, $\eta_p^2 = .15$. The main effects of Response-requirement, $F(1, 48) = 13.19$, $MSE = .03$, $p = .001$, $\eta_p^2 = .22$, and Cycle, $F(1, 48) = 8.16$, $MSE = .02$, $p = .006$, $\eta_p^2 = .15$, were also reliable but these effects were subsumed within a reliable three-way interaction between List-type, Response-requirement and Cycle, $F(1, 48) = 4.5$, $MSE = .03$, $p = .039$, $\eta_p^2 = .09$. This interaction reflects the fact that learning was most evident in the Hebb-test-recall-condition, that is, as predicted, recall for the Hebb sequence was better than the filler when the previous six instances of the Hebb sequence had been recalled but not when they had only been monitored. Moreover, this was not the case when the Hebb sequence had been rearranged (i.e., Hebb-transfer sequence). This interpretation is supported by the simple effects analysis shown in Table 2.6: The Hebb effect was significant in the Hebb-test-recall condition (*p* = .003). In contrast, while it would seem from Figure 2.7 that there may also have been a Hebb effect in the Hebb-transfer-monitor condition and the Hebb-transfer-recall condition, these effects missed, albeit narrowly, the standard level of statistical significance (*p* = .059 and *p* = .073, respectively). Moreover, these trends appear to be largely due to a reduction in recall of the filler-transfer sequences, which in the single voice condition should not differ from the other fillers, since all are presented in a single voice. Hence, when comparing the Hebb-transfer recall accuracy to that for the filler-test sequences, no Hebb effect is apparent. The full set of results is reported in Supplementary Table 8 of Appendix 3.

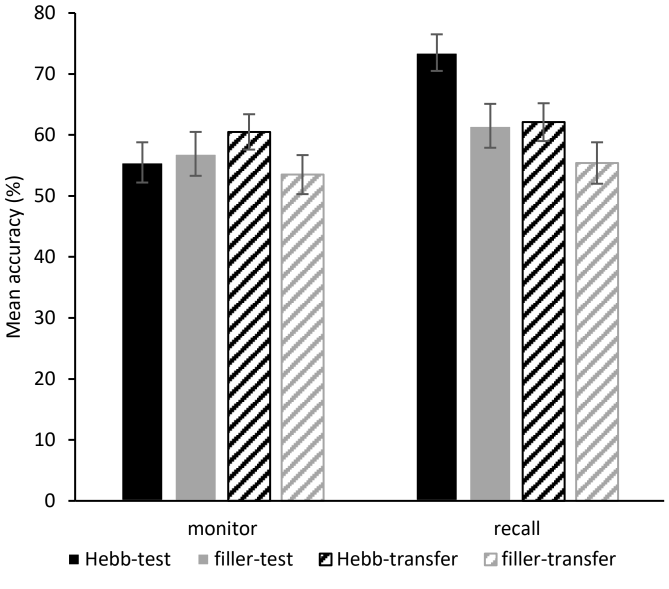


Figure 2.7. Serial recall accuracy according to List-type, Cycle, and Response-Requirement (collapsing over Similarity) in the single-voice condition in Phase 2 of Experiment 3.

Response-requirement	Cycle	Hebb (%)	Filler (%)	Magnitude of the Hebb effect (Hebb – Filler) (%)	*p*
Recall	test	73.5	61.5	12	.003
	transfer	62.1	55.4	6.7	.073
Monitor	test	55.5	56.9	-1.4	.749
	transfer	60.5	53.5	6.5	.059

Table 2.6. Hebb vs. Filler pairwise comparisons according to Response-requirement and Cycle in Phase 2 of Experiment 3.

Alternating voices. The results from the alternating voice condition (again collapsed across phonological similarity) are depicted in Figure 2.8. Again, a 2(List-type) × 2(Response-requirement) × (Cycle) within-subject ANOVA revealed a significant main effect of List-type, $F(1, 48) = 4.34$, $MSE = .02$, $p = .043$, $\eta_p^2 = .08$. Additionally, there was again a significant main effect of Cycle, $F(1, 48) = 10.91$, $MSE = .05$, $p = .002$, $\eta_p^2 = .19$, this time reflecting the expected benefit to serial recall of presenting the items in two successive voice-defined groups (in the transfer cycle) compared to presenting them in alternating voices (in the test cycle; Hughes et al., 2011). Of main interest is again a significant interaction of List-type, Response-requirement and Cycle, $F(1, 48) = 4.85$, $MSE = .03$, $p = .033$, $\eta_p^2 = .09$. As can be seen in Figure 2.8, and as predicted, when the Hebb sequence was presented in alternating voices across Phase 1, Hebb sequence learning, unlike with single voice sequences (see previous sub-section), transferred to the rearranged Hebb sequence but only when the Hebb sequences across Phase 1 had to be monitored and not when they had to be recalled. This interpretation was reinforced by a simple effects analysis reported in Table 2.7 in which it can be seen that the Hebb effect was only reliable when contrasting the Hebb-transfer-monitor condition compared to the Filler-transfer-monitor condition. One aspect of the results with alternating voices that appears odd, however, is the lack of a Hebb effect at ‘test’ when the Hebb sequence had been recalled throughout Phase 1. This is peculiar because, as evident from the analysis of learning across Phase 1 (as reported above), there was clear Hebb sequence learning despite alternating voices. There is no apparent reason why that learning effect should not also have been evident at the 'test' cycle therefore; this anomaly may be indicative of the rather noisy data that is to be expected when there is only one data-point per participant for each instance of the Hebb sequence. The full set of results is reported in Supplementary Table 9 of Appendix 3.

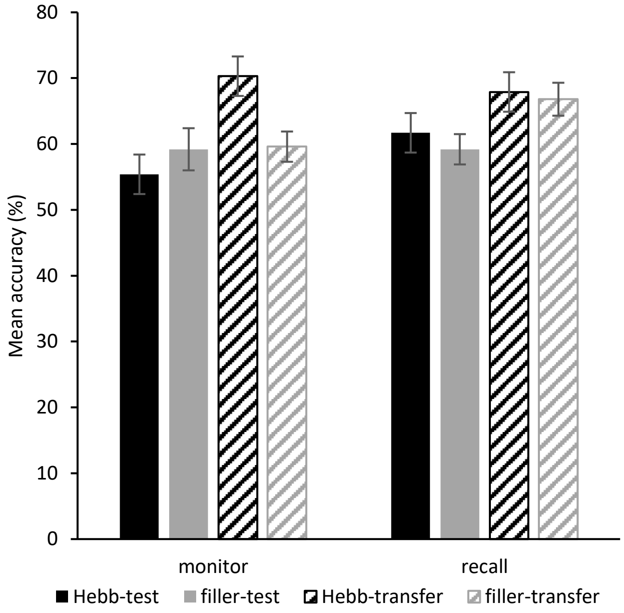


Figure 2.8. Serial recall accuracy according to List-type, Cycle, and Response-Requirement (collapsing over Similarity) in the alternating-voices condition in Phase 2 of Experiment 3.

Response-requirement	Cycle	Hebb (%)	Filler (%)	Magnitude of the Hebb effect (Hebb – Filler) (%)	*p*
Recall	test	61.7	59.2	2.5	.435
	transfer	67.9	66.8	1.1	.728
Monitor	test	55.4	59.2	-3.8	.201
	transfer	70.3	59.6	10.3	.002

Table 2.7. Hebb vs. Filler pairwise comparisons according to Response-requirement and Cycle.

2.4.2.5 Supplementary linear mixed-effect analysis of Hebb learning in Phase 2. Following the additional analysis procedure from Experiments 1 and 2, a secondary linear mixed-effects analysis was conducted. Again, the data were modified to utilise binary coding and the new data also included a factor for block-order. List-type, voice, response-requirement, and cycle were included as fixed effects. A slope for block-order for each participant was included as a random effect. Five models were constructed: The null model contained only the fixed effect for list-type, and the random effect. The main effects model contained all of the fixed effects and the random effect, but no interaction effects. The two-way interaction had the effects from the main effects model and an interaction of List-type and Cycle. The three-way interaction model contained a three-way interaction of List-type, Cycle and Response-requirement and the four-way interaction model contained a four-way interaction of List-type, Cycle, Voice and Response-requirement as well as the random effect. The main effects model was a better fit than the null model, $\chi^2(3) = 16.61$, $p < .001$, the two-way interaction model was a better fit than the main effects model, $\chi^2(2) = 6.03$, $p = .049$, and the three-way interaction was a better fit that the two-way interaction model, $\chi^2(3) = 11.24$, $p = .011$. The four-way interaction model was the best fit to the data of the tested models, $\chi^2(7) = 28.24$, $p < .001$. Although the four-way interaction was not, on its own, a significant fixed effect ($t < 1.96$), the chi-square indicated that the four-way interaction model can account for the data better than the other models. Thus, the four-way interaction model (accuracy ~ list-type × voice × response-requirement × cycle + (1 + block-order | participant) provided some support for the conclusions based on Figures 2.9 and 2.10, as the model points to the importance of including the four-way interaction terms in modelling the data. Estimates of effects, standard errors and t values for the interaction model are reported in Table 2.8. Within the model, the largest effect appear to be for the interaction of List-type and Response-requirement and the three-way interaction of List-type, Cycle and Response-requirement, which may largely account for the difference between the main-effects model and the interaction model. The three-way interaction appears related to the results from individual single-voice and alternating-voice ANOVAs, with the small estimated effect for the five-way interaction possibly related to the similar trending patterns observed in Figures 2.9 and 2.10.

Fixed effects	Estimate	SE	*t*
Intercept	0.61	0.06	9.85
List-type	0.1	0.08	1.18
Cycle	0.03	0.04	1.03
Voice	0.03	0.08	0.34
Response-requirement	0.06	0.08	0.73
List-type × Cycle	0.09	0.05	1.64
List-type × Voice	0.07	0.12	0.6
List-type × Response-requirement	0.28	0.12	2.33
List-type × Cycle × Voice	0.05	0.08	0.67
List-type × Cycle × Response-requirement	0.14	0.08	1.87
List-type × Voice × Response-requirement	0.07	0.17	0.42
List-type × Cycle × Voice × Response-requirement	0.01	0.11	0.09

Table 2.8. Estimates of fixed effects in the interaction model, with estimated differences from single to alternating voice, from test to transfer, from monitor to recall, and from filler to Hebb.

2.4.3 Discussion

The results of Experiment 3 further support the notion that motor planning and perceptual organisation are key processes in Hebb verbal sequence learning. The results relating to serial recall per se replicated not only the well-established phonological similarity effect but also the far less-studied phenomenon whereby performance is less accurate when a list is presented in alternating voices compared to a single voice (e.g., Greene, 1991; Hughes et al., 2009). Thus, these data add to the growing body of evidence highlighting the importance of the sensory-input properties for verbal serial recall performance. Interestingly, a weaker perceptual-alternation effect was observed than previously (e.g., Hughes et al., 2009, 2011, 2016). This is likely due to differences in item-presentation rate: Here, a slower presentation rate was used to encourage motor-planning of the sequences (in the recall condition) in order to increase the chances of witnessing its role in verbal sequence learning. At a faster presentation rate, it would be more likely that the influence of the auditory perceptual organisation on the motor-plan, and hence a greater cost of alternating voices, could be observed, than was seen in the current results. The previous findings had demonstrated an interaction of voice alteration and phonological similarity, whereby the two variables had non-additive effects on serial recall, presumably due to affecting the same, motor-planning, process (Hughes et al., 2009). The inability to replicate the finding here could be due to the smaller voice alternation effect that was produced in this experiment.

Turning to Hebb sequence learning, the findings align with, and further the findings supporting a key role for motor planning in verbal sequence learning. The effect of response-requirement is in line with the effect of articulatory suppression in Experiments 1 and 2. The fact that both articulatory suppression and the absence of a need for motor planning reduced sequence learning supports the interpretation of the articulatory suppression effect as impeding motor planning processes rather than exerting a general interference effect (cf. Oberauer et al., 2015). When motor planning processes are not engaged for the short-term reproduction of the sequence, learning can no longer occur as a by-product of such processes, providing converging support for the importance of motor planning. The absence of an effect of voice alternation on the Hebb effect in Phase 1, however, where the perceptual-motor mismapping of alternating-voice sequences would have been expected to affect both short-term serial recall and sequence learning, fails to support the hypothesis. However, the reduced magnitude of the voice alternation effect in serial recall—due perhaps to the longer presentation times—could account for the failure to observe an effect of voice alteration on verbal sequence learning, where the effect of mismapping may have been too small to be detected.

Interestingly, in the single voice condition, the analyses of the four different sequence types in Phase 2 indicates that when the sequence is presented as one coherent stream, learning occurs for sequences that have been recalled throughout the block, with learning attenuated for sequences that were only monitored and thus did not encourage the formation of a motor plan. At first glance, it appears that learning is not far from significance for either of the transfer-Hebb and transfer-filler comparisons regardless of response-requirement. However, this trend appears to be largely due to a reduction in performance for the 'transfer-filler', which in the single voice condition should not differ from the other fillers, since all are presented in a single voice. Hence, when comparing the Hebb-transfer recall accuracy to those for the filler-test sequences, no improvement is apparent.

In comparison, with alternating voices, transfer learning can be observed, particularly when the Hebb sequences have only required monitoring, and a motor plan corresponding to the true temporal order of the items would not have been required for successful task performance. The difference regarding the transfer learning in the monitor-condition between alternating and single voice sequences indicates that the perceptual properties influence both serial recall and Hebb sequence learning. When the perceptual properties encourage either a single auditory stream representation, or separate streams according to the properties shared by the alternating items, the resulting motor plan for reproducing the item order is affected. This appears to also suggest that when learning is reliant on passive auditory perceptual organisation, the partitioning of the sequence into the two smaller by-voice objects, compared to a single, larger object with a single voice presentation, is beneficial for learning. While these data indicate that learning with alternating voices is confined to the transfer sequences, likely only in the monitor-condition, the investigation of performance in the initial six cycles of Phase 1 did suggest that learning is also seen for the to-be-recalled Hebb sequences that are presented in alternating voices. Based on the data from Phase 1, it is concluded that the majority of the data suggest that learning of the motor plan will also occur for repeating sequences presented in alternating voices, when the formation of that motor-plan is required to facilitate the response that the task requires, that is, reproduction of the items in the correct order.

These findings show that response planning supports verbal sequence learning, directly contrasting the view that learning is not related to response-output planning or production processes, but can operate independently from them (Hitch et al., 2006; Kalm & Norris, 2016; Oberauer & Meyer, 2009). However, as noted earlier, in previous relevant studies the response-requirement was only specified after each sequence was presented (Kalm & Norris, 2016; Oberauer & Meyer, 2009), so the participant would have no way of knowing what response would be needed for successful task performance during the presentation of the sequence. As participants were able to recall the sequences when required, it suggests that they would have constructed a motor-response plan for the sequential production of the presented items. As these sequences did not, at the presentation stage, differ from those that did not require recall, participants would have been unable to predict which response was needed, and would have to be prepared for either. This would likely have resulted in the formation of a motor plan also for the sequences that did not ultimately require recall, which would be available for learning regardless of not having been executed for the immediate response. In contrast, this experiment used a paradigm where participants were informed of the response-requirement before presentation began, which should, provided that the instruction is attended to by the participant, reduce the likelihood of forming a sequential response plan for the lists that only required monitoring. This was expected to result in a clearer difference in the goal-orientation of performance-preparation between to-be-recalled and to-be-monitored sequences. This prediction was supported, demonstrated by the different learning patterns for to-be-recalled Hebb sequences and to-be-monitored sequences, particularly in the single voice condition. The finding that learning in this condition was confined to the repeated sequences that were recalled supports the need for motor planning for effective verbal sequence learning, and it is in line with the previous experiments where a clear distinction between Hebb sequences that did and did not require recall could be made (Cunningham, Healy, & Williams, 1984, Experiment 2; Glass, Krejci, & Goldman, 1989).

The interplay of perceptual organisation processes and motor planning comes apparent in similar ways as the effects of phonological similarity, as observed in Experiment 1, and the effect of alternating voices here. Phonologically similar sequences lead to originally less fluent, more error-prone, motor plans than phonologically dissimilar sequences due to being more likely to result in Spoonerism-like errors in transitions between the items of the sequence (Jones et al., 2004). Sequences presented in two alternating voices are more likely to result in two separate by-voice streams, compared to the true temporal order representation of a single stream resulting from items being presented on a common ground (i.e., single voice), resulting in a difficulty in extracting every other item for the motor-sequence plan from two different perceptual streams during the attempt to construct a motor-plan, particularly under tight time constraints. A particularly interesting comparison comes from the survival of some learning under articulatory suppression for repeated sequences that are auditorily presented and contain phonologically dissimilar items (as demonstrated in Experiment 1), and the observed learning for by-voice streams of items repeatedly presented in two alternating voices, with both suggesting that auditory perceptual organisation also supports verbal sequence learning. In short, both of these effects show a reduction to motor-plan fluency arising from characteristics of the perceptual input; a finding which has been demonstrated repeatedly in verbal short-term memory (e.g., Jones et al., 2004, 2006; Hughes et al., 2009, Macken et al., 2016). The present results suggest that these effects, in affecting the motor plan, also influence verbal sequence learning.

2.5 General Discussion

The present findings warrant a re-evaluation of the mechanisms supporting verbal sequence learning. In several converging ways, the data disconfirm major assumptions of the standard, phonological store-based, approach to verbal short-term memory and verbal sequence learning and support an alternative approach in which performance in these domains is parasitic on general-purpose motor planning and perceptual organisation processes. Experiment 1 replicated the key pattern of verbal serial recall performance that challenges the empirical basis of the phonological store construct in the context of short-term memory and which points instead to an alternative, perceptual-motor, account (e.g., Jones et al., 2006, 2004): The phonological similarity effect—the putative signature of a passive phonological store—is attenuated dramatically when motor planning is restricted (through articulatory suppression) and this, critically, is the case throughout most of the serial position curve regardless of whether the sequence is presented visually or auditorily. The survival of the similarity effect under suppression with auditory sequences is the survival of auditory recency, a phenomenon that no theorist, including proponents of the phonological loop model (Baddeley, 2007; Page & Norris, 1998), attributes to the action of a phonological store. Indeed, auditory recency and the dissimilarity effect found at this point in the curve are modulated by acoustic, not phonological, factors (Jones et al., 2006; Maidment & Macken, 2012; Nicholls and Jones, 2002). Thus, notwithstanding this acoustic-driven effect at recency with auditory sequences, the fact that the phonological similarity effect is eliminated under articulatory suppression regardless of modality suggests that the effect is a product of motor planning, obviating the need to posit a separate phonological store. In particular, the phonological similarity effect reflects speech-planning errors of the sort observed only infrequently in normal speech but exaggerated greatly in the context of the tongue-twister that is the phonologically-similar serial recall list (Acheson & MacDonald, 2009; Ellis, 1980; Jones et al., 2004; Page et al., 2007).

Another aspect of the serial recall data that supports the perceptual-motor account—not mentioned under ‘Experiment 1’ so as not to detract from the main thrust of the chapter at that point—concerns the fact that there is more than one modality effect (visual *vs.* auditory) in serial recall. Specifically, the rarely noticed inverted modality effect was replicated as well as the standard modality effect: The advantage for auditory sequences at recency was accompanied by a visual advantage at pre-recency (Beaman, 2002; Grenfell-Essam, Ward, & Tan, 2017; Macken et al., 2016). The differential interplay of articulatory suppression and each of these two modality effects was also replicated: Auditory recency remained very marked under articulatory suppression (cf. Hitch et al., 2009; Hughes et al., 2016; Jones et al., 2006, 2004; Macken et al., 2016) whereas visual pre-recency was eliminated (Macken et al., 2016). This pattern is in line with the characterisation of the classical modality effect (or auditory recency) as reflecting the opportunistic use of the strong perceptual-object boundary yielded by an auditory (compared to visual) sequence and the inverted modality effect as reflecting the greater facility with which visually-presented (and hence relatively temporally unbound) items can be assimilated into a motor-plan (Macken et al., 2016; see also Grenfell-Essam et al., 2017). Again, therefore, an appeal to perceptual and motor processes and their interplay can parsimoniously account for both the standard and inverted modality effects while proponents of the phonological-store construct explain the two effects through two different mechanisms, neither of which form part of that construct (e.g., an additional auditory-specific store to account for the modality effect, Page & Norris, 1998; the supplementary use of visual codes to account for the inverted modality effect; Baddeley & Larsen, 2007).

The output-RT data from Experiment 2 provided yet further convergent evidence for the role of motor planning in verbal serial recall (regardless of modality) as well as a role for additional passive perceptual organisation processes with auditory sequences: In the absence of articulatory suppression, output RTs aligned closely with the grouped timing of the to-be-remembered items regardless of presentation modality. However, with visually-presented lists, this alignment was greatly attenuated when motor planning was restricted by articulatory suppression. In contrast, the alignment of output- and presentation-grouping remained strong with auditory sequences despite articulatory suppression, pointing again to a contribution of passive perceptual processes to the recall of such sequences. Similarly, the serial recall data from Experiment 3 provided additional data on the reliance of verbal serial recall on motor planning and the role of perceptual-motor mapping in supporting recall. The separate-voices organised filler lists resulted in more accurate performance, where the two perceptual streams grouping the items within the sequences in a temporal manner useful for meeting the serial recall task goals supported the recall in true temporal order. In contrast, with alternating voice lists, with added perceptual variability resulting in the streaming of nonadjacent items together was detrimental for the recall of the true temporal order.

The foregoing lines of evidence for the role of motor planning and perceptual organisation in verbal serial recall provided a strong platform for testing the novel hypothesis that such general-purpose processes, rather than a bespoke phonological store (cf. Baddeley et al., 1998), also support verbal sequence learning. First, the first two experiments showed that restricting motor planning through articulatory suppression attenuates not only the short-term serial recall of a sequence but also the long-term learning of a repeating sequence, particularly a visual sequence in which there can be no contribution to learning of passive auditory perceptual organisation processes that bypass motor processes. This finding disconfirms a key assumption of phonological store-based models of Hebb sequence learning which posit that articulatory suppression impairs short-term serial recall by preventing the refreshing of decay-prone items in a phonological store but does not impair the mechanism by which a repeating sequence is learned (Burgess & Hitch, 2006; Page & Norris, 1998, 2009). In particular, the previously reported absence of a suppression effect on Hebb sequence learning has been taken, on these models, to support the idea that such learning relies on an abstract, non-motoric, learning mechanism. By the same logic, the fact that articulatory suppression does, after all, impair Hebb sequence learning points to a strong motoric involvement. A part of the differences in the current and previous findings in relation to the articulatory suppression manipulation, may be due to Experiments 1 and 2 finding an impeding effect of suppression on learning using a changing-state suppression (‘eight nine ten’), while Page et al. (2006) utilised a steady-state suppression (‘racket racket racket’), as did Hitch et al. (2009) (‘the the the’). Previous studies have found that changing-state suppression is more effective at disrupting motor-planning than changing-state suppression (Macken & Jones, 1995). The difference appears, then, to support the view that the extent to which the motor-plan is affected influences how well it can be utilised to aid both short-term verbal memory performance and verbal sequence learning. This is also corroborated by the findings of Experiment 3 which indicated that if the repeating sequence is merely observed but does not require reproduction, which would require motor planning, the learning of that sequence is attenuated. In particular, it transpires that verbal sequence learning is supported by the increasing accuracy and fluency of the motor-plan generated to retain and reproduce the sequence over the short term. Further support for this view comes from the impact of phonological similarity on the Hebb effect observed in Experiment 1. Given the evidence that the phonological similarity effect in short-term serial recall is due to speech-planning errors (Ellis, 1980; Jones et al., 2004; Page et al., 2007), the larger Hebb effect for phonologically similar sequences with visual presentation can be explained by supposing that the motor planning of such a sequence stands more to gain from repeated practice than that for an already relatively fluent phonologically dissimilar sequence (e.g., Heathcote et al., 2000). Again, phonological-store based models expressly deny that phonological similarity should modulate Hebb sequence learning due to their two-stage architectures (e.g., Hitch et al., 2009; Page et al., 2006).

A third converging line of support for the role of motor planning in Hebb sequence learning came from the relationship observed in Experiment 2 between the effect of temporal grouping on output-RTs during serial recall and grouping-inconsistency on the Hebb effect: With visual sequences—for which there was strong evidence for motoric grouping during serial recall that mimicked the presentation-grouping—grouping-inconsistency across the repetitions of the Hebb list eliminated the Hebb effect. Additionally, the goal-relevance of producing a motor-plan for the immediate reproduction of the repeating sequence is also important in comparing the findings of Experiment 3 on the effect of response-requirement and the effect of suppression demonstrated in Experiments 1 and 2. Both the requirement for articulatory suppression in the first two experiments, and the need to only monitor, rather than recall the sequences here, reduced or abolished verbal sequence learning. Both of these effects act on the motor-planning processes, but in different ways: While articulatory suppression reduces the availability of the motor-planning processes when they are required for successfully meeting current task goals, the requirement to only monitor makes motor-planning less relevant for the immediate task performance. Both of these aspects characterise motor-planning as an active, positive process used to generate behaviour and to increase fluency of current and future motor-behaviours, rather than as a process used to offset the negative effect of item-decay.

Turning to the additional contribution of passive auditory perceptual processes to verbal sequence learning with auditorily presented sequences, this was apparent in the finding that learning was in general attenuated to a lesser extent by suppression with such sequences compared to visually-presented sequences. There was also some indication of the added contribution of auditory perceptual organisation when motor planning was unrestricted however: First, phonologically dissimilar sequences—for which there is independent evidence for strong passive auditory order-encoding (e.g., Jones & Macken, 1995)—were particularly well learned with auditory compared to visual presentation (Experiment 1). Second, when perceptual organisation and motor planning co-occurred (as opposed to either operating in isolation), learning was resistant to the otherwise deleterious effect of an inconsistent temporal grouping of the Hebb sequence across repetitions (Experiment 2). This resilience of learning of an auditory sequence in the face of changes in its temporal structure poses yet another difficulty for phonological store-based models (Burgess & Hitch, 2006; Page & Norris, 2009): An interaction between input modality and grouping inconsistency is at odds with the assumption that the learning mechanism is insensitive to input modality. This finding does, however, align with the perceptual-motor account's characterisation of the auditory temporal grouping effect on short-term memory as an auditory-perceptual effect.

In Experiment 3, learning of the 'transfer-Hebb' sequence that was presented in alternating voices and merely monitored without requiring sequential recall was observed. Learning was indicated to occur for the nonadjacent items of a sequence when those items were grouped to two different perceptual streams according to a common voice identity. That both serial recall and sequence learning were affected by how the items are grouped into perceptual streams, with alternating voice sequences both diminishing serial recall according to true temporal order while also promoting the learning of the separate perceptual streams further indicates a role for perceptual organisation in verbal sequence learning. This is again, inconsistent with the modular models where, once the items enter the phonological store, they should be equal and stripped of their specific perceptual properties. If this were the case, learning of the nonadjacent items should not depend on whether the sequence was presented in a single or alternating voice, yet no reliable learning of the transfer sequence was observed with sequences presented in a single voice. The primacy model (Page & Norris, 2009), could possibly account for the learning of the by-voice sequence with the addition of voice as a front-end streaming parameter (Page & Norris, 1998). However, it is not clear why this would need to be a separate front-end process to a phonological store. A phonological store is not necessary, if learning is accounted for with an appeal to the perceptual organisation of auditory items into a stream according to their perceptual properties.

Taken together, the results of the present chapter suggest that learning verbal sequences could be explained as entirely or largely dependent on the formation of perceptual-object and motor-plans formed during short-term performance, potentially without the need to appeal to the operation of a dedicated storage module. They also highlight some of the contradictions in the explanations of the role of a phonological store in learning in different paradigms, namely how the explanations based on evidence from the Hebb repetition paradigm contrast with those from paired-associate learning and the relation between vocabulary size and nonword repetition ability. In the paired-associate paradigm, learning of nonwords is impaired by articulatory suppression, which, on the phonological store account, impairs the articulatory control process, and phonological similarity, which impairs the phonological store (Baddeley et al., 1998; Papagno & Vallar, 1995; Papagno et al., 1991). However, these findings would also support a perceptual-motor account of learning. Additionally, this conflicts with the models of Hebb sequence learning based on the phonological store: The learning mechanism, either a positional context signal (Burgess & Hitch, 2006) or a primacy gradient (Page & Norris, 2009), is considered separate from the phonological store, which only supports the learning by holding the items in the short term. This separation leads to the learning mechanism, and by extension verbal learning, being immune to the factors that affect the phonological store, apparently directly opposing the reasoning based on paired-associate learning.

The next chapter will turn to investigate some of these contradictions between the phonological store models' account of the Hebb repetition effect and paired-associate learning, and the contrasting explanations of word learning based on the phonological store-based and perceptual-motor accounts.

CHAPTER 3

WORD LEARNING IN THE PAIRED-ASSOCIATE LEARNING TASK: A RE-EXAMINATION

Abstract

That nonword learning in a paired-associate (*word—nonword*) learning task is impaired by phonological similarity constitutes a further key strand of evidence for the PS-LLD hypothesis. Yet this appears to contradict the classic *differential coding argument* for a separation of short- from long-term memory (where only short-term recall should exhibit phonological effects) as well as phonological store-based models of (Hebb) verbal sequence learning, in which factors that affect the phonological store should not affect learning. The possibility of a reprieve for the phonological-store account was explored by examining whether the phonological similarity effect in paired-associate learning is due to a difficulty in pairing the correct nonword to the known word rather than a difficulty in recalling the nonword-forms per se. This was done in Experiment 4 by examining whether a phonological similarity effect in nonword learning is also found in free recall (in which there are no paired-associates) and when similarity was manipulated within (as well as across) nonwords. This was indeed the case, leaving the contradictions for the phonological-store account in place. Experiment 5 demonstrated that while both paired-associate recall and free recall are impaired by articulatory suppression regardless of presentation modality (visual vs. auditory), learning was only impaired in the paired-associate task; it is speculated that this may have been due to response-phase motor planning disproportionately supporting learning in free recall. Experiment 6 replicated some of the key features of Experiment 4 but also, using functional near-

infrared spectroscopy (fNIRS), provided some evidence of differential activation of the pre-supplementary/supplementary motor areas of the brain as a function of across-nonword similarity, though ultimately the fNIRS data proved difficult to interpret and can only be viewed as preliminary. Overall, the results provide some further support for a key role for motor planning in verbal sequence learning but are, when considered together with theorising and evidence from the Hebb repetition effect, difficult to reconcile with a phonological store-based approach.

3.1 Introduction

The results of the experiments reported in Chapter 2 challenged the PS-LLD hypothesis in the context of Hebb sequence learning and instead supported key roles for motor planning and perceptual organisation processes in verbal sequence learning. In a similar vein, the experiments in this chapter re-examine the view that findings from the paired-associate learning paradigm also support the role of the phonological store in verbal sequence learning (Baddeley et al., 1998; Papagno et al., 1989; Papagno & Vallar, 1992).

Historically, the concept of a phonological short-term store emerged from, and in turn reinforced, the idea of separate short- and long-term memory systems (Atkinson & Shiffrin, 1968; Baddeley, 1966; Baddeley & Hitch, 1974; Broadbent, 1984). A key piece of evidence for the need to divide the cognitive architecture into a short-term store—and specifically a phonological short-term store—capable of holding information only for a very short period, and a long-term memory system storing information potentially for a lifetime was *differential coding*: that short-term memory tasks seemed to be performed using a phonological code while long-term remembering and learning seemed to rely on a semantic code (Baddeley, 1966, 1990; Baddeley & Dale, 1966). The key evidence for this was that short-term memory was found to be impaired by phonological similarity between memoranda but was relatively unaffected by semantic similarity, while the opposite pattern was evident in long-term memory performance (e.g., Baddeley, 1966).

The current view is that the phonological store evolved not for the short-term storage of a series of already-known phonological entities (e.g., words, letter-names) but for the learning of new ones (the PS-LLD hypothesis; Baddeley et al., 1998, 2012). And indeed, the fractionation of long-term and short-term memory systems persists in contemporary computational models of the phonological loop and its role in long-term learning that are based on the Hebb repetition paradigm (cf. Chapters 1 and 2). Here, effects considered to be empirical signatures of a short-term phonological store—such as the phonological similarity effect—are not thought to be characteristic of long-term learning (e.g., Hitch et al., 2009; Page et al., 2006). However, another major strand of evidence cited as support for the PS-LLD hypothesis—that from the paired-associate learning paradigm—appears to undermine the case for a separate phonological store based on differential coding (Baddeley, 1966, 1990) as well as the related supposition that long-term phonological learning should not be affected by the same variables as short-term phonological storage (Burgess & Hitch, 2006; Page et al., 2006). The evidence in question comes from a variant of paired-associate learning in which the learning of nonwords, i.e., their transfer into long-term memory, is examined by presenting a series of nonwords each of which is paired with a known word (e.g., Papagno & Vallar, 1992). After each presentation of a list of these pairs, participants are required to recall each nonword when cued by the presentation of its associated known-word. This requires each nonword to be recalled with the correct cue, otherwise even a correctly remembered nonword would not be scored as such. The same list of pairs is then repeatedly presented for recall over several cycles to examine the long-term learning of the nonwords. The paradigm is assumed to be a laboratory analogue of the phonological word-form learning component of language acquisition (Papagno et al., 1991; Baddeley, 2003).

Critical for the present discussion is that phonological similarity between the to-be-learned nonwords not only impairs their short-term recall—as indicated most clearly by performance on the first trial—but also results in poorer long-term memory performance as reflected in slower across-trial paired-associate learning. Specifically, Papagno and Vallar (1992) examined the effect of phonological similarity on associate learning of both *word—word* and *word—nonword* pairs. Phonological similarity was manipulated in terms of between target-item similarity. Thus, the target-items were either phonologically similar to each other while the cues were all dissimilar [e.g., *tetto* (‘roof’ in Italian) – *zibro* (a nonword target); *bocca* (‘mouth’) – zuro (a nonword target)], or the target-items as well as the cues were phonologically dissimilar to one another [e.g., *uva* (‘grapes’) – *raggo* (a nonword target); *pesca* (peach) – *patro* (a nonword target)]. It was found that between-target phonological similarity had a larger disruptive effect on the learning of *word—nonword* pairs than on the learning of *word—word* pairs. It was concluded that word-nonword associate learning relies on phonological short-term coding while the semantic representations already associated with known words circumvents the need for such coding in word-word learning.

The phonological similarity effect in the context of word-nonword associate learning has been taken as evidence that the phonological store supports new word learning (Baddeley, 2003; Baddeley et al., 1998). However, the finding that long-term memory performance is, after all, susceptible to phonological similarity appears to undermine the case for the existence of a distinct short-term phonological short-term store based on differential coding (Baddeley, 1990). It also appears that the effect of phonological similarity on paired-associate learning contradicts contemporary phonological store-based models of word-form learning based on the Hebb effect which predict no effect of phonological similarity on such learning (e.g., Hitch et al., 2009; Page et al., 2006).

One possibility that would at least resolve the apparent contradiction with phonological-store based models of Hebb sequence learning is that the phonological similarity effect in word-nonword paired-associate learning does not in fact demonstrate a role for the phonological store in word-*form* learning (e.g., Baddeley et al., 1998). Given that the dependent variable in the relevant word-nonword associate learning studies was the number of nonwords correctly paired with their cue words (Papagno & Vallar, 1992; Papagno et al, 1991), diminished learning of phonologically similar nonwords may have solely reflected a difficulty in learning which nonword was paired with which cue-word (cf. general similarity-based interference; e.g., McGeoch, 1932; Postman, 1961; Watkins & Watkins, 1975, 1976) rather than a difficulty in learning the forms of the nonwords per se. Indeed, the fact that phonological similarity has, to date, only been implemented *across* the nonwords as opposed to within each nonword reinforces this: the learning of the internal phonological structure of each nonword—which is what the phonological store is said to support (i.e., the learning of new word-forms)—has never been directly tested in this setting.

The present study, therefore, examined whether phonological similarity in the paired-associate learning task does indeed impair word-form learning specifically, as posited by the PS-LLD hypothesis. This was done in two ways. First, performance in the classic paired-associate learning task (e.g., Papagno & Vallar, 1992) was contrasted with that in a task in which the cue words were absent. That is, serial learning of word-nonword paired-associates was contrasted with the serial learning of nonwords presented for free recall. Free recall, rather than serial recall, was used so that the order of the nonwords could be randomised for each presentation, as is the case in paired-associate learning. The phonological similarity effect, though typically studied in serial recall, has also been observed in free recall (e.g., Coltheart & Langdon, 1998). If the phonological similarity effect found previously in paired-associate learning does indeed reflect poorer learning of the nonword-forms per se, and not simply cue-driven interference, then it should still be strongly apparent in serial learning when the nonwords are learned in a free recall task[5]. Second, across these two tasks, phonological similarity was manipulated in two ways orthogonally: The phonological similarity of the nonwords to one another (across-nonword similarity)—as has been implemented previously (e.g., Papagno & Vallar, 1992)—was crossed with a manipulation of the phonological similarity of the syllables *within* each nonword (within-nonword similarity). Surprisingly, the latter manipulation has not, to my knowledge, been implemented even though this would appear to be the more apposite manipulation of phonological similarity for testing the hypothesis that the phonological store is involved in word-form learning: phonological similarity between the syllables within a given nonword should affect the retrievability of the nonword and hence its long-term learning.

The implications of various possible outcomes from the foregoing manipulations may be summarised as follows. If a phonological similarity effect is found on learning in either the paired-associate or the free recall task and with either manipulation of phonological similarity, this would appear to undermine the case from differential coding for a fractionation of a (phonological) short-term store from a (non-phonological) long-term store. If instead a phonological similarity effect is only found in paired-associate learning (and not on learning in free recall) or only when phonological similarity is implemented across (as opposed to within) the

---

[5] Whilst using free recall, from some theoretical standpoints, may not entirely remove an effect of cue-driven interference (see, e.g., Neath & Surprenant, 2005), it seems reasonable to expect that removing the explicit cues would lessen the contribution of such interference.

nonwords, this would suggest that the phonological similarity effect in the paired-associate learning task is not an effect on word-form learning specifically. As such, this would resolve the contradiction with models of the phonological store's role in new word-form learning based on the Hebb effect. However, it would also mean that the phonological similarity effect observed previously in paired-associate learning could no longer be taken as evidence for the PS-LLD hypothesis. Finally, if a phonological similarity effect is found regardless of task and regardless of the particular manipulation of similarity, the contradiction with models of the Hebb effect would remain but it would leave the PS-LLD hypothesis intact (as applied within the paired-associate learning paradigm). However, in light of the reconceptualization of the phonological similarity effect in serial recall in perceptual-motor terms (Jones et al., 2006, 2004; cf. present Chapters 1 and 2), this latter outcome could not be taken as uniquely supporting the PS-LLD hypothesis. Specifically, as discussed in previous chapters, according to the perceptual-motor view, the phonological similarity effect in serial recall is primarily an articulatory effect rather than phonological, where 'phonological' similarity results in a less fluent, more error-prone, motor-plan, along with an additional acoustic similarity effect with auditorily presented material (e.g., Jones et al., 2004). Thus, to the extent that the phonological similarity effect is shown regardless of task or particular manipulation of similarity, this chapter will also investigate whether it can be understood alternatively in terms of the disfluency of the motor-plan for the articulation of the nonwords.

3.2 Pilot Experiments

Given that learning in the paired-associate recall task has not previously been compared directly to free recall, four pilot experiments were conducted to assess the feasibility of examining and comparing various implementations of phonological similarity across the two tasks. (Note that the method for these is not described here in full detail as it was carried forward to the experiments proper and so is described in detail later). The participants were students at Royal Holloway, University of London, and were given course credits for their participation. Pilot 1 had a 2 (Phonological similarity) × 2 (Recall-type) × 10 (Cycle) repeated-measures design. Phonological similarity was only manipulated across nonwords in this case (i.e., the nonwords sounded either similar or dissimilar to each other) as in previous studies (Papagno et al., 1991). In the paired-associate recall blocks, the nonwords appeared one at a time together with an English word, with each of the English words then used as recall cues for the nonwords. Participants were instructed to recall the nonword that had appeared together with the English cue word. In the free recall blocks, nonwords appeared on the screen one at a time without a paired word, and after all eight nonwords were presented, they were prompted with ‘recall’ to try and recall as many of the nonwords as they could in no particular order. Recall was spoken for all experiments reported in this chapter, including these pilot experiments, and recorded with the participants’ consent. Although only 5 participants took part (and the data not therefore analysed statistically), eyeballing the data was sufficient to suggest that learning—seen in the increased number of correctly recalled words across cycles (see Appendix 4, Supplementary Figure 1)—occurred in all four conditions and that, importantly, phonological similarity appeared to impair recall and learning in a similar way in both recall-type conditions.

The design of Pilot 1 was adapted a little for Pilot 2 (*n* = 4) to reduce the participants’ difficulty in keeping track of the number of responses given in each free recall trial, which may have accounted for the slightly lower overall accuracy in free recall compared to paired-associate recall in Pilot 1. This also served to broadly equate the timing of the responses across the two recall tasks: The free recall participants were now prompted with the numbers 1-8 which were matched with the pacing of the word-cues in the paired-associate recall task, but they were still free to recall any one of the nonwords for each number. The data suggested a slight increase in accuracy for the free recall task from Pilot 1 to Pilot 2 and informal inspection of the data again indicated an effect of phonological similarity regardless of recall task (see Appendix 4, Supplementary Figure 1).

Pilot 3 (*n* = 6) was conducted to get an initial indication of whether phonological similarity manipulated *within* the nonwords would exert an effect, using similar syllables to construct the nonwords in comparison to the dissimilar syllables within the nonwords in Pilots 1 and 2. The design was identical to Pilot 2, but now all the nonwords in each of the four conditions comprised two or three syllables with a common vowel sound (i.e., pir-lin-vin). The intention was to assess the effect of within-nonword similarity by informally comparing the data from this pilot to those from Pilot 2. Across-nonwords similarity, as implemented in Pilots 1 and 2, was manipulated as a within-participants factor, as before. Comparison of data from Pilots 2 and 3 indicated that both types of similarity impaired learning regardless of recall-type (see Appendix 4, Supplementary Figure 2).

Finally, Pilot 4 (*n* = 3) replicated Pilot 2 and was designed to check, again informally, whether at least the across-nonwords similarity effect on learning could be observed in both free and paired-associate recall with auditory and not just visual nonwords, which did indeed appear to be the case (see Appendix 4, Supplementary Figure 3). The relevance of establishing this will become clear in the context of Experiment 5.

The pilot experiments therefore established the viability of studying the effects of both within- and across-nonword similarity in both paired-associate and free recall, and across auditory and visual modalities.

3.3 Experiment 4

The first experiment proper in this chapter examined the possibility that the phonological similarity effect observed previously in paired-associate learning (e.g., Papagno & Vallar, 1992; Papagno et al., 1991) might be due to a greater difficulty in learning the word-nonword pairings with similar nonwords rather than an effect on the learning of the internal forms of the nonwords. As in Pilots 2 and 3, then, participants were required to free recall a list of eight nonwords (cued with the numbers 1-8) or to recall the nonwords when cued with a known-word paired with them during presentation (paired-associate recall). The nonwords could either be phonologically similar or dissimilar to one another (across-nonword similarity manipulation). Unlike in the pilot experiments, across-nonword phonological similarity was manipulated orthogonally with within-nonword phonological similarity in this experiment: the nonwords comprised syllables each of which had a different vowel sound (within-dissimilar) or contained at least two syllables that shared a vowel sound (within-similar).

A further issue addressed in the present experiment was the extent to which the learning of new words depends on the size of the learner’s current vocabulary. Positive correlations between vocabulary size and nonword repetition and learning measures have indicated a link between current vocabulary-size and the ability to repeat or recall new verbal items (e.g., Baddeley et al., 1998; Gathercole, 2006). Comparisons of participants with and without dyslexia and of monolingual and bilingual speakers have indicated that speaker-profiles have an impact on paired-associate learning, Hebb sequence learning and serial recall (Bogaerts et al., 2015; Duyck et al., 2003; Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995). In particular, it has been suggested that a larger vocabulary enhances learning of new words as it increases the likelihood that new words will share phoneme sequences and syllables with already known words. However, another plausible candidate mechanism for the multilingual advantage may be greater skill in 'online' novel motor-planning rather than vocabulary size per se (e.g., Gernsbacher, Sauer, Geye, Schweigert, & Goldsmith, 2008). The manipulation of both across- and within-phonological similarity as well as a comparison of monolinguals and multilinguals in the present experiment may help to determine whether the multilingual advantage is due to a larger vocabulary-size or whether it results from better novel motor planning. On the vocabulary-size hypothesis, across-similarity, by adding confusability between the nonwords, may not affect multilinguals as much as monolinguals as a larger vocabulary may provide more long-term reference-points for the across-similar nonwords. Better novel motor planning may have potential for reducing the effect of within-similarity between nonwords. Within-similarity may be expected to affect motor planning fluency more directly by introducing error-proneness within the nonwords. Multilinguals—to the extent that they enjoy greater novel motor-planning fluency—should be better equipped to deal with within-similar nonwords, and hence show a reduced within-similarity effect compared to monolinguals. Experiment 4 therefore included a comparison of monolingual English speakers and multilinguals who spoke at least one other language in addition to English.

3.3.1 Methods

3.3.1.1 Participants. Estimations of effects sizes in the paired-associate learning task appear to indicate large effect sizes for different manipulations in the paired-associate task from phonological similarity (Papagno & Vallar, 1992) to monolingual-multilingual comparisons and *word—word* compared to *word—nonword* pairs (Papagno & Vallar, 1995). However, exact Cohen's *d* values could not be calculated due to a lack of the required information in the reports of the relevant studies. Given this, I aimed for a sample size for each between-participants group of around 20-25 in order to resemble those from previous experiments that have successfully demonstrated interactions within the paired-associate task: 12-25 participants for detecting a phonological similarity effect (Baddeley, 1966; Papagno & Vallar, 1992) and 20 participants for detecting differences between monolinguals and multilinguals (Papagno & Vallar, 1995). For Experiment 4, then, fifty-two participants (4 males, 48 females), students at Royal Holloway, University of London, aged 18-36 years (mean 19.54, SD = 4.01) originally took part, 43 of whom were native English speakers and were included in the analyses (including 20 self-reported monolinguals). All participants reported their native language as well as all the other languages they spoke. Two participants reported having dyslexia. The 9 non-native English-speakers were not included in the analysis[6]. Participants received either course credits or a small honorarium for their participation.

3.3.1.2 Apparatus and Materials. The experiment was conducted using E-Prime software (Psychology Software Tools, Pittsburgh, PA) running on a PC. The visual stimuli were presented on a flat monitor and a microphone was used for

---

[6] Due to external constraints, participation could not be restricted to native English speakers. In opening the participation to more native monolingual English speakers, other additional participants had to be tested as well, increasing the number of both non-native English speakers and multilinguals.

recording participants' vocal recall responses. The to-be-learned nonwords consisted of a random ordering of nonwords that were either similar or dissimilar to one another (the Across-Nonword Similarity factor) and each nonword comprised syllables that were either similar or dissimilar to one another (the Within-Nonword Similarity factor). Two lists of eight English words and eight lists of eight pronounceable two- or three-syllable nonwords were constructed. Four of the nonword lists contained nonwords constructed from syllables that were dissimilar, such as *hawerty* (within-dissimilar), and the other four nonword lists contained nonwords in which the syllables were similar (shared a vowel sound), such as *pirlinvin* (within-similar). In each set of four lists, two sets were made up of nonwords that were dissimilar to each other, e.g., *hawerty* and *samtis* (across-dissimilar) and two sets in which the nonwords were similar to each other, e.g., *silaries* and *solories* (across-similar). The words and nonwords used are listed in Appendix 5, Supplementary Table 10. In the cued-recall condition, the word and nonword were presented simultaneously in the middle of the screen for 2 s (e.g., giraffe - hawerty) while in the free recall condition, only the nonword appeared in the middle of the screen. There was an inter-stimulus-interval of 1 s. After stimulus presentation, separated again by the 1 s interval, the recall cues (either the numbers 1-8 or the English words) appeared one at a time, for 4 s each, again separated by a 1 s interval.

3.3.1.3 Design. There were five factors: Recall-type (paired-associate, free), Across-similarity (whether the nonwords were phonologically similar or dissimilar to each other) and Cycle (referring to each of the ten successive trials consisting of the presentation and recall-test of the same 8 nonwords) were all within-participant factors while Within-similarity (whether the syllables within each nonword were similar or dissimilar) was manipulated between participants. Number of languages spoken was included as a further two-level factor (monolinguals, multilinguals). This left 10 monolinguals and 16 multilinguals for the Within-similar group, and 10 monolinguals and 7 multilinguals for the Within-dissimilar group. There were unequal numbers of participants across the between-participants groups due to difficulties recruiting enough monolingual English speakers, and because participants were randomly assigned to either of the Within-similarity groups. Each participant completed four blocks of ten trials in total, where all nonwords were either within-similar or within-dissimilar. The four blocks in each within-similarity condition were presented in four possible orders: A1-B2-A2-B1, A2-B1-A1-B2, B1-A2-B2-A1 or B2-A1-B1-A2, where A=paired-associate recall, B=free recall, 1=across-dissimilar, 2=across-similar. Any particular list of nonwords was presented either in the paired-associate recall or the free recall condition, counterbalanced across participants, and the two sets of English words were presented either with across-similar or across-dissimilar nonwords.

3.3.1.4 Procedure. Participants were tested individually in a quiet room. At the beginning of the experiment, participants gave informed consent and were then given task instructions. The instructions included a description of each of the recall task-types, explained that they would be required to respond vocally and that their responses would be recorded for subsequent scoring, and described the four-block structure of the experiment. Participants were also informed that they would have ten trials of each of the four lists, and that after each trial they should try to recall (in any order for free recall, and according to the associated known-word in paired-associate recall) as many of the nonwords as they could. The similarity manipulations were not mentioned in the instructions. Before each block, participants were informed whether the block would involve free or paired-associate recall. Each trial commenced with a button-press made by the participants. The nonwords or *word-nonword* pairs were presented in a different random order on each cycle, as were the word-cues in the recall phase of the paired-associate recall blocks. Following the experiment, participants were debriefed. The entire session lasted approximately 50 min.

3.3.2 Results

Each vocalised recall response was scored as correct if it contained each of the syllables of the presented stimulus in the correct order. In addition, for paired-associate recall, the spoken response had to have been produced following the correct known-word cue. Figure 3.1 shows the number of nonwords correctly recalled according to recall-type, within-similarity and across-similarity across the ten cycles collapsed across language groups. It is clear that both recall and learning was diminished in the presence of both within- and across-nonword phonological similarity and regardless of recall-type.

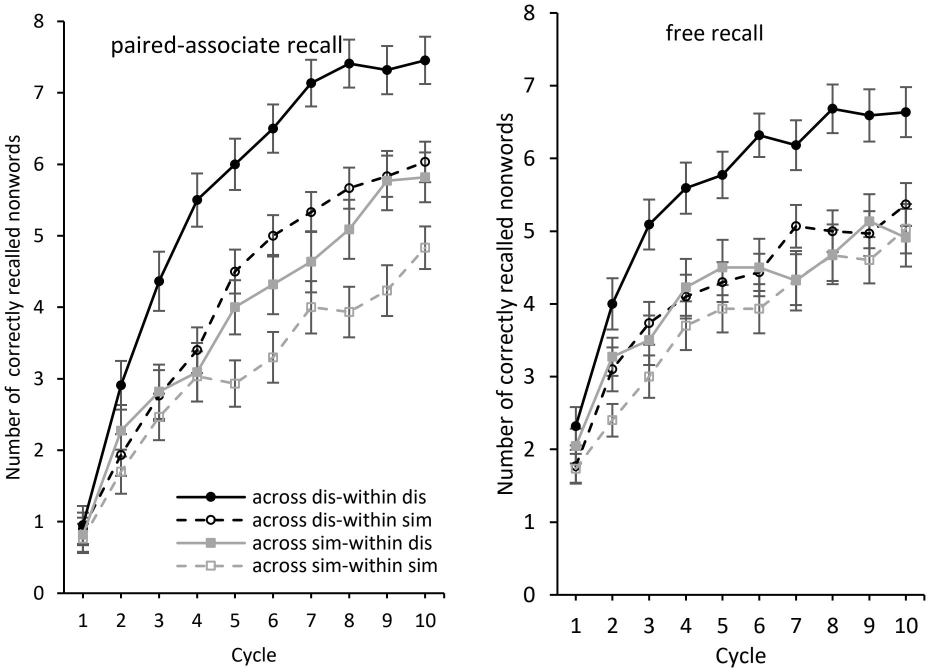


Figure 3.1. Recall accuracy collapsed across language-groups according to Recall-type, Across-similarity and Within-similarity across cycles in Experiment 4. Error bars represent standard error.

Recall accuracy was analysed with a 2 (Recall-type) × 2 (Across-similarity) × 2 (Within-similarity) × 2 (Language-group) × 10 (Cycle) ANOVA, where Within-similarity and language-group were between-participants factors. All main effects were significant, with a general improvement in recall across cycles, $F(5.3, 207.8) = 256.37$, $MSE = 2.04$, $p < .001$, $\eta_p^2 = .868$, higher recall accuracy in free recall compared to paired-associate recall, $F(1, 39) = 4.19$, $MSE = 6.63$, $p = .048$, $\eta_p^2 = .097$, and across-similar lists recalled more poorly than across-dissimilar lists, $F(1, 39) = 79.56$, $MSE = 7.27$, $p < .001$, $\eta_p^2 = .671$. The within-similar nonwords were also recalled more poorly than within-dissimilar nonwords, $F(1, 39) = 8.89$, $MSE = 42.23$, $p = .005$, $\eta_p^2 = .168$. While not shown in Figure 3.1, the main effect of language-group was also significant, with multilinguals recalling reliably more nonwords (M = 4.55, SD = .23) compared to monolinguals (M = 3.88, SD = .23), $F(1, 39) = 4.18$, $MSE = 42.23$, $p = .048$, $\eta_p^2 = .097$.

There was also a significant interaction between the two types of similarity, $F(1, 39) = 7.35$, $MSE = 7.27$, $p = .01$, $\eta_p^2 = .159$, with highest accuracy for dissimilar-dissimilar lists, followed by across-dissimilar lists with within-similar nonwords, then lists that were across-similar with within-dissimilar nonwords, and least accurate for similar-similar lists, with the two types of similarity having non-additive effects on performance. Thus, the interaction appears to reflect a smaller effect of within-nonword similarity with across-similar nonwords compared to across-dissimilar nonwords. Across-nonword similarity seems to have an independent effect on performance, dampening the effect of within-nonword similarity. Importantly, a significant interaction of Within-similarity and Cycle showed that within-similarity also impaired learning (and not just recall) whereby the difference in recall accuracy between the dissimilar and similar nonwords increased generally across cycles, $F(5.3, 207.8) = 4.18$, $MSE = 2.04$, $p = .001$, $\eta_p^2 = .097$. Also of particular importance is that learning was also impaired by across-similarity as indicated by the interaction of Across-similarity and Cycle, $F(9, 39) = 7.68$, $MSE = .99$, $p < .001$, $\eta_p^2 = .165$. There was also a trend towards a three-way interaction of Within-similarity, Across-similarity, and Cycle, $F(9, 39) = 1.89$, $MSE = .99$, $p = .053$, $\eta_p^2 = .046$. This is likely to reflect, at least in part, the effective floor-level performance for the initial cycle regardless of similarity and, in the later cycles, the diminution of the differences between the conditions as performance gets closer to a ceiling effect, with both similarity manipulations having the largest effect during the steepest parts of the learning curves, reflecting the independent effects of the two types of similarity. As the effect of within-similarity becomes more apparent as a function of cycle, this is enhanced for across-dissimilar lists compared to across-similar lists.

There was also a significant interaction of Recall-type and Cycle, mostly reflecting the enhanced recall with free compared to paired-associate recall at early cycles, but also possibly reflecting the advantage for paired-associate recall over free recall at the last few cycles, $F(9, 39) = 10.93$, $MSE = 1.06$, $p < .001$, $\eta_p^2 = .219$. This initial benefit for free recall is likely to reflect the fact that in this task the last few presented items can be output first (e.g., Murdock, 1962; Grenfell-Essam & Ward, 2012) whereas the output position of the nonwords in the paired-associate recall condition was determined extrinsically by the requirement to output the nonwords in accordance with their correct associates. Indeed, a supplementary analysis of the free recall data showed that the probability of the last-presented item being output first was reliably greater (.48, SD = .19) than the probability of the first-presented item being output first (.11, SD = .1), $t(36) = 8.53$, $p < .001$. The steeper learning curve for paired-associate compared to free recall at later cycles may reflect the increasing utility of the known-word cues in the former task as the pairings are learned.

There was no significant interaction between Recall-type and either type of similarity [Recall-type and Across-similarity, $F(1, 39) = 3.01$, $MSE = 14.35$, $p = .091$, $\eta_p^2 = .072$, and Recall-type and Within-similarity, $F(1, 39) = 1.36$, $MSE = 6.63$, $p = .25$, $\eta_p^2 = .034$], indicating that recall in the two tasks was not differentially affected by the manipulations of similarity. There was, however, a significant interaction between Recall-type, Across-similarity, and Cycle, $F(9, 39) = 2.19$, $MSE = .88$, $p = .023$, $\eta_p^2 = .053$, which appears to reflect the fact that the impairment of learning due to across-similarity was in general (when collapsing across levels of within-similarity) greater in the paired-associate task than in the free recall task. This suggests that some of the across-similarity effect in paired-associate recall may be attributable to cue-based interference, though a robust effect of across-similarity was also shown in free recall, indicating that the effect does also operate on word-form learning. The fact that this was particularly the case in the within-similar condition accounts for a further, four-way, interaction of Recall-type, Across-similarity, Within-similarity, and Cycle, $F(9, 39) = 2.34$, $MSE = .88$, $p = .014$, $\eta_p^2 = .057$. No other interactions were significant. Of particular note, language-group did not interact reliably with any other factor. Thus, whereas multilinguals, as reported, enjoyed a recall advantage, they did not exhibit enhanced learning. There was also no significant interaction of Language-group with either type of similarity. The full set of results is reported in Appendix 5, Supplementary Table 11.

3.3.3 Discussion

The results of Experiment 4 showed that the learning of nonwords is affected by their phonological similarity regardless of whether the nonwords need to be matched to their respective known-word associates (paired-associate recall) or do not (free recall). Moreover, this was the case regardless of whether phonological similarity was manipulated across the nonwords (as in previous experiments; Papagno et al., 1991; Papagno & Vallar, 1992) or within the nonwords. Additionally, regardless of recall-type, the nonwords were only scored as correct if the internal structure of the nonword produced by the participant matched closely to the structure of the presented nonword, differing from the scoring in previous studies of paired-associate learning (e.g., Papagno & Vallar, 1992), which did not appear to be so concerned with the accuracy of the internal structure of the nonword responses. These aspects of the results converge to suggest more clearly than hitherto that there is indeed a phonological similarity effect on the learning of the internal structure of nonwords. The results are consistent, therefore, with the PS-LLD hypothesis (Baddeley et al., 1998). Specifically, on this account, the phonological store holds a representation of a nonword while a longer-term representation of it is formed (e.g., Papagno et al., 1991).

However, as argued earlier, the existence of a phonological similarity effect in nonword learning would appear to contradict one of the empirical foundations of the idea of a separate short-term phonological store, namely, that there should be little, if any, effect of phonological similarity on long-term memory performance (i.e., the case from differential coding; Baddeley, 1966, 1990). A phonological similarity effect on nonword learning is also at odds with phonological store-based models of verbal sequence learning based on the Hebb effect which predict that variables that affect short-term recall such as phonological similarity should not impair such learning (Hitch et al., 2009; Page et al., 2006). It seems that a more parsimonious interpretation, therefore, is that the phonological similarity effect in both the paired-associate (and free recall) learning paradigm and the Hebb paradigm reflects the common role of motor planning in both short-term memory performance (e.g., Jones et al., 2004; present Chapter 2) and long-term sequence learning (cf. Chapter 2). From this perspective, motor planning underpins performance in the short-term task and learning occurs as the by-product of the need to plan and reproduce the nonwords repeatedly. From this standpoint, the effect of within-nonword phonological similarity on learning could be understood in terms of a greater number of speech-planning errors with similar nonwords (Dell, 1984; Shattuck-Hufnagel, 1992). There is ample evidence that when the vowels in a multisyllabic word are similar to one another, the surrounding consonants are more likely to be erroneously exchanged, especially consonants that occupy similar positions (e.g., the onset) in their respective syllable (e.g., Dell, 1984, 1995; Dell, Reed, Adams, & Meyer, 2000; Fowler, 1987). Though not formally analysed, it was observed that at least some of the recall errors in the current experiment were of this kind where, for example, *fenten* would be output instead of *tenfen* (see also Taylor, Macken, & Jones, 2015). In the case of the effect of across-nonwords similarity on learning, while general cue-based interference may account for some of this effect (even in free recall), it seems plausible that having whole nonword motor-plans share many elements could lead to migrations of elements from one motor-plan to another, resulting in errors. This could again be seen in the common error-types made, where a response could contain elements from two nonwords based on their shared properties: A response might contain the elements present in both of the nonwords, as well as distinguishing, unique, elements from both of the nonwords, such as the response *tredote* constructed from elements of *trudote* and *trelate*.

The comparison of the monolingual and bilingual speakers in this experiment also indicated better nonword recall for those knowing another language in addition to their native English. This may be due to greater linguistic knowledge or/and vocal-motor skill: Speakers of multiple languages would likely have a greater vocabulary-size, resulting in nonwords being more likely, on average, to be word-like for such participants compared to monolinguals. Multilinguals may also have enhanced motor planning skills in general which would also be expected to aid in nonword recall. However, the multilingual benefit did not extend to nonword learning, contradicting previous findings (e.g., Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995). The inability to replicate the bilingual advantage on learning may be related to the relatively heterogeneous sample of multilinguals recruited for this study, as they spoke a varied number of other languages and could have varying degrees of fluency in these languages. Most previous studies have used a more defined group of multilinguals (e.g., Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995): Generally, participants in these experiments all had the same native language and a similar amount of experience with their other language(s). The varied sample may have also contributed to the finding that the language groups were not differentially affected by either type of phonological similarity: Both long-term vocabulary knowledge and novel motor planning abilities would be likely to differ according to the additional language(s) spoken, though they would regardless be expected to have a larger vocabulary compared to monolinguals. This would be consistent with, for example, the finding that early and late bilinguals differ in their vowel production of their second language, with early bilinguals having an advantage in correctly pronouncing their second language compared to late bilinguals (e.g., Baker & Trofimovich, 2005; Piske, Flege, MacKay, & Meador, 2002).

3.4 Experiment 5

It was suggested that the data from Experiment 4 may be best explained in terms of the role of motor planning in verbal sequence learning. It follows that interfering with the availability of motor planning processes should also disrupt the learning of nonwords in this setting. Evidence was presented in Chapter 2 that verbal sequence learning as witnessed in the Hebb repetition paradigm is driven to a considerable degree by motor planning: a reduction in the possibility of, or the need for, engagement in motor planning attenuated the Hebb effect. In Experiment 5, therefore, the role of motor planning in nonword learning in the context of serial learning in both paired-associate and free recall was examined by restricting the availability of motor planning. Specifically, a requirement for articulatory suppression should interfere with the formation of an accurate and fluent motor-plan for the production of each of the nonwords, and should thus also interfere with their learning.

From the standpoint of the PS-LLD hypothesis, there are again contradictory views about whether or not articulatory suppression should affect learning. The phonological-store based models of the Hebb effect (cf. Chapter 2) posit that articulatory suppression, like phonological similarity, should only interfere with the short-term recall of, but not the learning of, verbal sequences (Hitch et al., 2009; Page et al., 2006). However, another finding cited as a key piece of support for the PS-LLD hypothesis is that verbal sequence learning witnessed in the word-nonword (or -foreign-word) version of the paired-associate learning task is indeed impaired by articulatory suppression (Baddeley, 2007; Baddeley et al., 1998). For example, Papagno et al. (1991) showed participants visually-presented lists of pairs of Italian words (to Italian-speaking participants) or *Italian-word—Russian-word* pairs (the participants did not understand Russian). Articulatory suppression was found to impair the learning of the Italian-Russian word pairs to a reliably greater extent than it did the learning of *Italian-word—Italian-word* pairs. This effect of articulatory suppression was replicated in a subsequent experiment with auditory presentation. Further experiments with (visually-presented) *English-word—nonword* pairs and *English-word—Finnish word* pairs (where the participants did not speak Finnish) again replicated the attenuating effect of articulatory suppression on nonword/foreign word learning.

The disruptive effect of articulatory suppression on nonword learning is taken to support a role for the phonological store in that learning. However, a difficulty with this reasoning (in addition to the contradiction with Hebb-based models) is that given that auditory stimuli are said to enter the store without the need for the articulation-based grapheme-to-phoneme conversion process (e.g., Baddeley et al., 1984), learning of auditory lists should be less affected by articulatory suppression than that of visual lists. But learning in both modalities was shown to be affected to the same degree by articulatory suppression (Papagno et al., 1991). To account for the fact that the effect of articulatory suppression was no worse with visual than with auditory presentation, Baddeley et al. (1998) posited that visual codes available with visual but not auditory stimuli might be expected to aid learning. But the notion that the obligatory access enjoyed by auditory (but not visual) stimuli and the visual codes available with visual (but not auditory) stimuli perfectly cancelled each other out under suppression is not entirely satisfactory, especially as the use of visual codes to retain verbal stimuli is not an integral tenet of the phonological loop model (see also Macken et al., 2016). Furthermore, the idea that learning of auditorily presented nonwords/foreign words is affected by articulatory suppression due to the blocking of articulatory rehearsal is at odds with the suggestion that word-form learning, at least in children learning their first language, is assumed to be supported not by articulatory rehearsal but by the passive phonological store. That is, on the PS-LLD hypothesis, articulatory rehearsal is deemed to only play a role in *second*-language learning in adults (or in children old enough to use articulatory rehearsal; Baddeley et al., 1998, p. 167). But if this is the case, it becomes questionable, on this account, whether any of the results from paired-associate learning studies—which have all involved adult participants (Baddeley et al., 1998; Papagno et al., 1991; Papagno & Vallar, 1992)—have a bearing on the core mechanisms involved in word-form learning.

From a perceptual-motor perspective, the effect of articulatory suppression indicates a primary role for a constructive motor planning process in performance (e.g., Jones et al., 2004) rather than a role for a process in which decaying representations in a separate, passive, store are refreshed or in which graphemic input is converted into phonological form (e.g., Baddeley, 1986). From this standpoint, articulatory suppression is argued to impair verbal sequence learning (both in serial learning and the Hebb paradigm) because such learning reflects the legacy of the motor planning process deployed to recall the nonwords over the short term. Experiment 5, then, revisits the effect on nonword learning of impeding motor planning through articulatory suppression. In contrast to the ambiguities identified above as to whether phonological-store based accounts predict an effect of articulatory suppression (e.g., Baddeley et al., 1998; Hitch et al., 2009; Page et al., 2006), the perceptual-motor account clearly predicts an effect of suppression on nonword learning and this regardless of whether the nonwords are presented auditorily or visually. While a role for passive auditory perceptual organisation in this kind of setting has not been established (unlike the case in serial recall and the Hebb effect, e.g., Hughes et al., 2009; Nicholls & Jones, 2002; present Chapter 2), interest centred also on whether learning of nonwords presented auditorily is impaired less by the restriction of motor planning than is the learning of visually-presented nonwords. Finally, the manipulations of articulatory suppression and modality were crossed with recall-type (i.e., paired associate vs. free recall). While recall-type was not of central interest in this particular experiment, it was included mainly for the sake of providing a continuity with Experiment 4.

3.4.1 Methods

3.4.1.1 Participants. Forty-two students (3 males, 39 females) from Royal Holloway, University of London, aged 18-24 years (mean 18.95, SD = 1.32), took part in the experiment in return for course credits. Twenty-seven of the participants were native English speakers (11 monolinguals). Given that Experiment 4 failed to detect meaningful differences between language groups on learning, however, this was not included as a factor in Experiment 5. Two participants reported having dyslexia.

3.4.1.2 Apparatus and Materials. The apparatus and materials were the same as in Experiment 4, with the following changes: Both auditory and visual nonwords were used, with the auditory stimuli presented via headphones. The auditory stimuli were recorded in a female voice of a native English speaker. New lists of English words and nonwords were generated, with the nonwords now all within-dissimilar and presented in across-dissimilar lists (that is, no type of similarity was manipulated in this experiment), with the same nonwords used for both auditory and visual conditions. A full list of the stimuli is provided in Appendix 6, Supplementary Table 12.

3.4.1.3 Design. The experiment had three repeated-measures factors: Recall-type (paired-associate, free), Articulatory suppression (without suppression, with suppression) and Cycle (referring to the ten successive trials with each list). Modality (auditory, visual) was a between-participants factor. Each participant completed four blocks in total in a counterbalanced fashion from four possible orders: A1-B2-A2-B1, A2-B1-A1-B2, B1-A2-B2-A1 or B2-A1-B1-A2, where A=cued recall, B=free recall, 1=no suppression, 2=with suppression. A particular list of nonwords or words could be presented in any of the four blocks, again in a counterbalanced fashion.

3.4.1.4 Procedure. The procedure was adopted from Experiment 4 with the following changes: After participants received extensive instructions for the task, they practised the articulatory suppression procedure with the Experimenter. Participants also completed two practice trials without suppression, and two with suppression before the experiment trials. Before each block, the participant was informed whether the block involved associate-cued or free recall, and whether articulatory suppression was required. In the articulatory suppression blocks, suppression was required during presentation but (necessarily) not during the vocal recall stage. The Experimenter monitored compliance with the articulatory suppression requirement throughout the relevant blocks.

3.4.2 Results

The data from two participants in the auditory group were not used due to their failure to follow the task instructions properly. As in Experiment 4, the data for the analyses were the spoken recall responses, with a response only scored as correct if it contained all of the syllables of the presented item in the correct order and, in addition for paired-associate recall, if the nonword was produced in response to the correct word cue. The data, shown in Figure 3.2, indicated that both free and paired-associate recall were clearly affected by articulatory suppression. Suppression also affected nonword learning, though this appears to have only been the case in the paired-associate task. A 2 (Recall-type) × 2 (Modality) × 2 (Articulatory suppression) × 10 (Cycle) mixed-effects ANOVA was conducted to analyse performance accuracy across the conditions. As in Experiment 4, more nonwords were correctly recalled in free recall compared to paired-associate recall, $F(1, 38) = 17.77$, $MSE = 10.41$, $p < .001$, $\eta_p^2 = .319$. There was also a significant interaction of Recall-type and Cycle, $F(9, 342) = 18.21$, $MSE = .99$, $p < .001$, $\eta_p^2 = .324$, which again appears to reflect the initial benefit of being able to recall the nonwords in any order in free recall (including starting with the last nonword presented), with a similar proportion of nonwords recalled in each task at later cycles. As in Experiment 4, there was indeed a greater probability of the response-output in free recall beginning with the last-presented item (.38, SD = .26) than the first-presented item (.11, SD = .07), $t(42) = 5.77$, $p < .001$.

Nonword learning was indicated by the significant main effect of Cycle, $F(3.4, 130.9) = 282.09$, $MSE = 3.37$, $p < .001$, $\eta_p^2 = .881$. A significant main effect of Articulatory suppression confirmed that fewer nonwords were recalled when suppression was required, $F(1, 38) = 72.95$, $MSE = 6.34$, $p < .001$, $\eta_p^2 = .657$. The interaction of Articulatory suppression and Cycle was also reliable, indicating that articulatory suppression also attenuated learning, $F(9, 342) = 2.05$, $MSE = .86$, $p = .033$, $\eta_p^2 = .051$. Additionally, while Recall-type and Articulatory suppression did not significantly interact, $F(1, 38) = .13$, $MSE = 8.36$, $p = .718$, $\eta_p^2 = .003$, there was an interaction of Recall-type, Articulatory suppression and Cycle, $F(9, 342) = 4.99$, *MSE* = .66, $p < .001$, $\eta_p^2 = .116$. This appears to reflect the fact that articulatory suppression impaired learning in the paired-associate task but not in the free recall task: Indeed, when the data from the two recall tasks were analysed separately, the interaction of Cycle and Articulatory suppression was significant in paired-associate recall, $F(9, 342) = 5.43$, $MSE = .82$, $p < .001$, $\eta_p^2 = .125$, but not in free recall, , $F(9, 342) = .82$, $MSE = .69$, $p = .6$, $\eta_p^2 = .021$.

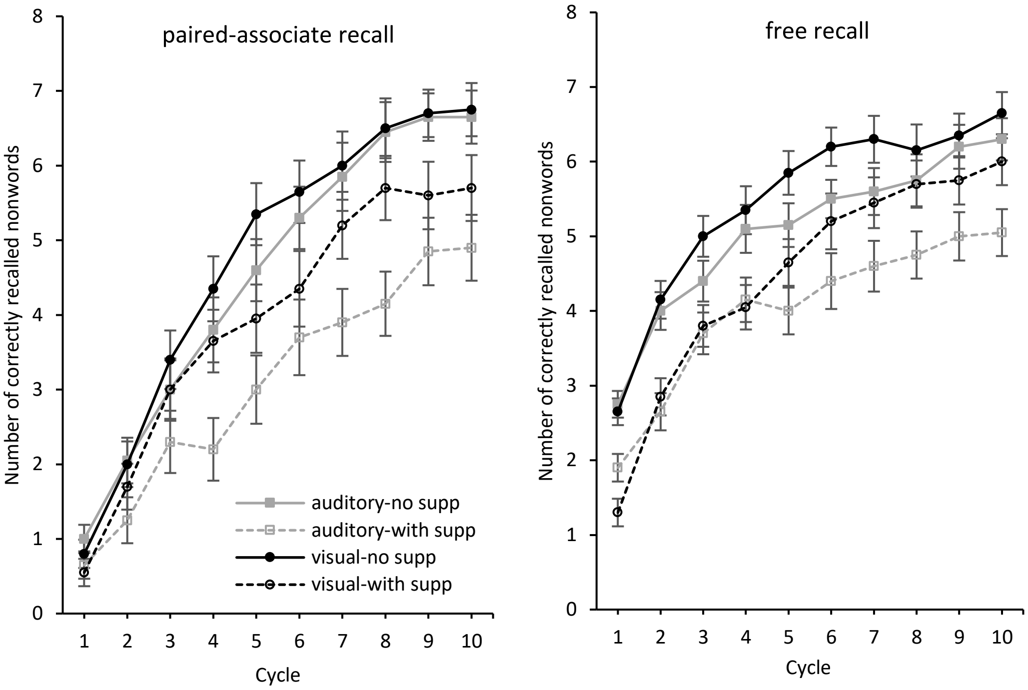


Figure 3.2. Performance accuracy in Experiment 5 according to Recall-type, Modality and Articulatory suppression conditions across the cycles, with error bars representing the standard errors.

While the main effect Modality was not significant, $F(1, 38) = 2.38$, $MSE = 38.3$, $p = .131$, $\eta_p^2 = .059$, there was a significant interaction of Modality and Cycle, $F(9, 342) = 2.99$, $MSE = 1.29$, $p = .002$, $\eta_p^2 = .073$, showing an advantage in the learning of visually- over auditorily-presented nonwords. This effect did not interact with either Articulatory suppression or Recall-type, and there were no other reliable interactions. These findings are inconsistent with the tentative hypothesis that auditorily-presented nonwords would be better recalled and learned than visually-presented items, at least under articulatory suppression. The full set of results is reported in Appendix 6, Supplementary Table 13.

One possibility as to why articulatory suppression did not affect learning in the context of free recall is that, unlike in the Hebb learning experiments in Chapter 2 using manual order-reconstruction, no suppression was required during the response phase. This would have allowed for motor planning to occur during the recall phase and this may have been sufficient to support word learning. This may have been particularly the case in free recall in which the order of recall—and hence inter-nonword motor planning—is likely to have been more consistent across cycles than in the paired-associates task (in which, by design, output order was randomised). Whilst it is difficult to determine whether this was the case, evidence consistent with it would be forthcoming if there was some regularity in the order in which items were output across cycles, including within the articulatory suppression condition. To examine this, the number of edits (additions, deletions, switching of adjacent items) that would be required for the output to match the output on the previous cycle was calculated (a similar measure has previously been used to examine error-learning in the Hebb repetition paradigm, see, e.g., Lafond et al., 2010). The analysis involved a comparison of the number of edits that had to be made when the actual order in which each participant produced the nonwords was summed (the 'actual-order condition') with the number of edits that had to be made when a random ordering of the items they recalled on each cycle was summed (a hypothetical 'random-order condition'). These values were then adjusted according to the number of nonwords produced on each cycle to give a proportional score. The first cycle was not included in the analysis, since there was no preceding output. Figure 3.3 indicates that while the number of edits remained relatively small in the actual-order condition, it increased across cycles in the random-order condition. A 2 (Modality) × 2 (Articulatory suppression) × 9 (Cycle) × 2 (Order: actual, random) ANOVA confirmed a significant main effect of Order, $F(1, 40) = 54.5$, $MSE = .03$, $p < .001$, $\eta_p^2 = .577$, with fewer edits in the actual-order condition. There was also a marginally reliable interaction of Order and Cycle, $F(8, 320) = 1.94$, $MSE = .03$, $p = .053$, $\eta_p^2 = .046$, reflecting the fact that the number of edits in the random-order condition increased across cycles. The number of edits in the actual-order condition did not increase systematically with cycle, even though the number of correctly recalled items increased across cycles. The proportional score can only partially account for the increasing number of responses made, as the scoring cannot reflect the relation between two recall attempts, simply scale the performance in the latter cycle to the earlier. The increasing number of edits in the random-order condition relates to the type of edits: In the actual-order condition, in addition to the new items (compared to previous cycle) that are recalled, increasing the number of edits, many of the edits constitute the switching of adjacent items, so that two items change places and are counted as a single edit. This is less the case in the random-order condition, resulting in edits which could have been switches changing to a deletion and an addition, increasing the total number of edits. In effect, therefore, there was a greater similarity of output-order across successive cycles than would be expected by chance, in line with the notion of relatively consistent inter-item planning across cycles. No other terms were significant. Of particular relevance is that the pattern evident in Figure 3.2 did not vary reliably according to articulatory suppression; there was no significant interaction between Order, Cycle and Articulatory suppression, $F(8, 320) = 1.18$, $MSE = .03$, $p = .311$, $\eta_p^2 = .029$.

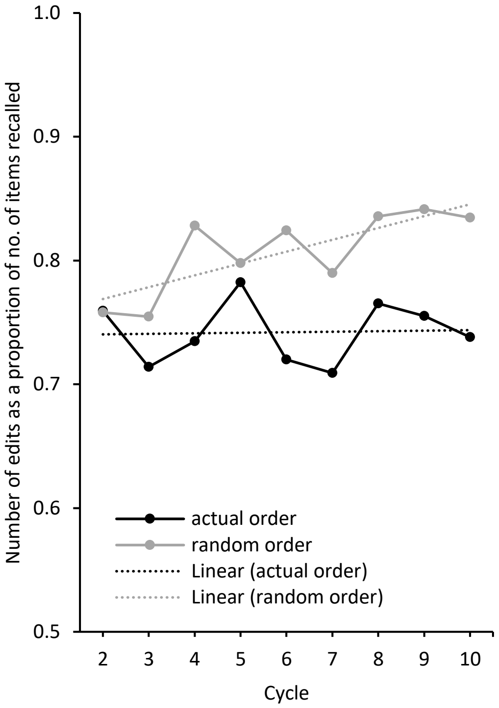


Figure 3.3. Proportion of edits to the number of items recalled in the actual- and random-order 'conditions' according to cycle, averaged across all other variables.

3.4.3 Discussion

The results of Experiment 5 replicated nonword learning in both paired-associate and free recall tasks. Similar to Experiment 4, being able to recall the nonwords in any order resulted in higher accuracy in free recall compared to paired-associate recall in the initial cycles, with the advantage disappearing in the later cycles as the pairings in the paired-associate task were learned. The data also showed that articulatory suppression reduced recall accuracy regardless of presentation modality or recall task. Articulatory suppression also reduced learning but this was only the case in paired-associate recall. In the free recall task, the effect of suppression remained comparatively stable across cycles, indicating that while recall was impeded, learning was not affected. The follow-up analysis of recall-order in the free recall task did, however, provide some indication of output-matching across successive cycles, consistent with the hypothesis that learning could have occurred during the recall attempts. To elaborate, a tentative proposal is that learning of the nonwords occurred during responding, resulting in the previously-recalled nonwords requiring less motor planning at the next output opportunity, with performance able to rely on the previously constructed motor-plans as well as inter-nonword plans. This advantage may then have led to more capacity within the task constraints to deal with additional nonwords (i.e., those not yet recalled) on the next cycle. This interpretation was supported by the analysis of edits between cycles, which showed that, regardless of suppression condition, participants tended to recall the previously recalled nonwords together, with previously unrecalled nonwords being output first or the last within the particular cycle, rather in the middle of previously-recalled nonwords. This may suggest that the newly recalled nonwords benefit from the most recent presentation or some available motor-planning capacity, when the production of previously recalled nonwords can rely on the already learned motor-plans.

While it is clear that learning in the paired-associate recall task was reduced when articulatory suppression was required, there was clearly still learning despite suppression, appearing to suggest that while motor planning supports learning, it is not necessary. However, it is again possible that motor-planning based learning of the nonwords during the response stage played at least some role in nonword learning in the paired-associate task too.

The effect of articulatory suppression on learning is consistent with both the perceptual-motor account as well as the phonological store account as both would predict that articulatory suppression should impair learning. The findings do, however, highlight again the inconsistency between the PS-LLD hypothesis based on the paired-associate learning paradigm (Baddeley et al., 1998; Papagno et al., 1991; Papagno & Vallar, 1992) and the phonological-store based account of the Hebb repetition effect (Hitch et al., 2009; Page et al., 2006), which predict opposite outcomes regarding the effect of articulatory suppression: In the paired-associate task, the phonological store account, similar to the perceptual-motor account, predicted an effect of articulatory suppression on learning, while the phonological store-based models of Hebb sequence learning predict no such effect.

The faster learning of visually presented lists compared to auditory lists appears, at first, inconsistent with both the phonological store and the perceptual-motor accounts, as both accounts would predict better learning with auditorily presented sequences, at least under suppression. However, the advantage in learning visually-presented nonwords over auditory lists replicates some earlier findings (Papagno et al., 1991). This was ascribed to the additional availability of visual codes (Baddeley, 2003). Another explanation, though speculative, would be that if motor planning processes are the key processes that underpin the learning of the new material, the longer presentation times associated with this paradigm compared to the classic serial recall/Hebb paradigm can, with visual stimuli, allow for more accurate matching of the motor-plan to the stimuli, which can be read several times during presentation, allowing the motor-plan for the response to be corrected during the relatively long exposure time. As modality did not interact with articulatory suppression or recall-type, it further implies that performance in both recall tasks is more reliant on modality-general motor planning processes, rather than modality-specific processes, and that these processes may have, despite suppression during presentation, been able to support learning during the response-phase.

In the next, and final, experiment, the range of methodological tools used thus far in the thesis was extended considerably through the introduction of a brain imaging component, specifically, functional near-infrared spectroscopy (fNIRS; Kleinschmidt, Obrig, Requardt, Merboldt, Dirnagl, Villringer, & Frahm, 1996) was used to examine the role of brain regions associated with motor planning in nonword recall and learning.

3.5 Experiment 6

If nonword learning is reliant on a phonological store, such learning could be associated with a number of brain regions. Though the exact location of the putative phonological store remains unclear, studies of patients with brain damage have often been used to support the notion of separate memory systems (e.g., Baddeley, 2007), and neuropsychological data have also been cited to support more specifically the existence of a discrete phonological store in the brain. Data from patients showing a pure deficit in functions that the phonological store is thought to serve, in the absence of other general cognitive deficits, have been taken to indicate the special role of the phonological store in verbal short-term memory (e.g., Baddeley et al., 1988).

Specific brain regions where the phonological store or other parts of the working memory system might reside have been proposed based on data from patients with brain lesions, with the phonological store associated with the inferior parietal lobule and the superior and middle temporal gyri and the articulatory control process with subcortical and premotor regions (Vallar & Baddeley, 1984; Vallar et al., 1997). Cerebellar regions have also been suggested as supporting the articulatory control process (Silveri et al., 1998), and the seminal positron emission tomography experiment of Paulesu, Frith and Frackowiak (1993) located the phonological store in the left inferior supramarginal gyrus. However, the proposed brain correlates in the posterior parietal lobe, located outside the areas associated with processing auditory-verbal material, appear inadequate for accounting for the obligatory access of auditory-verbal items into the phonological store. Moreover, the other suggested sites in the superior temporal lobe have the opposite problem: Localising the store in this auditory-verbal processing area would suggest the use of perceptual organisation-based representations, rather than specifically indicating a store (Buchsbaum & D'Esposito, 2008). The patient data indicating a pure auditory verbal short-term memory deficit, initially the key evidence for a distinct phonological store with a specific brain correlate, has also since been considered to be equally interpretable as a deficit in sensory-motor integration processes (Morey, Rhodes, & Cowan, 2019). The damage to the left temporoparietal areas, which can result in aphasia, has previously been taken to indicate damage to the phonological store, but now appears more congruently explained as damage to articulatory motor functions. Patients with damage in these brain regions show a complex pattern of impairments related to sensory-to-articulatory transitions in the context of complex sound sequences such as those characteristic of speech and music, with the aphasia symptoms explicable in terms of impairments to the auditory-to-articulatory sequence mapping processes (Buchsbaum et al., 2011).

Additionally, evidence of the involvement of sensory-specific processing in parietal and temporal regions during short-term memory processing has indicated the involvement of perceptual processing in these tasks, which appears to point to distributed networks supporting processing (Fuster, 1990; Tresch et al., 1993). It has also been argued that much of the early work defining lesions as resulting in pure phonological store deficits or pure articulatory control deficits (e.g., Shallice & Vallar, 1990) involved appeals to data that are inconsistent with the behavioural hallmarks of the phonological store account (Caplan, Waters, & Howard, 2012). For example, some patients described as having a phonological store–specific impairment continue to show a phonological similarity effect with auditory sequences (Caplan et al., 2012), which should not occur if the phonological similarity effect is a hallmark of the store (Baddeley, 1968). From a functional neuroanatomy viewpoint, a discrete, autonomous brain unit dedicated to temporarily holding verbal items, separated from the perceptual and motor processes responsible for perception of the items and their motor -based maintenance and production appears implausible, given that lesion and neuroimaging studies appear unable to converge on a location for such a unit (e.g., Paulesu et al., 1993; Postle, 2006; Vallar & Baddeley, 1984). Instead, functional neuroimaging seems to suggest that the retention of verbal items may be a component of a network that forms an auditory-motor interface, binding auditory and articulatory speech (Buchsbaum & D’Esposito, 2008). In lesion studies where a specific impairment is indicated, there may be processes involving the temporal or parietal cortex that are recruited for auditory-to-articulatory translations or other articulatory rehearsal processes which, when interrupted by a lesion, result in an impairment in verbal short-term memory behaviours (Buchsbaum & D’Esposito, 2008).

If motor planning is a key part of nonword learning, as proposed by the perceptual-motor account, brain regions associated with motor planning would be expected to be active during nonword learning. In particular, when the fluency of motor planning for a new verbal sequence increases, this may result in more automated processing that relies on a long-term memory for that motor-plan, which could be observable in a diminished requirement for active, short-term, motor planning to produce the correct output. If the verbal sequence learning process involved cortical regions associated with short-term motor-planning, this transition from using short-term motor planning to the retrieval of established/long-term motor plans may be reflected in a decrease in the activity of those brain regions. Interest in Experiment 6, therefore, centred on the supplementary motor area (SMA) and the pre-supplementary motor area (pre-SMA) as they have been shown to be involved in a variety of tasks such as language production, gesture-learning and motor sequencing (e.g., Hatakenaka et al., 2007; Koziol et al., 2014; Hirano et al., 1996). More specifically, the SMA, for instance, has been shown to be involved in the networks that are recruited for planning and producing vocalisations of native language, which require verbal sequencing (Ackermann & Riecker, 2010; Hirano et al., 1996). When participants are required to read out loud visually presented words in their native language, the SMA has been implicated in networks also including areas of the cerebellum (Hirano et al., 1996). The SMA has also been implicated in other aspects of language processing such as prosody, further supporting the role of motoric processes associated with these areas in verbal tasks (Hertrich et al., 2016). This suggests that during speech, performance is planned, controlled, and monitored during output in a network involving the SMA and the cerebellum, which is widely considered to be involved in monitoring and coordinating movements as well as in learning (Koziol et al., 2014).

Of particular interest here is that the pre-SMA and the SMA have also been implicated in motoric learning, such as the learning of hand-movements, with activation moving from the pre-SMA to the SMA as learning progresses, coupled with an overall decrease in activation across the two areas (Hatakenaka et al., 2007). This suggests processes involving motor planning that result in learning could be observed in the pre-SMA and SMA. These contributions of motor areas of the brain to verbal tasks and to gesture-learning tasks in neuroimaging studies is commensurate with the hypothesis that motor planning supports verbal serial short-term memory and learning, which has, to date, been tested largely via behavioural methods (e.g., Chapter 2; Hughes et al., 2009, 2016; Jones et al., 2004, 2006;

Macken et al., 2014; Maidment et al., 2013). Thus, observing that these areas are active also during verbal sequence learning—nonword learning in particular—would provide further evidence consistent with the perceptual-motor account. In Experiment 6, therefore, the design of Experiment 4 was replicated but now with the inclusion of fNIRS to investigate the possible involvement of the SMA/pre-SMA in non-word recall and learning. Note that both the SMA and pre-SMA were examined as a 'single' brain area of interest in the present experiment.

Functional NIRS allows for the level of activity in certain brain areas to be measured using near infrared wavelengths of light to detect changes in oxygenated haemoglobin concentration, as near infrared light can easily penetrate living organisms. As oxygenated and deoxygenated haemoglobin have different rates of absorption of the near infrared wavelengths of 700-900 nm, concentrations of both can be calculated. Oxygenated haemoglobin (HbO) is primarily indicative of inflow of oxygen to the tissue, and deoxygenated haemoglobin (HbR) is associated with the absorbed amount of oxygen.

When used for brain imaging, the light can penetrate approximately 20 mm from the surface of the head, making the method suitable for imaging the surface of the cortex. In fNIRS, near infrared light is sent from one optode through the head, where the light is partially absorbed and scattered, and some of the scattered light will reach a receiver optode, and converted to an electrical signal, which can be used to calculate the extent to which the light was absorbed by oxygenated and deoxygenated haemoglobin (Cui, Bray, Bryant, Glover, & Reiss, 2011). Comparisons of fNIRS and fMRI methods have indicated that while fNIRS has a poorer spatial resolution and signal-to-noise ratio, results from the two methods are well correlated (Cui et al., 2011; Lu et al., 2010). Functional NIRS also has the advantage of ease of administration, and as fine spatial resolution was not essential for current purposes, the method appears suitable for investigating the role of motor planning in nonword learning through examining the possible activation of a particular, but relatively large, area of the cortex. Resting-state functional connectivity between sensory and motor areas, connecting the primary processes of interest in the present thesis, perceptual organisation and motor planning, have also been indicated in fNIRS investigations (Zhang et al., 2010).

There are already some data indicating the involvement of the SMA/pre-SMA in verbal serial short-term memory as measured using fNIRS. Marsh, Richardson, Barker, and Hughes (2019) showed that these motor regions are more active in a verbal short-term memory task that requires sequencing than one that only requires item recall, which is in line with the account of verbal serial short-term memory as based on the formation of a motor sequence-plan. If the SMA/pre-SMA is also involved in the formation of motor-plans for nonwords, and nonword recall and learning are supported by such motor plans, activation of this area would be expected during the presentation of the nonwords, that is, well before they need to be produced overtly. In addition, a diminution of activity in SMA/pre-SMA is predicted as the nonwords are learned across cycles, that is, as reliance on short-term novel motor-planning gives way to reliance on the retrieval of already-established motor-plans (cf. Hatakenaka et al., 2007). Finally, it was conjectured that SMA/pre-SMA activation might be greater generally under more difficult motor-planning conditions, that is, with either within- or between-nonword phonological similarity. Previous experiments have indicated that more difficult motor-planning conditions are associated with more active motor planning (e.g., Breitling, Guenther, & Rondot, 1986; Halgren, 1991; Krings, Töpper, Foltys, Erberich, Sparing, Willmes, & Thron,

2000) and increased activation in the pre-SMA/SMA has also been shown in high motor-task difficulty conditions (Winstein, Grafton, & Pohl, 2016).

3.5.1 Methods

3.5.1.1 Participants. Forty-six (26 males, 20 females) participants, aged 19-42 years (mean 26.72, SD = 5.81) were recruited from University of Gävle, 43 of whom were native Swedish speakers (the other 3 were not included in the analyses). Twenty-two (10 males, 12 females) of the 43 native Swedish participants, aged 21-48 years (mean 27.86, SD = 6.15), underwent fNIRS recording during the experiment. All participants reported being also fluent in English. One participant reported having dyslexia. Participants received a small honorarium for their participation.

3.5.1.2 Apparatus, Materials, and Design. These aspects of the method were the same as Experiment 4.

3.5.1.3 Procedure. The procedure followed that of Experiment 4 but participants received task instructions in both English and Swedish. The participants also completed an additional questionnaire regarding their fluency in, and use of, English as a second language. The participants who were also included in the imaging during the experiment received additional instructions to remain as still as possible during the experiment. Before the experiment began, they were presented with a fixation cross for 30 s, during which they were asked to stay still and look at the cross so that a baseline level of activation in the SMA/pre-SMA could be recorded for later comparisons in the imaging analysis.

3.5.1.4 Data acquisition and processing (fNIRS). For 22 of the participants (11 from each within-similarity condition), neural activation in the SMA/pre-SMA was also measured during nonword presentation by measuring changes in HbO and HbR using an Oxymon Mk III (Artinis Medical Systems BV, The Netherlands) continuous wave NIRS system with a sampling rate of 50 Hz, with changes in light attenuation measured at two wavelengths (845 nm and 763 nm). The optodes were placed in a 2 × 4 configuration on an elasticised cap with a distance of 2.5 cm between the centre of each optode and the next. The optodes were placed between the frontal and the middle part of the cap to approximate a location over Brodmann area 6, based on a previous fNIRS study implicating the SMA/pre-SMA in verbal serial short-term memory (Marsh et al., 2019). Measures of HbO and HbR were collected for each 0.02 s of the period of interest, defined as the 24 s during which the nonwords were being presented, as well as a 30 s baseline period during which participants looked at a fixation cross. The channel arrangement is depicted in Figure 3.4. The fNIRS data were processed with MATLAB (The Math Works Inc., Natick, MA) and Homer2 (Boas, Dubb, & Huppert, 2012) to convert the raw data to haemoglobin concentrations of HbO and HbR levels during presentation of the nonwords. The data were converted into changes in optical density with hmrIntensity2 OD in Homer2, and artefacts were assessed with hmrMotionArtifacts in Homer 2 (with values tMotion: 0.5, tMask: 1.0, STDEVthresh: 50.0, AMPthresh: 10.0, based on Becerra et al., (2016). The Homer2 functions enStimRejection (-2.0 – 24.0) and hmrMotionCorrectSpline were also used for removing artefacts. In order to remove physiological signals such as variations in blood pressure and cardiac signals, the fNIRS signals were bandpass-filtered using the values 0.01-0.5 Hz (cf. Sibi, Balters, Mok, Steinert, & Ju, 2017). For each of the source-detector pairs, changes in optical density were converted to haemoglobin concentrations using the hmrOD2conc function in Homer2, using partial pathlength factors of 6 for both wavelengths. Due to difficulties gaining a strong enough signal for the data to be analysed by cycle, the data had to be averaged over the ten cycles in each condition. Unfortunately, this meant that it was not possible to examine the involvement of the SMA/pre-SMA in learning as planned, but only its involvement in nonword *recall* as a function of the eight conditions. Thus, the variation in HbO and HbR concentration within each of the eight conditions [four for each participant: 2 (across-similarity) × 2 (recall-type), as within-similarity was a between-participants factor] were averaged across cycles using hmrBlockAvg (tRange -5.0 – 24.0).

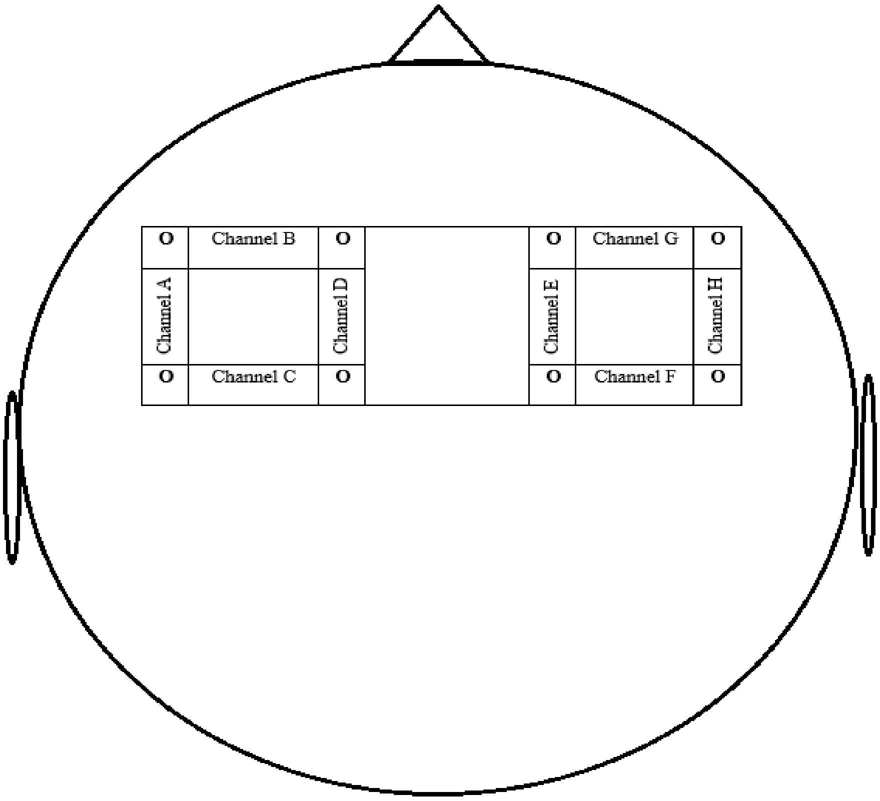


Figure 3.4. Schematic of the location of the channels between each of the eight optodes (O), with channels B and G located most frontally on the cap, placed to capture activation across the SMA/pre-SMA.

3.5.1.5 Data analysis (fNIRS). Analyses of the fNIRS data were carried out in R 1.1.383 (R Core Team, 2012), using lme4 (Bates, Maechler & Bolker, 2012) to carry out linear mixed-effect models to investigate the effects of across-similarity, within-similarity and recall-type on HbO and HbR concentrations (separately for each) during the presentation of the nonwords. The data were then pre-processed (see the above section on fNIRS data processing) to provide a measure of the concentration of haemoglobin for each participant for the 24 s presentation period averaged across the ten cycles, for each of the eight conditions. In the event, technical problems prevented the retrieval of the resting state (baseline) recordings. Interest centred, therefore, on comparing levels of activation in the SMA/pre-SMA as a function of the various experimental manipulations.

3.5.2 Results

3.5.2.1 Behavioural data. Two participants (who did not undergo fNIRS recording) were excluded from the analysis for not complying with all task instructions. Responses were scored in the same way as for Experiments 4 and 5. The data, presented in Figure 3.5, were analysed with a 2 (Recall-type) × 2 (Across-similarity) × 2 (Within-similarity) × 10 (Cycle) ANOVA, where within-similarity was the only between-participants factor (Within-dissimilar N=21, Within-similar N=19). As in Experiment 4, there were significant main effects of Across-similarity, with across-dissimilar lists recalled better than across-similar, $F(1, 38) = 48.64$, $MSE = 5.95$, $p < .001$, $\eta_p^2 = .561$, and a main effect of Cycle, reflecting nonword learning, $F(4.3, 164.3) = 233.97$, $MSE = 2.38$, $p < .001$, $\eta_p^2 = .860$. The main effects of Recall-type, $F(1, 38) = 2.83$, $MSE = 17.6$, $p = .101$, $\eta_p^2 = .069$, and Within-similarity were not significant however, $F(1, 38) = .001$, $MSE = 46.76$, $p = .97$, $\eta_p^2 < .001$. Thus, unlike Experiment 4, recall was not better in free recall than in the paired associate task and there was no impairment of recall overall with within-similar lists (but see below).

An interaction of Across-similarity and Cycle showed that across-similarity not only affected recall but also learning, $F(6.5, 245.6) = 6.43$, $MSE = 1.17$, $p < .001$, $\eta_p^2 = .145$, replicating the pattern from Experiment 4. There was also a significant interaction of Recall-type and Cycle, with an advantage for free over paired-associate recall in the first few cycles, and an advantage for paired-associate recall in the last two cycles, $F(5.6, 211.44) = 13.28$, $MSE = 1.63$, $p < .001$, $\eta_p^2 = .259$. As suggested in relation to Experiments 4 and 5, this probably reflects the greater likelihood of a free recall response beginning with the last presented item (.49, SD = .25) compared to the first presented item (.12, SD = 15), $t(42) = 6.36$, $p < .001$.

As in Experiment 4, there were no significant interactions between Recall-type and either of the similarity manipulations, indicating that recall in the two tasks was not, overall, differentially affected by either Across-similarity, $F(1, 38) = .74$, $MSE = 10 .28$, $p = .394$, $\eta_p^2 = .019$, or Within-similarity, $F(1, 38) = .38$, $MSE = 17.6$, $p = .384$, $\eta_p^2 = .02$. However, while there was no significant interaction of Within-similarity and Cycle, $F(5.6, 211.4) = 1.35$, $MSE = 1.14$, $p = .208$, $\eta_p^2 = .03$, there was a fairly strong numerical trend for a three-way interaction of Recall-type, Within-similarity and Cycle, $F(5.6, 211.4) = 1.98$, $MSE = 1.63$, $p = .075$, $\eta_p^2 = .05$. When the data are averaged across levels of across-similarity (see Figure 3.6), it can be seen that there is some evidence of a within-similarity effect on learning in free recall but little effect (or a slight reverse effect) of within-similarity on learning in the paired-associate task. Finally, while the pattern in Figure 3.5 suggests that there may have been somewhat greater impairment of learning due to across-similarity in paired-associate recall compared to free recall, this was not confirmed, $F(5.6, 211.4) = .54$, $MSE = .8$, $p = .845$, $\eta_p^2 = .01$. The four- way interaction was not reliable on this occasion, $F(5.6, 211.4) = 1.19$, $MSE = .8$, $p = .302$, $\eta_p^2 = .03$. The full set of results is reported in in Appendix 7, Supplementary Table 14.

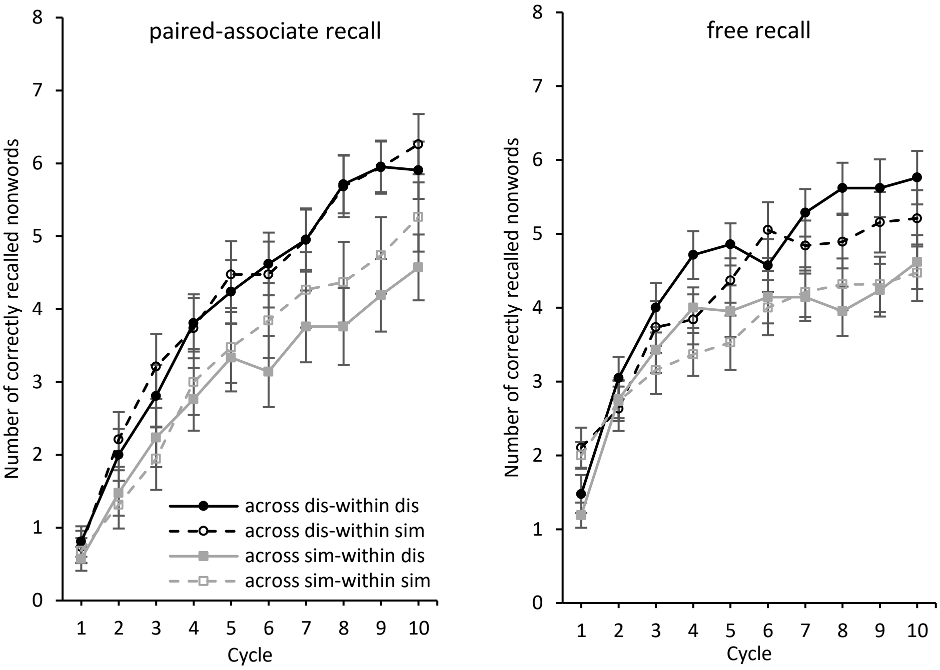


Figure 3.5. Performance accuracy in Experiment 6 according to Recall-type, Across-similarity and Within-similarity across cycles, with error bars representing the standard error of the mean.

In order to check that the group of participants who underwent fNIRS recording did not happen to differ in terms of their behavioural data from those who only participated in the behavioural experiment, the analysis was re-run with Group (with-fNIRS, non-fNIRS) as a between-participants factor. Group was not found to have a significant effect, $F(1, 37) = .62$, $MSE = 47.3$, $p = .437$, $\eta_p^2 = .016$, and neither did it interact with any other factors.

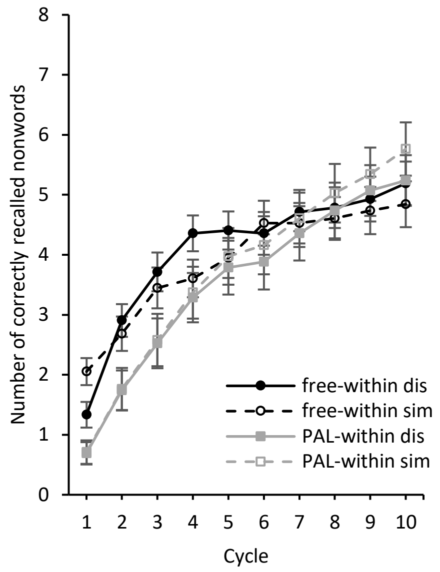


Figure 3.6. Recall accuracy in free and paired-associate recall (PAL) for within-similar and within-dissimilar nonwords across cycles, averaged across across-similarity. Error bars represent the standard error of the mean.

3.5.2.2 fNIRS data. HbO and HbR data are expected to show the opposite pattern across the recording period and hence a contrast between the two can be used to check the reliability of the measurements. Comparing the individual curves for the HbO data and the HbR data across time, here in Figure 3.7, there should be a general trend such that as HbO increases, HbR decreases and vice versa (Cui et al., 2011). As HbO data have been shown to be a more reliable measure, the analysis will use the HbO data for those channels that conform to the expected contrasts between HbO and HbR data (Dravida, Noah, Zhang, & Hirsch, 2017). Figure 3.7 appears to broadly confirm the expected inverse pattern for channels A, B and G, while the pattern is less clear for the other channels. Based on the better reliability of the recordings for channels A, B and G, the measurements recorded via these channels were selected for the following analyses.

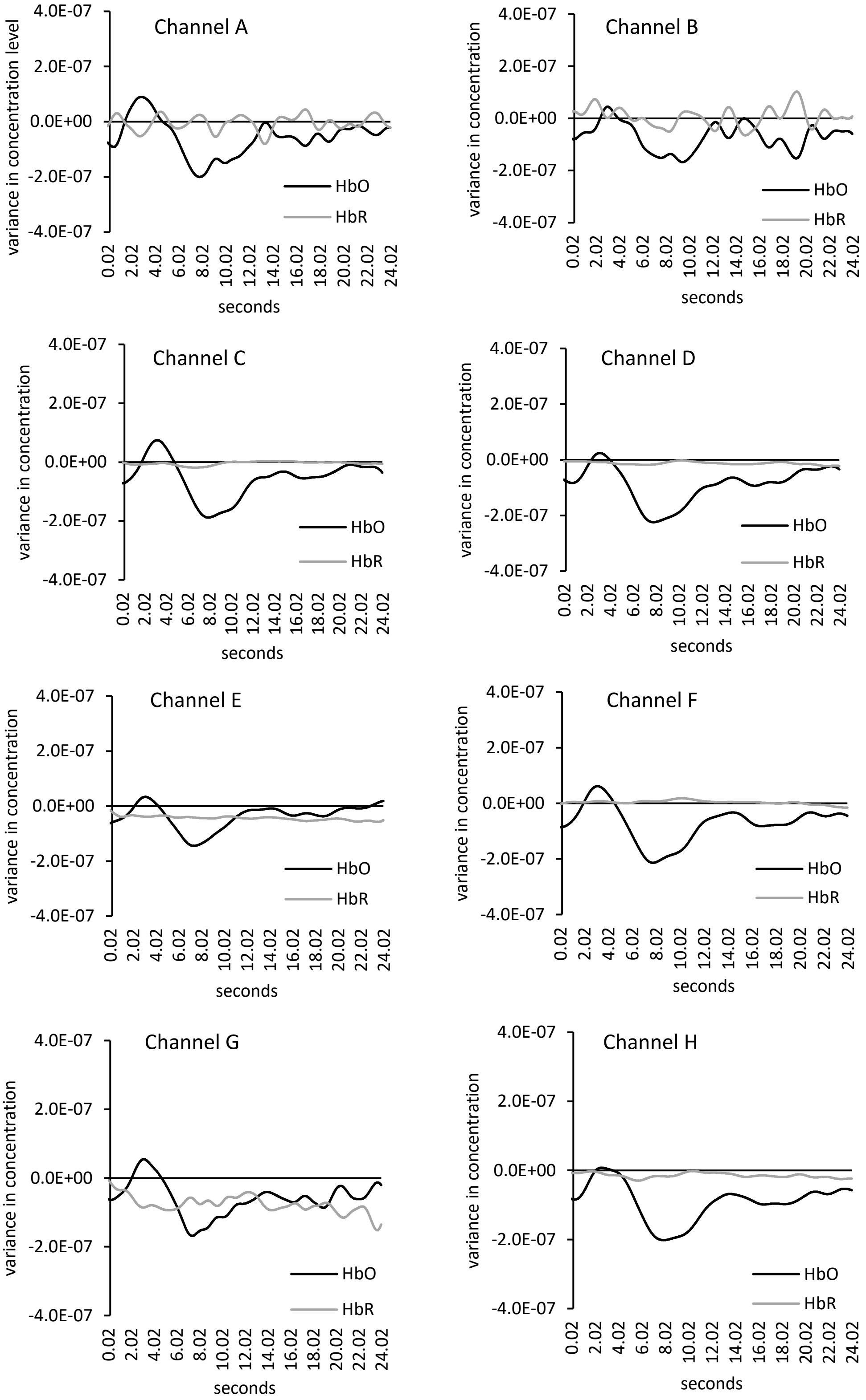


Figure 3.7. HbO and HbR concentration across the 24 s stimulus presentation period in each channel, collapsed across all experimental factors.

Figure 3.8 shows HbO concentration level for each of the eight [2 (within-similarity) × 2 (across-similarity) × 2 (recall-type)] conditions, averaged across the 24 s time period of nonword presentation, across cycles, and across channels A, B, and G. Figure 3.8 suggests that across-similarity interacted with within-similarity but differentially so in the two recall tasks: For paired-associate recall, lower HbO concentration was evident with across-similar compared to across-dissimilar lists, particularly with within-similar nonwords. In free recall, there was again some evidence—though not as great—for lower HbO for across-similar lists but, in contrast to the pattern in paired-associate learning, this was particularly or possibly only the case with within-dissimilar nonwords. Thus, in sum, while there was some clear evidence of a modulation of HbO concentration as a function of phonological similarity—particularly across-similarity in paired-associate learning—the pattern did not align with the expectation of higher activation in phonologically similar (and hence, it was presumed, motorically more demanding) conditions.

HbO concentration during the presentation of the nonwords in each condition was modelled with linear mixed effects models. Data for each participant for each of the four conditions in which they took part was averaged across the ten cycles of that condition. R (R Core Team, 2012) and lme4 (Bates, Maechler & Bolker, 2012) were used to perform a linear mixed effects analysis of the relationship between HbO, within-similarity, across-similarity, and recall-type. The data from the three channels for which the recordings appeared to be most reliable were combined and modelled, with the models of interest including the factors Recall-type, Across-similarity and Within-similarity. The random effects used in the models were the random slopes for time in seconds for each participant, time in seconds over each channel and for channels for each participant as sources of potential variance in the HbO signal, in order to account for variation in concentrations due to individual differences. Due to recording problems, data for one participant in the within-similar condition could not be included in the analyses.

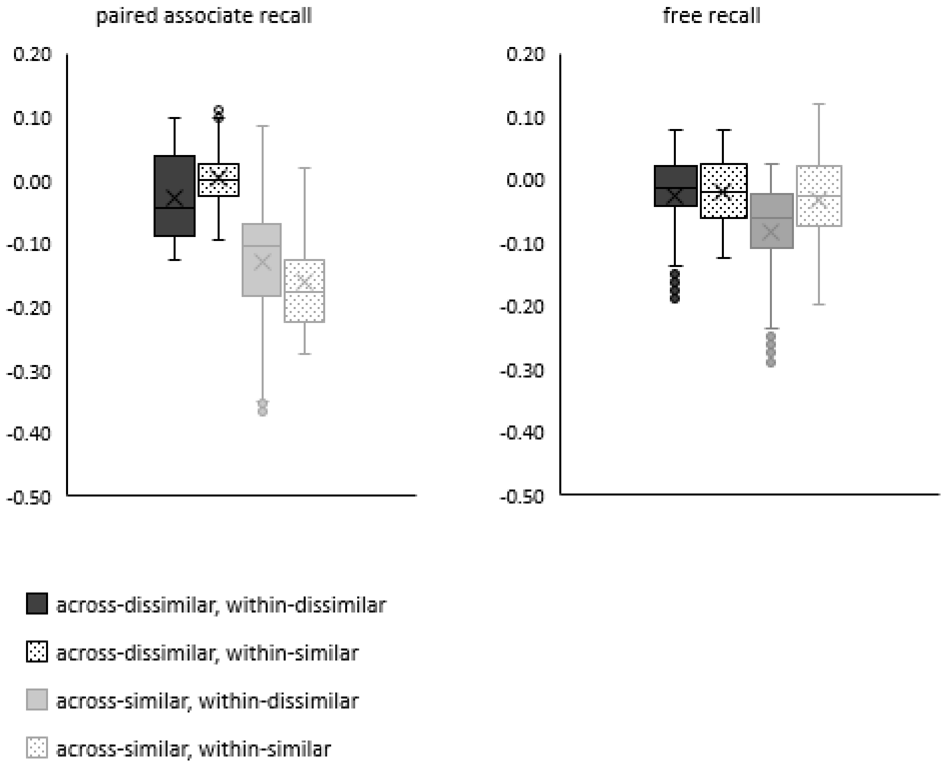


Figure 3.8. HbO concentration averaged across the 24 s stimulus presentation period according to Recall-type, Across-similarity and Within-similarity collapsed over Cycle and Channels A, B and G. The mean (indicated by the cross), median (the middle line dividing the box), the first and third quartiles (the lower and the upper parts of the box), the minimum and maximum values (indicated by the error bars), and the outliers (circles), are calculated from the data from each of the 0.02 s time-points across the 24 s presentation period.

Three models were constructed: Model 1 contained the fixed effects of Across-similarity, Within-similarity and the random effects. Model 2 contained the same effects as Model 1 but with an added term for the interaction of Across-similarity and Within-similarity. Finally, Model 3 contained the fixed effects of Across-similarity, Within-similarity, Recall-type (with their three-way interaction), and the random effects. Significance was tested by likelihood ratio tests of each model of interest against the simpler model with fewer factors, i.e., testing Model 3 against Model 2, to determine whether a model with the factors of interest was a better fit to the data. The two types of similarity interacted, indicated by the significantly better fit of Model 2 compared to Model 1, $\chi^2(1) = 209.12$, $p < .001$. Contrasting with the behavioural findings, the model containing an interaction of Across-similarity, Within-similarity and Recall-type (Model 3) was a better fit to the data than Model 2 (which did not contain a term for Recall-type), $\chi^2(4) = 14997.6$, $p < .001$. Model 3, indicated by a likelihood ratio test to be the best fit to the data out of the constructed models, was modelled with the formula: HbO ~ recall-type × across-similarity × within-similarity + (1 + seconds | participant) + (1 + channel | participant) + (1 + seconds |channel). Estimates of effects, standard errors and *t* values for this model are reported in Table 3.1. The modelling of the data was particularly dependent on the three-way interaction, as suggested by the *t* values.

Fixed effects	Estimate	SE	*t* value
Intercept	-1.04e-07	6.09e-02	-1.71
Recall-type	5.01e-08	2.42e-09	20.63
Across-similarity	1.34e-07	4.07e-08	3.28
Within-similarity	9.4e-07	2.43e-09	38.71
Recall-type × Across-similarity	-1.62e-07	3.43e-09	-47.17
Recall-type × Within-similarity	-1.57e-07	3.52e-09	-44.56
Across-similarity × Within-similarity	-2.09e-07	3.52e-09	-59.31
Recall-type × Across-similarity × Within-similarity	3.42e-07	4.97e-09	68.86

Table 3.1. Estimates of fixed effects in the final model for HbO for each Channel, with estimated directions of difference from paired-associate recall to free recall and from dissimilar to similar (for both Across-similarity and Within-similarity).

3.5.3 Discussion

In terms of the behavioural results, Experiment 6 replicated several key features of those found in Experiments 4 and 5. Recall type was shown to affect learning and, regardless of recall-type, phonological similarity across nonwords was again shown to reduce both recall accuracy and learning. In free recall, there was a recall accuracy benefit in the early cycles, with paired-associate recall later reaching a similar level of accuracy as free recall, confirming the observation in Experiments 4 and 5. However, there were also features that were not replicated: Experiment 4 showed that across-similar lists, regardless of within-similarity, were recalled more poorly than across-dissimilar lists, with within-similarity reducing accuracy for both across-dissimilar and across-similar lists, particularly in the middle cycles, accounting for the interaction with cycle. In Experiment 6, there was no main effect of within-similarity, the interaction of the two types of phonological similarity was not replicated, nor was there an interaction of either type of similarity with cycle. Further, Experiment 4 indicated that the effect of across-similarity was particularly evident in paired-associate recall, which again was not replicated in Experiment 6.

The lack of a main effect of within-similarity in the present experiment suggests, from the perspective of the perceptual-motor account, that this particular sample's motor-planning fluency was either not affected by whether the syllables in the nonwords were similar or that motor planning did not play a role at all. However, the marginal three-way interaction between within-similarity, recall-type and cycle may indicate instead that the effect of within-similarity depends on the task context, with within-similarity only affecting accuracy at early cycles, particularly for free recall. The effect of across-similarity on nonword recall accuracy appears rather more in line with the data from Experiment 4 regarding the effect of across-similarity on nonword learning, and can be observed in the interaction of across-similarity and cycle. These data suggest that the multilingual sample in this experiment may have been less affected by the phonological similarity implemented within the nonwords, while across-similarity appears to have had a similar effect in both experiments.

They may have been able to call upon a stronger, more diverse set of prior word-knowledge or/and enhanced novel motor planning, as also suggested by the higher recall performance by the multilingual group compared to the monolingual group in Experiment 4. The differences between the native languages of the two multilingual samples in the two experiments also further suggests that the characteristics of the multilingual sample may influence performance, which may relate to the inability to detect a learning benefit in the multilingual sample in Experiment 4.

The initial benefit across cycles of free recall over paired-associate recall was also observed, replicating the pattern from the two previous experiments, and supporting the conclusion that the specific task requirements will affect how the responses are produced. As the performance accuracy in the paired-associate task caught up to the accuracy with free recall, it also suggests that once the material is sufficiently learned, performance begins to rely more on the long-term representations of the nonwords, which negates the benefit in free recall of being able to use the most recently presented nonwords as the first to be recalled. As less active short-term novel motor planning is required for task performance, a larger proportion of the to-be-recalled material is available through long-term recognition.

Turning to the fNIRS data, although activation in the SMA/pre-SMA could not, as had inititially been intended, be examined in relation to learning, some differences in activation according to experimental conditions were observed on nonword recall. The data across the 24 s recording period showed that, for three of the recording channels, increases in HbO concentration tended to be associated with decreases in HbR concentration and vice versa, giving an indication of consistency in the recording and thus the reliability of the observations based on the data for these channels. The analysis of activity in the SMA/pre-SMA based on HbO concentration level collapsed across the 24 s presentation period suggested some possible differences in the motor processing of the nonwords as a function of task demands, across-nonword similarity and within-nonword similarity. However, the particular pattern of differences in HbO concentration levels was not predicted and due to problems recording activity during the rest-period, the concentration levels during the experimental conditions could not be compared to a baseline. This made it very difficult to ascertain whether relatively high HbO concentration represented an increase in activation relative to resting levels or whether, instead, a relatively low HbO level represented a reduction in activation compared to resting levels or some combination of the two. From comparing the levels of HbO across the four conditions (within each task), the clearest effect observed appeared to be some reduction in activation for across-similar lists. There was certainly no clear evidence of an *increase* in activation in the phonologically similar conditions as hypothesised.

A comparison of the behavioural and fNIRS results also indicated some apparent contradictions: While across-similarity appeared to have a larger effect on recall with paired-associate recall compared to free recall in both the behavioural and fNIRS data, this interaction was only marginally reliable for the behavioural data. Moreover, as noted, the effect in the fNIRS data was for *lower* activation with (across) phonologically similar lists, contrary to expectations. While some of the differences may be due to the fNIRS recording taking place during presentation rather than recall, where further motor planning would have occurred once the recall-stage was reached, it may also indicate the difficulty in mapping neural activation onto performance across a relatively long task duration with multiple components. Thus, it must be concluded at present that the fNIRS results are moot with regard to the role of motor planning and fluency in nonword recall. Nevertheless, the fact that there were some clear, reliable, differences in preSMA/SMA activation as a function of the current experimental manipulations indicates that the technique in general may, with further work, be a useful convergent method of addressing the aims of the present thesis, particularly if the issues regarding baseline and individual cycle-level measurement were resolved.

There were some further limitations to the fNIRS component of this experiment that also renders the results preliminary. It was assumed that motor planning would have occurred during stimulus presentation based, for example, on the fact that articulatory suppression required only during presentation markedly impairs recall (e.g., current Experiment 5; Hughes & Marsh, 2017). However, this was not directly tested within Experiment 6, and a comparison of activation during mere stimulus perception with and without the need for motor planning (cf. present Experiment 3) could not be made , leaving it somewhat unclear whether the activation observed here was specific to motor planning. Another possible problem was the size of the sample: With a large number of experimental factors, as was the case here, larger sample sizes have been preferred in recent fNIRS work, particularly in the language-domain (Bajaj, Drake, Butler, & Dhamala, 2014; Fu, Wan, Baker, Montgomery, Evans, & Gillam, 2016; Moriai-Izawa et al., 2012; Sugiura et al., 2018; Zhang, 2012), suggesting that a larger-scale future study may be required for a reliable and valid test of the present hypotheses. This is closely related to the problem observed for the ‘rest’ period recording, where a problem with recording enough data prevented the comparison or correction of the HbO and HbR concentration during the time-period of interest to the resting-state recording. The relatively small amount of data also prevented another analysis-of-interest regarding potential changes in the activation across cycles, as the nonwords become learned and when their recall should be increasingly less reliant on motor planning due to increasing support from long-term memory.

Regarding choosing only three channels for the analysis, the other channels were excluded due to possible unreliability due to the fact that the HbR data appeared not to show the inverse pattern to the HbO data as should be the case. Though a strictly equal degree of fluctuation or concentration might not be expected, clearly opposing patterns in the plotted HbO and HbR data should be apparent to indicate reliable recording, which has consistently been observed in studies of motor control and motor imagery (e.g., Abdalmalak, Milej, Diop, Shokouhi, Naci, Owen, & Lawrence, 2017; Chiarelli et al., 2007; Walsh, Tian, Tourville, Yucel, Kuczek, & Bostian, 2017). However, other findings have provided some evidence that the HbO measure might be more reliable than HbR, in that the systemic components in the HbO signal can increase the reliability of the HbO signal compared to the HbR signal (e.g., Dravida et al., 2017). Despite these limitations and the difficulties in interpreting the results, the findings suggest differential activation associated with nonword processing according to phonological similarity in the pre-SMA/SMA, which encourages further experimentation on this topic.

3.6 General Discussion

To summarise the results of the experiments in this chapter, Experiment 4 indicated that nonword learning in both free and paired-associate recall is attenuated by phonological similarity regardless of whether that similarity is implemented across the nonwords or within the nonwords. The results also showed a tendency for free recall responses to begin with the last presented item, resulting in an early benefit for free compared to paired-associate recall, which disappeared in the later cycles as more of the nonwords were learned. While Experiment 4 did indicate a recall benefit for multilingual speakers, it failed to replicate the previous findings which have also demonstrated a learning benefit (e.g., Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995). This could be expected to be related to the relatively heterogenous sample of multilinguals recruited for this study, as the characteristics of their language profiles were not restricted beyond whether they were able to speak at least one other language. Experiment 5 showed that while nonword recall accuracy was impeded by articulatory suppression regardless of presentation modality or recall-type, learning appeared to only be affected by suppression in the paired-associate task. There were, however, some indications that learning may have occurred during the recall phase in the free recall task, where participants could most benefit from learning from their own responses, as their order was free, rather than restricted as it was in the paired-associate recall task. The learning from output may have circumvented the effect articulatory suppression during presentation in this task. Experiment 6 replicated the behavioural results regarding the across nonwords phonological similarity effect, though the effect of within-similarity was not as clear. The findings from the fNIRS analysis indicated differences in HbO concentration levels in the pre-SMA/SMA during nonword recall according to phonological similarity, particularly across-nonword similarity. At a general level, this converges with the notion that phonological similarity affects motor planning, in line with the reconceptualisation of the effect as primarily an articulatory similarity effect rather than a phonological similarity effect (Jones et al., 2004; current Chapter 2). However, the particular pattern of activation was not as predicted and indeed difficult to interpret generally, for reasons already discussed earlier.

Before embarking on further discussion of the implications of the present chapter’s findings, it is important to highlight an issue with the design of

Experiments 4 and 6 that came to light during the writing-up stage of the thesis: The co-manipulation of within- and across-nonword similarity was not optimal in that the way across-similarity was implemented changed with the levels of the within-similarity factor. Specifically, within-similarity was implemented by having the syllables within a nonword share a vowel and this was the case both in across-dissimilar lists (e.g., *ponlon, temrem*) and across-similar lists (e.g., *lenven, nenhen*). However, across-similarity with within-dissimilar nonwords, in line with previous studies (Papagno & Vallar, 1992), was implemented mainly in terms of the nonwords having consonants in common (e.g., across-similar: *saries, somas* vs. across-dissimilar: *hawerty, samtis*) whereas with within-similar nonwords, across-similarity was, inadvertently, implemented in terms of vowel similarity (e.g., across-similar: *tinfinpil, rindingil* vs. across-dissimilar: *humdum, dantanlam*). (See Appendix 5, Supplementary Table 10, for the full list of nonwords in each condition.) A more appropriate implementation of across-similarity in the latter condition would have involved the nonwords in the across-similar lists again having consonants in common with one another but not vowels (e.g., *tinfinpil, tenfenpel*). This imbalance in the design complicates any attempt to interpret the interaction of the two similarity effects observed in Experiment 4. However, it does not alter the key conclusions to be drawn from the experiments because that interaction—or more complex interactions involving the two similarity manipulations—was not of central interest. What is of greater importance is that, despite this design issue, there remains clear evidence of independent effects of the two types of similarity. For example, in Experiment 4, there were effects of across-similarity on recall and learning regardless of the level of the within-similarity factor (and hence regardless of how across-similarity was implemented) as well as effects of within-similarity regardless of the level of the across-similarity factor. Thus, for the similarity effects, the focus will be on their independent effects rather than their interaction.

The current findings leave unresolved the apparent inconsistencies between the PS-LLD hypothesis, the differential coding argument, and the assumptions of the phonological store–based models of the Hebb effect. Given that phonological similarity was found to affect learning in free recall as well as the paired-associate task in Experiments 4 and 6, it would be difficult to argue that the phonological similarity effect in the latter task could be reascribed to general cue-based interference. As such, the differential-coding case for separate short-term and long-term memory systems is undermined: As phonological coding should be a unique signature of the action of the short-term system, phonological similarity effects should not be evident when the dependent measure is not short-term recall performance per se but the increasing use of the long-term system in support of that performance (i.e., learning across cycles).

On the other hand, it must be acknowledged that the fact that phonological similarity effects were observed in both free and paired-associate recall, and particularly that within-similarity affected nonword learning is, in itself, supportive of the PS-LLD hypothesis. That is, in principle, a within-nonword similarity effect, more clearly than an across-similarity effect, supports the notion that the store supports word-form learning by holding the phonological representations that make up a nonword temporarily while a long-term representation is constructed. This is because it can be argued that theoretical interest in the paired-associate learning task—as far as the goal to study word-form learning is concerned—does not centre on the recall or learning of the *list* of nonwords—across which similarity has previously been implemented (Papagno & Vallar, 1992)—but on the recall or learning of the form of each individual nonword. Indeed, it is the within-nonword similarity manipulation that maps conceptually onto the standard phonological similarity manipulation in the serial recall task: While the sequence of interest in serial recall is the novel sequence of letters, digits, or words, the sequence of interest in paired-associate learning is the novel sequence of elements making up a given nonword. However, there remains a contradiction between citing a phonological similarity effect in nonword learning as support for the PS-LLD hypothesis and phonological store-based models of verbal sequence learning in the Hebb repetition paradigm: The views based on the paired-associate paradigm cite the existence of a phonological similarity effect on learning as support for the PS-LLD hypothesis, while the phonological store-based models of the Hebb effect predict no effect of similarity on learning (e.g., Baddeley et al., 1998; Hitch et al., 2009; Page et al., 2006).

Furthermore, notwithstanding the contradiction with phonological store-based models of the Hebb effect, while the phonological similarity effects observed in the present chapter can be taken to support the PS-LLD hypothesis, they cannot be taken to uniquely do so: They are also consistent with the perceptual-motor account. In this view, the phonological similarity effect is primarily an articulatory similarity effect, where the similarity results in less fluent and error-prone motor planning. However, this account is also weakened somewhat by an empirical discrepancy, namely, that between the poorer nonword learning of phonologically similar visually-presented nonwords observed in this chapter (regardless of the particular manipulation of similarity) and the greater learning observed for phonologically similar visually-presented sequences in Experiment 1 using the Hebb repetition paradigm (from Chapter 2). There are of course numerous differences between the tasks used in the current chapter and the Hebb paradigm: For example, in the latter case, a single repeating sequence is learned while the present tasks involved the learning of multiple sub-sequences. Indeed, this may in part account for the much lower recall performance at early cycles in the current tasks than was the case in the Hebb experiments. In addition, unlike the Hebb paradigm, the present tasks did not involve ‘filler’ sequences; as such, learning may be more implicit in the Hebb task than it is in the current tasks. It is not clear at present, however, how any of these methodological differences could explain the opposite effects of phonological similarity on learning in the two paradigms. A possible way of trying to understand the nature of the phonological similarity effect on verbal sequence learning further would be to increasingly match the features of the tasks used here to the Hebb task, with a view to discovering which feature or set of features causes the opposite effects of phonological similarity on learning in the two paradigms.

Like phonological similarity, articulatory suppression has been shown to impair the learning of *word—nonword* pairs, but not *word—word* pairs (Papagno et al., 1989; Papagno & Vallar, 1992), which again has been taken to indicate the involvement of the phonological store, together with articulatory rehearsal, in nonword learning (Baddeley, 2007). However, this appears to be in contradiction with the argument that such learning is supported by the phonological store and not articulatory rehearsal (Baddeley et al., 1998), as well as the supposition of phonological store–based models of Hebb sequence learning that articulatory suppression, like phonological similarity, should not affect verbal sequence learning (e.g., Hitch et al., 2009; Page et al., 2006).

In sum, the effects of phonological similarity and articulatory suppression observed here on paired-associate learning are in line with previous studies that have been cited as support for the PS-LLD hypothesis (Papagno et al., 1989; Papagno & Vallar, 1992). However, the findings do not help to resolve the contradictions between phonological store–based accounts of Hebb sequence learning and paired-associate learning, the misalignment of the PS-LLD hypothesis and the notion of a phonological short-term store separate from a semantic long-term memory system, and they are at odds also with the contention that it is specifically the store, not articulatory rehearsal, that serves as the learning device. The presence of both the phonological similarity effect and the articulatory suppression effect in both the Hebb repetition paradigm and in paired-associate learning is accounted for more coherently by the perceptual-motor account in which the locus of both effects in both the short- and long-term aspects of both kinds of task is the same motor planning process: Articulatory suppression engages the motor-planning processes in a task-irrelevant activity while phonological similarity increases the error-proneness of motor planning. While the opposite effects of phonological similarity in the Hebb repetition paradigm and in the current experiments do present a challenge for the perceptual-motor account, they do also appear to point to the hypothesised importance of task requirements and goals in shaping performance.

CHAPTER 4

GENERAL OVERVIEW AND THEORETICAL IMPLICATIONS

4.1 Overview of Empirical Findings

The empirical chapters of this thesis have presented evidence that supports a view of verbal sequence learning as embodied rather than as a function of a specialised module, adding to the increasing body of evidence questioning the modularity of cognitive functions in general. The findings converge to support an embodied approach to cognition that highlights the importance of perceptual-input and motor-output processing in 'central' cognitive functions. The results suggest that verbal sequence learning, clearly a 'cognitive' function essential for language acquisition, may be usefully reconceptualised in terms of general-purpose motor planning and perceptual organisation processes. These processes are exploited to meet the behavioural goals of the task, in this case the reproduction of a novel verbal sequence, and the involvement of these processes in short term performance also supports the long-term learning of that sequence.

An important aspect of the experiments reported in Chapter 2 was the replication of, and in some cases the extensions to, some of the key findings within the context of verbal serial recall that have suggested a re-conceptualisation of verbal serial short-term memory performance in terms of embodied motor planning and perceptual organisation functions (Jones et al., 2006, 2004; Hughes et al., 2009, 2016). Experiment 1 demonstrated that, regardless of the modality in which to-be-recalled sequences are presented, performance relies on the motor-plan formed to meet the task goals: When the motor-planning processes are restricted via articulatory suppression, accuracy is reduced. A motor-plan formed for the recall of a phonologically dissimilar sequence is less error-prone than the response plan for a phonologically similar sequence. As the similarity effect is removed when motor-planning is restricted, the similarity effect is therefore articulatory, rather than being based on abstract phonological representations residing within a passive store. These patterns implicate a key role for motor planning processes in serial recall performance. A role for perceptual organisation was evident in the comparison of the recall of auditorily and visually presented sequences. If the presentation modality is auditory, the perceptual streaming of the auditory sequence can support recall performance. This can result in more accurate performance for auditory compared to visual sequences, the latter being more reliant on motor planning given the absence of a supporting passively-derived perceptual-object representation. This modality effect was restricted mainly to the end-boundary of the sequence and, in line with the notion that it results from passive processes, survived articulatory suppression. Thus, the pattern of results replicated that which has undermined the main empirical basis for postulating a passive phonological store separate from motor processes (e.g., Jones et al., 2004).

The serial recall data from Experiment 2 further supported a key role for motor planning by demonstrating that the motor plan encompasses paralinguistic characteristics of the sequence, in particular the temporal structure of items, as well as item identities. When performance was reliant primarily on motor planning (i.e., with visual presentation), the restriction of such planning by articulatory suppression resulted in the marked attenuation of the way the response-output closely mimicked the temporal grouping of the sequence as presented. If order in visual sequences was represented independently of motor processes , the temporal structure should not be lost when motor-planning is impeded (cf. Farrell & Lelièvre, 2012). Only with auditory sequences did the output grouping still strongly mimic the input-grouping when motor planning was impeded, in line with the view that passive auditory perceptual organisation processes yields a direct representation of order.

The main goal of the present thesis was to examine whether motor planning and perceptual organisation processes may also support long-term verbal sequence learning, and as such raise further questions about the need to invoke a dedicated short-term phonological module in models of cognition. The results of Experiments 1 and 2 relating to the Hebb repetition effect disconfirmed several predictions of the phonological store-based models of verbal sequence learning (e.g., Hitch et al., 2009; Page et al., 2006) and provided support for a perceptual-motor approach. These models propose that the learning mechanism can be linked up to, but is independent of, the phonological store (e.g., Burgess & Hitch, 2006; Page et al., 2009) and as such predict that factors deemed to affect phonological storage, in particular phonological similarity and articulatory suppression, should not affect learning. Both experiments showed, however, that when motor planning is impeded by articulatory suppression, learning of verbal sequences, particularly visually presented ones, is indeed reduced (and can even be eliminated; cf.Experiment 2), demonstrating that articulatory motor-planning of the to-be-recalled sequence is important not only for short-term serial recall but also sequence learning.

Further evidence in line with a motor account was that the initial fluency of the motor plan for a given sequence was also shown to be important: Experiment 1 showed a larger learning effect for a visually-presented phonologically-similar sequence compared to a phonologically-dissimilar sequence. It was suggested that the less fluent motor plan associated with the similar sequences had more to gain from repeated practice compared to the already relatively fluent dissimilar sequences.

This finding, again, contradicts the phonological store-based models, which predict no effect of phonological similarity on learning, as the effect of phonological similarity, on these models, is restricted to the phonological store rather than affecting the learning mechanism (e.g., Hitch et al., 2009; Page et al., 2006).

Experiment 2 included a condition in which inconsistent temporal grouping of the items in the Hebb sequence across repetitions was used to induce variability in the motor-plan. The importance of motor planning was most clearly demonstrated with visual sequences, where no additional processes are available to support learning: Learning was not only attenuated under articulatory suppression but also in the inconsistent-grouping condition. While the attenuation of learning with inconsistently grouped visually-presented sequences is consistent also with at least one of the phonological-store based models (Hitch et al., 2009), problematic for both such models (Hitch et al., 2009; Page & Norris, 2009) was that learning was still observed with inconsistent grouping if the presentation modality was auditory.

The importance of motor planning for verbal sequence learning was also demonstrated in the comparison of tasks that either did or did not require motor planning for the short-term task goal to be met. Experiment 3 showed that verbal sequence learning occurred when the repeating sequence had to be recalled after each presentation (and hence presumably planned), but learning was not seen generally for a repeating sequence that only had to be monitored (and hence presumably not assembled into a motor sequence-plan). This finding is also problematic for the phonological store-based models as they posit that production, or a motor-plan to support a potentially required production, is not necessary for sequence learning (Hitch et al., 2009; Page et al., 2006).

In addition to motor planning, a role for passive auditory perceptual organisation was suggested by several findings across Chapter 2. The survival of learning of auditorily-presented sequences under articulatory suppression in Experiment 1 and (to a lesser extent) in Experiment 2 suggested that if presentation modality is auditory, the repeatedly presented and perceptual organised auditory object is learned, and can support learning to some extent when motor planning is restricted. The comparison of visual and auditory sequences with an inconsistent temporal grouping in Experiment 2, with learning present for auditory, but not for visual sequences, also suggested that just as in verbal serial short-term memory performance, auditory perceptual organisation of the input into an auditory object can to some extent also support the learning of that sequence.

In Experiment 3, single vs. alternating-voice presentation was used to promote different perceptual organisations of the sequence. With single-voice presentation, no learning was observed for a transfer-Hebb sequence where the Hebb sequence presented on previous cycles was rearranged so that the items from odd and even positions were now grouped together. However, critically, learning was indeed observed for the transfer-Hebb sequence when that sequence had previously been presented in two alternating voices. This suggests that, with alternating-voice presentation, the items in a given voice had been organised into a single stream and learning operated on the basis of this perceptual grouping not (just) according to true temporal order. The fact that this was only the case when the sequence was monitored, rather than requiring recall in the true temporal order, suggests that the involvement of the passive perceptual organisation of auditory stimuli in the learning of auditorily-presented verbal sequences is more apparent when a motor-plan assembled according to true temporal order does not contradict it. These patterns are, again, in contrast with the phonological store account, where the identity of the items should lose all acoustic characteristics once they have entered the store (Baddeley, 2007). If this were the case, learning should not be affected by voice alternation, yet the results demonstrate that the perceptual properties of the material do affect verbal sequence learning. Arguably, the phonological store-based models might be able to account for the learning of the by-voice sequences observed in the alternating voice conditions of Experiment 3 by adding a front-end streaming parameter (see, e.g., Page & Norris, 1998), but it is not apparent that this would need to be a front-end process to a phonological store. An appeal to auditory perceptual organisation of items into streams according to their acoustic characteristics, together with motor planning, appears to account for the learning without the need to invoke a phonological store inbetween.

Chapter 3 re-evaluated the evidence for the PS-LLD hypothesis drawn from the paired-associate learning task (e.g., Papagno et al., 1991). In particular, a contradiction between the differential coding argument on which the existence of a separate phonological store is partly based and the finding that phonological similarity affects long-term retrieval was investigated by examining whether phonological similarity in the paired-associate task affects word-form learning per se, or instead affects merely the pairing of the nonword and its cue word. If the latter, the contradiction would have been resolved but at the same time this would undermine the use of the phonological similarity effect in paired-associate learning as evidence for the PS-LLD hypothesis. The potential effect of phonological similarity on word-form learning was also investigated more clearly than hitherto by a novel manipulation whereby phonological similarity was implemented within the nonwords. Experiment 4 therefore included manipulations of recall-type by contrasting paired-associate recall with free recall, and of two types of similarity (within and across the nonwords). The results demonstrated that regardless of recall type, nonword learning is reduced by both across- and within-nonword similarity. The effect of within-similarity in particular suggests that word-form learning specifically is reduced when individual nonwords contain phonologically similar syllables. These particular results are therefore in line with the PS-LLD hypothesis but are equally explicable in terms of the action of constructive motor planning on nonword learning, where similarity (of either kind) would be expected to affect motor planning.

Further evidence for a role for motor planning was that articulatory suppression had a detrimental effect on nonword learning (Experiment 5), converging with the observations made in Chapter 2 in the context of Hebb sequence learning. The effect of articulatory suppression on learning in the paired-associate/free recall task does, however, appear to be smaller than that on Hebb sequence learning. This may be related to the different response requirements in the two paradigms: During the Hebb task, responses were made via mouse-clicks on images of the stimuli on the screen order reconstruction, meaning that participants could be (and were) required to continue suppression throughout the recall period. In contrast, during the paired-associate task, participants were required to recall the nonwords vocally, meaning that articulatory suppression could only be required during stimulus presentation. This left the motor planning processes involved in producing the spoken response more available for also supporting the learning of that response. Indeed, it was suggested that this would have been especially likely in the free recall task in Experiments 4-6 where order of output was unconstrained. The findings suggested that participants may have taken advantage of the unconstrained recall order, which could have promoted the contribution of consistent output-ordering to recall and thus potentially also learning. Accordingly, there was no evidence that articulatory suppression attenuated learning in this case. The notion that response-period motor planning contributed more to paired-associate learning than Hebb sequence learning is lent further plausibility by the fact that, in the former setting, the same material is to be recalled again in a relatively short amount of time and without intervening trials, unlike in the Hebb paradigm in which filler material is interspersed among the presentation of the repeating sequence. In line with the previous findings, articulatory suppression was not found to affect learning or recall differentially according to presentation modality, which may be problematic both for phonological loop-based approach to paired-associate learning and the perceptual-motor account. If articulatory suppression affects an articulatory rehearsal process that is subservient to a phonological store, visually presented lists should have been more affected by suppression compared to auditory lists, which did not appear to be the case. A further complicating issue for the PS-LLD hypothesis is that it emphasises the role of the store, rather than articulatory rehearsal, in word-form learning generally (Baddeley et al., 1998) and yet emphasises articulatory rehearsal when seeking to explain the effect of articulatory suppression on paired-associate learning (e.g., Papagno & Vallar, 1992). The perceptual-motor account would have also expected auditory lists to be less affected by suppression compared to visually presented lists, due to a possible contribution of passive auditory perceptual organisation. However, this may not have been evident in this task, as the object boundaries during presentation differ from those in the Hebb repetition paradigm.

The functional near-infrared spectroscopy results from Experiment 6 provided some support that the SMA/pre-SMA, an area associated with motor planning, is differentially activated during nonword recall depending on phonological similarity. When modelling the HbO data, the models that included the factors that are expected to influence motor planning fluency were a better fit than the simpler models. This also supports the proposal that subvocal motor planning of the to-be-recalled nonwords occurs during presentation. However, the activation in the SMA/pre-SMA could not be investigated in term of nonword learning due to technical difficulties with the recordings. While the findings suggest that the SMA/pre-SMA may be involved in preparing verbal responses for nonword production, inferences regarding learning could not be made, and the problems with the resting state recordings leave it unclear how the activation during the proposed period of subvocal rehearsal would compare to resting state activation. The fNIRS results must be seen as preliminary due to the relatively small sample size and problems with the resting state recordings, and further work will be required to test the hypothesised contribution of the SMA/pre-SMA to nonword learning.

Together, the empirical findings indicate a key role in verbal sequence learning for the motor processes involved in the planning and production of the relevant verbal output in a number of converging ways: First, from impaired performance when these processes are engaged in task-irrelevant behaviour; second, from the influence of articulatory similarity and motor-plan variability across sequence repetitions; and third, from the absence of sequence learning when the short-term output does not require articulatory sequencing processes.

4.2. Wider Theoretical Applications

4.2.1 The Coherence of the PS-LLD Hypothesis

An apparent mismatch exists between different strands of evidence for the phonological store’s role in language learning, in particular between the evidence from the Hebb sequence learning paradigm on the one hand and the evidence from paired-associate learning and nonword repetition paradigms on the other. The mechanism responsible for Hebb learning according to phonological store-based models—either a positional context signal (Burgess & Hitch, 2006) or primacy gradient (Page & Norris, 2009)—is immune to the manipulations that affect the phonological store. In the paired associate paradigm, both phonological similarity and articulatory suppression impair learning, as replicated in the present thesis, and this has been taken to indicate the phonological store's involvement (Papagno & Vallar, 1995; Papagno et al., 1991). The need to invoke a learning process separate from the store appears to be related to the architecture of the phonological store, which in itself contains nothing that would support learning of verbal sequences or even the short-term recall of verbal sequences, which was the original key function ascribed to the store. The store holds representations of individual verbal items that decay quickly. The phonological store-based modelling of Hebb sequence learning (and sequencing generally) has therefore involved adding a sequencing mechanism: a primacy gradient (Page & Norris, 2009) or a positional context signal (Burgess & Hitch, 2006). It could be argued that the articulatory component attached to the phonological store is inherently sequential in its operation, and thus could act as a substrate for sequence encoding and learning. And indeed, this is precisely the reason that articulatory processes are given such prominence in the perceptual-motor account of verbal serial short-term memory and, in the present thesis, in verbal sequence learning. However, the articulatory control process does not owe its inclusion in the phonological loop model to its sequential nature but to its use of phonology; that is, its function is to convert individual items into phonological form and thereafter to refresh individual phonological item-representations in the face of a decay process (Baddeley, 1986, 2007). As such, the order representation and learning mechanisms in models of Hebb sequence learning operate separately from the phonological store and take a domain-general, non-phonological, form (Burgess & Hitch, 2006; Page et al., 2006). Indeed, if the learning mechanisms in these models were specifically phonological, separate mechanisms would need to be assumed to account for sequence learning in non-verbal domains, including a specific mechanism for sequence learning of finger and hand movements (e.g., in learning to write), another for visual sequences, or indeed for any other nonverbal skill. This would clearly not be a sustainable system, and result in a massive proliferation of modules for all the cognitive functions that would need to be supported by a specialised module. As it is the added, non-specifically-phonological, sequencing mechanism that supports learning in these models, it is difficult to sustain the idea that the phonological store per se is a ‘language learning device’. The current findings suggest that there is no need, in any case, to invoke a specific store for the short-term maintenance of verbal material in order to account for verbal sequence learning.

Another important strand of evidence for the PS-LLD hypothesis is the positive correlation between a child’s ability to repeat back a heard nonword (nonword repetition) and their vocabulary size, with the capacity of the store thought to determine the ability to acquire vocabulary (Baddeley et al., 1998; Gathercole, 2006). Like serial recall, nonword repetition requires a reproduction of a verbal sequence, and performance in both exhibits a bowed serial position curve and is affected by the length of the sequence, and performance in one task correlates positively with that in the other (Archibald & Gathercole, 2007; Gupta, Lipinski, Abbs, & Lin, 2005). Both nonword repetition and serial recall are deemed to rely on a phonological store, but nonword repetition has been suggested to be a purer measure of the store (as opposed to the phonological loop as a whole): “nonword repetition provides a measure of the phonological store, not phonological rehearsal” (Baddeley, Gathercole, & Papagno, 1998, p. 168). However, results are emerging showing that nonword repetition is impaired by both phonological similarity (between syllables) and articulatory suppression (Harvey & Hughes, 2019). While the presence of the phonological similarity is consistent with nonword repetition as reliant on the phonological store, the effect of articulatory suppression goes against the notion that nonword repetition is a pure measure of the phonological store. Both findings, particularly when considered together, are better accounted for in terms of a role for motor planning in nonword repetition. Thus, to the extent that the correlation between nonword repetition and vocabulary size cannot be explained solely in terms of a large vocabulary supporting better nonword repetition (e.g., Bowey, 2001; Melby-Lervåg et al., 2012; Metsala, 1999), the available data suggest that it may be short-term motor-planning ability that is predictive of vocabulary acquisition, not passive phonological storage capacity.

4.2.2 A Perceptual-Motor Approach to Verbal Sequence Learning

The perceptual-motor account has previously been shown to provide a parsimonious explanation of verbal serial short-term memory phenomena (e.g., Jones et al., 2006, 2004; Jones & Macken, 2018), which can now be extended to verbal sequence learning. The current findings point to verbal sequence learning as reliant on the general-purpose motor-planning and perceptual organisation processes that are exploited according to the task goals, rather than requiring a specialised phonological store. Given that the task goal in both verbal serial short-term memory tasks and in verbal sequence learning is centred around the production of a verbal sequence, the processes utilised in these tasks appear to be those that would be involved in the production of other similar behaviours, e.g., the production of a novel sentence, which also requires the on-the-fly planning of speech-elements in an appropriate order. Behaviour in the task is also affected by the degree to which the passive organisation of the perceptual input aligns with the behavioural goal: Auditory material such as speech, which is, by its nature, sequential (Bregman & Rudnicky, 1975), provides a stronger perceptual organisation-derived representation of a verbal sequence than visually presented material, which is more typically spatially rather than temporally organised (Wagemans et al., 2012). Such passive perceptual organisation can either facilitate recall (e.g., for recalling the order of items near the end-boundary of a sequence; Nicholls & Jones, 2002) or hinder recall (e.g., when the perceptual organisation is at odds with the motor planning most suitable in the context of the particular task demands; Hughes et al., 2009, 2016; Macken et al., 2016).

The similarities in non-verbal motor short-term memory performance to verbal short-term memory phenomena (Wilson & Fox, 2007) also support the notion of general rather than separate, distinct processes supporting memory and linguistic processes (MacDonald, 2016). This in turn supports the applicability of the perceptual-motor account to the Hebb effect observed with nonverbal stimuli: It could be expected that nonverbal sequence learning would rely on the perceptual and motor processes that are involved in the production of the behaviour in question. An account employing perceptual and motor planning functions as the mechanism for verbal sequence learning appears better equipped to account for the resemblances between such learning and motor sequence learning effects (Sakai, Kitaguchi, & Hikosaki, 2003) and visuo-spatial learning (Couture & Tremblay, 2006). For example, the learning of a hand-movement sequence would, from this standpoint, be supported by the nonverbal motor-planning functions that are involved in the production of hand-gestures.

Verbal sequence learning is supported by the increasing accuracy and fluency of the motor-plan generated to retain and reproduce a particular verbal sequence for short-term output production. In particular, there is evidence that practice with the co-articulation of successive items improves serial recall accuracy independently of the frequency of exposure to the items being co-articulated (Woodward et al., 2008). Motor-skill learning studies have also identified chunking and sequencing as particularly important, where individually familiar sub-components of the new behavioural sequence are formed into larger units through practice, improving the performance in that new behaviour (Sakai et al., 2003). Additionally, temporal patterns, such as rhythm, have been indicated to have an important influence on the learning of complex new motoric skills (Sakai et al., 2004). The temporal pattern supports the flexible and fluent way of producing the new behaviour, leading to improved performance (Wymbs et al., 2012). This was also observed in Experiment 2, where a consistent temporal grouping of an auditory sequence was better suited to support the learning of that sequence, and also in the way in which the temporal pattern of the participants' responses mimicked the temporal grouping of the sequence presentation.

It appears that while there are processes that are particularly suited for short-term memory tasks (Norris, 2017), a clear separation to short-term and long-term memory functions may not be supported. The processes involving the use of perceptual systems (Bregman, 1990, Jones et al., 2004) or long-term skills such as motor planning, or processes that utilise long-term knowledge (Bowey, 2006, Jones

& Macken, 2018), are likely to be flexibly employed to support both long-term and short-term elements that are required by the current task goals.

4.2.4 Embodied Cognition and Verbal Sequence Learning

This thesis used verbal sequence learning as an example behaviour to critically examine embodied and modular approaches to cognition, with the findings favouring the embodied approach: even the complex, ‘higher level’ process of verbal sequence learning appears to be explicable parsimoniously by recourse to processes evolved for goal-relevant motor output planning and production and the perceptual organisation of sensory input. From this perspective, short-term retention and its translation into learning is supported by the recruitment of general-purpose processes that have traditionally been considered ‘lower level’, together with the use of appropriate long-term knowledge and skills (Macken et al., 2015). The results presented here contrast with the theoretical position whereby perceptual and motor processes are insufficient for long-term learning (e.g., Craik & Watkins, 1973; Lockhart & Craik, 1990), while fitting well with the wider embodied cognition literature. Development of new behaviours and skills has been shown to be based on the interaction of initially random and spontaneous movements and the changing context, where the repeated interaction results in an emergent, organised product, which holds promise for generalisability of the embodied approach to development and other functioning (e.g., Smith & Thelen, 1994). This would eliminate the need for pre-wired modules that result in new behaviours with maturation. The many embodied, or grounded, approaches to cognition have re-defined cognitive phenomena, such as considering vision as action-guiding environmental processing, where bodily sensation and feedback relating to any goal-oriented movement in the environment is closely related to the visual processing, rather than vision simply constituting the re-coding of sensory data into internal representations for manipulations in an abstract format (O’Regan & Noe, 2001). This is seen also in the similar activations of motor-related neural networks when visually observing others perform movements and when producing those same movements (Rizzolatti & Craighero, 2004).

The findings of the present thesis show that articulatory planning is involved in verbal sequence learning. This suggests that it is the nature of the material that is to be learned that determines the processes that will support that learning (Morris, Bransford, & Franks, 1977): when the learning relates to relatively surface, verbal forms, articulatory planning can support it. The current findings align well with concepts from theories of skilled motor learning. In motor control and motoric learning, three stages produce the behaviour: First, selecting the task-relevant targets from the environment, then sequencing the required behaviour targets, and then producing the commands for the movements. Learning may improve performance by affecting one or more of these stages, either by improving the efficacy of selecting goal-relevant objects-for-actions, or by improving the sequencing processes (Willingham, 1998). Studies of motor-skill learning have pointed to the particular importance of chunking and sequencing, where the individual parts of the new movement are formed into larger units through learning, where repetition of the parts of the new movement together produce improved performance in the new action (Sakai et al., 2003). The current data show the same pattern in articulatory sequence learning, where individual components are formed into a unit through repetition. Bigger improvements have been observed for motor-skills, which are, in the beginning, harder (e.g., Grantcharov et al., 2003). This observation from other motor domains was conceptually replicated in the current verbal sequencing data in the larger improvement observed in Experiment 1 for similar compared to dissimilar sequences. Findings from other motoric domains align with our findings from Experiment 2, where temporal grouping affected both the short-term response and learning. Previous findings have indicated that rhythm, or temporal patterns, emerge in complex motor behaviours with learning (Sakai et al, 2004). This temporal patterning supports the fluent, flexible and automated ways of producing the learnt behaviours.

4.2.5 Motor Planning and First-Language Acquisition

Native-language acquisition in children has been proposed to rely on at least partially different mechanisms from those supporting second-language learning; in particular, it has been suggested that articulatory rehearsal is not involved in infant language acquisition (e.g., Baddeley et al., 1998). This view has been based largely on the finding that the phonological similarity effect is not observed with visual sequences in young children (e.g., Henry, 1991), suggesting that they do not have the articulatory-based capacity to convert the graphemic input into phonological form (e.g., Baddeley & Larsen, 2007). Additionally, rehearsal as a means to refresh decay-prone phonological representations has been considered to only become a viable strategy when articulation rate, which is comparatively low in children, increases (Hulme, Thomson, Muir, & Lawrence, 1984). Indeed, some have argued that a qualitative change may occur in the rehearsal process before children can begin to use it to support short-term recall (Flavell, Beach, & Chinsky, 1966). These findings led to the suggestion that children do not rehearse verbal material to support recall until they are around seven years old (e.g., Baddeley et al., 1998; Gathercole, 1998). Given that children learn, at a conservative estimate, more than 2000 words by the age of five, with vocabulary increasing by 3000 words per year for the following few years (Nagy & Herman, 1987; Smith, 1926), this would leave a considerable part of the vocabulary to other learning mechanisms. Specifically, infant language acquisition is thought to be supported by the passive phonological store component of the phonological loop system (e.g., Baddeley et al., 1998).

If articulatory rehearsal is not involved in infant language learning, this would question the involvement of motor planning in language learning as one of the core mechanisms together with auditory input processing. However, if different learning mechanisms were involved in first and second language learning, it would also appear to be problematic for the phonological store account. The suggestion that word-form learning in children learning their first language is supported not by articulatory rehearsal but by the passive phonological store is at odds with the idea that learning of auditorily presented nonwords or foreign words is affected by articulatory suppression due to the blocking of articulatory rehearsal, as suggested by the phonological loop theory–based explanations of paired-associate learning in adults (e.g., Papagno & Vallar, 1992). That is, on the PS-LLD hypothesis, articulatory rehearsal is deemed to only play a role in second-language learning in adults (or in children old enough to use articulatory rehearsal; Baddeley et al., 1998, p. 167). But if this is the case, it becomes questionable, on this account, whether any of the results from paired-associate learning studies—which have all involved adult participants (Baddeley et al., 1998; Papagno et al., 1991; Papagno & Vallar, 1992)—have a bearing on the core mechanisms involved in word-form learning.

However, research has suggested that children may rehearse much earlier than originally suggested. Canonical babbling, a form of rehearsal for articulatory movements for producing sounds, generally appears by 7-8 months of age, with patterns of babbling corresponding to the first produced words (Davis &

MacNeilage, 1995). Indeed, developmental research has indicated that language learning is preceded by imitation of speech sounds, as infants attempt to repeat the speech sounds they hear (e.g., Perez-Pereira, 1994; Ramer, 1976; Rodgon & Kurdek, 1977). Given that the Hebb effect can be observed in very young children (Mosse & Jarrold, 2008; Yanaoka, Nakayama, Jarrold, & Saito, 2019), it would be consistent with repeated opportunity to rehearse the sequence playing a role in native-language acquisition. The acoustic consequences of a vocalisation—be it babbling or a word—can be used to model the refinement of the articulatory–based learning of how the sound is produced, and that both babbling and imitation are essential early steps, followed by the refinement of articulation and the correct rhyming of speech (Bailly, 1997; Philippsen, Reinhart, & Wrede, 2014). This is consistent with the observation that infants begin to learn to control their motor activities based on auditory feedback around the same time as they begin to babble in the canonical way associated with the language they hear (Ejiri, 1998). Additionally, the marker for articulatory rehearsal for recall—a phonological similarity effect with visually presented lists—has also been observed in children younger than seven, implying that there is no qualitative change in rehearsal such as that appealed to previously to suggest that rehearsal is not used by younger children (Jarrold & Citroën, 2013). The foregoing considerations, then, suggest that motor planning may indeed play a role in first-language learning and this, together with auditory perceptual organisation processes, seems to obviate the need to invoke a separate passive phonological store to account for such learning.

4.3 Limitations and Further Work

Further work will be required to determine the applicability of the perceptual-motor account to other types of learning, as currently its explanatory capability has only been directly tested with verbal serial short-term memory and verbal sequence learning, though promising evidence from other domains indicate a key involvement of motoric and perceptual processes also in non-verbal gestural learning (Adolph & Hoch, 2019; Sakai et al., 2003). From the perceptual-motor perspective, learning of other new behaviours would be reliant on the processes that can be recruited to support elements of that behaviour over the short term. The more fluent knowledge or skills that are applicable to the new behaviour, the better they should be suited to support the learning. Further work will be required to show how this applies to learning other skilled behaviours.

Another key limitation is the extent of the conclusions that can be drawn from the single fNIRS experiment regarding the role for motor processes in the SMA/pre-SMA during the task phase involving the presentation of the to-be-recalled stimuli. In addition to the specific limitations discussed in the discussion section of Experiment 6, a key consideration for general theorisation is related to the proposed wide network of systems supporting behaviour (e.g., Postle, 2006; Buchsbaum & D’Esposito, 2008), and the limited cortical area captured by the fNIRS method that was used here. As clear changes in the HbO concentrations were measured throughout the 24 s recording period, differential activity across a wider network could be postulated, but remains for future work to establish. Though a detailed cognition-to-brain-activation correspondence was beyond the scope of the present thesis, it appears to be a promising avenue for further investigation of how the relevant perceptual and motor processes are recruited to support higher-level cognitive functions such as verbal sequence learning.

The current work has limits also in direct applicability. Though the theoretical standpoint relating to embodied processes being involved in verbal sequence learning has received support, an important avenue for further work will be to examine its applied implications. Given the wide application of many modular cognitive models to education, explanations of deficits in cognitive function, and to development, both normal and abnormal (e.g., Gathercole et al., 2004; Gathercole & Baddeley, 2014), future work will be needed to investigate whether and how these applications can be maintained and possibly extended by reference to embodied processes and without recourse to dedicated modules.

4.4 Conclusions

The general arguments supported through the experiments in this thesis centre around an embodied processes-based view of memory: The particular demands of the task determine the processes used to meet the task goals. The processes applied to complete complex goals involve the organisation of the sensory information and the planning and production of appropriate behavioural output. Learning and long-term effects of memory are determined by the task demands, when same or similar requirements are encountered again (e.g., Craik & Lockhart, 1972; Kolers & Roediger, 1984; Morris et al., 1977). When articulatory motor processes are applied to meet the immediate task goals, such as in short-term serial recall, they can incidentally influence the long-term effects on performance that requires the same articulatory processes. In a similar vein, the auditory perceptual organisation processes involved in processing the sensory input can have effects not only on short-term recall but on long-term retrieval when the products of those processes happen to be useful for meeting subsequent task demands.

REFERENCES

Abdalmalak, A., Milej, D., Diop, M., Shokouhi, M., Naci, L., Owen, A. M., & Lawrence, K. S. (2017). Can time-resolved NIRS provide the sensitivity to detect brain activity during motor imagery consistently?. *Biomedical Optics Express, 8*(4), 2162-2172.

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*(1), 50-68.

Ackermann, H., & Riecker, A. (2010). The contribution (s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Structure and Function, 214*(5-6), 419-433.

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology, 70*, 141-164.

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence, 149*(1), 91-130.

Archibald, L. M., & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic Bulletin & Review, 14*(5), 919-924.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, 6(2), 115-116.

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

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

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

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

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

Baddeley, A. D. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 544-549.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189-208.

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 & 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.

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.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575-589.

Baddeley, A., & Larsen, J. D. (2003). The disruption of STM: A response to our commentators. *The Quarterly Journal of Experimental Psychology Section A, 56*(8), 1301-1306.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *The Quaterly 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., Cocchini, G., Della Sala, S., Logie, R. H., & Spinnler, H. (1999). Working memory and vigilance: Evidence from normal aging and Alzheimer's disease. *Brain and Cognition, 41*(1), 87-108.

Bajaj, S., Drake, D., Butler, A. J., & Dhamala, M. (2014). Oscillatory motor network activity during rest and movement: an fNIRS study. *Frontiers in Systems Neuroscience, 8*, 1-12.

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech, 48*(1), 1-27.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(617-645).

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences, 22*(4), 637-660.

Barsalou, L. W., Breazeal, C., & Smith, L. B. (2007). Cognition as coordinated non-cognition. *Cognitive Processing, 8*(2), 79-91.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: linear mixed-effects models using S4 classes. R package version 0.999375-42.2011.

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

Behrmann, M., & Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends in Cognitive Sciences, 17*(5), 210-219.

Boisgontier, M. P., & Cheval, B. (2016). The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews, 68*, 1004-1005.

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 & 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.

Buller, D. J., Hardcastle, V. (2000). Evolutionary psychology, meet developmental neurobiology: Against promiscuous modularity. *Brain and Mind, 1*(3), 307-325.

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.

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

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

Breitling, D., Guenther, W., & Rondot, P. (1986). Motor responses measured by brain electrical activity mapping. *Behavioral Neuroscience, 100*(1), 104.

Bregman, A. S. (1990). Auditory Scene Analysis (MIT Press, Cambridge, MA).

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89*(2), 244-249.

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

Broadbent, D. E. (1984). Modules in models of memory. *Behavioral and Brain Sciences, 7*(1), 86-94.

Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities*, *43*, 106-122.

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology, 80*(2), 1-17.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*(3), 441-469.

Caplan, D., Waters, G., & Howard, D. (2012). Slave systems in verbal short-term memory. *Aphasiology, 26*(3-4), 279-316.

Carruthers, P. (2003). Moderately massive modularity. *Royal Institute of Philosophy Supplement, 53,* 67-89. doi:10.1017/S1358246100008274

Carruthers, P. (2005). The case for massively modular models of mind. *Contemporary Debates in Cognitive Science ,* ed. R. Stainton, 205-25.

Carruthers, P. (2006). *The architecture of the mind*. Oxford, UK: Oxford University Press.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences, 3*(9), 345-351.

Chiarelli, P. A., Bulte, D. P., Gallichan, D., Piechnik, S. K., Wise, R., & Jezzard, P. (2007). Flow‐metabolism coupling in human visual, motor, and supplementary motor areas assessed by magnetic resonance imaging. *Magnetic Resonance in Medicine, 57*(3), 538-547.

Chomsky, N. (2011). Language and other cognitive systems. What is special about language? *Language learning and development, 7*(4), 636-278.

Cohen, R. L., & Johansson (1967). The activity trace in immediate memory: A re-evaluation. *Journal of Verbal Memory and Verbal Behavior, 6*, 139-143.

Cona, G., & Semenza, C. (2017). Supplementary motor area as key structure for domain-general sequence processing: A unified account. *Neuroscience & Biobehavioral Reviews*, *72*, 28-42.

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

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In Hirschfeld, L., & Gelman, S. (Ed.), *Mapping the Mind: Domain specificity in cognition and culture*. NY: Cambridge University Press.

Cosmides, L., & Tooby, J. (1995). Cognitive adaptations for social exchange. In Barkow, J. H., Cosmides, L., & Tooby, J. (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture*. NY: Oxford University Press.

Couture, M., Lafond, D., & Tremblay, S. (2008). Learning correct responses and errors in the Hebb repetition effect: Two faces of the same coin. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(3), 524.

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34*(8), 1720-1729.

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory and Language, 46*, 153-177.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

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

Cunningham, T. F., Healy, A. F., & Williams, D. M. (1984). Effects of repetition on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 575-597.

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *Neuroimage, 54*(4), 2808-2821.

Dell, G. S. (1995). Speaking and misspeaking. *Cognitive Science*, *1*, 183-208.

Dell, G. S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(2), 222-233.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: a study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1355-1367.

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. . (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition, 102*(1), 139-149.

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal working memory is involved in associative word learning unless visual codes are available. *Journal of Memory and Language, 48*(3), 527-541.

Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron, 44*(1), 109-120.

Elliot, E. M., Hughes, R. W., Briganti, A., Joseph, T. N., Marsh, J. E., & Macken, B. (2016). Distraction in verbal short-term memory: Insights from developmental differences. *Journal of Memory and Language, 88*, 39-50.

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.

Ellis, N. C. (1996) Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology Section A, 49*(1), 234-250.

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences, 15*(5), 202-209.

Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology, 34*(5-6), 301-307.

Farah, M. (1994). Neuropsychological inference with an interactive brain: A critique of the "locality" assumption. *Behavioral and Brain Sciences, 17*, 43-104.

Farrell, S., & Lelièvre, A. (2012). The dynamics of access to groups in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(6), 1659-1674.

Fendrich, D. W., Healy, A. F., & Bourne, L. E. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 137–151.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition, 77*(1), 1-79.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fournet, N., Juphard, A., Monnier, C., & Roulin, J. L. (2003). Phonological similarity in free and serial recall: The effect of increasing retention intervals. *International Journal of Psychology, 38*(6), 384-389.

Fowler, C. A. (1987). Consonant-vowel cohesiveness in speech production as revealed by initial and final consonant exchanges. *Speech Communication, 6*(3), 231-244.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24*(2), 200-209.

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 469.

Fu, G., Wan, N. J., Baker, J. M., Montgomery, J. W., Evans, J. L., & Gillam, R. B. (2016). A Proof of Concept study of function-based statistical analysis of fNIRS Data: Syntax comprehension in children with specific language impairment compared to typically-developing controls. *Frontiers in Behavioral Neuroscience, 10*, 108-124.

Fuster, J. M. (1990). Inferotemporal units in selective visual attention and short-term memory. *Journal of Neurophysiology, 64*(3), 681-697.

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

Gathercole, S. E., & Adams, A. (1993). Phonological working memory in very young children. *Developmental Psychology, 29*, 770-778.

Gathercole, S. E., & Adams, A. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language, 33*, 672-688.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*(1), 65-77.

Glass, A. L., Krejci, J., & Goldman, J. (1989). The necessary and sufficient conditions for motor learning, recognition, and recall. *Journal of Memory and Language, 28*(2), 189-199.

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill Goldsmith, H. (2008). Infant and toddler oral-and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry, 49*(1), 43-50.

Geschwind, N. (1970). The organization of language and the brain. *Science, 170*(3961), 940-944.

Gibbs, R. W. (2005). Embodiment in metaphorical imagination. In D. Pecher, & Zwaan R. A. (Ed.), *Grounding Cognition: The role of perception and action in memory, language, and thinking* (pp. 65-92). UK: Cambridge University Press.

Gibbs, R. W., & Van Orden, G. C. (2010). Adaptive cognition without massive modularity. *Language and Cognition, 2*(2), 149-176.

Grantcharov, T. P., Bardram, L., Funch-Jensen, P., & Rosenberg, J. (2003). Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *The American Journal of Surgery, 185(*2), 146-149.

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.

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.

Guerrette, M. C., Saint-Aubin, J., Richard, M., & Guérard, K. (2018). Overt language production plays a key role in the Hebb repetition effect. *Memory & Cognition, 46*(8), 1389-1397.

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

Gupta, P., Lipinski, J., Abbs, B., & Lin, P. H. (2005). Serial position effects in nonword repetition. *Journal of Memory and Language*, *53*(1), 141-162.

Halgren, E. (1991). Firing of human hippocampal units in relation to voluntary movements. *Hippocampus, 1*(2), 153-161.

Hanakawa, T., Dimyan, M. A., Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: A time-course study with functional MRI. *Cerebral Cortex, 18*(12), 2775-2788.

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*, 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.

Harnishfeger, K. K., & Pope, R. S. (1996). Intending to forget: The development of cognitive inhibition in directed forgetting. *Journal of Experimental Child Psychology, 62*(2), 292-315.

Hatakenaka, M., Miyai, I., Mihara, M., Sakoda, S., & Kubota, K. (2007). Frontal regions involved in learning of motor skill—a functional NIRS study. *Neuroimage, 34*(1), 109-116.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207.

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

Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience & Biobehavioral Reviews, 68*, 602-610.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., ... & Konishi, J. (1996). Cortical speech processing mechanisms while vocalizing visually presented languages. *Neuroreport, 8*(1), 363-367.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 116-139.

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.

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.

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 in short-term memory: Testing a perceptual organization-based account. *Memory & Cognition, 39*(8), 1435–1447.

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.

Hurley, S. L. (1998). *Conciousness in action*. Cambridge, MA: Harvard University Press.

Hurley, S. (2001). Perception and action: Alternative views. *Synthese, 129*(1), 3-40.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446.

Jones, D., Alford, D., Bridges, A., Tremblay, S., & Macken, B. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 464.

Jones, D. M., Beaman, C. P., & Macken, W. J. (1996). The object-oriented episodic record model. *Models of short-term memory*, 209-238.

Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory & Cognition*, *46*(2), 216-229.

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. 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. *The Quarterly Journal of Experimental Psychology*, *60*(4), 505-511.

Kalm, K., & Norris, D. (2016). Recall is not necessary for verbal sequence learning. *Memory & Cognition*, *44*, 104-113.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302-4311.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705-710.

Kleinschmidt, A., Obrig, H., Requardt, M., Merboldt, K. D., Dirnagl, U., Villringer, A., & Frahm, J. (1996). Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy. *Journal Of Cerebral Blood Flow And Metabolism, 16*, 817–826.

Koziol, L. F., Budding, D., Andreasen, N., D’Arrigo, S., Bulgheroni, S., Imamizu, H., ... & Pezzulo, G. (2014). Consensus paper: the cerebellum's role in movement and cognition. *The Cerebellum, 13*(1), 151-177.

Koziol, L. F., Budding, D. E., & Chidekel, D. (2011). Sensory integration, sensory processing, and sensory modulation disorders: Putative functional neuroanatomic underpinnings. *The Cerebellum, 10*(4), 770-792.

Krings, T., Töpper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., & Thron, A. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neuroscience letters, 278(3),* 189-193.

Lackner, J. R., & Goldstein, L. M. (1974). Primary auditory stream segregation of repeated consonant—vowel sequences. *The Journal of the Acoustical Society of America*, *56*(5), 1651-1652.

Lacourse, M. G., Orr, E. L. R., Cramer, S. C., & Cohen, M. J. (2005). Brain activation during execution and motor imagery of novel and skilled sequential hand movements. *NeuroImage, 27*(3), 505-519.

Lafond, D., Tremblay, S., & Parmentier, F. (2010). The ubiquitous nature of the Hebb repetition effect: Error learning mistaken for the absence of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 515.

Larsen, J., & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology*, *56A*, 1249-1268.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley.

Lewandowsky, S., Geiger, S. M., & Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language*, *59*(2), 200-222.

Lu, C. M., Zhang, Y. J., Biswal, B. B., Zang, Y. F., Peng, D. L., & Zhu, C. Z. (2010). Use of fNIRS to assess resting state functional connectivity. *Journal of Neuroscience Methods*, *186*(2), 242-249.

MacCandliss, B. D., Cohen, L., Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293-299.

Machery, E. (2008). The folk concept of intentional action: Philosophical and experimental issues. *Mind & Language, 23*(2), 165-189.

MacDonald, M. C. (2016). Speak, act, remember: The language-production basis of serial order and maintenance in verbal memory. *Current Directions in Psychological Science, 25*(1), 47-53.

Macken, B., Taylor, J. C., & Jones, D. M. (2014). Language and short-term memory: The role of perceptual-motor affordance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1257-1270.

Macken, B., Taylor, J., & Jones, D. (2015). Limitless capacity: a dynamic object-oriented approach to short-term memory. *Frontiers in Psychology, 6*(293), 1-15.

Maidment, D. W., & Macken, W. J. (2012). The ineluctable modality of the audible: Perceptual determinant of auditory verbal short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 989-997.

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 93*(2), 95-119.

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. *Cognition, 107*(2), 395-419.

Maybery, M. T., Parmentier, F. B., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory and Language, 47*(3), 360-385.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352–370.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological Bulletin, 138*(2), 322-352.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Memory and Language*, *2*(1), 1-36.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3.

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience, 4*, 1-11.

Morey, 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.

Moriai-Izawa, A., Dan, H., Dan, I., Sano, T., Oguro, K., Yokota, H., Tsuzuki, D., & Watanabe, E. (2012). Multichannel fNIRS assessment of overt and covert confrontation naming. *Brain and Language, 121*(3), 185-193.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Memory and Language, 16*(5), 519-533.

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *The Quarterly Journal of Experimental Psychology, 61*(4), 505-514.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*(5), 482-488.

Murray, A., & Jones, D. M. (2002). Articulatory complexity at item boundaries in serial recall: The case of Welsh and English digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 594-598.

Nairne, J. S., & Kelley, M. R. (1999). Reversing the phonological similarity effect. *Memory and Cognition, 27*, 45–53.

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50*(2), 113-133.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition, 1*, 1-55.

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*, 12–28.

Norman, D. A. (Ed.) (1970). Models of human memory. New York: Academic Press.

Norris, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin, 143*(9), 992-1009.

Norris, D., Baddeley, A., & Page, M. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(5), 1093-1105.

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.

Nudo, R. J., Plautz, E. J., & Frost, S. B. (2001). Role of adaptive plasticity in recovery of function after damage to motor cortex. *Muscle & Nerve, 24*, 1000-1019.

Näätänen, R., Lehtokoski, A., Lennest, M., Luuki, A., Alliki, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432-434.

Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and comples memory span. *Memory & Cognition, 43*(6), 852-865.

Oberauer, K., Lewandowsky, S., Avh, E., Brown, G. D., Conway, A., Covan, N., ... & Ma, W. J. (2018). Benchmarks for models of short term and working memory. *Psychological Bulletin*.

Oberauer, K., & Meyer, N. (2009). The contributions of encoding, retention, and recall to the Hebb effect. *Memory, 17*(7), 774-781.

O'Reagan, J. K., Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939-973.

Oxenham, A. J. (2018). How we hear: The perception and neural coding of sound. *Annual Review of Psychology, 69*, 27-50.

Page, M., 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.

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, 364*, 3737-3753.

Page, M. P. A., & Norris, D. (2009b). Is there a common mechanism underlying word-form learning and the Hebb repetition effect? Experimental data and a modelling framework. In A. Thorn, & Page, M. (Ed.), *Interactions between short-term and long-term memory in the verbal domain*. Hove, East Sussex, UK: Psychology Press.

Page, M. P., 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.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*(3), 331-347.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology Section A, 44*(1), 47-67.

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

Pinker, S. (1994). *The language instinct: How the mind creates language*. UK: Penguin.

Pinker, S. (2005). So how does the mind work?. *Mind and Language, 20*(1), 1-24.

Piske, T., Flege, J. E., MacKay, I. R., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica, 59*(1), 49-71.

Perez, T. M., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 111*(4), 708-723.

Prinz, J. J. (2005). Passionate thoughts: The emotional embodiment of moral concepts. In D. Pecher, & Zwaan R. A. (Ed.), *Grounding Cognition: The role perception and action in memory, language, and thinking* (pp. 93-114). UK: Cambridge University Press.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139(*1), 23-38.

Postman, L. (1961). The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior* (pp. 152–179). New York: McGraw-Hill.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80*(1), 1-24.

Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R. S., Gati, J. S., ... & Kim, S. G. (2000). Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *Journal of Cognitive Neuroscience, 12*(2), 310-320.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M-P., Dold, G., & Hallett, M. (1996). Activation of primary visual cortex by Braille reading in blind subjects. *Nature, 380*, 526-528.

Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences, 8*(12), 547-553.

Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research, 152*(2), 229-242.

Salame, 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.

Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. *Neuropsychological impairments of short-term memory*, 11-53.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*(2), 261-273.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition, 42*, 213–259.

Sibi, S., Baiters, S., Mok, B., Steiner, M., & Ju, W. (2017). Assessing driver cortical activity under varying levels of automation with functional near infrared spectroscopy. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 1509-1516). IEEE

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.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343-348.

Spurgeon, J., Ward, G., & Matthews, W. J. (2014). Why do participants initiate free recall of short lists of words with the first list item? Toward a general episodic memory explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 469–488.

Stadler, M. A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory & Cognition*, *21*(6), 819-827.

Sternberg, S. (2011). Modular processes in mind and brain. *Cognitive Neuropsychology, 28*(3-4), 156-208.

St-Louis, M. È., Hughes, R. W., Saint-Aubin, J., & Tremblay, S. (2018). The resilience of verbal sequence learning: Evidence from the Hebb repetition effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *45*(1), 1-17.

Sugiura, L., Ojima, S., Matsuba-Kurita, H., Dan, I., Tsuzuki, D., Katura, T., & Hagiwara, H. (2011). Sound to language: different cortical processing for first and second languages in elementary school children as revealed by a large-scale study using fNIRS. *Cerebral Cortex, 21*(10), 2374-2393.

Schwartz, J. L., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics, 25*(5), 336-354.

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *The 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. A., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language, 67*, 342-354.

Thelen, E. & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. The MIT Press, London, England.

Thomas, N. J. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science, 23*(2), 207-245.

Thulborn, K. R., Carpenter, P. A., & Just, M. A. (1999). Plasticity of language-related brain function during recovery from stroke. *Stroke, 30*(4), 749-754.

Tresch, M. C., Sinnamon, H. M., & Seamon, J. G. (1993). Double dissociation of spatial and object visual memory: Evidence from selective interference in intact human subjects. *Neuropsychologia, 31*(3), 211-219.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A centure of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138,* 1172-1217.

Walsh, B., Tian, F., Tourville, J. A., Yücel, M. A., Kuczek, T., & Bostian, A. J. (2017). Hemodynamics of speech production: an fNIRS investigation of children who stutter. *Scientific Reports, 7*(1), 4034.

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*, 430-447.

Warren, R. M., Obusek, C. J., Farmer, R. M., Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science, 164*, 586-587.

Weisberg, J., van Turennout, M., & Martin, A. (2007). A neural system for learning about object function. *Cerebral Cortex, 17*(3), 513-521.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558-584.

Wilson, F. A. W., O'Scalaidhe, S. P., Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260,* 1955-1958.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636.

Wilson, M., & Fox, G. (2007). Working memory for language is not special: Evidence for an articulatory loop for novel stimuli. *Psychonomic Bulletin & Review, 14,* 470–473.

Wilson, R. A., & Clark, A. (2009). How to situate cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.

Winstein, C. J., Grafton, S. T., & Pohl, P. S. (1997). Motor task difficulty and brain activity: investigation of goal-directed reciprocal aiming using positron emission tomography. *Journal of Neurophysiology, 77*(3), 1581-1594.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences, 5*(11), 487-494.

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.

Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron, 74*(5), 936-946.

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), 573-590.

Zhang, H., Zhang, Y. J., Lu, C. M., Ma, S. Y., Zang, Y. F., & Zhu, C. Z. (2010). Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements. *Neuroimage, 51*(3), 1150-1161.

APPENDICES

Appendix 1. Supplementary material for Experiment 1

Supplementary Table 1.

Full set of results from the Modality × Phonological similarity × Articulatory suppression × Serial Position ANOVA on serial recall accuracy in Experiment 1

	F	*MSE*	*p*	η_p^2
Modality	1.2	.07	.278	.02
Similarity	42.9	.09	< .001	.47
Suppression	176.12	.12	< .001	.78
Position	231.18	.08	< .001	.83
Modality × Similarity	5.58	.04	.022	.10
Modality × Suppression	9.07	.05	.004	.16
Similarity × Suppression	24.69	.06	< .001	.34
Modality × Position	33.54	.012	< .001	.41
Similarity × Position	14.97	.01	< .001	.23
Suppression × Position	1.88	.02	.084	.04
Modality × Similarity × Suppression	2.83	.04	.099	.06
Modality × Similarity × Position	16.08	.01	< .001	.25
Modality × Suppression × Position	1.31	.01	.253	.03
Similarity × Suppression × Position	4.39	.01	< .001	.08
Modality × Similarity × Suppression × Position	3.13	.01	.005	.06

Supplementary Table 2.

Full set of results from the List-type × Modality × Phonological similarity × Articulatory suppression × Cycle ANOVA conducted to assess Hebb sequence learning in Experiment 1

	F	*MSE*	*p* value	$\eta_p{}^2$
List-type	59.25	.19	< .001	.55
List-type × Modality	.28	.10	.602	.01
List-type × Similarity	< .001	.09	.993	.00
List-type × Suppression	7.8	.15	.007	.14
List-type × Cycle	7.64	.04	< .001	.14
List-type × Similarity × Cycle	1.73	.04	.064	.03
List-type × Suppression × Cycle	3.16	.04	< .001	.06
List-type × Modality × Similarity	5.06	.15	.029	.09
List-type × Modality × Suppression	.40	.15	.528	.01
List-type × Similarity × Suppression	.07	.12	.795	.00
List-type × Modality × Cycle	.29	.04	.987	.01
List-type × Modality × Similarity × Cycle	1.39	.04	.174	.03
List-type × Modality × Suppression × Cycle	1.37	.04	.184	.03
List-type × Suppression × Similarity × Cycle	.87	.04	.570	.02
List-type × Modality × Similarity × Suppression	4.6	.06	.037	.09
List-type × Modality × Similarity x Suppression × Cycle	.62	.04	.817	.01
Modality	.35	.29	.558	.01
Similarity	44.32	.27	< .001	.48
Suppression	220.24	.41	< .001	.82
Cycle	9.48	.04	< .001	.16
Modality × Similarity	12.47	.14	.001	.20
Modality × Suppression	8.5	.2	.005	.15
Similarity × Suppression	17.08	.24	< .001	.26
Modality × Cycle	2.77	.04	.002	.05
Similarity × Cycle	1.78	.04	.055	.04
Suppression × Cycle	.81	.04	.631	.02
Modality × Similarity × Suppression	.58	.12	.449	.01
Modality × Similarity × Cycle	1.36	.04	.188	.03
Modality × Suppression × Cycle	.87	.04	.569	.02
Similarity × Suppression × Cycle	2.62	.04	.003	.05
Modality × Similarity × Suppression × Cycle	1.08	.04	.372	.02

Appendix 2. Supplementary material for Experiment 2

Supplementary Table 3.

Full set of results from the Modality × Articulatory suppression × Grouping × Serial position ANOVA conducted to assess the output RTs in Experiment 2

	F	*MSE*	*p*	η_p^2
Modality	27.54	2189694.6	< .001	.21
Suppression	7.97	1097956.9	.006	.07
Grouping	1.02	248171.6	.407	.01
Position	116.36	387491.4	< .001	.53
Suppression × Modality	5.97	1097956.9	.016	.01
Grouping × Modality	.416	248171.6	.838	.01
Position × Modality	11.74	387491.4	< .001	.10
Suppression × Grouping	1.36	233087.3	.237	.01
Suppression × Position	8.68	379563.2	< .001	.08
Grouping × Position	11.81	245722.7	< .001	.10
Suppression × Grouping × Modality	.345	233087.3	.885	.00
Suppression × Position × Modality	1.88	379563.2	.096	.02
Grouping × Position × Modality	2.1	245722.7	.001	.02
Suppression × Grouping × Position	2.65	255978.9	< .001	.03

Supplementary Table 4.

Full set of results from the List-type × Modality × Articulatory suppression × Grouping-consistency × Cycle ANOVA conducted to assess Hebb sequence learning in Experiment 2

	F	MSE	*p*	$\eta_p{}^2$
List-type	20.93	.03	< .001	.17
Grouping-consistency	1.97	.08	.164	.02
Suppression	497.71	.09	< .001	.83
Cycle	.921	.04	.467	.01
List-type × Modality	2.32	.07	.131	.02
List-type × Cycle	.83	.04	.529	.01
List-type × Suppression	8.71	.04	.004	.08
Suppression × Modality	3.11	.18	.081	.03
Grouping-consistency × Modality	.03	.08	.868	.00
Suppression × Cycle	2.31	.04	.043	.02
Cycle × Modality	.69	.04	.632	.01
Grouping-consistency × Cycle	1.51	.04	.185	.02
List-type × Grouping-consistency	.23	.06	.635	.01
Grouping-consistency × Suppression	.57	.1	.452	.01
Grouping-consistency × Cycle × Modality	.67	.04	.65	.01
Grouping-consistency × Suppression × Cycle	.39	.04	.854	.01
List-type × Suppression × Cycle	3.57	.04	.003	.03
List-type × Grouping-consistency × Cycle	.86	.04	.509	.01
Suppression × Cycle × Modality	.75	.04	.588	.01
List-type × Cycle × Modality	.25	.04	.938	.00
List-type × Suppression × Modality	2.53	.07	.115	.02
List-type × Grouping-consistency × Modality	.01	.06	.935	.00
List-type × Grouping-consistency × Cycle × Modality	.22	.04	.955	.00
Grouping-consistency × Suppression × Modality	1.08	.10	.301	.01
List-type × Grouping-consistency × Suppression	.07	.08	.789	.00
List-type × Grouping-consistency × Suppression × Modality	.02	.08	.871	.00
List-type × Suppression × Cycle × Modality	1.11	.04	.354	.01
Grouping-consistency × Suppression × Modality × Cycle	1.15	.04	.332	.01
List-type × Grouping-consistency × Suppression × Cycle	.76	.04	.576	.01
List-type × Grouping-consistency × Suppression × Modality × Cycle	4.38	.02	.014	.04

Supplementary Table 5.

Simple-effects analysis of the interaction of List-type, Modality, Articulatory suppression, Grouping-consistency and Cycle in Experiment 2

Modality	Grouping	Articulatory suppression	Cycle	Hebb – filler (SE)	*p*
Visual	Consistent	No suppression	1	-7.7 (4.2)	.071
			2	3.7 (4.6)	.421
			3	0.1 (4.0)	.973
			4	5.5 (3.8)	.151
			5	8.8 (3.8)	.023
			6	10.2 (4.2)	.017
		With suppression	1	4.3 (4.0)	.289
			2	1.9 (3.4)	.569
			3	2.3 (3.9)	.553
			4	-1.1 (4.0)	.785
			5	0.8 (4.3)	.847
			6	0.8 (4.0)	.836
	Inconsistent	No suppression	1	-0.1 (4.4)	.999
			2	1.2 (4.3)	.775
			3	6.6 (4.4)	.137
			4	4.8 (4.2)	.252
			5	1.8 (4.6)	.701
			6	3.8 (3.6)	.288
		With suppression	1	6.0 (3.8)	.117
			2	0.7 (3.3)	.837
			3	-1.0 (4.2)	.819
			4	-4.8 (4.2)	.259
			5	2.5 (4.0)	.536
			6	1.2 (3.6)	.735
Auditory	Consistent	No suppression	1	1.4 (4.2)	.746
			2	10.3 (4.6)	.027
			3	14.3 (4.0)	.001
			4	9.2 (3.8)	.017
			5	4.7 (3.8)	.223
			6	7.8 (4.2)	.066
		With suppression	1	3.0 (4.0)	.451
			2	-1.1 (3.4)	.744
			3	-8.0 (3.9)	.045
			4	4.3 (4.0)	.292
			5	10.4 (4.3)	.016
			6	1.1 (4.0)	.782
	Inconsistent	No suppression	1	5.9 (4.5)	.187

		2	1.8 (4.3)	.680
		3	7.6 (4.4)	.089
		4	8.2 (4.2)	.051
		5	12.8 (4.6)	.007
		6	10.7 (3.6)	.004
	With suppression	1	4.0 (3.8)	.300
		2	2.9 (3.3)	.388
		3	-3.4 (4.2)	.414
		4	-1.4 (4.3)	.746
		5	-1.8 (4.0)	.655
		6	0.6 (3.6)	.875

Appendix 3. Supplementary material for Experiment 3

Supplementary Table 6.

Simple-effects analysis of the interaction of Voice, Similarity and Position in serial recall in Experiment 3

Similarity	Position	Single – Alternating voice (SE)	*p*
Dissimilar	1	4.11 (1.96)	.041
	2	-3.3 (2.59)	.209
	3	-6.9 (6.78)	.313
	4	6.83 (2.58)	.011
	5	10.33 (3.69)	.007
	6	5.11 (2.8)	.074
	7	5.27 (2.42)	.034
Similar	1	7.26 (2.19)	.002
	2	.93 (2.64)	.726
	3	6.98 (3.27)	.038
	4	7.62 (2.81)	.009
	5	4.11 (2.66)	.129
	6	3.35 (2.76)	.23
	7	.74 (2.48)	.769

Supplementary Table 7.

Results from the List-type × Voice × Phonological similarity × Cycle ANOVA conducted to assess accuracy in Phase 1 of Experiment 3

	F	*MSE*	*p*	η_p^2
List-type	34.35	.083	<.001	.417
Voice	15.17	.12	<.001	.24
Similarity	39.27	.27	<.001	.45
Cycle	2.6	.05	.026	.051
List-type × Cycle	4.34	.05	.001	.083
List-type × Voice	1.13	.1	.294	.023
List-type × Similarity	.01	.1	.921	<.001
Voice × Cycle	.496	.05	.779	.01
Voice × Similarity	2.09	.12	.155	.042
Similarity × Cycle	.17	.04	.974	.004
List-type × Voice × Similarity	1.31	.11	.259	.026
List-type × Voice × Cycle	.96	.05	.442	.02
List-type × Similarity × Cycle	.08	.04	.996	.002
Voice × Similarity × Cycle	2.15	.04	.06	.043
List-type × Voice × Similarity × Cycle	1.65	.05	.147	.033

Supplementary Table 8.

Results from the List-type × Phonological similarity × Response-requirement (shortened to 'Response' in the table) × Cycle ANOVA conducted to assess sequence learning in the single voice conditions of Phase 2 of Experiment 3.

	F	*MSE*	*p*	$\eta_p{}^2$
List-type	8.19	.04	.006	.146
Response	13.19	.03	.001	.216
Cycle	8.16	.02	.006	.0145
List-type × Response	2.73	.04	.105	.054
List-type × Cycle	0.16	.04	.687	.003
Response × Cycle	7.13	.03	.01	.129
List-type × Response × Cycle	4.5	.03	.039	.086

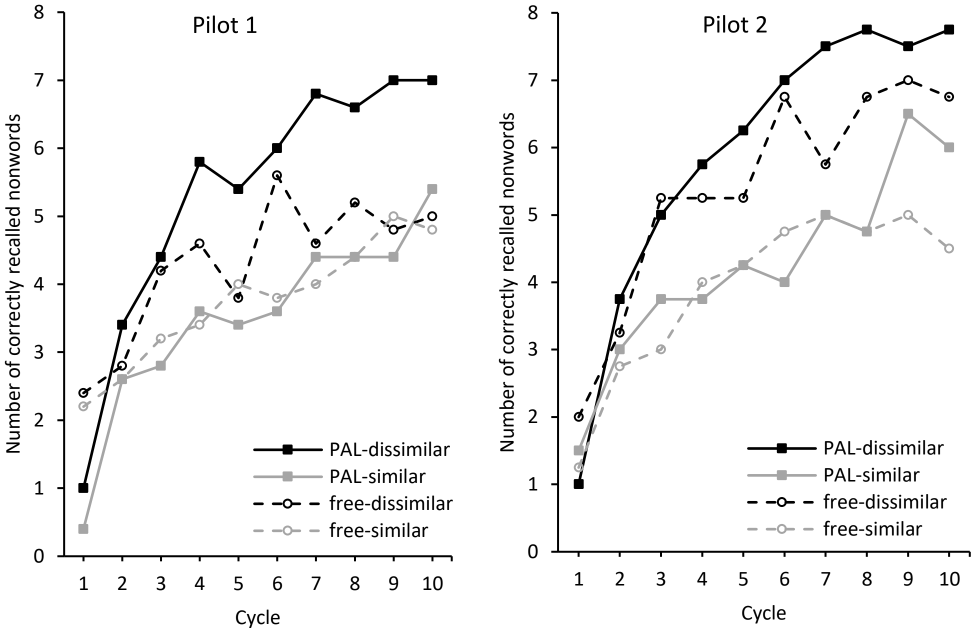
Supplementary Table 9.

Results from the List-type × Phonological similarity × Response-requirement × Cycle ANOVA conducted to assess accuracy across the four main list-types in the alternating voice conditions of Phase 2 of Experiment 3.

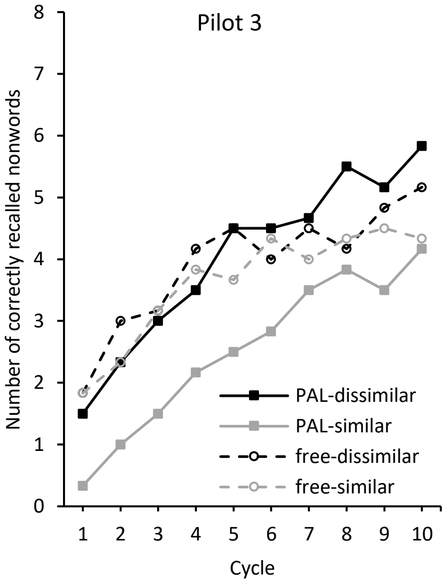
	F	*MSE*	*p*	$\eta_p{}^2$
List-type	4.34	.02	.043	.083
Response	2.96	.03	.092	.058
Cycle	10.91	.05	.002	.185
List-type × Response	0.25	.03	.618	.005
List-type × Cycle	4.12	.03	.048	.079
Response × Cycle	0.05	.03	.828	.001
List-type × Response × Cycle	4.85	.03	.033	.092

Appendix 4. Supplementary material for pilot experiments in Chapter 3

Supplementary Figure 1.
Recall accuracy data in Pilots 1 and 2 according to Recall-type (free, paired-associate [PAL]), Across-similarity and Cycle, using within-dissimilar nonwords.

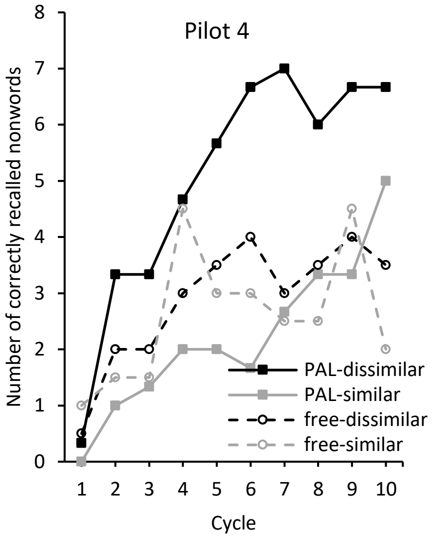


Supplementary Figure 2.
Recall accuracy data in Pilot 3 according to Recall-type, Across-similarity and Cycle, using within-similar nonwords.



Supplementary Figure 3.

Recall accuracy data in Pilot 4 according to Recall-type, Across-similarity and Cycle, using within-dissimilar nonwords and auditory presentation.



Appendix 5. Supplementary material for Experiment 4

Supplementary Table 10.
The lists of words and nonwords used in Experiments 4 and 6.

English 1	English 2	Across-dissimilar, within-dissimilar 1	Across-dissimilar, within-dissimilar 2	Across-similar, within-dissimilar 1
telephone	letter	hawerty	lamser	saries
road	staircase	samtis	panooper	solories
trousers	giraffe	prection	corraty	siras
water	engine	pomlonal	berous	setories
window	lake	beimaty	selgra	sefas
clock	curtain	remipper	pemmible	sasomies
strawberry	tunnel	dimolt	sibate	silaries
poster	flower	putine	voinant	somas

Across-similar, within-dissimilar 2	Across-dissimilar, within-similar 1	Across-dissimilar, within-similar 2	Across-similar, within-similar 1	Across-similar, within-similar 2
trelate	humdum	pirlinvin	tinfinpil	tenfen
trupogate	dantanlam	hunnunbul	rindingil	ferhenmen
trixopate	ponlon	nesmes	pinlin	germenpen
trudote	raitailain	rastasdan	finsin	lenven
trecopate	risdisdim	rilsil	ginminfil	serdenpen
trigote	tynwyn	derrer	winmin	nenhen
trabate	folnongon	pannan	minpin	wersennen
tramonate	temrem	tusruslun	sinhinbil	bennen

Supplementary Table 11.

Full set of results from the Recall-type × Across-similarity × Within-similarity × Cycle ANOVA on recall accuracy from Experiment 4

	F	*MSE*	*p*	$\eta_p{}^2$
Recall-type	4.19	6.63	.048	.10
Across-similarity	79.56	7.27	<.001	.67
Within-similarity	8.89	42.23	.005	.19
Cycle	256.37	2.04	<.001	.868
Language-group	4.18	42.23	.048	.10
Language-group × Cycle	.71	2.04	.628	.02
Language-group × Within-similarity	.12	42.23	.727	<.01
Recall-type × Within-similarity	1.36	6.63	.250	.03
Recall-type × Language-group	<.01	6.63	.999	<.01
Within-similarity × Across-similarity	7.35	7.27	.01	.159
Across-similarity × Language-group	1.66	7.27	.206	.04
Recall-type × Across-similarity	3.01	14.35	.091	.07
Recall-type × Cycle	10.93	1.06	<.001	.22
Across-similarity × Cycle	7.68	.99	<.001	.17
Recall-type × Across-similarity × Within-similarity	.04	14.35	.837	<.01
Recall-type × Language-group × Across-similarity	.01	14.35	.94	<.01
Language-group × Across-similarity × Within-similarity	.02	7.27	.9	<.01
Language-group × Within-similarity × Cycle	.22	2.04	.96	.01
Recall-type × Within-similarity × Cycle	.75	1.06	.661	.02
Recall-type × Language-group × Cycle	.34	1.06	.961	<.01
Recall-type × Language-group × Within-similarity	.37	6.63	.546	.01
Across-similarity × Within-similarity × Cycle	1.89	.99	.053	.05
Language-group × Across-similarity × Cycle	1.48	.99	.154	.04
Recall-type × Across-similarity × Cycle	2.19	1.28	.043	.05
Recall-type × Language-group × Across-similarity × Within-similarity	.13	14.35	.718	<.01
Language-group × Across-similarity × Within-similarity × Cycle	.84	.99	.581	.02
Recall-type × Language-group × Within-similarity × Cycle	.61	1.06	.787	.02

Recall-type × Across-similarity × Within-similarity × Cycle	2.34	.88	.014	.06
Language-group × Recall-type × Across-similarity × Cycle	.47	.88	.895	.01
Language-group × Recall-type × Across-similarity × Within-similarity × Cycle	.83	.88	.59	.02

Appendix 6. Supplementary material for Experiment 5

Supplementary Table 12.
The lists of words and nonwords used in Experiment 5.

Practice, English	Practice, free recall	Practice, cued nonwords	English 1	English 2
car	poltion	fumal	mirror	bracelet
popsicle	ferty	lorres	violin	pencil
wolf	nical	posgar	wing	nest
triangle	sural	golate	giraffe	bottle
market	sador	tarint	cloud	compass
fork	favrint	rista	mountain	paper
milk	lobate	vorite	hammer	chair
anchor	pascul	sasom	sand	book

Nonwords 1	Nonwords 2	Nonwords 3	Nonwords 4
vurpano	conlit	pundamy	plumity
suction	detlasien	redolic	noncal
catenfir	vavement	nenderfel	cormuner
nilonts	stecal	systacle	ramution
nosilla	vonfercul	fembal	wirmon
mitive	boalen	lalmian	sutny
frixodile	ricomper	foxtie	hortworn
pusela	ponfection	vomanty	vaxrie

Supplementary Table 13.

Full set of results from the Recall-type × Modality × Articulatory suppression × Cycle ANOVA on recall accuracy from Experiment 5

	F	*MSE*	*p*	$\eta_p{}^2$
Recall-type	17.77	10.41	<.001	.32
Modality	2.38			
Articulatory suppression	72.95	6.34	<.001	.66
Cycle	282.09	13.37	<.001	.881
Recall-type × Articulatory suppression	.13	8.36	.718	<.01
Articulatory suppression × Modality	1.93	6.34	.173	.05
Articulatory suppression × Cycle	2.05	.86	.033	.05
Recall-type × Modality	.12	10.41	.735	<.01
Recall-type × Cycle	18.21	.99	<.001	.88
Modality × Cycle	2.99	1.29	.002	.07
Recall-type × Modality × Articulatory suppression	.972	8.36	.718	<.01
Recall-type × Articulatory suppression × Cycle	4.99	.66	<.001	.12
Modality × Articulatory suppression × Cycle	1.73	1.24	.112	.04
Recall-type × Modality × Cycle	1.09	1.46	.367	.03
Recall-type × Modality × Articulatory suppression × Cycle	.65	.66	.752	.02

Appendix 7. Supplementary material for Experiment 6

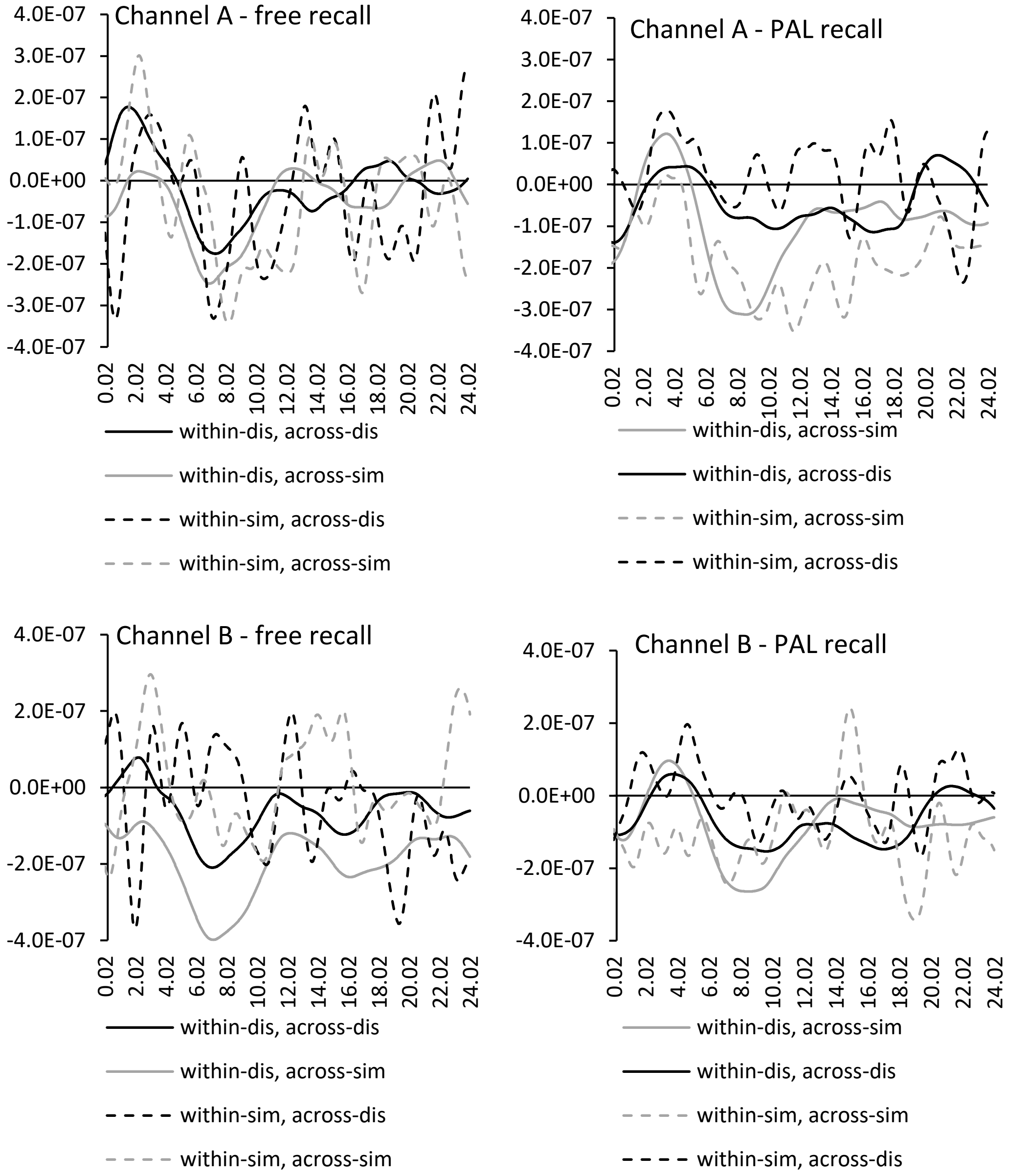
Supplementary Table 14.

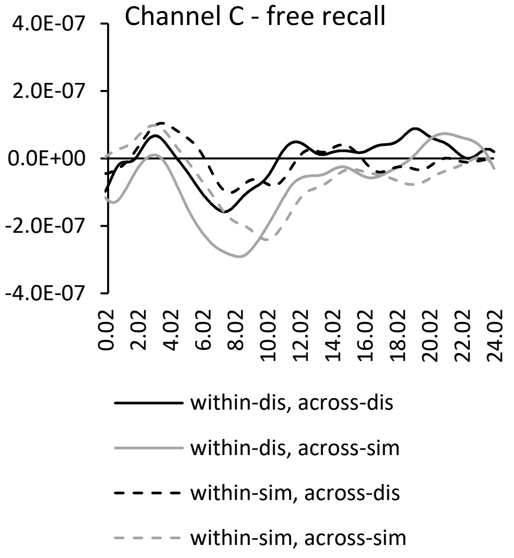
Full set of results from the Recall-type × Across-similarity × Within-similarity × Cycle ANOVA on recall accuracy from Experiment 6

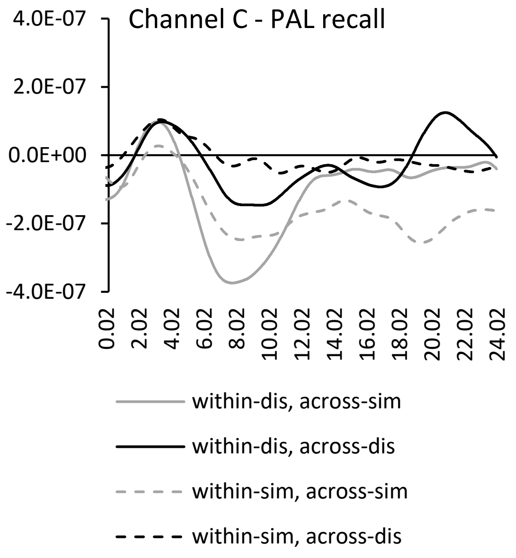
	F	*MSE*	*p*	$\eta_p{}^2$
Recall-type	2.83	17.6	.101	.07
Across-similarity	48.64	5.95	<.001	.56
Within-similarity	.001	46.76	.97	<.01
Cycle	233.97	2.38	<.001	.86
Recall-type × Within-similarity	.38	17.6	.384	.02
Within-similarity × Across-similarity	1.05	5.95	.313	.03
Within-similarity × Cycle	1.35	1.14	.208	.03
Recall-type × Across-similarity	.74	10.28	.394	.02
Recall-type × Cycle	13.28	1.63	<.001	.26
Across-similarity × Cycle	6.43	1.17	<.001	.15
Recall-type × Across-similarity × Within-similarity	.01	10.28	.929	<.01
Recall-type × Within-similarity × Cycle	1.98	1.63	.075	.05
Across-similarity × Within-similarity × Cycle	1.39	.84	.19	.04
Recall-type × Across-similarity × Cycle	.54	.8	.845	.01
Recall-type × Across-similarity × Within-similarity × Cycle	1.19	.8	.302	.03

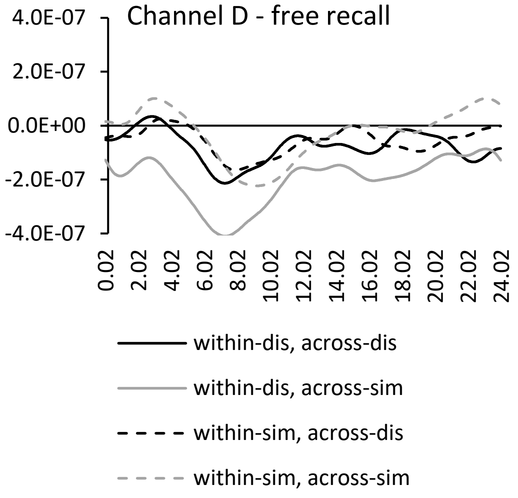
Supplementary Figure 4.

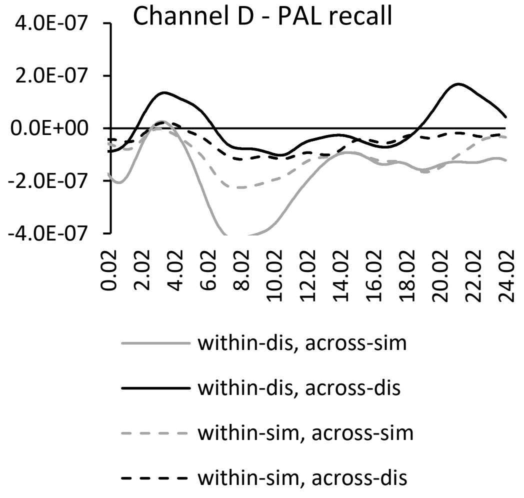
HbO for each of the eight channels across time, plotted for all of the eight conditions from Experiment 6.

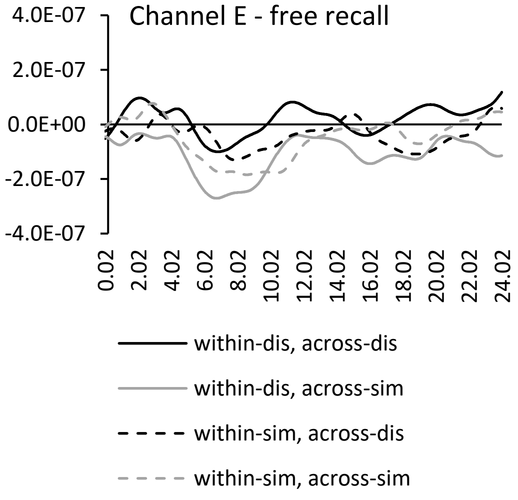


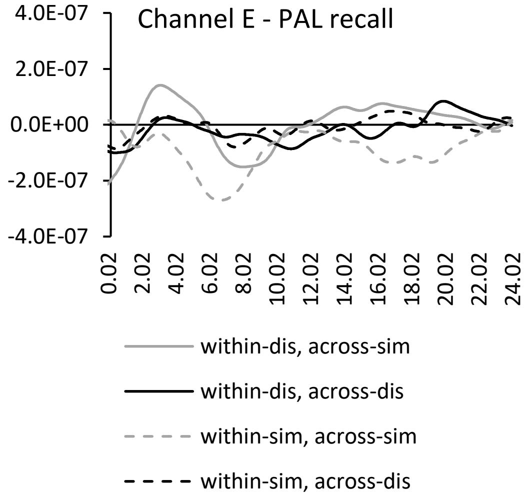

Channel C - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim

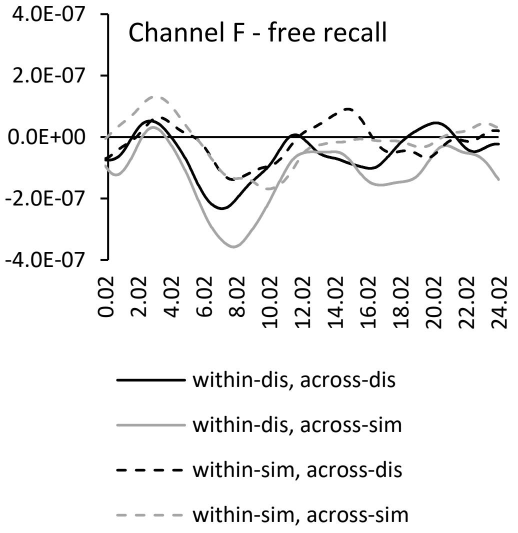

Channel C - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis

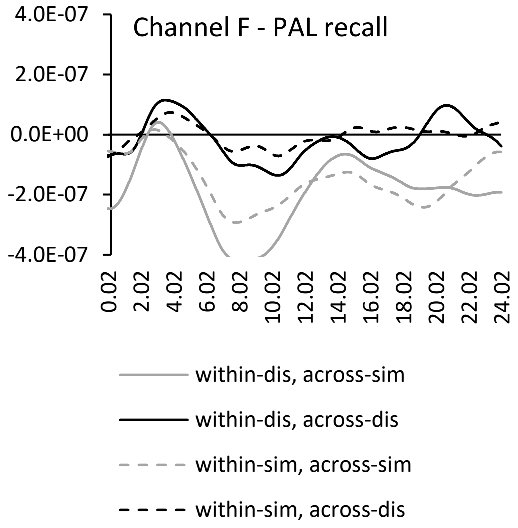

Channel D - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim

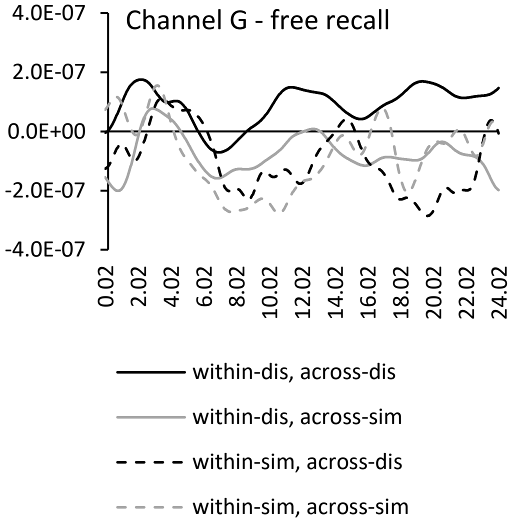

Channel D - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis

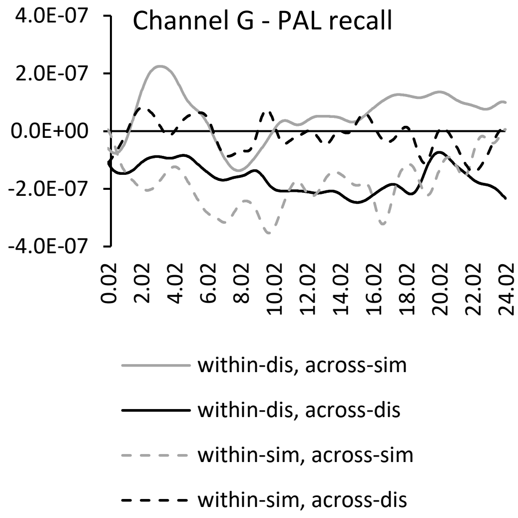

Channel E - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim

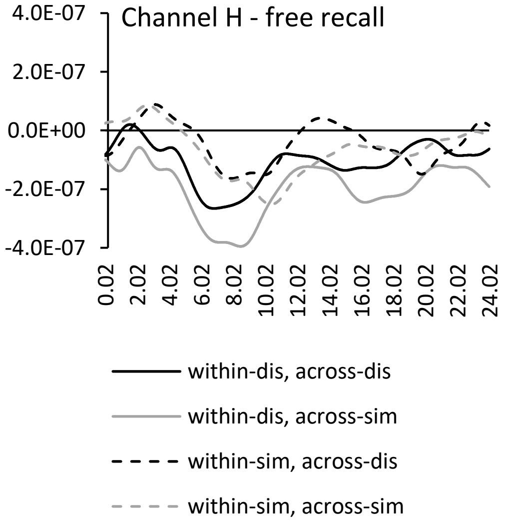

Channel E - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis

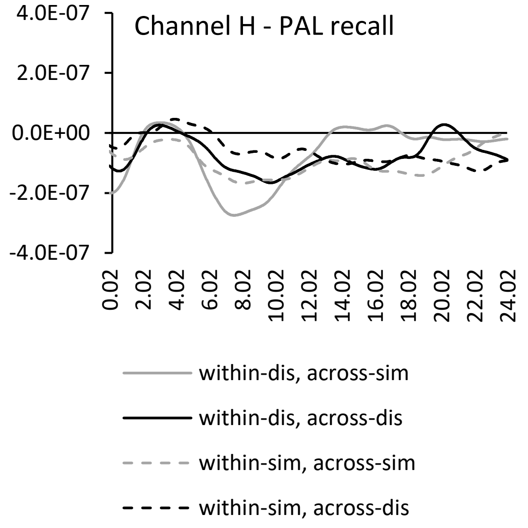

Channel F - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim


Channel F - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis


Channel G - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim


Channel G - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis


Channel H - free recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-dis
within-dis, across-sim
within-sim, across-dis
within-sim, across-sim


Channel H - PAL recall
4.0E-07
2.0E-07
0.0E+00
-2.0E-07
-4.0E-07
0.02
2.02
4.02
6.02
8.02
10.02
12.02
14.02
16.02
18.02
20.02
22.02
24.02
within-dis, across-sim
within-dis, across-dis
within-sim, across-sim
within-sim, across-dis