VERBAL SHORT-TERM MEMORY: COGNITIVE AND NEUROSCIENTIFIC TESTS OF A PERCEPTUAL-GESTURAL ACCOUNT

Michail Dmitrievic Kozlov

Thesis submitted to Cardiff University
for the degree of
Doctor of Philosophy

School of Psychology
Cardiff University
June 2012

DECLARATION AND STATEMENTS

DECLARATION
This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ………………………………………… (candidate)
Date…26/06/2012

STATEMENT 1
This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …………………………………………(candidate)
Date …26/06/2012

STATEMENT 2
This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references. The views expressed are my own.

Signed…………………………………………(candidate)
Date… 26/06/2012

STATEMENT 3
I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ………………………………………(candidate)
Date…26/06/2012

CONTENTS

PREFACE

The following work was undertaken while at the School of Psychology, Cardiff University, between the years 2008-2012 under the supervision of Dr. Robert Wyn Hughes and Prof. Dylan M. Jones.

The work was carried out under a studentship from the Biotechnology and Biological Sciences Research Council (BBSRC), and the School of Psychology, Cardiff University.

Research presented in chapter 2 has been published in:

Kozlov, M. D., Hughes, R.W., & Jones, D. M. (2011). Gummed Up Memory: Chewing Gum Impairs Short-Term Recall. *Quarterly Journal of Experimental Psychology, 65*, 501-513**.**

Research reported in chapter 3 was presented at the Wales Institute of Cognitive Neuroscience (WICN) Biennial Conference in January, 2011 and the BPS Cognitive Section Conference in September, 2011. Methodological aspects of the research were included in:

Maizey, L., Allen, C. P., Dervinis, M., Verbruggen, F., Varnava, A., Kozlov, M. D., Adams, R. C., Stokes, M., Klemen, J., Bungert, A., Hounsell, C. A., Chambers, C. D. (2012). *Comparative incidence rates of mild adverse effects to transcranial magnetic stimulation.* Manuscript in revision.

Research reported in chapter 4 was presented at the 5th International Conference on Memory (ICOM5) in August, 2011.

ACKNOWLEDGEMENTS

I am especially grateful to my supervisors Dr. Robert W. Hughes and Prof. Dylan M. Jones, for their constant encouragement and support.

I also would like to thank my friends and colleagues at Cardiff University for making my time at uni yet more rewarding.

Finally, I would like to thank my family, my parents, my grandparents and my brother, to whom I would like to dedicate this thesis, for believing in me even when my own faith would falter.

I would also like to acknowledge the Biotechnology and Biological Sciences Research Council and the School of Psychology, Cardiff University for their financial support during the time of my studies.

Моим бабушкам,

Папе, маме и брату.

С любовью и почтением.

SUMMARY

It has often been suggested that verbal short-term memory, the ability to maintain verbal information for a brief period of time, is based on the upload of to-be-remembered material into passive, dedicated, information stores. Alternatively, it has been claimed that all information is remembered but that access to it gets obstructed because of interference by subsequent similar material. The aim of the present thesis was to challenge both these approaches and to examine the viability of a different, perceptual-gestural, view of information buffering over the short term. This approach conceptualizes verbal short-term storage as an active process that emerges from, and is defined by, the recruitment of receptive and (speech) productive mechanisms. In Experiments 1-3, the significant impact of non-verbal concurrent motor tasks on verbal short-term memory suggests an active involvement of productive mechanisms. These experiments also cast doubt on the proposal that forgetting occurs because of interference by similar content. Experiment 4 expands upon this challenge of the interference-based view by showing that a temporary lesion of a brain area involved in speech planning (Broca’s area), induced with transcranial magnetic stimulation, affects verbal short-term memory performance in the absence of any additional potentially interfering verbal input. Further, challenging the store-based view, the virtual lesion of Broca’s area also attenuated the phonological similarity effect, a hallmark effect of the function of the hypothetical language-independent store. Finally, Experiments 5-9 sought to determine the origin of variations in recall performance as a function of sensory-modality of input. It is concluded that only the perceptual-gestural approach can offer an account of presentation type-based differences in verbal list recall that goes beyond a redescription of the observed effects. The thesis closes with an outline of a neurological model of active storage of verbal information over the short-term.

CHAPTER 1: HIC EST ANIMUS

"HC SVNT DRACONES!" This Latin phrase, translated as "Here be dragons" was often used on ancient maps in order to denote unexplored regions. Naturally, if a region is unknown it is impossible to conclude that dragons do not live there. Whilst in medieval times this logic might have been sound, nowadays hardly anyone seriously believes that there are dragons living in the few unmapped regions of the globe. This assumption could be based on the inductive conclusion that since most of the earth has now been cartographed and dragons were not found, dragons won't exist in the remaining regions. However, black swans were also unheard of before Australia was discovered. Perhaps then it is premature to dismiss the existence of dragons. Yet, the nonbelief in dragons is founded on more than just the past inability to find them. There are also no fossil records, and it is difficult to conceive of a gland that would enable a living creature to spit fire out of its mouth. Dragons do not exist.

Just like dragons were proposed by medieval cartographers as explanations for why travellers rarely returned from their perilous journeys to foreign lands, many modern theories of verbal short-term memory (STM), the ability to maintain verbal information in an active state over a brief period of time, often postulate bespoke mechanisms in order to explain STM phenomena. This is partly because current knowledge about the human brain in many ways resembles medieval knowledge about the globe. Many brain functions have already been mapped but there is still much debate about the function of certain regions and the pathways connecting these regions are also yet poorly understood. Alas, adequate research techniques to map brain regions like functional magnetic resonance imaging or transcranial magnetic stimulation are still comparatively new and underdeveloped. It is therefore still possible to propose credible theories of mental entities like verbal STM without giving much thought to whether it is plausible that those entities could be found in the brain. After all, perhaps the neural correlates of these entities have simply not been discovered yet. Surely, the label “Hic est animus”, “Here be memory” should be appropriate somewhere in the brain.

As with dragons, in order to show that verbal STM does not exist as a separate entity two paths can be pursued: The first approach concerns mapping the brain. If it is impossible to find the STM construct anywhere in the brain then it is narrow to assume that it does not exist. However, it could always be countered that the memory entity, like a black swan rather than like a dragon, is hiding in an undiscovered region. Until all regions and their connecting pathways are excluded as the seats of STM this objection is difficult to refute.

The other approach is to demonstrate that the functions of already-mapped brain regions can sufficiently or better account for processes that are cited as evidence for the existence of a bespoke STM system. By analogy, if it can be shown that other mundane dangers lie behind the horizon, it is not necessary to believe that foreign lands are dangerous because dragons live there. Thus, if existing perceptual and gestural processes that are clearly mapped onto brain regions can explain STM phenomena, then parsimony dictates that it is not necessary to invoke bespoke short-term stores to explain the data. Note that in this case parsimony refers to evolutionary and not necessarily explanatory parsimony. In order to achieve explanatory parsimony for STM phenomena it is easy enough to invent additional constructs such as an as-yet-unmapped memory store. This however is not much different from “parsimoniously” explaining that celestial bodies move because it is God’s will. In contrast to this, evolutionary parsimony refers to how evolutionary pressures necessitate that any organ or appendage that does not serve a vital or reproductive purpose is selected against. This is because an appendage or brain region with no purpose would still require nutrients, which would have to be obtained by the organism through additional and potentially dangerous foraging. Incidentally, this is also the reason why the popular belief that humans only use 5% of their mental capacity is nonsensical (Anderson, 1995). Clearly, therefore, a brain region whose functions can be performed by other, more functionally generalized, mechanisms should have a very low likelihood of ever evolving. Hence, if it can be shown that basic perceptual and motor planning mechanisms can explain STM phenomena, the assumption of additional memory storage entities becomes evolutionary unparsimonious.

The question of whether a dedicated verbal STM system that can be mapped onto the brain exists is not trivial. All human communication is based on the assembly of meaningful utterances form distinct and individually arbitrary gestures and symbols. Clearly, in order to write or speak or gesture a meaningful sentence in sign language, it is necessary that a long cascade of gestures is assembled into a coherent routine. Since the routine is issued over a temporally extended period it is necessary for some buffering to take place: Later words in a sentence have to be maintained in an active state while earlier words are produced. Thus, the process of STM is vital for any meaningful verbal communication. Moreover, any goal-directed behaviour that goes beyond stimulus-response associations requires the ability to meaningfully organize single actions into sequences (Lashley, 1951). Because the most sophisticated manifestation of such sequential behaviour in humans is speech, an understanding of the formation and execution of speech sequences should be informative of the formation and execution of meaningful action sequences in general.

An important distinction needs to be drawn here between the process of STM and the mechanism(s) of STM. If STM were accomplished by a dedicated system, then the process of STM and the mechanism of STM would be equivalent. Damage of this mechanism—as a result of a stroke for example—would necessarily impair the verbal STM process. However, if the short-term buffering of verbal information is an emergent property of general receptive (e.g., perceptual) and productive (e.g., vocal-articulatory) mechanisms, then it would only make sense to speak of the STM process but not of a STM mechanism or system.

Another distinction between the conceptualization of verbal STM as an emergent property of general language-related processes and the conceptualization of verbal STM as the product of a dedicated STM mechanism is the way in which the two approaches permit generalizing verbal STM processes to other cognitive domains. If short-term buffering of verbal information is an emergent property of the interconnectivity of neurons in the verbal system then it stands to reason that similar buffering can emerge from the interconnectivity of neurons inside the visual-spatial system, the kinesthetic system, or even inside a petri dish (c.f., Vishwanathan, Bi & Zeringue, 2011). STM performance in these domains would then resemble verbal STM performance because in either case STM would emerge from the same substrate. Indeed, some verbal STM-like phenomena have been found outside the verbal domain (e.g., Logie, Della Sala, Wynn, & Baddeley, 2000; Tremblay, Parmentier, Guerard, Nicholls, & Jones, 2006). If, however, the verbal STM process were dependent upon the function of a verbal short-term store, then additional stores would have to be invoked for every domain in which STM behaviour is observed. This is implausible, because it is doubtful that the human brain could accommodate a bespoke store for every single activity that might need buffering. The alternative would be that a single or a limited number of stores accomplish all STM functions in all cognitive domains. This assumption is also implausible. Arguably a single or even a handful of stores would be overwhelmed by the buffering demands of the entire cognitive system. Nonetheless, if it should prove impossible to ascribe all verbal STM phenomena to the function of general receptive and productive mechanisms, then the concept of a bespoke short-term store would have to be invoked in the verbal domain and in other domains, too.

Classical Approaches to Verbal STM

Store-based approach

Perhaps the most prominent instantiation of the view that STM function is supported by a bespoke system is the Working Memory model (Baddeley, 1986, 2003). The model postulates the existence of three storage systems: A phonological loop for STM of verbal material, a visuo-spatial sketchpad for visual and spatial STM and an episodic buffer that acts as a short-term store for any information that is not stored by the other two systems. It should be noted that the episodic buffer is a relatively recent addition to the model (Baddeley, 2000) and its properties remain somewhat under-specified. It is also not yet clear why additional specialized verbal and visuo-spatial buffers should be needed if the brain already contains an episodic buffer that could perform their functions. Indeed, of the three buffer systems proposed by the Working Memory model, the phonological loop has received the greatest amount of research attention (Baddeley, 2003), and therefore lends itself particularly well for the discussion of the merits of a store-based approach to STM.

The phonological loop is divided into two components, a passive phonological store and an active articulatory control process akin to sub-vocal speech (Baddeley, 2003). As its name indicates, the phonological store holds to-be-remembered verbal items in a modality-neutral (and hence abstract) phonological code. Within the store, these phonologically-coded items decay rapidly, and need to be refreshed via the articulatory control process (i.e., articulatory rehearsal), or else they are lost. Another critical feature of the phonological store is that despite the fact that its unit of currency is a modality-independent phonological representation, the route into the store differs according to the modality of input: If the to-be-remembered items are presented in the auditory modality, they gain direct, obligatory, access to the store. In contrast, visually presented material can only be uploaded into the phonological store indirectly via the articulatory control process. Thus, the articulatory control process has a dual function: To refresh decay-prone phonological representations (regardless of modality of input) and to convert graphemes into phonemes (for visual items).

It is a credit to the Working Memory model, that it identifies specific brain regions onto which the articulatory control process and the passive phonological store can be mapped, and thus avoids falling into the “dragon trap”. Thus it does not postulate that the dedicated STM mechanism is out there somewhere, but makes specific and testable predictions about its location. The phonological store is suggested to be located in the left temporoparietal region of the brain, Brodmann Area (BA) 40, whereas the articulatory control process is mapped onto Broca’s area (BA 44) (Baddeley, 2003). Indeed, the Phonological Loop model even addresses the issue of evolutionary plausibility, arguing that the phonological loop evolved to facilitate language acquisition (Atkins & Baddeley, 1998, Baddeley, 2003, Baddeley, Gathercole, & Papagno, 1998).

The Phonological Loop model is capable of explaining a broad range of verbal STM phenomena. The standard test of verbal STM is the verbal serial recall task.

Participants are typically presented with a brief list of words, letters or digits, which have to be reproduced immediately or very shortly after presentation, usually in their original order. This has long been the standard test of STM because it is assumed to tap into the ability to organize sequences of actions, an ability that is central to much of goal-directed animal and human behaviour, from locomotion, through reaching and grasping, to language use and the control of logical reasoning (Lashley, 1951). Variations of the to-be-remembered material, the modality of presentation and the addition of various concurrent tasks have revealed some very robust patterns of performance. These patterns are thought to reveal the limitation of human cognitive functioning, and any credible theory of STM must therefore adequately account for these limitations.

Perhaps the most crucial benchmark phenomenon of verbal serial recall performance and one that is pivotal to the Phonological Loop model is the *phonological similarity effect* (PSE): It is more difficult to serially recall a list of similar “sounding” items, e.g. “B”, “C”, “V”, etc. than it is to serially recall a list of dissimilar “sounding” items, e.g. “X”, “Y”, “Q”, etc. (Conrad, 1964). The reason “sounding” is placed in quotation marks here is that, critically, this effect does not depend on whether the items are presented auditorily or visually (Conrad & Hull, 1968), an observation that is central to the claim that the phonological store is indeed phonological: On the Phonological Loop model, the PSE occurs because verbal items, regardless of input modality, gain access to an abstract-phonological store; the more similar these phonological representations, the more easily confused they are during retrieval from the store (Baddeley, 1992).

Another key phenomenon of verbal serial STM is that the concurrent articulation of an irrelevant speech utterance like “the, the , the…” reduces serial recall performance markedly (also often termed ‘articulatory suppression’; Baddeley, 1986; Baddeley, Lewis, & Vallar, 1984; Macken & Jones, 1995; Nairne, 1990). According to the Phonological Loop model, concurrent articulation blocks the loop’s articulatory control process. Since this control process is needed to refresh decaying item representations in the store, concurrent articulation results in loss of items from the store and hence an overall performance reduction (Baddeley, Thomson, & Buchanan, 1975).

Crucial to the fractionation of the phonological loop into a passive phonological store and an articulatory control process is a three-way interaction between the PSE, concurrent articulation, and input modality. If to-be-remembered material is presented visually then in addition to reducing overall performance, concurrent articulation reduces or abolishes the PSE (Baddeley & Larsen, 2007, Jones, Macken, & Nicholls, 2004, D. J. Murray, 1968). In contrast, with auditory presentation, concurrent articulation still negatively affects overall performance, but the PSE is not fully abolished. Particularly in recency, that is, the last few items of a to-be-remembered list, the PSE tends to be preserved under concurrent articulation (Jones et al., 2004, Jones, Hughes, & Macken, 2006). The Phonological Loop model accommodates this observation by pointing to the different pathways through which visual and auditory information gains access to the short-term store. Since auditory stimuli are obligatorily uploaded into the store, the phonological similarity of to-be-remembered items still determines the likelihood of inter-item confusion, even when the articulatory control process is suppressed. In contrast, if item presentation is visual, then suppression of the control process prevents items from accessing the store, so that their phonological similarity is immaterial for recall success.

A yet further canonical finding in STM research is the effect of input modality. This *modality effect* manifests itself primarily as a recall advantage in recency on lists containing an auditory component when compared to pure visual lists. It is thus observable with pure auditory items (e.g., Jones et al., 2004), vocalized visually-presented items (Conrad & Hull, 1968), and visually presented items that are read to the participants (Crowder, 1970). Whilst the Phonological Loop model does not directly account for the modality effect, it is conceivable that the preferential access of auditory information to the phonological store proposed by the model somehow benefits recall on the auditory list. Another approach taken by store-based STM theories is to propose the existence of a precategorical acoustic store, an additional low capacity store dedicated exclusively to auditory information (Crowder & Morton, 1969).

Numerous findings from studies with brain damaged patients can also be accounted for by the Phonological Loop model. For example, in patients with a speech planning impairment, like apraxia of speech, the PSE is reduced for visually, but not auditorily, presented material (Waters, Rochon, & Caplan, 1992). In contrast, patients with peripheral motoric speech production impairments, like anarthria or dysarthria show a normal PSE (Bishop & Robson, 1989). The Phonological Loop model suggests that this is because the articulatory control process is based primarily on speech planning mechanisms associated with BA 44. Therefore, a pathological disruption of the speech planning mechanism disrupts the articulatory control process, which limits or prevents the access of visual information to the phonological store. As is the case under concurrent articulation, if items do not get access to the store then confusions between items on the basis of phonological similarity cannot take place, and the PSE is reduced. Peripheral motoric speech impairment does not affect the control process and hence visual and auditory items still get access to the store where similar items are liable to be confused (Baddeley, 2003).

Finally, it has been observed that patients with damage to the left temporoparietal brain region tend to show selective verbal STM deficits, in the absence of an immediately detectable impairment in speech fluency (Vallar & Baddeley, 1984; Vallar, 2006). This observation is central to the Phonological Loop model as it seems to clearly indicate that verbal STM capacity is dissociable from general language-related processes. If verbal STM can be selectively impaired through damage to a specific brain region then this suggests both that the existence of a specialized language-independent verbal STM mechanism is likely, and that this mechanism is located in the damaged region.

Alas, it is questionable whether the possibility of language impairments in “pure” STM patients can be ruled out completely. It might always be the case that the language impairment is substantial enough to have a knock-on effect on verbal short-term storage, but not substantial enough to be detected by conventional tests of linguistic ability. Furthermore, neuroimaging studies with healthy volunteers have so far failed to confirm a specific region in the left temporoparietal brain area as the seat of the language-independent phonological store (Buchsbaum & D’Esposito, 2008). Hence, the existence of a neurological equivalent of a phonological store remains debatable.

The absence of unequivocal neurological evidence for a language-independent phonological store, however, begs the question whether it is necessary at all to invoke the theoretical concept of a short-term store as an entity in order to explain verbal STM phenomena. One prominent alternative approach is to discard the idea that memory requires a dedicated store coupled with a separate active-refreshing process.

Instead, it is suggested that all information is obligatorily remembered but that access to this memory is prone to interference.

Interference-based approach

Interference-based models of STM are based on the assumption that there are two kinds of memories; a secondary memory in which all of a person's experience is stored, and a primary memory in which currently active representations are held (James, 1890). Thus, no information is ever truly forgotten in secondary memory, but access to that information from primary memory is often occluded by interference. A prominent instantiation of interference-based STM models is the Feature Model (Nairne, 1990, Neath, 2000). According to this model each to-be-remembered item is composed of a number of modality-dependent features. Modality dependent features are physical features of the item, like its visual shape or the voice in which it is spoken. When an item is encoded its features are simultaneously uploaded into primary and secondary memory. An additional set of internal modality-independent features is appended to the representation of an item in either memory. These modality-independent features arise from internal item-categorization processes. For example, if the same digit is presented twice, once auditorily and once visually then the two memory representations of the digit will have many overlapping modality-independent features, but no overlapping modality-dependent features (Nairne, 1990). According to this model, forgetting occurs because items interfere with earlier items (retroactive interference) in primary memory. Specifically, a given feature of item n is overwritten if that feature is also present in item n +1. Modality-dependent features interfere with modality-dependent features only, and the same applies to modality-independent features. Since retrieval depends on accessing the correct item from secondary memory given item features present in primary memory, the more degraded the representation of an item in primary memory, the more difficult it is to access the correct item in secondary memory.

The Feature Model demonstrates how an interference-based model can successfully account for many verbal STM effects without having to invoke the existence of a dedicated short-term buffer: For example, on this model, the PSE occurs because similar items have more overlapping features, so that more retroactive interference between to-be-remembered items takes place in primary memory, making accurate retrieval of the correct items from secondary memory more difficult (Nairne, 1990). The modality effect is also easily explained by the Feature Model, without any necessity for an additional acoustic store: Visually presented items are represented more heavily in terms of modality-independent features and auditorily presented items more heavily in terms of modality-dependent features. Since the modality-independent features are related to internally generated activity, it is more likely that the last visual item will be overwritten by internal activity like task irrelevant thoughts or indeed by subvocal rehearsal of early list items (Nairne, 1990). It is noteworthy that from the interference-based perspective active rehearsal of list items is thus in fact considered somewhat detrimental for recall success. The Feature Model explains the effect of concurrent articulation by arguing that features of the irrelevant utterance are adopted into primary memory where they distort the modality-independent features of the to-be-remembered material. This leads to a reduction in STM performance. If the to-be-remembered material is highly confusable to begin with, such as when to-be-remembered items are phonologically similar, then the additional interfering features introduced by concurrent articulation will have less of an impact than when the to-be-remembered material consists of phonologically dissimilar items. This is how the

Feature Model accounts for the reduction of the PSE in the presence of concurrent articulation. Since it is assumed that concurrent articulation generates primarily modality-independent based interference and that auditory items are encoded to a greater extent in terms of modality-dependent features, suppression does not affect the PSE as much if the to-be-remembered material is presented auditorily (Neath, 2000).

Thus it seems that an interference-based account of STM is also capable of explaining the effects of concurrent articulation, phonological similarity and modality on verbal STM without invoking the concept of bespoke short-term buffers. Nevertheless, a severe limitation of the interference-based approach to STM is that very little concern is given to specifying what neurological equivalents there might be for entities like primary or secondary memory. For once, this makes the concepts of primary and secondary memory very “draconic”. Clearly, these mental entities do not exist outside of the brain, i.e. on a metaphysical plane, yet without any specification of their location inside the brain it is impossible to falsify the existence of these constructs. Believing in non-falsifiable entities is, alas, not much different than believing in dragons. In addition, the lack of specification of neurological correlates of the Feature Model constructs reduces the utility of the interference based approach for predicting the effects of neurological disorders, or indeed for explaining data associated with these. It is for example unclear from an interference-based perspective why damage to specific brain areas, like Broca’s area, as observed in apraxia of speech (Ogar, Salama, Dronkers, Amici, & Gorno-Tempini, 2005), should affect verbal STM performance in ways similar to concurrent articulation. It is difficult to see how damage to specific brain areas might, like concurrent articulation, introduce irrelevant item features that would interfere with items in primary memory. Indeed, if memory is conceived of as a passive process so that active maintenance of the to-be-remembered information through, for example, rehearsal is unnecessary or indeed disruptive, it is unclear how disruption of any brain mechanism might negatively affect memory performance.

Perceptual-gestural approach

The perceptual-gestural approach to verbal STM (Hughes, Marsh, & Jones, 2009, 2011; Jones et al., 2004, 2006) is not as far removed from the store-based approach as interference-based models. Like the store-based account the perceptual-gestural perspective argues that active maintenance of the to-be-remembered material needs to take place. A crucial distinction between the store-based and the perceptual-gestural perspectives, however, is that the former considers the maintenance process to be in service to a passive store whereas the latter rejects the idea of a bespoke short-term storage entity altogether. Instead, it is argued that verbal STM is primarily an emergent property of the function of mechanisms that are not specifically mnemonic but ones involved in general perception and production processes. For example, the store-based tradition would explain the relationship between verbal STM task performance and second language acquisition (Atkins & Baddeley, 1998) with a bespoke STM mechanism having evolved to facilitate language acquisition. In contrast, the perceptual-gestural account would suggest that the human ability to use language has evolved to facilitate human co-habitation, and that this ability can be recruited for short-term retention of verbal material. This is not to say, however, that the perceptual-gestural approach considers language as indispensable for verbal STM or STM in general. Language is just exemplary of a very sophisticated ability which relies heavily on perceiving and gesture planning. Furthermore, it stands to reason that if information is categorized as verbal, that it will be maintained in a verbal way. This is because any living organism needs to be economical with its energy expenditure (Anderson, 1995). Hence, if information is categorized as verbal then the linguistic neural path is likely to be the most well-trodden, and hence least effortful, for processing that information. For example, whereas an illiterate person might process a written letter as a visual token, a skilled reader is likely to encode the letter verbally.

In order to explain verbal STM processes without invoking bespoke stores the perceptual-gestural perspective emphasizes the planning of articulatory gestures and processes of (auditory) perceptual organization. In contrast to the Phonological Loop model which only sees a role for articulatory mechanisms in refreshing the decaying representations of items in a passive phonological store, the perceptual-gestural perspective proposes that the articulatory plan itself serves as the repository of verbal information. Specifically, in order to maintain verbal information an articulatory motor plan is assembled wherein the to-be-remembered material is maintained as a series of articulatory gestures. The assembly and maintenance of the planned gesture sequence is however not flawless and transposition errors between to-be-remembered items are possible. These transpositions are akin to Spoonerisms and are more likely between items that require similar articulatory actions (Jones et al., 2006). Additionally, the perceptual-gestural account draws attention to the high degree of sophistication and automation with which the perceptual system meaningfully organizes incoming pieces of information, particularly when the information is sequential. In the visual domain the principles of perceptual organization have been described in the Gestalt literature (e.g. Koffka, 1935). For example, it has been observed that the visual system tends to process continuous entities as cohesive objects (Spelke, 1990). These objects are, amongst other things, defined by their perceived edges, which constitutes the Gestalt principle of the figure-ground contrast.

Similar principles apply in the auditory domain (see Bregman, 1990), where certain characteristics of the auditory input like pauses or changes in voice are perceived as markers of distinct auditory perceptual objects or “streams” (Frankish, 1989; Hughes et al., 2009, 2011). Evidently, any series of to-be-remembered events is subjected to a considerable amount of categorization and segregation, which is likely to influence its recall. Thus perceptual organization influences the memory process before the to-be-remembered material could possibly reach any dedicated storage system. The perceptual-gestural account acknowledges this by proposing that the way the to-be-remembered information has been organized by the perceptual system influences the nature of the articulatory motor plan generated to maintain to-be-remembered information (Hughes et al., 2009, 2011).

According to the Phonological Loop model, verbal input is represented in a modality-neutral phonological code. In contrast the emphasis on general acoustic and gestural processing of the perceptual-gestural approach suggests the codes are more peripheral and modality specific than ‘phonological’. Thus, it is argued that the PSE arises from a greater articulatory and not phonological confusability between items. This articulatory confusability leads to more frequent transpositions of the articulatory gestures through which the items are cohered into a sequence and maintained for serial recall (A. W. Ellis, 1980; Jones et al., 2006).

The perceptual-gestural account also redefines the impact of concurrent articulation. According to this account, suppression does not prevent the refreshing of decaying phonological item representations residing in some separate passive store but rather disrupts the formulation and maintenance of a gesture sequence assembled with the purpose of correct output of the to-be-remembered verbal material (Jones et al., 2004, 2006). This is because concurrent articulation itself requires the planning and production of verbal utterances, and thus limits the ability to recruit the articulatory planning system for the formulation and retention of a sequence-output plan. If the to-be-remembered material is not processed through the articulatory system, however, then the articulatory similarity of to-be-remembered items will have little impact on the likelihood of correct recall. Thus the perceptual gestural account explains why the PSE is reduced under concurrent articulation.

Acknowledging the sophisticated perceptual streaming of an auditory list, the perceptual-gestural account is also capable of explaining the modality effect without invoking an additional store dedicated to retention of acoustic items. If in the auditory domain silences can serve as object-defining boundaries (cf. Bregman, 1990), then it follows that the silence at the end of auditory to-be-remembered list presentation will act as such a boundary. The perceptual-gestural view suggests that the silence will thus act as a perceptual anchor thereby facilitating the recall of the end of the auditory list (Jones et al., 2006). It should be noted here that this principle does not apply with similar sounding items. Because similar items are less perceptually distinct, the transitions between the items are relatively indistinct, too. Thus, the perceptual boundary at the end of an auditory similar item list constitutes a less salient order cue, and auditory similar item list recall does not show a recency advantage. Moreover, because in the visual domain objects tend to be defined through spatial as opposed to temporal boundaries (Bregman, 1990), the cessation of the presentation of the visual list is not as salient as it is for an auditory list and does not serve as such a strong anchor. Hence performance in recency is superior for auditory lists.

Clearly, if the perceptual processes responsible for the modality effect are independent of processes responsible for the PSE, that is, articulatory planning processes, then it is not surprising that these two effects should be observed independently of each other. Thus, for visual lists when the PSE is attenuated with concurrent articulation recall of ‘phonologically’ similar and dissimilar visual lists will be equal, because the end of the visual dissimilar lists does not constitute a perceptual anchor that improves recall. If to-be-remembered lists are presented auditorily, however, dissimilar lists will be recalled under suppression better than similar lists, but only because the auditory advantage in recency will not be affected by suppression. While this might seem like the PSE is preserved under concurrent articulation, it is the perceptual-gestural view that the superior performance on auditory dissimilar lists in recency under suppression constitutes an acoustic not a phonological similarity effect (Jones et al., 2006).

Importantly, findings with brain damaged patients can also be accommodated without postulating a dedicated STM mechanism. Generally, since the perceptual-gestural account, like the Phonological Loop model and in contrast to interference based models, argues that short-term retention of verbal information requires an active process it is conceivable that a lesion of the brain would impede that process and produce a STM impairment. Thus, for example, the observation that patients with speech planning impairments like apraxia of speech perform similarly to non-clinical experimental participants under concurrent articulation (Waters et al., 1992) is in line with the perceptual-gestural account. The account clearly predicts that if the speech planning mechanism is impaired, then recall performance will be generally reduced, and the articulatory similarity between items should have no bearing on how well they can be recalled. This is because the account postulates that a speech plan needs to be assembled to maintain the to-be-remembered information. Thus, damage to the speech planning mechanism is seen as direct impairment of the verbal STM process rather than as an impediment of the process through which items are refreshed for storage in a separate bespoke system.

Finally, regarding the evidence for the existence of a brain area that could be regarded as the neurological equivalent of the phonological store (Vallar & Baddeley, 1984; Vallar, 2006), it has been observed that the left temporoparietal region, the most probable location of the store, might instead be responsible for the integration of speech perception and production (Hickok, 2009). If verbal STM emerged, primarily, from the function of speech perception and production processes, as postulated by the perceptual-gestural account, then it is clear that damage to an area responsible for the integration of these processes would result in a substantial verbal STM deficit. The cause of this deficit would, however, not be the dysfunction of a bespoke storage mechanism, but rather the inability to upload the perceived verbal information stream into an articulatory motor plan for maintenance. The selectivity of the verbal STM impairment (cf. Vallar & Baddeley, 1984; Vallar, 2006) can also be thus accommodated: Selective impairment of a region integrating speech perception and production processes would not necessarily affect the discrete abilities to either perceive or produce speech. Only when integration of these abilities is required, as is the case when an articulatory motor plan is assembled from a perceived list of verbal tokens that needs maintaining, would a selective STM deficit become apparent. Thus, while the left temporoparietal region might show the properties of a short-term buffer, it is clearly not an area that is language-independent and specifically dedicated to verbal STM.

Rationale for Empirical Work

It appears that store-based, interference-based and perceptual gestural approaches to explaining verbal STM are about equally capable of accounting for the effects of phonological similarity, modality and concurrent articulation, and the interactions between them. Clearly behavioural manipulation of these factors was not sufficient to clearly adjudicate between the three accounts. One possible way to address this issue is to look towards the evolutionary plausibility of the constructs that each theory postulates. From this perspective, it seem that of the three presented accounts the perceptual-gestural is the most promising, because it postulates fewer dedicated system than the store based approach, and yet, in contrast to the interference-based view, enables a clear mapping of its constructs onto brain mechanisms. Evolution, however, can often have unexpected results and generate surprising adaptations (i.e. black swans). The likelihood of evolution of its constructs can therefore not be the sole criterion for dismissing a cognitive theory, in particular if the theory offers ways how its postulated constructs might be evolutionary plausible, like proposing that the phonological loop is a language learning device (Baddeley et al., 1998).

Another approach is to empirically test the predictions of the STM theories against new types of experimental manipulations. One particularly promising type of manipulation that will receive special attention in this thesis is the induction of brain lesions with transcranial magnetic stimulation (TMS). This technique makes it possible to temporarily reduce the activity of the stimulated region. Thus it is possible to conduct lesion studies on healthy volunteers. Lesions studies can reveal a lot about the adequacy of STM theories. For example, the observation that patients with damage to speech planning areas have severe impairments of verbal STM (e.g. Waters et al., 1992), suggests that some active articulatory process is involved in maintaining verbal information. This is in line with the store-based and perceptual-gestural, but not with the interference-based view, as it is not clear how damage to articulatory planning should impair the function of primary or secondary memory. Lesions studies with patients are inherently problematic, however, because patients with appropriate lesions are rare, the brain damage is rarely selective, making it difficult to establish clear correlations between a single brain region and its function, and patients are often capable of compensating for their impairments. Inducing temporary lesions with TMS circumvents many of these problems.

Following these deliberations, the aim of the research presented in this thesis was to test the predictions of the perceptual-gestural, the interference-based and the store-based accounts using new behavioural manipulations and in particular to attempt temporary disruption of constructs proposed by these accounts with TMS.

Preview of Empirical Chapters

Chapter 2: The impact of non-verbal concurrent tasks on verbal STM

A clear distinction that can be drawn between the interference-based, store-based and perceptual-gestural approaches to verbal STM is the importance each approach attributes to articulatory-motoric planning processes. The perceptual-gestural account considers articulatory gesture planning processes as heavily involved in normal verbal STM function. The Phonological Loop model also considers these processes as important albeit regarding them as subservient to a passive short-term store. Both accounts therefore predict that even peripheral motoric impairment of articulatory processes should impede verbal STM.

In contrast to this, from an interference-based perspective articulatory processes only play a role in the short-term maintenance of verbal information if they generate verbal representations. Thus concurrent articulation impairs verbal STM because the modality-independent features of the irrelevant verbal utterance distort the representations of the to-be-remembered material (Neath, 2000). Non-verbal impairment of articulatory motor processes should hence have little effect on verbal STM performance. The first empirical chapter of this thesis (Chapter 2) attempts to adjudicate between the contrasting predictions of the interference-based account on one hand and the perceptual-gestural and store-based accounts on the other, by investigating whether and how a non-verbal constraint on articulation, chewing gum, impedes verbal short-term memory.

Chapter 3: Theta Burst Stimulation of Broca's area modulates verbal STM

Chapter 3 continues to evaluate the interference-based, store-based and perceptual-gestural approaches to verbal STM, by empirically addressing the varying predictions the accounts make about the consequences of speech planning impairment. Because the Phonological Loop model identifies BA 44, Broca's area, (Baddeley, 2003) as the location of the articulatory control process, the model predicts that a lesion to the area should reduce visual verbal STM performance and attenuate the PSE. Without the articulatory control process, visual material should not gain access to the phonological store. Thus visual items would not be maintained irrespective of their phonological similarity. Overall, reduced function of Broca's area should have effects similar to concurrent articulation.

At first glance, the perceptual-gestural account makes similar predictions about the consequences of damage to Broca’s area for verbal STM. The account suggests that inhibition of the articulatory planning mechanism should impair the ability to assemble an articulatory plan for visual to-be-remembered list recall. Consequently, recall performance should be reduced as should the likelihood of articulatory confusions and hence the PSE. Yet, it should be noted that the perceptual-gestural account does not explicitly specify Broca’s area as the seat of the speech-plan assembly mechanism. Indeed, given how the account usually emphasizes the interaction of perceptual and speech planning processes, which involve a large number of brain areas, it seems more in line with the account to consider Broca’s area merely a component of a distributed mechanism capable of generating an articulatory plan. This means that, according to the perceptual-gestural view, a selective lesion of Broca’s area might produce a very selective impairment of the speech-plan assembly process, reducing, for example, only the likelihood of articulatory confusions or only the likelihood of correct recall. Finally, from an interference-based perspective, a selective lesion of Broca’s area should have no effect on STM performance because, in contrast to concurrent articulation, a lesion would not introduce interfering item representations to the memory traces of to-be-remembered items. In Chapter 3 these varying predictions are addressed empirically by applying repetitive TMS to the pars opercularis of the left inferior frontal gyrus in order to induce a temporary lesion of Broca’s area in healthy volunteers.

Chapter 4: A new approach to modality effects in verbal serial recall: Meeting the challenge of explaining a visual mid-list advantage

In Chapter 4 the focus of the thesis shifted towards comparing and evaluating the predictions of the three prominent verbal STM accounts in regards to the effect of perceptual factors on verbal short-term recall. In the past, the centre of such a discussion would be the standard modality effect, the auditory advantage in recency when comparing visual and auditory list recall. Indeed the standard modality effect can be accommodated by any of the three prominent verbal STM accounts. Thus the interference-based account postulates that auditory items are encoded more in terms of modality-dependent features which are not prone to interference from internal activity at list end. The store-based view explains superior performance in recency on auditory lists with preferential access of auditory information to the phonological store, or invokes an additional low capacity buffer dedicated to storing auditory items. The perceptual-gestural view argues that the silence at the end of the auditory list constitutes a perceptual anchor that improves recall in recency.

However, a review of literature presented in chapter 4 reveals that a considerable number of previous studies indicate that the traditional modality effect is often matched by a visual advantage at early and mid-list portions of the serial position curve. The aim of Chapter 4 was to determine to what extent this hitherto neglected phenomenon—the inverted modality effect (Beaman, 2002),—might be accommodated by each of the three STM accounts.

CHAPTER 2: THE IMPACT OF NON-VERBAL CONCURRENT TASKS ON VERBAL STM

Abstract

The store-based and perceptual-gestural accounts of verbal STM suggest that any impairment (e.g. by concurrent oral activity) of articulatory planning/production processes will also impair verbal STM. In contrast, the interference-based Feature Model argues that a concurrent oral activity is only disruptive for STM performance because it introduces irrelevant verbal item features, which interfere with the internal representations of the to-be-remembered material. This chapter reports the first studies to show that chewing gum, a non-verbal constraint on articulation, impairs verbal short-term recall of both item-order and item-identity. Experiment 1 showed that chewing gum reduces serial recall of letter lists. Experiment 2 indicated that chewing does not simply disrupt vocal-articulatory planning required for order retention: chewing equally impairs a matched task that required retention of list item identity. Experiment 3 demonstrated that manual tapping produces a similar pattern of impairment to that of chewing gum. These results pose a problem for verbal STM theories asserting that forgetting is based on domain-specific interference.

Introduction

As described in the previous chapter, both the store-based approach (e.g., Baddeley, 2003) and perceptual-gestural approach (Hughes et al., 2009, 2011) to verbal STM postulate that, particularly for verbal serial recall tasks, the classic test of verbal STM, speech planning mechanisms are utilized covertly, either to refresh decaying phonological representations in a labile short-term store (e.g., Baddeley, 2003) or to bind the grammatically and semantically unconstrained sequence into a coherent motor-plan for action (e.g., Hughes et al., 2009). In contrast, the interference-based approach—at least as exemplified by perhaps the most prominent account of this type, the Feature Model (Nairne, 1990; Neath, 2000)—postulates that verbal STM does not require an active articulatory process. Instead, verbal information is obligatorily and passively encoded as a set of modality-dependent features (based on perceptual processing) and modality-independent features (based on internal processing) in an interference-prone primary memory and a secondary memory representing the compendium of all experience. The challenge of correct retrieval arises from the need at recall to find an adequate match between the potentially degraded representation of an item in primary memory and its stable counterpart in secondary memory. Mismatches constitute forgetting, which becomes more likely if more interfering features enter primary memory.

In line with the store-based and perceptual-gestural accounts is the frequently observed negative impact on serial recall of concurrent repetition of an irrelevant verbal utterance (e.g. “the…the...the...”)—i.e., concurrent articulation (e.g., Baddeley, 1986; Jones et al., 2004; D. J. Murray, 1968). Both the store-based and the perceptual-gestural accounts argue that concurrent articulation thereby impedes articulatory planning processes, thus either preventing refreshing of information inside a dedicated short-term store (Baddeley, 2003) or preventing the assembly and upkeep of a motor plan for maintaining the to be remembered sequence. However, the Feature Model can also accommodate the effect of concurrent articulation: Whilst the model denies that the repetition of an irrelevant utterance impairs use of the articulatory system to support recall (e.g., Neath, 2000), it supposes that the modality-independent features of the irrelevant verbal utterance distort the representations of the verbal to-be-remembered (henceforth: TBR) material in primary memory.

To adjudicate between the interference-based approach on the one hand and the store-based and perceptual-gestural accounts on the other it seems necessary to look for alternative concurrent activities. Specifically, an activity is needed that may be expected to impair articulation, but, from the perspective of the Feature Model (Neath, 2000), would not produce irrelevant modality-independent features which might distort TBR item representations in primary memory. An oral activity that may lend itself well in this respect is gum chewing. Like concurrent articulation, chewing gum has also been argued to involve complex movement of the jaw and tongue muscles (Sakamoto, Nakata, & Kakigi, 2009), yet, according to the Feature Model, chewing, in contrast to concurrent articulation, should not interfere with the TBR material. Indeed, the Feature Model eschews the claim that non-verbal tasks generate features that are adopted into representations of TBR verbal items in primary memory (cf. Guerard, Jalbert, Neath, Surprenant, & Bireta, 2009; Neath, 2000). Hence, according to the Feature Model, while chewing gum might limit articulatory fluency, it should have minimal impact on verbal STM performance. In contrast, from the standpoint of the store-based and perceptual-gestural accounts, verbal STM should be impaired by any process that obstructs speech planning, including chewing gum.

The only studies to date that have examined the impact of chewing gum on short-term recall suggest that, if anything, chewing gum *enhances* performance (Baker, Bezance, Zellaby, & Aggleton, 2004; Wilkinson, Scholey, & Wesnes, 2002). To anticipate: The current series of experiments provides evidence that such a conclusion is unwarranted, as this chapter shows for the first time that fundamental aspects of STM—recall of both order and item identity—are in fact impaired by gum chewing.

Chewing Gum and Short-Term Memory

In the first study to investigate the effects of gum chewing on STM, participants were either given a mint flavoured gum, asked to mimic chewing movements in the absence of gum or did not engage in any chewing movements (Wilkinson et al., 2002). Cognitive abilities were assessed with the Cognitive Drug Research (CDR) computerized battery (for details, see Kennedy, Scholey, & Wesnes, 2000). It was found that when chewing gum, participants performed better on spatial item-recognition memory and short-term old/new number and word recognition tasks. Additionally, when participants were only pretending to chew gum, their number recognition performance was still higher than that of the control group. However, on most other CDR tasks—whether dependent on STM or not—their performance was worse (for similar results, see Stephens & Tunney, 2004). Beneficial effects of chewing gum have also been found for free recall of a relatively long list of words (fifteen items; Baker et al., 2004; Johnson & Miles, 2008). It has been suggested that the facilitative effects of chewing gum on memory may be mediated by an increase of blood flow to fronto-temporal brain regions due to the mastication process (Wilkinson et al., 2002). Others suggest that the effects might at least partly reflect a context effect, to which the flavour of the gum contributes rather than have to do with chewing or gum per se (Baker et al., 2004; Johnson & Miles, 2008).

At first glance the lack of impairment as a result of chewing gum appears to be in line with the Feature Model, and poses a problem for the store-based and perceptual-gestural accounts, which predict that any constraint on articulation should reduce verbal STM performance. However, none of the previous studies that have examined the effects of chewing gum on STM have employed serial recall, the bedrock on which theories of STM have been built (e.g., Baddeley, 1986, 2003; Conrad, 1964). The aim of the experimental series presented in this chapter therefore was to dissociate between the conflicting predictions of the interference-based approach on the one hand and the store-based and perceptual-gestural approaches on the other through investigating the impact of non-verbal concurrent oral activity— chewing gum— on verbal STM.

Experiment 1a

The first experiment tested the effects of chewing (flavourless) gum on serial recall. Participants were presented with lists of TBR letters whilst chewing or not chewing gum. Based on theories of STM that appeal to speech-planning mechanisms (e.g., Baddeley, 2003; Jones et al., 2004), it was expected that the tongue, mouth, and jaw movements involved in the task-extraneous activity of chewing would impair STM performance. In contrast, given that it would seem implausible to suppose that chewing gum would produce modality-independent features, which would be adopted into representations of TBR items, the lack of a negative effect of chewing would be interpreted as confirming the Feature Model.

As an additional means of examining the possible similarity of action between chewing gum and concurrent articulation on serial recall, phonological similarity was included as an additional variable. As was described in the previous chapter, phonologically similar items (P, V, B...) are recalled more poorly than phonologically dissimilar items (H, Q, L...), and previous research showed that this PSE is attenuated by concurrent articulation (Baddeley & Larsen, 2007; Jones et al., 2004; D. J. Murray, 1968). It was hence of interest to see whether the PSE would also be attenuated by a non-verbal constraint on articulation.

Method

Participants.

Forty-six Cardiff University native English-speaking students (32 females), aged between 18 and 37 years (mean: 21.8 years) participated in the experiment.

Materials, Design & Procedure.

To be comparable to previous research examining the effects of concurrent articulation, the experiment was modelled closely on the visual list conditions from Jones et al. (2004). The experiment was a 2 (gum chewing) x 2 (phonological similarity) x 7 (serial position) within-participant design. On each trial, 7 randomly ordered letters were presented visually, in black *Times New Roman* 72-point font on a 17 inch monitor. The letters were either phonologically similar (P, V, B, C, D, G, T) or dissimilar (H, Q, L, R, K, X, Y). Each letter was presented for 250 ms with an inter-stimulus interval (offset to onset) of 750 ms. At the end of each trial, seven buttons featuring the letters presented on the trial appeared on screen. Participants were to click on the letters in the order in which they occurred in the just-presented list, by operating the mouse with their dominant hand. Each button could only be clicked once, and all buttons had to be clicked in order to proceed with the experiment.

There were two blocks of 28 trials, one block in which participants were required to chew gum (*Wrigley's* flavourless gum; see Johnson and Miles, 2008) and one in which they were not. The blocks immediately followed each other. The order of blocks was randomized across participants. In the chewing gum condition, the participants were instructed to chew the gum more vigorously during the presentation of the TBR items but could reduce their pace of chewing somewhat during response output. The experiment lasted approximately 30 min. It was conducted in a sound attenuated booth and, with their permission, participants were monitored via a video link to ensure compliance with the instructions.

Results and Discussion

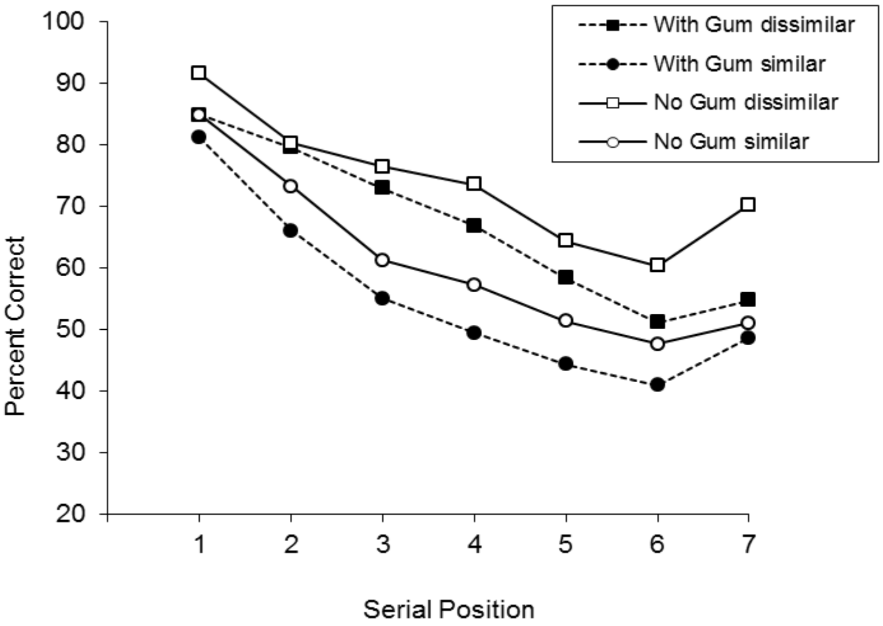
As per convention, performance was measured by assessing for each TBR item whether it had been recalled in its correct serial position. Average correct recall for the four conditions is plotted in Figure 1a.[1]

The first aspect of the results to note is that the PSE was replicated: A 2 (gum) by 2 (similarity) by 7 (serial position) repeated measures ANOVA revealed that recall was poorer for similar compared to dissimilar letters, $F(1, 45) = 54.52$, $MSE = 0.08$, $p < .01$, $\eta_p^2 = .548$. The novel feature of the results, however, is that serial recall—regardless of phonological similarity—was also significantly poorer whilst chewing

---

[1] To check whether or not block-order (chewing or non-chewing) had any influence on the results, the sample was initially split into two groups depending on whether the chewing condition was the first or second condition (22 participants in gum-first group). A 2 (group; gum condition first or second) by 2 (gum) by 2 (similarity) by 7 (serial position) mixed ANOVA revealed no significant difference between the groups, $F(1, 44) = 0.44$, $MSE = 0.64$, $p = .51$, $\eta_p^2 = .01$, and no significant interaction between group and gum, $F(1, 44) = 0.7$, $MSE = 0.07$, $p = .41$, $\eta_p^2 = .02$, or between group and any other variable. Thus, the order of the chewing/non chewing blocks did not have a bearing on the results.

gum, *F*(1, 45) = 22.25, *MSE* = 0.07, $p < .01$, $\eta_p^2 = .331$. There was also a main effect of serial position, *F*(6, 270) = 113.135, *MSE* = 0.03, $p < .01$, $\eta_p^2 = .715$, reflecting the classic serial position curve. Note in particular that chewing did not alter the magnitude of the PSE: The interaction between phonological similarity and chewing was not significant, *F*(1, 45) = 0.79, *MSE* = 0.05, $p = .38$, $\eta_p^2 = .02$.



<u>Figure 1a</u>: Mean percentage of items correctly recalled in order with phonologically similar and dissimilar lists as a function of chewing or not chewing gum and serial position.

The present experiment establishes that chewing gum reduces verbal serial STM performance. These findings are in line with the hypothesis that mouth/jaw movements that are not dedicated to the articulatory planning of the TBR list should impair memory (Baddeley, 2003; Jones et al., 2004). From this standpoint, chewing movements may either disrupt encoding and refreshing of decay-prone phonological item representations (cf. Baddeley, 1986) or the assembly and maintenance of a motor sequence-plan (see, e.g., Hughes et al., 2009). In any case, it seems that, against the predictions of the Feature Model, a non-verbal constraint on articulation can also impair verbal STM.

The results of Experiment 1a also indicate that the previous assertion that chewing gum is beneficial for STM (e.g., Wilkinson et al., 2002) must be qualified with an important caveat: In contrast to previous research in this area, when the task involves STM for sequences of events as opposed to short-term item recognition or free recall (i.e., Baker et al., 2004; Johnson & Miles, 2008; Stephens & Tunney, 2004; Wilkinson et al., 2002), a clear reduction in performance is found as a result of gum chewing. Before accepting this caveat, however, it seems prudent to check whether the fact that previous studies showing benefits of chewing gum involved instructing participants to ‘chew naturally and constantly’ (cf. Wilkinson et al., 2002) as opposed to chewing ‘vigorously’ during item presentation had any bearing on the results. It is possible that it was the apprehension of the need to chew vigorously as opposed to the act of chewing itself that impaired performance in the chewing gum condition. Indeed, from the perspective of the Feature Model, it could be argued that this apprehension might have perhaps been internally vocalized, and might thus have produced interfering modality-independent features. Experiment 1b therefore replicated Experiment 1a except participants were instructed to chew ‘naturally and constantly’ throughout the chewing block.

Experiment 1b

Method

Participants.

Twenty-three Cardiff University native English-speaking students (aged 18-27, mean: 21.04; 9 males) participated in this experiment.

Materials, Design & Procedure.

This experiment was a replication of Experiment 1a with the only difference being that participants were now instructed to chew naturally and constantly throughout the chewing block.

Results

Average performance across conditions is depicted in Figure 1b. A 2 (chewing) by 2 (similarity) by 7 (serial position) ANOVA revealed that, as in Experiment 1a, there was a main effect of chewing gum, $F(1, 22) = 9.64$, $MSE = 0.05$, $p < .01$, $\eta_p^2 = .31$, and chewing did not interact with phonological similarity, $F(1,22) = 0.14$, $MSE = 0.06$, $p = .71$, $\eta_p^2 = .01$.

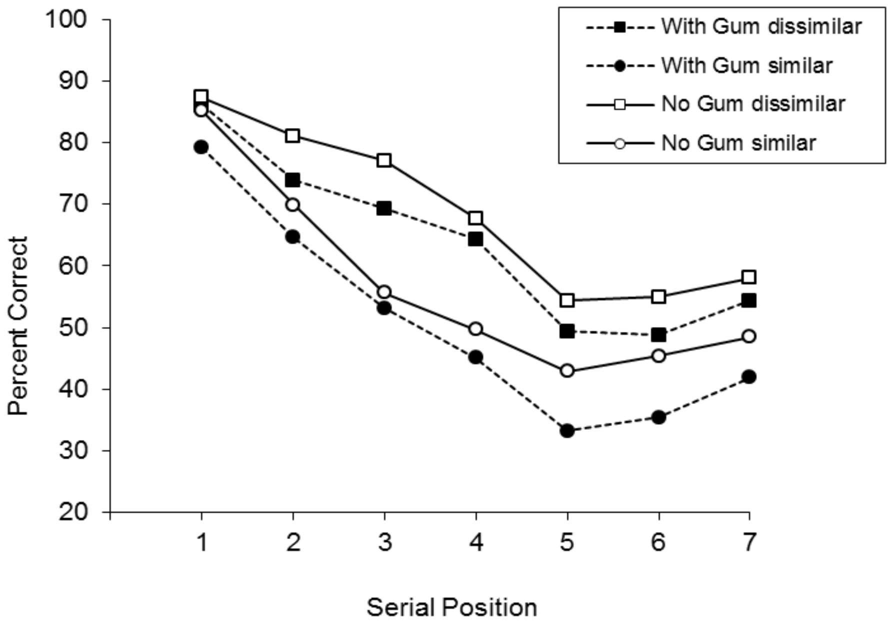


Figure 1b: Mean percentage of items correctly recalled in order with phonologically similar and dissimilar lists as a function of chewing or not chewing gum and serial position under instructions to chew ‘naturally and constantly’ rather than vigorously (cf. Experiment 1a).

Discussion

It appears that it does not matter whether people are instructed to chew vigorously during item presentation or are free to chew naturally: In both cases chewing has an overall adverse effect on serial recall. Nevertheless, in the context of serial recall, the instruction to chew vigorously during TBR list presentation makes the paradigm more comparable to other concurrent tasks used in STM research, like concurrent articulation, which usually have to be performed during a certain stage in each trial but rarely throughout the entire experiment or constantly throughout a block of trials (cf., Baddeley, 1986; Jones et al., 2004; D. J. Murray, 1968). Thus, in the subsequent experiments of the series, the instruction to ‘chew vigorously’ during item presentation was used.

In sum, Experiment 1 suggests that chewing gum, a non-verbal constraint on articulation, can reliably reduce verbal STM performance. The effects of chewing thus resemble the effects of concurrent articulation, and pose a challenge for the Feature Model. Yet, it might be possible to align the current results with the Feature Model, if it were argued that the act of chewing in and of itself, like concurrent articulation, somehow introduced modality-independent features that were adopted into the representation of the TBR list in primary memory. Such an explanation is unlikely, however, given that from the perspective of the Feature Model feature adoption is not generally considered to take place on non-verbal tasks (Guerard et al., 2009; Neath, 2000). Furthermore, if such an explanation were given, then the Feature Model would also predict that chewing, like concurrent articulation, should impact the PSE as well as general performance. This effect was not observed.

Concededly, however, even from the perspective of accounts that emphasize the role of articulatory-motoric planning processes in STM (Baddeley, 2003; Jones et al., 2004), if chewing gum were entirely like concurrent articulation, then it should also reduce the PSE. It therefore seems somewhat premature to dismiss the Feature Model based on the present findings without further investigating to what extent chewing and concurrent articulation are actually comparable.

One possibility is that the main effect of gum is simply not of sufficient strength to have the more subtle impact on the magnitude of the PSE, with the main effect of concurrent articulation being typically much greater (cf., Baddeley, 1986; Jones et al., 2004; D. J. Murray, 1968). However, Experiment 2 provides another, arguably stronger, test of whether the action of chewing gum is similar to that of concurrent articulation. Several studies have observed that concurrent articulation has a particularly strong impact on *serial* STM tasks, when compared to matched tasks not requiring memory for order (cf. Beaman & Jones, 1997; Macken & Jones, 1995). If the effects of gum were to match the effects of concurrent articulation, gum should also have a stronger impact on serial memory. Experiment 2 addresses this suggestion by comparing the effect of chewing gum on a task requiring STM for order with that on a matched task that requires the retention of item identity but not order.

Experiment 2

A test of verbal STM for a list of items that is devoid of the need to retain their serial order is the ‘missing item’ task (e.g., Beaman & Jones, 1997; Buschke, 1963; LeCompte, 1996). Here, participants are required to identify a missing item from a randomly ordered fixed set of items (e.g., ‘7’ is missing from the list ‘28149365’ taken from the digit-set 1-9). Thus, each item presented must be retained so as to identify the item that is not. However, the serial order of the list items is immaterial and the task is not thought therefore to rely on sequence planning but rather on a judgment of item familiarity (e.g., Buschke, 1963). Corroborating this, compared to serial STM tasks, the missing item task has been shown to be far less affected by factors that are thought to act upon sequence planning including talker variability (Hughes et al., 2011), temporal grouping (Klapp, Marshburn, & Lester, 1983), changing-state irrelevant sound (Beaman & Jones, 1997; Macken & Jones, 1995) and, of particular relevance here, concurrent articulation (Klapp et al., 1983).

A serial STM task that is—other than the need to retain serial order—well matched to the missing item task is the probed order task (Beaman & Jones, 1997; Hughes et al., 2011). Here, participants are again presented with the randomized fixed set of items but at test are re-presented with one of the presented items (the probe) and required to indicate which item followed it in the list. This ensures that the missing item and the probed order tasks are matched on the stimuli and output requirements. If chewing gum is like concurrent articulation, then it should disrupt particularly tasks that require *serial* STM. It should therefore adversely affect the probed order task more than the missing item task.

This experiment also provides a test of whether the chewing effect observed in Experiment 1 is one that specifically affects the initial encoding of the TBR stimuli, rather than one that acts on vocal-articulatory rehearsal. This was achieved by manipulating whether the TBR lists were presented visually or auditorily. As discussed in Chapter 1, several theories of STM suggest that auditory and visual items are encoded differently, with auditory items having direct access to a phonological store (Baddeley, 2003), being obligatorily processed through automatic perceptual organization processes (e.g., Jones et al., 2004), or being subject to obligatory processing by brain regions responsible for speech planning (Hickok, 2009). This contrasts with the more active, deliberate, recoding of visually-presented items into a phonological (Baddeley, 2003) or articulatory code (Hughes et al., 2009). However, these theories also suggest that the use of the articulatory system to serially rehearse a to-be-remembered list is the same regardless of modality, as indicated by the fact that concurrent articulation impairs both visual-verbal and auditory-verbal order recall (e.g., Jones et al., 2004). Thus, if the effect of chewing differs from that of concurrent articulation and is one that operates instead at an early stimulus-encoding stage, its effect should be greater with visual lists.

Method

Participants.

Twenty-eight Cardiff University native English-speaking students (24 females), aged between 18 and 23 (mean: 19.86) participated in the experiment.

Materials, Design & Procedure.

The same type of flavourless gum was used as in Experiment 1. The TBR lists comprised eight digits selected randomly from the 9-item set *1-9.* In the visual condition, they were presented in the same fashion as in Experiment 1. For the auditory condition, the digits were recorded in a male voice with a 16-bit resolution, at a sampling rate of 48 kHz, and compressed digitally to 250 ms using *Audacity 1.3.12 (Beta)* software (http://audacity.sourceforge.net), without altering acoustic features such as pitch, and presented with a gap of 750 ms between the digits. On each trial, the TBR items were presented in a quasi random order with the constraint that that there were no more than two ascending or descending runs of two or more digits (e.g., 2-3 or 7-6) within a given list and that there were no runs of 3 or more digits.

The experiment was a 2 (gum chewing) x 2 (task) x 2 (modality) within-participant design. Participants encountered in a random order a chewing and a non-chewing block. In each of these blocks there were four randomly ordered 18-trial blocks, one for each modality (auditory vs. visual) and each task (probed order vs. missing item). Each trial block was preceded by two practice trials. On each trial a random digit from 1-9 was omitted. The trial blocks were arranged so that each digit from the set 1-9 would be missing twice. On missing item trials participants were required to indicate on an array of buttons 1-9 which digit was missing on a given trial. In the probed order condition participants were presented with a digit from the TBR list and had to indicate which digit immediately followed it. As only 7 serial positions could thus be probed, each serial position was probed twice in a random order across trials, and then another 4 randomly selected serial positions were probed to match the number of trials in the missing item condition. The procedure was the same as Experiment 1a, with the experiment lasting approximately 50 minutes.

Results

Figure 2 shows the percentage of correctly identified missing items and correctly recalled probed items across the eight conditions. As suggested by the pattern evident in Figure 2, a 2 (task) by 2 (modality) by 2 (chewing gum) repeated measures ANOVA revealed a main effect of task, with performance on the missing item task being better than on the probed order task, F (1, 27) = 40.73, MSE = 0.03, $p < .05$, $\eta_p^2 = .6$. There was also a main effect of modality: Recall was better with auditory than visual lists, F (1, 27) = 5.52, MSE = 0.02, $p < .05$, $\eta_p^2 = .17$. Furthermore, and of greater interest, there was a main effect of chewing gum, F (1, 27) = 25.11, MSE = 0.02, $p < .05$, $\eta_p^2 = .48$, and this detrimental effect of chewing gum was found regardless of task or modality, as indicated by the absence of any significant interaction terms: F (1, 27) = 0.38, MSE = 0.02, $p = .54$, $\eta_p^2 = .01$, F (1,

27) = 0.12, *MSE* = 0.01, $p = .73$, $\eta_p^2 = .01$ and *F* (1, 27) = 2.08, *MSE* = 0.02, $p = .16$, $\eta_p^2 = .07$, for chewing and task, chewing and modality and the three-way interaction, respectively.

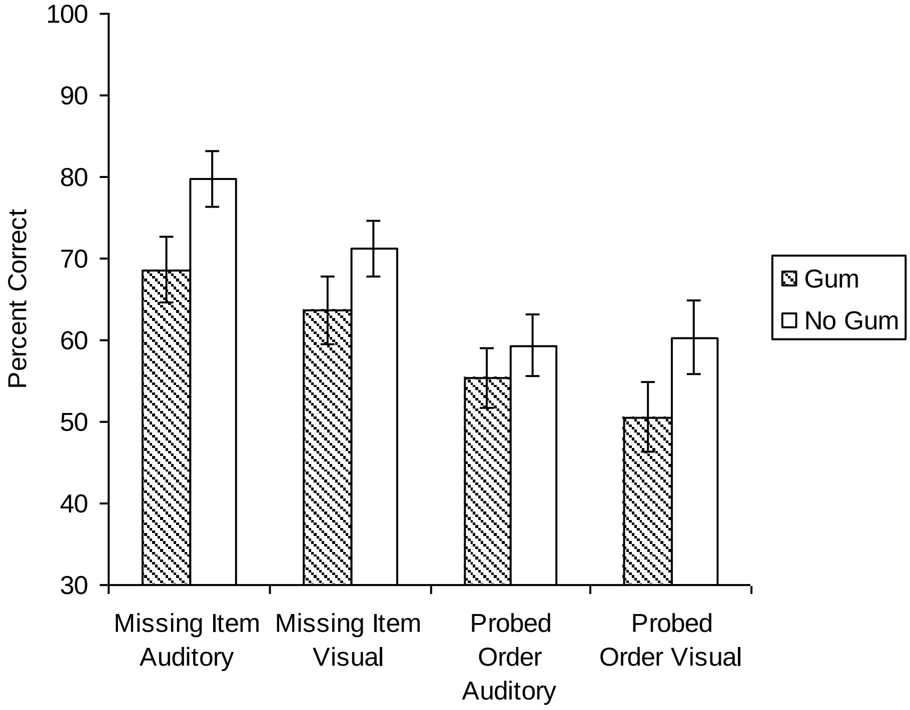


Figure 2: Mean percentage of missing and probed items recalled with auditorily and visually presented lists in the presence or absence of chewing gum. Error bars represent +/-1 standard error.

Discussion

Experiment 2 showed that, as in Experiment 1, chewing gum significantly impaired STM for order as measured on this occasion by its disruption of probed order recall. However, Experiment 2 also demonstrated that this adverse effect extends to memory for item identity: Missing-item recall was compromised to a comparable degree to that of probed order recall. The adverse effects of chewing on STM do not appear to be limited, therefore, to tasks that have typically been more strongly associated with articulatory sequencing (Baddeley, 2003; Hughes et al., 2009). The results also show that chewing impairs STM of visually and auditorily presented lists to a similar extent. This suggests that chewing is not impairing the kind of deliberate encoding often associated with visual as compared with the obligatory encoding of auditory lists (e.g., Baddeley, 1986; Hickock, 2009; Hughes et al., 2009). Instead, chewing seems to exert its effect at a more central stage potentially concerned with maintenance of the TBR material.

The results of Experiment 2 provide mixed evidence regarding the implications of the effects of chewing gum for theories of verbal STM. On the one hand, chewing, like concurrent articulation, consistently reduces STM performance. Further, its effects are clearly not limited to a simple impediment of encoding; if this was so, it would not affect recall of visually and auditorily presented material to a similar extent. This is in line with STM theories that invoke a key role for speech mechanisms and thus predict a negative impact of any task constraining articulation (e.g., Baddeley, 2003; Jones et al., 2004). However, there is a discrepancy between the predictions of these theories and the present results insofar as they predict that impairment of speech planning mechanisms that serve to maintain order information should impair the probed order task more than the missing item task. Moreover, the fact that concurrent articulation, by preventing rehearsal, reduces (indeed usually abolishes) the PSE with visual presentation (Baddeley et al., 1984) but, as was noted earlier, chewing gum does not (Experiment 1), also militates against a simple account in terms of an impairment of speech mechanisms.

There are indications in the pattern of data reported thus far that the effects of chewing resemble more the effects of manual tapping than they do concurrent articulation. The tapping task traditionally involves the repeated placement of one or several fingers on a hard surface in a steady and rhythmic fashion. Chewing and tapping have both been suggested to promote cognitive abilities by releasing excessive muscle tension (Freeman, 1940). This assertion is challenged, however, by numerous studies demonstrating the adverse effects of tapping on STM (e.g., Guerard et al., 2009; Saito, 1994). Tapping has also been contrasted with chewing: Chewing was found to increase and tapping to decrease reaction speed in an auditory oddball paradigm (Sakamoto et al., 2009). Yet, the effects of tapping and chewing on STM have, to my knowledge, never been compared in the same study. However, as was observed with chewing gum in Experiment 1, there is some evidence that simple tapping impairs serial recall without affecting the PSE (Guerard et al., 2009). Furthermore, it seems that order recall and missing item recall are not differentially affected by simple tapping (Macken & Jones, 1995), which mimics the effect of chewing found in Experiment 2. However, it is difficult to draw firm conclusions from the study of Macken and Jones (1995) because performance in the absence of a secondary task was not assessed and the TBR lists were presented only visually. Thus, Experiment 3 replicates Experiment 2 in all respects except that chewing was substituted by simple tapping. If the two activities affect STM through a similar mechanism, then the same pattern of results should be observed as in Experiment 2.

Experiment 3

Method

Participants.

The participants were 23 Cardiff University native English-speaking students (19 females), aged between 18 and 23 (mean: 19.52), who had not participated in Experiment 2.

Materials, Design & Procedure.

The method was similar to Experiment 2 except that participants were required to tap their fingers rather than chew vigorously. Participants were to tap the table with their fourth, then third and then second finger of their non-dominant hand at a pace of 3 taps per second. In line with previous STM studies involving tapping (e.g. Guerard et al., 2009)—as well as concurrent articulation (e.g. Jones et al., 2004)—participants were only required to engage in the secondary activity (tapping) during list presentation.[2]

Results

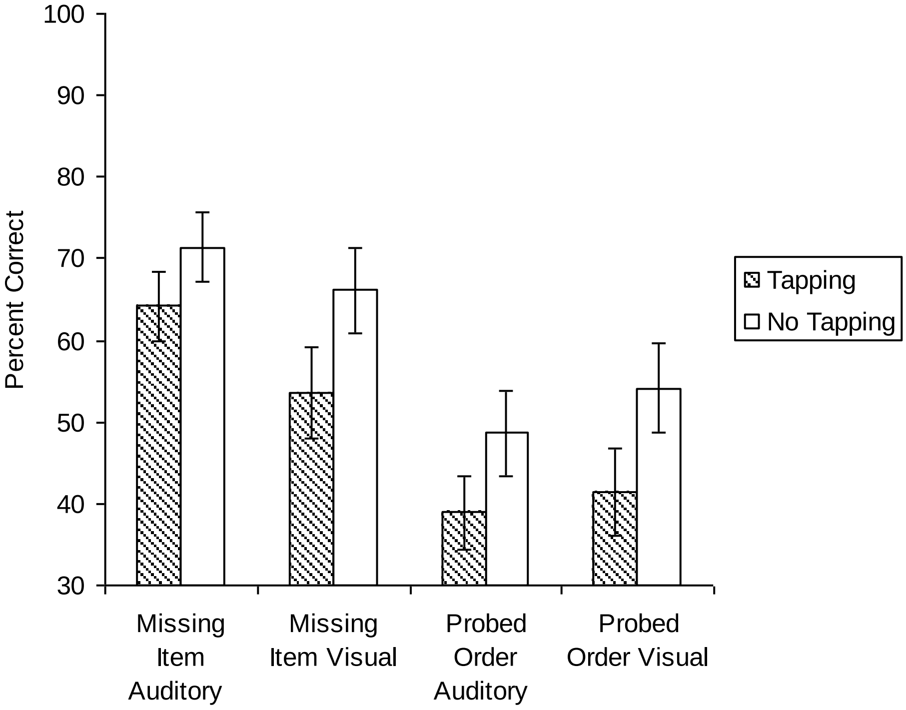
Figure 3 depicts the percentage of correctly identified items in the missing item and probed order tasks. The overall pattern of performance resembles that of Experiment 2. A 2 (tapping) by 2 (modality) by 2 (task) repeated measures ANOVA showed that there was a significant main effect of task: performance was significantly better in the missing item task, F (1, 22) = 12.47, MSE = 0.1, $p < .05$, $\eta_p^2 = .36$. Most importantly, as with chewing in Experiment 2, there was also a significant reduction in performance during tapping, F (1, 22) = 13.16, MSE = 0.05, $p < .05$, $\eta_p^2 = .37$. Tapping did not significantly interact with any other factor, with the interaction terms for tapping and task, tapping and modality and the three-way interaction being: F (1, 22) = 0.12, MSE = 0.01, $p = .73$, $\eta_p^2 = .01$, F (1, 22) = 0.55, MSE = 0.01, $p = .47$, $\eta_p^2 = .01$ and F (1, 22) = 1.16, MSE = 0.01, $p = .29$, $\eta_p^2 = .05$, respectively.

The pattern deviates somewhat from Experiment 2, however, insofar as there was no significant effect of modality, F (1, 22) = .42, MSE = 0.03, $p = .52$, $\eta_p^2 = .02$,

---

[2] It seems unlikely that having to chew ‘naturally’ during the recall phase in Experiment 2 (but not continue tapping during the recall phase in Experiment 3) would make comparison of the impact of the two forms of activity problematic, especially given that the recall phase involved only a single keypress response.

but instead a significant task by modality interaction, *F* (1, 22) = 4.87, *MSE* = 0.02, *p* < .05, η_p^2 = .18. An additional simple effects comparison between the average visual and auditory condition performance on each task reveals that this interaction reflects significantly higher performance on the auditory condition in the missing item task, *F*(1, 22) = 5.19, *MSE* = 0.01, *p* < .05, η_p^2 = .19. There was no difference in performance between the two modalities on the probed order task, *F* (1, 22) = .62, *MSE* = 0.02, *p* = .44, η_p^2 = .03. Note that these discrepancies between the present data and the results of Experiment 2 do not involve the tapping manipulation and so are not of primary concern here.



<u>Figure 3</u>: Mean percentage of missing and probed items recalled with auditorily and visually presented lists in the presence or absence of tapping. Error bars represent +/-1 standard error.

To directly compare the effects of chewing to the effects of tapping, the average differences between performance in the presence and in the absence of the concurrent motor tasks were calculated for each condition in the present dataset and the data from Experiment 2. The average impact of chewing on each task in each modality was then compared to the average impact of tapping in a 3-way mixed ANOVA. The two within-participant variables were modality and task and the between-participants variable was concurrent motor task (chewing or tapping). This comparison yielded no significant main effect of concurrent motor task, F (1, 49) = .93, MSE = 0.06, p = .34, η_p^2 = .02, indicating that both chewing and tapping had a similar impact on STM. There was also no significant interaction of concurrent motor task with any other variable, with the interaction terms for concurrent task and modality, concurrent task and STM task and the three-way interaction being: F (1, 49) = 0.1, MSE = 0.03, p = .76, η_p^2 = .002, F (1, 49) = 0.46, MSE = 0.04, p = .5, η_p^2 = .01 and F (1, 49) = 3.11, MSE = 0.03, p = .09, η_p^2 = .06, respectively. This indicates that the effects of tapping and chewing were equivalent for both STM tasks independently of presentation modality.

Discussion

It appears that there is no difference between the effects that tapping and chewing have on short-term item and order recall. The lack of a significant interaction between concurrent task and presentation modality further indicates that the adverse effects of neither tapping nor chewing are due to an impairment of item encoding. Rather, it seems that these peripheral motor tasks disrupt some modality unrelated process involved in the maintenance of items in a list regardless of whether the retention of their order is required. Because tasks that are thought to rely on vocal-articulatory sequencing to different extents (order-based tasks such as serial recall and probed order recall compared to the missing-item task; e.g., Beaman & Jones, 1997;

LeCompte, 1996) are equally impaired by chewing and tapping, this maintenance process seems to be independent of such articulatory sequencing.

General Discussion

The present series has shown for the first time that chewing gum—a non-verbal oral activity—has an overall negative impact on verbal STM tasks, both serial and non-serial. In Experiment 1, it was demonstrated that chewing has an adverse effect on visual-verbal serial recall, the most commonly used test of STM capacity. In Experiment 2, it was shown that this observation extends to a different short-term order recall task, to auditory lists, and to a task that is not thought to depend on articulatory sequence planning: A task requiring short-term retention only of item identity was also reduced by chewing. Finally, Experiment 3 yielded results that were consistent with the hypothesis that the detrimental effects of chewing on STM are akin to those of simple manual tapping (e.g., Guerard et al., 2009; Saito, 1994).

Chewing Gum and STM Theories

At first glance the present data challenge STM accounts that postulate that verbal STM is a (negative) function of domain-specific interference. For example, one prominent model of this type—the Feature Model (e.g., Guerard et al., 2009; Nairne, 1990; Neath, 2000)—suggests that concurrent irrelevant articulation reduces memory performance by generating task-irrelevant verbal representations that corrupt the representations of the (also verbal) TBR items. Clearly, if this were the case then a non-verbal constraint on articulation like chewing should have no impact on memory.

However, proponents of the Feature Model may appeal to a free parameter included in the model representing a general attentional resource (parameter ‘a’).

Thus, non-verbal concurrent tasks can impair STM because they increase general task demands and deplete attentional resources needed for successful item retrieval (Guerard et al., 2009; Neath, 2000). Tapping and chewing may therefore both simply be general distracters. Indeed, both concurrent tasks had very similar effects on verbal STM, and neither of the tasks produced results that would usually be associated with a constraint on articulation, like reduction of the PSE or selective impact on memory for order. It seems therefore on second glance that the Feature Model offers in fact the best explanation for the current data.

That said, the “general distraction” explanation seems questionable. First, the Feature Model does not seem to offer a way to determine a priori to what extent a task should deplete the general attentional resource. Indeed, invoking the Feature Model, it would be impossible to predict that the simple tapping task should produce the same amount of distraction as the, physically very different, chewing task (cf. Jones & Tremblay, 2000). Moreover, by invoking the parameter ‘a’, the model implies that concurrent tasks that convey phonological features, like concurrent articulation, and concurrent tasks without a phonological component, like tapping or chewing (or, according to the Feature Model, irrelevant sound, cf. Hughes & Jones, 2005; Neath, 2000), impact verbal STM through different mechanisms. Inspecting the present results in conjunction with related literature it then becomes unclear why tasks with and without a phonological component produce such similar results. For example, complex tapping, like concurrent articulation, can reduce the PSE (see Guerard et al., 2009), and steady-state concurrent articulation, like simple tapping or chewing, does not have a distinctive impact on serial memory tasks (Macken & Jones, 1995). These similarities between verbal and non-verbal impairments of STM suggest that the degree of impairment is unitarily determined by the complexity of the planned gestures involved in the concurrent task. The Feature Model's suggestion of two separate processes, namely the increase in attentional demands for non-verbal concurrent tasks and feature adoption for verbal concurrent tasks, appears untenable.

With the Feature Model not offering a satisfying explanation for the present data another glance at verbal STM accounts that invoke a key role for language planning/production processes (Acheson & MacDonald, 2009; Baddeley, 2003; Jones et al, 2004) is in order. The majority of these accounts (e.g., Baddeley, 2003; Jones et al., 2004), though not all of them (cf. Acheson & MacDonald, 2009), differentiate between constraints on articulatory planning and those on articulatory production. Indeed it has been demonstrated that patients with anarthria, an impairment of the neuromuscular mechanisms required for articulation, show no reduction of the PSE (Baddeley & Wilson, 1985). Only when patients show speech planning deficits, as opposed to pure production deficits (such as in apraxia of speech), is a clear reduction of the PSE observed (Waters et al., 1992). Similarly, steady-state suppression, that is, concurrent articulation with low speech planning demands, like the concurrent repetition of a single letter, reduces performance on the missing item task and the probed order task to a comparable extent (Macken & Jones, 1995). Only changing-state suppression—concurrent repetition of a sequence of, say, three letters—reduces performance on the serial memory task more than on the missing item task. Thus, accounts that see a central role for language planning/production processes (Baddeley, 2003; Jones et al., 2004) can be reconciled with the present findings if it is assumed that both chewing and tapping impair articulation at a peripheral level. At that level the concurrent activity reduces overall performance but does not differentially affect performance on phonologically similar and dissimilar lists, nor differentially affect performance on order and item recall tasks. Thus, from this standpoint, tapping and chewing are not simply distracters: They are peripheral impairments placed on the production aspect of the articulatory planning and production network needed to either refresh decaying item representations in a short-term store (e.g., Baddeley, 2003) or to assemble a coherent motor-plan for action (e.g., Hughes et al., 2009, 2011; Jones et al., 2006).

Implications for Research on Chewing Gum and Cognition

The present findings also clearly warrant a re-evaluation of the assertion that chewing benefits STM (e.g., Wilkinson et al., 2002). The discrepancy between the current study and previous research on the effects of chewing on STM could be associated with the absence of flavour in the gum used in the present study. Flavour has previously been suggested as one factor underpinning the beneficial effects of gum, by creating a context in which encoding of the items would be promoted (Baker et al., 2004; Johnson & Miles, 2008). It is feasible that there could be an evolutionary advantage to better encode one’s environment in the presence of a palatable stimulus to be able to later recreate the circumstances in which the stimulus was found. Thus, in the present study it is possible that a flavoured gum would have enhanced encoding and would thus have offset the negative effects of the concurrent motor task. However, because chewing gum usually loses its flavour after several minutes of chewing, with flavourless gum being potentially quite unpalatable, it seems advisable especially in light of the current findings, that chewing gum is only considered a performance enhancer as long as its flavour lasts. Thereafter, the adverse effects on cognition, as demonstrated in the present study, might outweigh the beneficial ones. Establishing the exact tradeoffs between the cognitive advantages and disadvantages of chewing flavoured and flavourless gum is beyond the scope of the present thesis but could be a worthwhile avenue for further research. However, the absence of flavour could not have been the main reason why chewing reduced performance in the present experiment, because tapping produced similar results to chewing. Clearly, both chewing and tapping involve a motor component, and if the adverse effect of chewing were to do with the absence of flavour, in addition to or instead of it being a motor impairment, it seems likely that chewing would have had a different effect from tapping.

Another possible reason for the negative effect of chewing observed in the present study might be the rigorous control that was implemented to ensure that the participants did indeed chew during item presentation. Even in Experiment 1b in which participants were instructed to chew naturally they were still monitored to make sure they were chewing. Previous studies, however, (e.g. Baker et al., 2004; Wilkinson et al., 2002) are somewhat vague about how it was ensured that the participants were indeed chewing. As Experiment 3 of the present study demonstrates, a motor activity needs to be present in order for a decline in performance to occur. If participants in some of the previous gum studies failed to follow instructions and ceased chewing, one cannot be certain which aspect of having chewing gum in their mouth might have influenced their performance. Furthermore, the present study employed tasks in which encoding and reproducing the TBR stimuli took place over the course of a few seconds. The comparatively long trials of some previous studies (e.g., Baker et al., 2004; Johnson and Miles, 2007) might have enabled participants to compensate for any motoric disruption caused by chewing. Finally, it should be noted that many studies in fact failed to find a beneficial effect of chewing on memory (Johnson and Miles, 2007, 2008; Miles and Johnson, 2007; Overman, Sun, Golding, & Prevost, 2009; Tucha, Mecklinger, Maier, Hammerl, & Lange, 2004), despite using methods similar to the studies that did find a benefit of gum (i.e. Baker et al., 2004; Wilkinson et al., 2002). Their number is likely to be conservative due to the difficulty of publishing null results. Thus, it seems that whatever beneficial effect chewing might have on memory, it is not very robust.

The finding that chewing and tapping have comparable effects on cognitive performance also has implications for chewing gum in the academic setting. There is some evidence that the efficacy of repeatedly tapping fingers in a predetermined order—the tapping task used in the current Experiment 3—is related to phonological decoding skills required for reading (Carello, LeVasseur, & Schmidt, 2002). If tapping, reading, and, as the present study suggests, chewing rely on some of the same mechanisms, then engaging in one of these tasks would interfere with the other. Clearly, more research is needed to determine how chewing gum might interact with phonological decoding and reading.

Conclusions

The experiments reported in this chapter establish that some fundamental aspects of STM—memory for list order and item identity—are adversely affected by peripheral motoric tasks like chewing gum or tapping. This is informative for theories of verbal STM as it challenges predictions of models postulating a central role for domain-specific interference. Instead, accounts postulating the involvement of peripheral motoric processes in verbal STM are supported. Previous applied research in this area, which postulates a generally beneficial effect of gum, is also challenged by the present findings. Indeed, the disruption produced by chewing might, like tapping, affect performance on other everyday tasks such as reading.

CHAPTER 3: THETA BURST STIMULATION OF BROCA’S AREA MODULATES VERBAL STM

Abstract

There is a long established tradition that assumes that the retention and reproduction of a sequence over the short-term relies on bespoke short-term memory stores. For example, the Working Memory model postulates that to-be-remembered visual-verbal material is uploaded via an articulatory control process into a language-independent phonological store. Accordingly, the phonological similarity of the items is a key determinant of the success with which they can be retrieved from that store. The study described in the present chapter is the first of its kind: Activity of a brain region associated with the articulatory control process, Broca’s area (localized through a combination of structural and functional methods), was inhibited with theta burst transcranial magnetic stimulation. According to the Phonological Loop model, this temporary lesion should reduce access of visual-verbal material to the store, resulting in a deficit in overall STM performance. However, this was not observed; rather, there was a selective attenuation of the phonological similarity effect. This dissociation of the effect of transcranial magnetic stimulation (TMS) of Broca’s area on overall performance and on the impact of phonological similarity seems more readily accommodated by accounts that emphasize a primary role for articulatory processes in serial short-term memory rather than ones that regard such processes as peripheral to a dedicated store. Note that alternative approaches that emphasize domain-specific interference in STM also struggle with the present data because they do not predict any effect of inhibiting an articulatory planning area.

Introduction

The present chapter continues to evaluate the store-based, interference-based and the perceptual-gestural accounts of verbal STM, by testing the adequacy of the role these accounts ascribe to articulatory-motoric processes. The studies reported in Chapter 2 revealed, against the predictions of the interference-based perspective, that even peripheral motoric constraints on articulatory processes can impair STM. The present chapter elaborates on this finding by examining the effects of more central gesture planning impairments through a temporary lesion of Broca's area induced via TMS.

To re-cap, the most prominent store-based STM account, the Working Memory model (Baddeley & Hitch, 1974; Baddeley, 2003), posits that in the verbal domain short-term retention is accomplished through the action of a phonological loop, which comprises two components: a bespoke, passive, language-independent *phonological store* in which phonological representations of verbal input last for one or two seconds before decaying, and an *articulatory control process,* a rehearsal mechanism analogous to subvocal speech, which serves to reactivate the stored items, thus preventing their decay. The articulatory control process is also the means by which visually presented verbal material gains access to the phonological store. Auditory items, on the other hand have direct access to the store (Baddeley, 2003).

These key propositions of the Phonological Loop model account for a wide range of empirical phenomena, chief among them the PSE (Baddeley, 1986; Conrad, 1964). According to the Phonological Loop model phonologically similar items are more readily confused inside the store (Baddeley, 1992), leading to poorer recall. Another canonical effect explained by the Phonological Loop model is the impact of concurrent articulation: The model suggests that the concurrent production of an irrelevant verbal sequence disrupts the articulatory control process, which means that the phonological representations of the TBR items cannot be refreshed, and are more readily lost. Moreover, for visual TBR material, disruption of the articulatory control process impairs access to the phonological store. Thus, the most empirically obvious impact of concurrent articulation with visual lists is an impairment of overall recall performance (e.g., Baddeley et al., 1975; D. J. Murray, 1968). A secondary consequence of the impairment of access of visual material to the phonological store is the reduction of the PSE with visual lists during concurrent articulation (e.g. Baddeley, 2000, 2007; Baddeley & Larsen, 2007; Baddeley e al., 1984): When items cannot enter a store designed specifically to hold verbal items, verbal recall is impaired generally (otherwise it is unclear why such a store would have evolved at all), but this impairment is especially pronounced for items that would, through being phonologically discriminable, have particularly benefitted from gaining access to the store.

In recent decades, neuroimaging and neuropsychological evidence has been brought to bear on the Phonological Loop model. Studies with speech-impaired patients have shown that peripheral motoric impairments of speech production, as observed in anarthric and dysarthric patients, do not impact upon effects like the PSE (Bishop & Robson, 1989). However, apraxic patients—those with a deficit in speech *planning*—lack a PSE for visual but not auditory lists (Waters et al., 1992), exhibiting a pattern of performance similar to that of non-clinical participants under concurrent articulation. Thus the articulatory control process within the Phonological Loop model has been pinpointed to Broca’s area (BA 44), the area that is commonly damaged in apraxic patients (Ogar et al., 2005) and which has been repeatedly implicated in speech planning (Amunts et al., 2004; Davis et al., 2008; Friedman et al., 1998). It is also the area that Working Memory model-inspired imaging studies find to be active during tasks that supposedly tap into the function of the articulatory control process of the phonological loop (Paulesu, Frith, & Frackowiak, 1993). Furthermore, the passive phonological store component of the model has been mapped onto BA 40 (Baddeley, 2003) based on brain-damaged patients in whom damage to this area seems to have resulted in a selective, “pure” impairment in verbal STM tasks in the absence of a substantial general language impairment (see Vallar, 2006, for a review).

In combination, the cognitive and neurological aspects of the Phonological Loop model allow clear predictions about the function of several brain areas and the consequences of lesions to these areas. Thus the model predicts that lesions to BA 40, the phonological store, will result in a reduction in verbal STM performance. A further consequence of such impairment is a reduction of the PSE: Without a mechanism to store phonological representations of TBR items, the phonological similarity of the items ceases to be relevant for recall success. Selective lesion of BA 44, that is, damage to the articulatory control process, should have similar results but only for visually presented items, because the control process is the pathway through which these items gain access to the phonological store. Given that it should be immaterial whether access to the store is blocked because the store itself is damaged (lesion of BA 40) or because access to an (intact) store is constrained (lesion of BA 44), either form of selective impairment should lead to a reduction in STM performance and a reduction of the PSE, at least for visual TBR material.

As pointed out previously, the interference-based approach has been rather silent with respect to predictions flowing from the effects of brain lesions on STM performance. In particular, as was argued in Chapter 2, the interference-based approach struggles to predict an impairment of STM due to a constraint on articulation. Thus, according to the Feature Model (Nairne, 1990; Neath, 2000), the ability to actively rehearse TBR items is immaterial for recall success. Hence, the interference-based account already struggles with data obtained from patients with lesions of Broca’s area, the speech planning region (e.g. Waters et al., 1992), showing that these patients’ verbal STM is impaired.

In contrast to this, the perceptual-gestural account, like the Phonological Loop model, predicts that an impairment of speech planning should lead to verbal STM impairment. However, in contrast to the store-based Phonological Loop model, the perceptual-gestural account rejects the idea that a bespoke short-term store accomplishes the processes associated with verbal STM. Instead it is postulated that STM emerges from the function of general receptive and productive mechanisms (Hughes et al., 2009, 2011; Jones et al., 2004, 2006). In verbal STM these mechanisms are primarily speech-related. While it is in line with the perceptual-gestural view that these general mechanisms have neurological correlates, it is difficult from the perceptual-gestural perspective to pinpoint distinct STM-related constructs onto specific brain regions because, according to this perspective, one would not expect functionally isolable regions of the brain to relate to STM specifically. Clearly, the assembly of perceived information into a motor plan assembled with the purpose of maintaining and reproducing the TBR material is likely to involve a large number of brain regions from the auditory or visual cortex (depending on the modality of presentation) to speech planning regions like Broca’s area to the oral motor cortex. Selective impairment of any of these regions, including BA 40 and BA 44 could, according to the perceptual-gestural view, plausibly result in some kind of verbal STM impairment, because all these regions together contribute to the process of STM. However, because no single area constitutes a bespoke storage mechanism, lesion to a single area would be very unlikely to obliterate STM, as the Phonological Loop model would predict. Instead selective damage is likely to produce a highly selective impairment of STM that would be a consequence of the impairment of general receptive and productive processes that are normally accomplished by the damaged region. This also means that, contrary to the predictions of the Phonological Loop model, a selective impairment of verbal STM in absence of a general speech-related impairment cannot exist.

Indeed, evidence for the existence of a “pure” verbal STM impairment is equivocal. Patients with allegedly selective STM impairments are extremely rare and thus the possibility that their impairment is due to a peculiarity in their neuropathology cannot be excluded (Buchsbaum & D’Esposito, 2008). Furthermore, the extent to which the minor language impairments these patients often show (see, e.g., Shallice & Butterworth, 1977) are dissociable from an impairment of the supposedly language-independent phonological short-term store is debatable. Additionally, some of these patients, like patient PV (Vallar & Baddeley, 1984), still exhibit a preserved PSE for auditory lists (Baddeley, 2003; Vallar & Baddeley, 1984) despite having substantial damage to BA 40 and beyond (Basso, Spinnler, Vallar, & Zanobio, 1982); an outcome that is at odds with the notion that the patients lacked a store in which phonologically similar items could be confused. Finally, attempts to identify a specific area in the left temporoparietal region that exhibits the properties of a phonological store and is language-independent in the non-clinical population using brain imaging techniques have been met with limited success (Buchsbaum & D’Esposito, 2008). Indeed, the region that seems to be the most likely candidate for the seat of the phonological store—the Sylvian parietal temporal region, which is indeed located within BA 40—is also associated with the integration of speech perception and production (Hickok, 2009).

From the store-based perspective, the objection could be raised that the difficulties with finding clear neurological correlates of the phonological loop might be inherent to the interpretational difficulties associated with single-case neuropsychological data and correlational imaging data. One way in which these difficulties might be circumvented, however, is to use TMS. In particular, given that this technique can be used to temporarily induce lesions in the brains of healthy volunteers it is possible to test a sample drawn from a known population that is homogenous, thereby avoiding the risk of sampling error as might be brought about by a range of other factors, such as medication, socio-economic class, age, and so on. A further major advantage of TMS is that participants can serve as their own controls (Romero, Walsh, & Papagno, 2006). Moreover, the potentially confounding effects of auditory and tactile artefacts that accompany TMS can be avoided by using an offline protocol in which TMS is applied prior to the performance of a behavioural task. In the study reported in this chapter, therefore, I used the technique of continuous theta-burst stimulation (TBS; Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005).[3] This high frequency, low intensity, protocol produces a suppressive aftereffect on cortical excitability for up to 1 h and beyond (e.g. Verbruggen, Aron, Stevens, & Chambers, 2010). This is, to my knowledge, the first time TBS has been applied within the field of STM.

Previous studies that have used TMS to study STM have, like imaging and neuropsychological studies, encountered some difficulties in inducing a selective reduction in memory performance when stimulating the supposed site of the

---

[3] From here on, the term “TMS” will be used to refer to transcranial magnetic stimulation in general, and the term “TBS” will be used to refer to the specific continuous theta burst TMS protocol introduced by Huang et al. (2005).

phonological store. For example, Romero et al. (2006) applied TMS (separately) to region BA40 and region BA44. It was expected that stimulation of BA 40 would selectively affect performance on a STM task that required the retention of a digit sequence, but that—given the supposed language-independence of the phonological store—it would not affect performance on two phonological judgment tasks that did not require STM. However, it was found that stimulating either area equally affected performance on both types of task. Thus BA 40 could not be confirmed as the purported site of the language-independent store.

Another fMRI-guided TMS study found that stimulation within BA 40 reduced the PSE for visually-presented non-words (Kirschen, Davis-Ratner, Jerde, Schraedley-Desmond, & Desmond, 2006). This was interpreted as confirmation of the involvement of phonological storage. However, the task used was a forced-choice item recognition test and not the traditional serial memory task that has been used to investigate STM; indeed, all the main constructs within the Phonological Loop model are predicated upon phenomena that are quintessentially serial STM memory phenomena. Furthermore, the pattern of performance did not, in any case, match expectations based on the phonological storage account: TMS produced an increase in the speed with which false phonologically similar lures were rejected. The Phonological Loop account, however, posits that inhibition of the phonological store reduces the PSE because such inhibition impacts STM performance and decreases any potential recall benefit of items that are phonologically discriminable. It does not improve performance on items that are confusable. In sum, previous TMS studies have yielded equivocal results regarding the role of the parietal cortex in STM and phonological storage in particular. However, it has been reliably demonstrated that repeated TMS of Broca’s area can produce a selective impairment of the ability to plan speech (Aziz-Zadeh, Cattaneo, Rochat, & Rizzolatti, 2005; Stewart, Walsh, Frith, & Rothwell, 2001). According to the Working Memory model (Baddeley, 2003), a reduction of speech planning ability (or the articulatory control process in the parlance of the Phonological Loop model) should have a similar effect to that of concurrent articulation in healthy participants or the effect of apraxia (Waters et al., 1992): A reduction in the overall ability to recall visually presented lists coupled with a reduction of the PSE. From the perceptual-gestural standpoint a similar outcome could be expected, however, the observed impairment might not be as severe, since Broca’s area is only considered as contributing to the assembly of the motor plan and not as the seat thereof. In contrast, the interference-based approach would struggle with any effects of TMS of a speech planning area on STM. This is because the account does not consider constraints on articulation to have an impact on verbal STM aside from a potential depletion of an attentional resource (see Chapter 2), which hardly applies to TMS of an articulatory area.

Participants undertook the probed order task (Beaman & Jones, 1997; Hughes et al., 2011; Murdock, 1968; Experiment 2 of present thesis), an often-used adapted version of the serial recall task that requires just a single response. This task was chosen to minimize the potential problem that TMS could plausibly interfere with the relatively great overt-motor demands of outputting a series of responses. Any such interference should have little bearing on a single non-speeded response, but might have a negative knock-on effect if reproduction of more than one TBR list item was required. Given that the key prediction of the present experiment related to the impact of TMS on the PSE it was important to establish first that the effect could be observed using the probed order task (which, to my knowledge, has not previously been examined). To this end, a pilot study was conducted and indeed confirmed a significant recall advantage for phonologically dissimilar (vs. similar) lists presented visually, and that this advantage was attenuated under concurrent articulation (see Appendix for a fuller report). According to the Phonological Loop model, TBS of BA 44 should, like concurrent articulation, reduce participants' overall performance on the visual probed order recall task and reduce the PSE.

Experiment 4

Method

Participants

The participants were 18 volunteers from Cardiff University, all screened for contraindications to TMS or MRI. They were all native English speakers, with normal or corrected-to-normal vision and hearing. All participants were right-handed, thus increasing the likelihood of their speech centre being located in the left cerebral hemisphere (cf. Epstein et al., 1996).

Behavioural task

Seven phonologically similar (P, V, B, C, D, G, T) or seven phonologically dissimilar letters (H, Q, L, R, K, X, Y) were presented on a screen in a 72 Times New Roman font in a different random order for each trial. A computer program written in Python was used to present the TBR stimuli and record participants' responses. Each trial started with a blank grey screen, in the middle of which, after 1 s, the seven TBR items were presented. Each item was presented for 250 ms and was followed for 750 ms by a blank screen. Thereafter a response screen appeared, which, in the top part of the screen, featured the question: "Which letter came after letter … ?", with the blank space occupied by a probe letter, that is, one of the letters presented on that trial. Beneath the question seven buttons corresponding to the seven letters presented on the trial were placed in alphabetical order from left to right. Participants were to respond by clicking on the appropriate button, operating the mouse with their left hand. The next trial began 7 s after the appearance of the response screen, independent of whether a response was made. Thus, the overall duration of a single trial was always 15 s. The program recorded whether the clicked letter did indeed follow the probe letter on the given trial and the serial position of the probed letter.

The behavioural task comprised 216 trials. Phonologically dissimilar and similar lists were presented in a quasi-random order, with no more than two trials from the same condition presented in immediate succession. These trials were preceded by a dissimilar and a similar practice trial. The 216 trials were grouped into six 9 min long trial-blocks of 36 trials. Within each block, each serial position that could be probed (positions 2-7) was probed 3 times in a random order for both phonologically similar and dissimilar lists. There were 2 min pauses between each 36 trial block, to not overburden the participants.

Transcranial magnetic stimulation

Cortical stimulation was administered using a 70 mm figure-of-eight induction coil. The coil was oriented at an angle of 45 degrees to the midline for acquisition of resting motor threshold (MT; using the abbreviated distance-adjusted MT procedure described by Stokes et al. 2007) and horizontally for all speech arrest related stimulations (see below). Stimulation was administered using a Magstim Rapid 2 biphasic system.

A combination of structural and functional localization methods was adopted to define Broca’s area. Initially, the region of Broca’s area was defined as the pars opercularis of the left inferior frontal gyrus (IFG) in each participant, based on 1x1x1 isotropic anatomical MRI scans. The closest location to the area on each participant’s scalp was then calculated and co-registered using a magnetic tracking device (miniBIRD 500, Ascension Tech). This anatomical localization of Broca’s area was then used to guide the functional localization of the speech-planning hotspot. Eight additional locations in a 1x1 inch grid were marked and stimulated on the participant’s scalp around the anatomically identified speech planning area. The scalp-cortex distance in that region was used to adjust the MT value. Each participant’s average counting speed was then established by asking them to count briskly from one to ten repeatedly and noting how many cycles the participant went through in 4 seconds. This was repeated twice. The previously marked potential locations of the speech planning hotspot were then stimulated in a random order using 140% of the adjusted MT intensity. At each location, participants were first given 2 single pulses to ensure comfort. If the stimulation intensity was perceived as uncomfortable in a specific location, that location was discarded. If the stimulation intensity was deemed as overall uncomfortable, it was reduced. For four participants, stimulation was reduced to 130% MT and for one participant to 120% MT. Afterwards, 20 stimuli at 5 Hz using the highest comfortable stimulation intensity were administered twice to each marked location. Each time, participants were instructed to count from one to ten. The number of clearly pronounced digits was recorded for each location. Each time after stimulating three locations, two sham stimulations were administered to the centre of the grid whilst the participant was once again counting, to account for practice or fatigue effects. After each repetitive stimulation, participants provided a further rating (1-7) of the amount of facial muscle contractions they felt. Facial muscle contractions are usually uncomfortable and the aim was to stimulate a region that would induce a feeling of “not getting the word out”, which is associated with speech planning impairment, as opposed to reduce speech fluency because of discomfort or uncontrolled muscle contractions (cf. Stewart et al., 2001). Each participant's vocal outputs during the entire speech planning region localization session were recorded.

The region that produced the strongest speech impairment was selected as the speech region. If the same amount of speech arrest could be induced in more than one region, the region with the lowest facial contraction score was selected. If this site differed from the anatomically identified site, it was registered with the magnetic tracking device, photographed, and its distance from prominent locations on the participant's head, such as the left ear was measured. Then the scalp-cortex distance of the new location was calculated and the MT adjusted accordingly. If the site was identical with the anatomical site, its location was simply measured and photographed. The average normalized stereotactic space coordinates (Montreal Neurological Institute [MNI]) of the identified speech planning areas were: X=-63, Y=16, Z=20 (SD: X=1.9, Y=6.1, Z=12.7). This location is depicted in Figure 4.

At the beginning of each TBS session, the previously identified speech planning site was localized and stimulated using continuous TBS (Huang et al. 2005). This protocol includes 3 pulses of stimulation administered at 50 Hz, repeated every 200 ms for 40 seconds at 80% intensity of the adjusted MT (600 pulses in total). Previous research (e.g. Verbruggen et al., 2010) suggests that TBS should reduce cortical excitability of the stimulated area. Thus TBS should inhibit any function that the stimulated area normally fulfils. In a separate Sham TBS session, the same protocol was administered with the coil in a sham orientation, i.e. with the coil pressed perpendicularly against a participant's head so that the direction of the magnetic flux was at a right angle to the surface of the to-be-stimulated area. At the end of the TBS session, after the participants finished the behavioural task, they were instructed to count twice for 4 seconds and then another two times, each time receiving 20 TMS stimuli at 5 Hz. This post-test was undertaken to ensure that the correct location had been stimulated with TBS, which was indeed the case in each of the 18 participants.

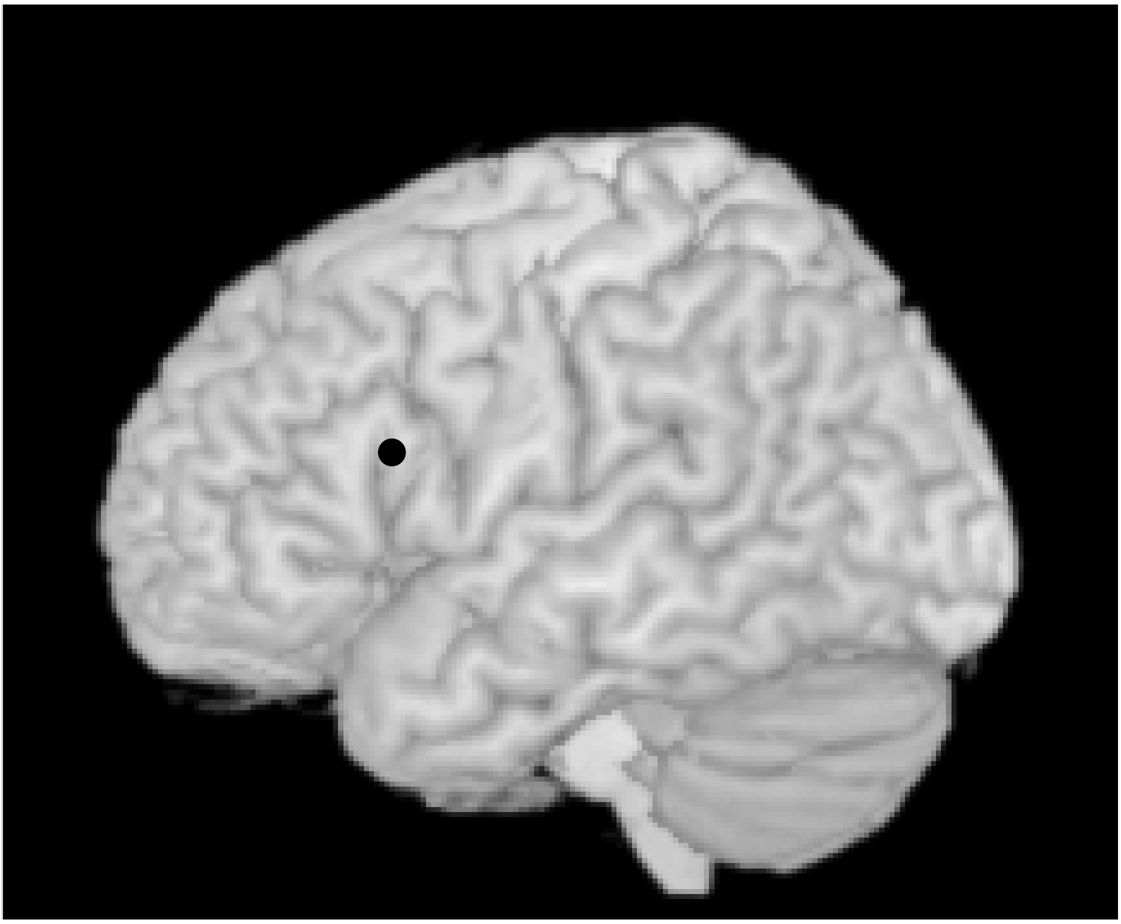


Figure 4. The black dot marks the location of the average speech arrest hotspot on a normalized MNI template brain.

Design

The study was a 2x3x6 within participant design with the independent variables being phonological similarity, interferer, and serial position. They were operationalized by visually presenting seven-item lists (note the first item cannot of course be probed hence the six levels of the serial position factor) comprising either phonologically similar or dissimilar letters, during concurrent articulation or following TBS of the speech planning area or sham TBS. In the Concurrent Articulation condition, participants were asked to whisper the digits 8, 9 and 10 during the presentation of the TBR lists, at a rate of approximately 1 cycle per second (see Jones et al., 2004). In order to ensure compliance with the instructions the whispering was monitored by the experimenter. Participants stopped repeating the digits during the response phase of each trial. In the TBS and Sham conditions the appropriate stimulation was administered at the beginning of a session, before participants commenced the behavioural memory task. The order of these three conditions was fully counterbalanced across participants. The dependent measure was recall at each of the six serial positions that could be probed.

Procedure

The experiment was carried out across several sessions. During the first session (30 min - 1 h) the initial screening took place. Participants were introduced to the experimental paradigm, and were given a brief version of the behavioural task (36 trials, no TMS or concurrent articulation). If their average performance on the task across serial positions was below 25% then they would be excluded from the experiment proper. All 18 participants performed above that level. Experimental sessions involving TBS or concurrent articulation took 1.5 - 2 hours each, with the order of conditions counterbalanced across participants. Figure 5 further illustrates the experimental paradigm.

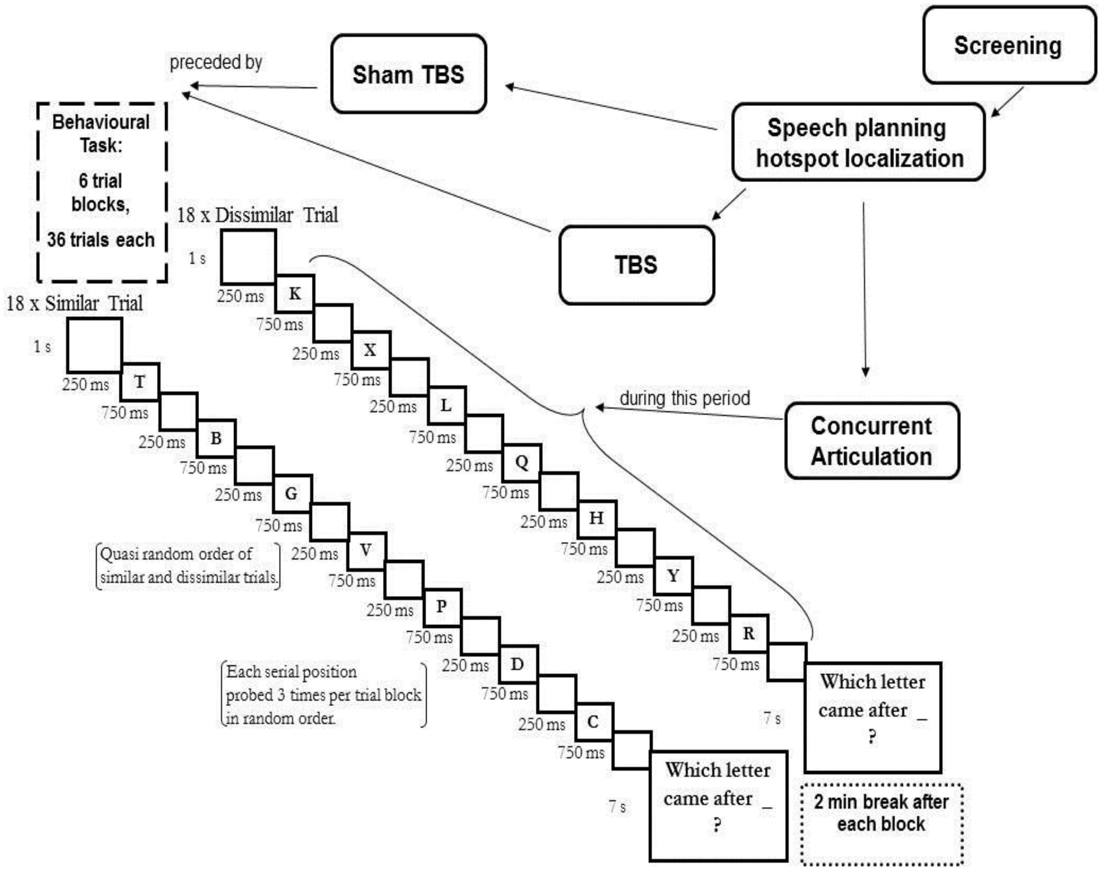


Figure 5. Illustration of the experimental procedure. After an initial screening session, participants' speech arrest hotspots were localized. Then, in an order counterbalanced across participants, they encountered the concurrent articulation, TBS and Sham TBS condition. In the TBS and Sham condition a burst of stimulation was administered before the onset of the behavioural task. In the Concurrent Articulation condition participants repeated an irrelevant utterance during the list presentation phase on each behavioural task trial (cf. Method).

Results

Participants' averaged recall performance in each condition is shown in Figure 6. An analysis of the overall results revealed a main effect of serial position, *F* (5, 85) = 8.35, *MSE* = 0.1, $p < .05$, $\eta_p^2 = .33$. Furthermore, there was a significant effect of interferer, *F* (2, 34) = 85.91, *MSE* = 0.07, $p < .05$, $\eta_p^2 = .84$. Figure 6 reveals that performance in the Concurrent Articulation condition was clearly poorer than in the other conditions. There was also a significant effect of phonological similarity, *F* (1, 17) = 18.08, *MSE* = 0.05, $p < .05$, $\eta_p^2 = .52$, and a significant interaction between similarity and interferer, *F* (2, 34) = 28.35, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .63$. Figure 6 indicates that dissimilar items were recalled better than similar items in both the TBS and the Sham condition, but not in the Concurrent Articulation condition. Additional simple effects comparisons confirmed that there was no significant difference between phonologically similar and dissimilar lists in the Concurrent Articulation condition, *F* (1, 17) = 3.79, *MSE* = 0.04, $p = .07$, $\eta_p^2 = .18$. An additional comparison between the Sham and TBS conditions revealed that there was a significant advantage for dissimilar items, *F* (1, 17) = 31.37, *MSE* = 0.04, $p < .05$, $\eta_p^2 = .65$, which was equally present in both conditions, as there was no significant interaction between phonological similarity and interferer, *F* (1, 17) = 0.42, *MSE* = 0.01, $p = .53$, $\eta_p^2 = .02$. There was also no significant main effect of interferer in this comparison, *F* (1, 17) = 0.32, *MSE* = 0.05, $p = .58$, $\eta_p^2 = .02$, indicating that TBS did not significantly affect overall performance.

At first glance, these results suggest that applying inhibitory TBS to Broca’s area does not affect short-term order recall, nor reduce the PSE. However, there are several reasons for suspecting that this conclusion may be premature. First, the duration of the aftereffect of TBS in different experimental contexts is not fully understood. Whereas initially it was observed that motor evoked potentials can be inhibited after TBS for up to 1 h (Huang et al., 2005) later studies have only observed effects lasting about 30 min (Hubl et al., 2008; Nyffeler et al., 2006, 2008). It is possible therefore that in the present analysis we might have failed to capture an otherwise significant effect of TBS because its effect was relatively short-lived (e.g., the first 30 min). We therefore conducted a second analysis restricted to the first 108 trials, which, with pauses, were completed in 31 min.

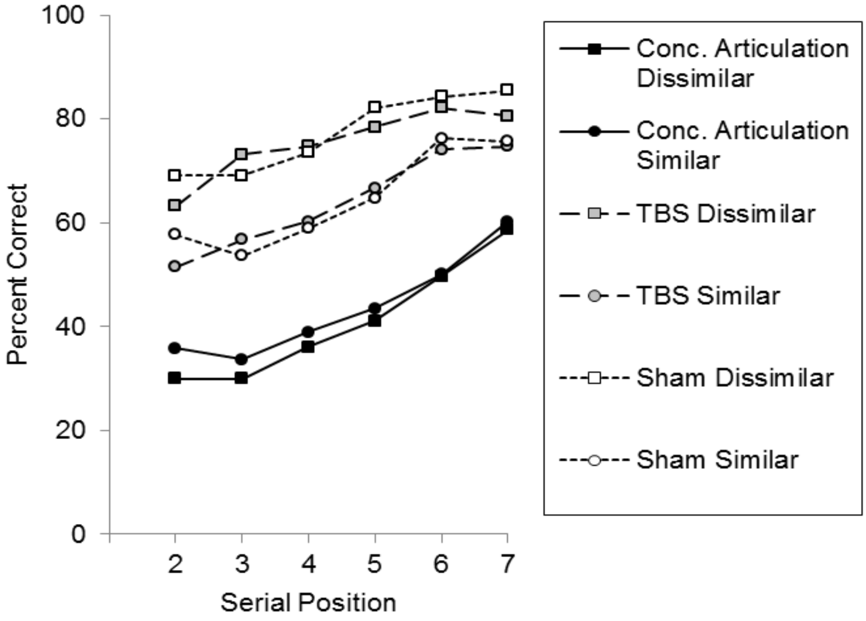
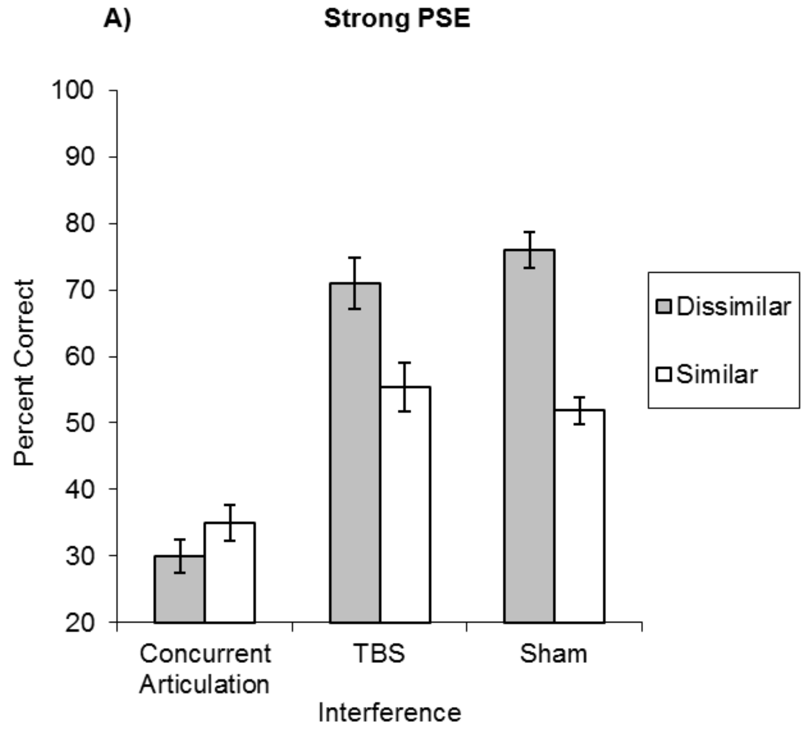
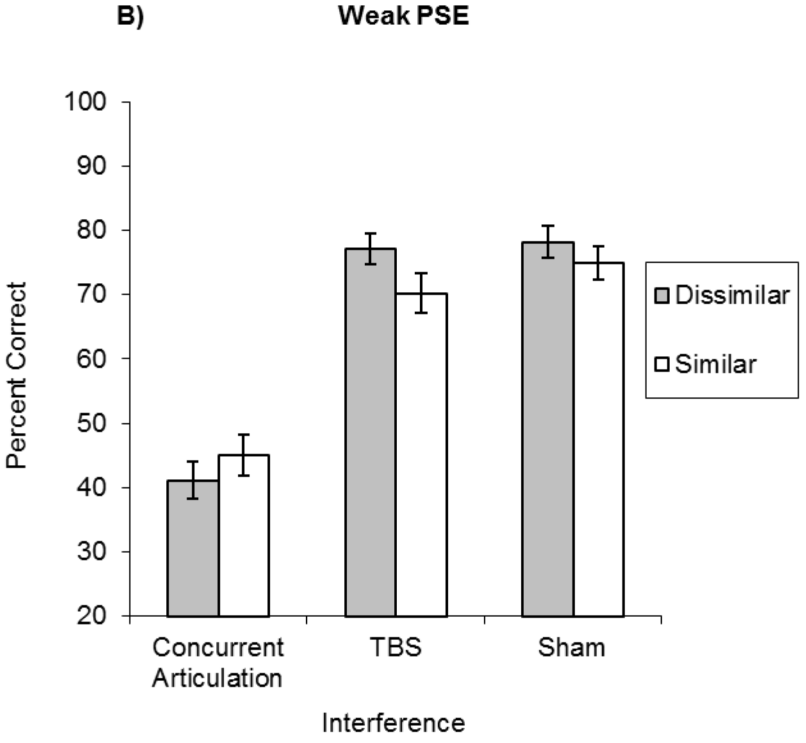


Figure 6. Accuracy of recall in each condition at each serial position.

Another possible reason why our initial analysis may have been relatively insensitive is the well-established variation in the degree to which individuals exhibit the PSE (e.g. Beaman, Neath, & Surprenant, 2007; Della Sala & Logie, 1997; Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996). Indeed, up to 33% of participants do not show a PSE for visual items (Della Sala & Logie, 1997). Note that according to the Phonological Loop model, these participants must be adopting strategies that do not involve the phonological store (Baddeley, 2000, 2003; Baddeley and Larsen, 2007). To address the possibility that TBS did not reduce the PSE according to our initial analysis due to some participants not exhibiting a sufficiently strong PSE to begin with, the participant sample was divided, via a median split, into two groups: Participants who showed a relatively strong PSE and participants who showed a relatively weak (or non-existent) PSE in the baseline (i.e., Sham) condition. In sum, therefore, this second analysis examined the influence of interferer on phonological similarity in participants showing either a strong or weak PSE during the first half of the experiment, during which we can be more confident that the TBS was exerting an effect. Serial position was initially included as a factor, and it did exert a significant main effect, F (5, 80) = 9.36, $MSE = 0.11$, $p < .05$, $\eta_p^2 = .37$, but given that this factor did not interact with any other, performance was collapsed across serial position. Figure 7 depicts performance in all experimental conditions for participants who showed a strong PSE (Figure 7A) and participants who showed a weak (Figure 7B) PSE in the Sham condition.





<u>Figure 7</u>. Accuracy of recall of phonologically similar and dissimilar lists in each interferer condition for participants who showed A) a strong and B) a weak PSE in the Sham condition. Error bars are +/-1 within-participant standard error.

A 2 (group; weak vs. strong PSE in Sham) by 2 (similarity) by 3 (interferer) mixed ANOVA revealed no significant main effect of group, F (1, 16) = 3.24, MSE = 0.11, p = .09, η_p^2 = .16. However, there was a significant interaction between the three factors, F (2, 32) = 12.16, MSE = 0.002, $p < .05$, η_p^2 = .43. Two additional repeated measures ANOVAs comparing the Sham and the TBS levels of the interferer factor at either level of the group variable indicated that for the weak PSE group there was a significant main effect of similarity, F (1, 8) = 10, MSE = 0.002, $p < .05$, η_p^2 = .56, no significant effect of interferer, F (1, 8) = 0.52, MSE = 0.01, p = .49, η_p^2 = .06, and no significant interaction between similarity and interferer, F (1, 8) = 1.71, MSE = 0.002, p = .23, η_p^2 = .17. For the strong PSE group, however, there was a main effect of similarity, F (1, 8) = 16.24, MSE = 0.01, $p < .05$, η_p^2 = .67, but also, critically, a significant interaction between interferer and phonological similarity, F (1, 8) = 5.57, MSE = 0.003, $p < .05$, η_p^2 = .41: The PSE was significantly reduced by TBS in this group. However, against the prediction of the Phonological Loop model, there was no main effect of interferer for this group, F (1, 8) = 0.02, MSE = 0.02, p = .88, η_p^2 = .003. TBS did not affect overall performance but rather impaired recall of dissimilar lists while improving recall of similar lists. [4]

Additional post-hoc comparisons between the Concurrent Articulation and the Sham condition indicate that in both the strong and the weak PSE group, concurrent articulation significantly reduced overall performance, F (1, 8) = 96.5, MSE = 0.01, $p < .05$, η_p^2 = .92, and F (1, 8) = 52.54, MSE = 0.02, $p < .05$, η_p^2 = .87, respectively. There was also a significant interaction between Interferer (with only Articulation vs.

---

[4] This result does not simply reflect a regression to the mean. If the results could be thus explained, then the PSE of participants who show a large PSE in the TBS condition should also be significantly reduced in the Sham condition. This was not observed: when comparing performance of the nine participants who showed a large PSE in the TBS condition there was no significant interaction between interferer (Sham, TBS) and similarity (similar, dissimilar), F (1, 8) = 0.02, MSE = 0.004, p = .88, η_p^2 = .003.

Sham levels included) and Similarity, for both strong and weak PSE groups, F (1, 8) = 92.04, MSE = 0.002, $p < .05$, $\eta_p^2 = .92$, and F (1, 8) = 6.37, MSE = 0.002, $p < .05$, $\eta_p^2 = .44$, respectively. Thus, even in the weak PSE group concurrent articulation reduces the PSE.

General Discussion

The initial analysis of the experimental data did not reveal any significant effect of TBS. Yet, given the novelty of the experimental technique, the size of its effect on the function of various brain areas is not yet fully established. Because at the time when the experiment was conducted, there had not been any previous studies looking at the effect of TBS on Broca's area, difficulties with discerning a potentially mild specific effect were to be anticipated. It is for this reason that an additional post hoc analysis of the experimental data was conducted looking at performance only during the first 31 minutes after the stimulation, that is, when the experimental technique was most likely to have an effect (cf. e.g. Hubl et al., 2008). Furthermore, data from participants displaying a strong PSE and participants displaying a weak PSE in the baseline condition were analysed separately, to better detect any potential effect TBS might have on the PSE specifically. Indeed, with these constrains in place, and despite the fact that there were only 9 participants left in each group (weak vs. strong PSE in Sham), which should have reduced statistical power, it was observed that TBS significantly reduces the PSE in verbal STM, so long as the PSE is relatively strong in the first place (i.e., in the Sham TBS condition). Importantly, the reduction of the PSE due to TBS was not accompanied by impairment in overall performance of the task. While these findings are unquestionably post hoc, their considerable implications for cognitive and neuroscientific research into STM and speech merit a careful albeit reserved consideration.

The Phonological Loop model predicts that selective disruption of an area responsible for speech planning will result in a reduction of visual verbal STM performance accompanied by a reduction of the PSE. Several aspects of the present results are at odds with this prediction. At first glance, it seems problematic for the model that the disruption of Broca's area only had a significant effect for participants showing a strong baseline PSE, as one would expect that disrupting the articulatory control process, and hence limiting access to the phonological store, should reduce the PSE for all participants. The Phonological Loop model (Baddeley, 2000, 2003; Baddeley & Larsen, 2007) can however accommodate this finding: From the perspective of this model, the absence of the PSE is interpreted as an empirical signature of non-reliance on phonological storage for recall. Thus, impeding the articulatory control process of the phonological loop had no effect on the PSE in participants who showed a weak baseline PSE because they were not utilizing the phonological loop, or only utilizing it to a lesser extent, to begin with.

Another aspect of the present results is more difficult to reconcile with the Phonological Loop model: Whereas TBS reduced the PSE, it did not reduce overall performance. While performance on dissimilar lists decreased, performance on similar lists increased. This finding echoes previous research demonstrating that TMS of a region assumed to be the seat of the phonological store reduced the PSE by improving performance on phonologically similar items (Kirschen et al., 2006). However, if, as the Phonological Loop model postulates, the articulatory process acts as a gateway through which visual-verbal material gains access to the store, then inhibiting the articulatory process with TBS should restrict the gateway to the store and thus reduce the ability to upload TBR material into it. With access to the mechanism that is indispensable for short-term maintenance of verbal material restricted, the primary consequence should be a reduction of STM performance. Reduction of the PSE would then arise as a secondary consequence, because a restriction of access to the phonological store would limit any advantage TBR items might gain though being phonologically discriminable. Clearly, this store-based account struggles with the present observation that the reduction of the PSE and the reduction in verbal STM performance can be dissociated.

At this point one might wonder why in previous behavioural (e.g. Baddeley et al., 1984; D. J. Murray, 1968) and neuropsychological (e.g. Waters et al. 1992) paradigms, reduction of the PSE was always accompanied by a reduction in overall STM performance. A possible reason for this discrepancy with previous research is that the effects of TMS are much more subtle and localized than the effects of either concurrent articulation or of an extensive brain lesion. For example, speech production involves activation of a broad network of regions, including the cerebellum (Ackermann, 2008), Broca’s area, and mouth motor areas (e.g., Stewart et al., 2001). Indeed, the previous chapter of this thesis demonstrated that even simple concurrent motor activity of the jaw can reduce performance on STM tasks. It is therefore clear that the utterance of an irrelevant speech sequence is likely to disrupt much more than just the function of a single speech planning region. Similarly, the lesions of patients with speech impairments usually extend beyond a single region and are likely to vary across patients. It is therefore not surprising that these methods show no dissociation between the PSE and overall performance.

Whilst the current results present a challenge for the Phonological Loop model, it is possible that they may be accommodated within the broader framework of the Working Memory model. For example, it might be argued that TBS of the control process caused participants to abandon the use of the phonological loop and rely on other mechanisms. Within the confines of the Working Memory model they could have thus recruited the visuo-spatial sketchpad, a short-term buffer for visuo-spatial information, like graphemic representations of the TBR items (Baddeley and Hitch, 1974, Baddeley, 2000). Participants could have also recruited the episodic buffer, a universal storage device invoked to explain short-term storage that cannot plausibly be accomplished by the phonological store or the sketchpad (Baddeley, 2003). Thus TBS could have induced a tendency to abandon phonological processing of the TBR lists. If the TBR lists are not encoded phonologically then it follows that the impact of phonological similarity on recall would be reduced, but, if the alternative mechanisms are equally efficient for serial recall, overall performance might remain intact. This would explain why TBS reduced the PSE but not overall performance for those participants that showed a strong PSE in the baseline Sham condition. Furthermore, this would explain why participants showing a weak baseline PSE performed just as well as people in the strong baseline PSE group: Although they were using the phonological loop to a lesser extent, they compensated with alternative mechanisms that enabled them to perform just as well. Furthermore, because these mechanisms do not rely on articulatory rehearsal, TBS had no effect on this group.

The problem with the phonological store-abandonment idea, however, is that it raises the question of why there should be a bespoke mechanism for phonological short-term storage if relying on other multi-purpose mechanisms is equally efficient. Furthermore, one has to wonder whether, in patients with a speech planning impairment, their lesion also generally affects the ability to recruit alternative mechanisms for STM, and if it does not, why these patients are unable to compensate for their impairment of the phonological loop, like the participants in the present study. In sum, the results of the present study present substantial challenges to the Working Memory model, and particular its phonological loop construct. It is for this reason that we look to alternative accounts of STM to explore whether they offer a better fit for the present data.

Following the arguments put forward in Chapter 2, it is clear that the interference-based approach to verbal STM, at least as instantiated in the prominent Feature Model (Beaman et al. 2007; Nairne, 1990; Neath, 2000) is also incapable of accounting for the present data. The key point for present purposes is that the Feature Model assumes that concurrent articulation interferes with performance because additional irrelevant item representations are introduced into primary memory and interfere with the representations of the TBR items (Neath, 2000). The articulatory action itself is argued to have little bearing on the memory trace. The present study, however, clearly suggests otherwise: STM performance was modulated with TBS of a speech planning area in the absence of any additional item features being introduced to the memory trace. Note that the argument that was put forward in the previous chapter to account for the impact of non-verbal constraints on articulation from an interference-based perspective, namely that such constraints might deplete an attentional resource, also does not account for the present results: It is difficult to see how TBS administered at the beginning of an experimental session to the speech planning area should deplete attention. Moreover, depletion of an attentional resource could not account for the observation that performance on dissimilar items decreased, while performance on similar items increased. It seems therefore, that interference based models cannot account for the present data any better than the Working Memory model.

Finally, according to the perceptual-gestural account (Hughes et al., 2009, 2011; Jones et al., 2004, 2006), it is not necessary to invoke bespoke short-term buffers to account for serial STM phenomena. For example, it has been demonstrated recently that concurrent articulation reduces the PSE for auditory lists just as much as for visual lists, except for the last few items in the auditory list (Jones et al., 2004, 2006). This recency advantage, however, is not due to obligatory phonological storage of the auditory list, but is based on sensory-acoustic factors governing the sequential perceptual organization of the auditory list that are not in play in the case of visual lists (e.g., Jones et al., 2004; for a dialectic on this issue, see Baddeley & Larsen, 2007; Jones, Hughes, & Macken, 2007). That the key signature of the phonological store—the PSE—is absent regardless of modality when rehearsal is impeded by concurrent articulation obviates the need to posit an additional passive store to which auditory information has preferential access. Instead, the phenomena of verbal STM such as the PSE are, primarily, products of the articulatory planning process itself. In this view, the PSE results from exchanges between articulatorily similar elements, akin to Spoonerisms, during the speech-planning process (Jones et al., 2006; see also Acheson & McDonald, 2009; A. W. Ellis, 1980). Indeed, without a default assumption of phonological storage, the present finding that the PSE was reduced as a result of inhibiting articulatory planning would suggest a clear link between the similarity effect and articulatory processes.

By not invoking a dedicated storage mechanism, the perceptual-gestural account need not be committed to the idea that verbal STM is associated with any single brain area. Instead the STM process is likely to be distributed across areas involved in perception, action planning and production, and the integration of perception and action. Given the nature of the material, for verbal STM the areas that are recruited are generally language-related. This includes BA 44, and BA 40, areas that are considered to be the locations of the articulatory control process and the phonological store, respectively, by the Phonological Loop model. Thus, it is in line with the perceptual-gestural view, as much as it is in line with a store-based view, that damage to BA 44 or BA 40 should impair verbal STM. Since BA 44 is associated with speech planning and BA 40 is associated with the integration of perception into speech-related action (Hickok, 2009), both are important areas for the assembly of a motor-plan for reproducing TBR verbal material. Yet, to assume that selectively impairing either area should be sufficient to disrupt the entire verbal STM process seems too restrictive and localized to be in line with the picture of verbal STM as an emergent property of receptive and productive mechanisms drawn by the perceptual-gestural account (e.g. Jones et al., 2006). If Broca's area is considered merely one of many areas contributing to the verbal STM process and the verbal STM process is not predicated on the function of the area (e.g. because it is the primary pathway for visual-verbal material into a bespoke storage mechanism) then it is conceivable that selective inhibition of the region could have a very selective effect on the STM process. Thus selective impairment of Broca's area could plausibly affect the verbal STM process in a way that would simultaneously improve performance on similar items and reduce performance on dissimilar items. It might be possible, therefore, to account for the results of the present study from the perspective of the perceptual-gestural view although further research will be needed to identify the details of such an account.

It is important to emphasize that the perceptual-gestural account suggests that the recruitment of articulatory mechanisms for short-term recall is task-driven and opportunistic. Thus, it is likely that some participants could opt for less articulation-dependent strategies to maintain the TBR list. Whilst these participants would show a reduced PSE it does not need to be associated with an overall reduction in performance, the pattern observed in the present study. This is because these participants would, according to the perceptual-gestural account, simply choose a different strategy for list maintenance and not, as the Phonological Loop model would claim, abandon a bespoke mechanism required for short-term storage.

The present findings also clearly speak to the debate concerning the function of Broca's area. This region has been implicated in speech production (Amunts et al., 2004), speech perception (Watkins & Paus, 2004), and, more controversially, in STM, with some studies clearly linking the area to STM (e.g., Romero et al., 2006; Waters et al. 1992) but others arguing against this position (Grodzinsky & Santi, 2008). In the present study, repetitive TMS administered to the left pars opercularis of the IFG (or not further than 1.27 cm [0.5 inch] from its centre) induced observable speech arrest in 18 out of 20 tested participants. In 3 participants, speech production was almost entirely abolished. This is clear evidence for the involvement of Broca's area in the speech process. Furthermore, TBS of the area reduced the PSE observed over the first 31 min, which further suggests that it is indeed involved in STM contrary to some previous claims (Grodzinsky & Santi, 2008). Nonetheless, further research is needed in order to clarify the function of Broca's area, and what aspects of the speech production process it might accomplish in order for its inhibition to reduce the PSE selectively.

In summary, the present study is to my knowledge the first to utilize TBS to study STM processes. Although the initial analysis did not reveal any effects of TBS on STM, a more detailed post hoc scrutiny of the data revealed that TBS significantly reduced the PSE in participants showing a strong baseline PSE. This finding suggests that the stimulated Broca’s area is involved in STM. Furthermore, the post hoc analysis showed that the reduction of the PSE was not accompanied by an overall reduction in performance. This indicates that overall performance and the PSE might be dissociable and that the PSE might be associated with articulatory planning as opposed to phonological processes. This pattern is difficult to reconcile with the prominent store-based Phonological Loop model of verbal STM (Baddeley, 1986, 2007) or with item-interference type models (e.g., Nairne, 1990). Accounts of STM that appeal to a primary role for motor planning without invoking an additional passive store (Jones et al., 2004; see also Hickok, 2009) seem better suited to explain the present neurologically-based findings as well as other recent experimental-behavioural results (e.g., Hughes et al., 2009, 2011; Jones et al., 2006; Chapter 2).

It is clear, however, that, given the difficulty of the present experimental paradigm to detect the effect of TBS on STM, firm conclusions are difficult to make and replication of the current results appears necessary. Hereby it is highly recommended to avoid behavioural performance measurements that take longer than 30 minutes, as the inhibiting effects of TBS seem to wear off after this period. Perhaps then, the effects of TBS on Broca’s area could be contrasted with the effects of stimulating additional brain areas such as the Sylvian parietal temporal region, which in its activity resembles a phonological store, whilst being also involved in the integration of auditory perception and vocal tract gestures (Buchsbaum & D’Esposito, 2008, Hickok, 2009). Another cautious expansion upon the current paradigm might be an investigation of the impact of TBS, or TMS in general, on recall of lists presented in different modalities. If the locus of the PSE is primarily the speech-planning process—not an ancillary phonological store—then TBS of Broca’s area should reduce the PSE throughout even for an auditorily-presented list, except at recency as previously observed with concurrent articulation (Jones et al., 2004). In conclusion, tenuous though the link between Broca's area and STM that was established in this study might be considered, its finding does represent an important step in the endeavour to develop a more precise neurological model of verbal STM.

CHAPTER 4: A NEW APPROACH TO MODALITY EFFECTS IN VERBAL SERIAL RECALL: MEETING THE CHALLENGE OF EXPLAINING A VISUAL MID-LIST ADVANTAGE

Abstract

Several accounts of verbal STM postulate a special role for auditorily presented material. Auditorily presented sequences are either thought to have exclusive access to a bespoke acoustic store (store-based view), to be encoded in a code less prone to interference than visual (interference-based view), or to be perceptually organized into objects so that the silence at the end of a sequence can serve as an order-disambiguating boundary (perceptual-gestural view). Each of these theories is confirmed by the robust finding that if item sequences are presented auditorily as opposed to visually, recall of the end of the sequence is particularly strong. The current chapter thematizes challenges to the assumption of a special role for auditory material, in particular the observation of strong end-list performance in the absence of auditory input, and, chiefly, the often observed but rarely commented upon finding of superior performance on visual sequences in mid-list, the inverted modality effect. In this context it is scrutinized to what extent the assumptions of the different verbal STM accounts can be plausibly modified to accommodate the challenges to the proposition of hardwired auditory recall supremacy. The discussion favours the perceptual-gestural account, as it only needs minor adjustments in order to accommodate the evidence against a ubiquitous auditory advantage: It is proposed that sequences with clear boundaries (onset and cessation of auditory input or sequence accompanying gestures) are obligatorily encoded in their entirety, which

leads to a recall advantage at the list edges but also an auditory mid-list recall disadvantage.

Introduction

The studies reported in the preceding empirical chapters sought to adjudicate between different accounts of verbal STM, in particular the store-based Working Memory model, the interference-based Feature Model, and the perceptual-gestural account, by focusing on the differing predictions these accounts make in regards to the role of articulatory processes. The interference-based approach was found wanting when confronted with data showing that non-verbal constraints on articulation, which are unlikely to produce domain-specific interference, can impair verbal STM (Chapter 2). Alternative explanations derived from the Feature Model, such as non-verbal constraints on articulation depleting a central attentional resource fail to explain the effects of TMS of what is commonly considered the speech planning area, Broca's area, on verbal serial STM: Chapter 3 described how a selective lesion of Broca's area induced with TBS reduces the PSE without reducing overall STM performance. It is unclear how attentional resource depletion could account for such an outcome. The store-based Phonological Loop model is also challenged by the findings of Chapter 3. The model associates Broca's area with the articulatory control process, the pathway by means of which visual-verbal information gains access to the bespoke short-term store. The primary consequence of inhibiting Broca's area should therefore have been a reduction of visual-verbal STM performance, which was not observed. In contrast, the perceptual-gestural account was seemingly able to accommodate the results of Chapter 3: While considering Broca's area important for the assembly of an articulatory motor plan to maintain the TBR item sequence, the area is not considered to be the motor plan assembly centre, but rather only a part of a larger neural network dedicated to speech planning. Thus it is plausible from the perceptual-gestural perspective to expect only selective effects like the reduction of PSE in the absence of a general STM performance reduction as a consequence of selective inhibition of Broca’s area.

In the present chapter the focus shifts towards comparing the predictions that various verbal STM accounts make about the impact of presentation modality on memory for item sequences. Many verbal STM theories propose an inherent advantage for recall of auditorily presented sequences. For example, the store-based Working Memory model (Baddeley, 1997) suggests that auditory-verbal items have direct access to a bespoke phonological short-term store. The store-based perspective has also proposed the existence of bespoke *acoustic* stores, like the Precategorical Acoustic Store (Crowder & Morton, 1969), a limited capacity buffer dedicated exclusively to the retention of auditory information. It has also been claimed that a sequence of auditory items is encoded with greater positional resolution (Henson, 1998). The interference-based Feature Model (Nairne, 1990; Neath, 2000) also assumes an inherent advantage for auditorily presented material. According to the Feature Model, TBR items are represented in memory in terms of a mixture of modality-dependent physical features and modality-independent features arising from internal processing of the items. The model argues that auditorily presented items are encoded primarily in terms of modality-dependent features. Thus representations of auditory items in primary memory are less prone to interference from internal processes. The perceptual-gestural account (Hughes et al., 2009, 2011; Jones et al., 2004, 2006), points out—with reference to findings on the perceptual organization of sound into auditory streams (Bregman, 1990)—that an auditory-verbal TBR item sequence tends to be perceived as a temporally-extended object, with the silence at the end of the sequence demarcating the object boundary. Evidence shows that memory is particularly high at the boundary of such objects (Swallow, Zacks, & Abrams, 2009), presumably because these edges constitute violations of expectations (cf. Vachon, Hughes, & Jones, 2012; Zacks, Speer, Swallow, Braver, & Reynolds, 2007), and are hence evolutionarily important (e.g. appearance of a predator). Finally, additional accounts of verbal STM that focussed primarily on explaining the modality differences have argued that auditory TBR material is encoded in an acoustic code so that its maintenance requires less allocated attention than the maintenance of visual material (Penney, 1975, 1989). There have also been claims that auditory items are encoded with better temporal resolution than visual items (Glenberg & Swanson, 1986).

All these theories are confirmed by the robust and frequent observation (see Penney, 1989, for a review) that, particularly at the end of a TBR list (i.e., at ‘recency’), items with an auditory component are remembered better than visual items: the *modality effect*. This effect can be observed if participants are required to read visually presented items out loud (Conrad & Hull, 1968), if visually presented items are read to the participant (Crowder, 1970), or if items are purely auditory (e.g. Jones et al., 2004). At first glance it seems therefore, that there is indeed, as many verbal STM accounts claim, a hardwired benefit to memory if presentation of TBR material is auditory.

There are, however, several stumbling blocks for theories claiming an inherent memory advantage for auditory presentation. One is that a recency advantage is also often obtained with TBR lists that do not contain an acoustic component, like lists of visually presented verbal items that are silently mouthed (that is, gestured without being vocalized), or lip-read lists (Greene & Crowder, 1984). Such findings raise the question of whether the recency advantage might be associated with the way TBR sequences are processed, as opposed to the modality in which they are presented per se.

Another obstacle for the assumption of a hardwired auditory advantage, and the key focus of the present chapter, is the inverted modality effect (henceforth: IME): The auditory advantage in recency is often matched by a visual advantage in pre-recency, the early to middle portion of the serial position curve. This effect has been often overlooked, having been the object of research in only one study (Beaman, 2002). If the effect is genuine, and the auditory recency advantage is indeed often matched by a visual pre-recency advantage, then this calls into question any claim for a dedicated cognitive or neurological system that is hardwired to promote recall of auditory material. It seems that one would either have to assume yet another process or store specifically to account for a visual mid-list advantage, or to seek alternative explanations that might also be capable of explaining a recency advantage in sequences without an auditory component, like silently mouthed or lip-read lists. Before delving into a discussion about how established verbal STM accounts could plausibly be modified to accommodate the IME, however, it is important to establish that the effect is real and robust.

The inverted modality effect (IME)

The observation that visually presented verbal items can sometimes be recalled better than auditory items, particularly in the pre-recency portion of the serial position curve is not a new one, nor is it uncommon. For example, if the temporal and spatial order of TBR items are orthogonal variables, then spatial order is more easily retained with visually-, as opposed to auditorily-presented, lists (Metcalfe, Glavanov, & Murdock, 1981). A recall advantage in pre-recency for visual lists compared to lists with an auditory component has also been demonstrated in serial recall. Thus a small recall advantage at the beginning of the serial position curve has been established for pure visual lists when compared to lists that are presented visually but are vocalized by the participant (N. R. Ellis, 1969; Greene & Crowder, 1984; Conrad & Hull, 1968; Crowder, 1970). These demonstrations of the IME have, however, been dismissed as relatively trivial, with the argument that vocalization of visually presented items impairs rehearsal—akin to concurrent articulation—and has thus a detrimental effect particularly on recall of early-list items (e.g., N. R. Ellis, 1969; Penney, 1975). Yet, such an explanation clearly cannot account for an early-list recall advantage of pure visual lists over visual lists that were vocalized by the experimenter (Crowder, 1970) or lists of exclusively auditorily-presented verbal items (e.g., Maylor, Vousden, & Brown, 1999; Penney & Blackwood, 1989). Inspection of the serial position curves in both these studies (i.e., Maylor et al., 1999; Penney & Blackwood, 1989) reveals that the auditory advantage in recency was offset by a visual advantage in pre-recency. Curiously, in neither study do the authors comment on this IME. Further studies have replicated the IME contrasting pure visual with visual-vocalized and purely auditory lists (Baddeley & Larsen, 2007; Harvey & Beaman, 2007; Jones et al., 2004; Tremblay et al., 2006) but again the effect was not targeted for much if any discussion.

In the only study to date devoted to examining the IME, an explanation for the effect was offered based primarily on data from a non-standard, split list, serial recall setting (Beaman, 2002). In this study participants had to reproduce a list in serial order but start with the last few items. It was speculated that “with visual presentation participants rely upon a visual code that supports recall of early items when recall of those items is delayed” (Beaman, 2002, p. 387), implying that the visual superiority in pre-recency was peculiar to the split-list design. However, although the IME has been observed in several split-list recall studies (Beaman, 2002; Cowan, Saults, & Brown, 2004), as noted, it is also observed in strict forward serial recall (e.g., Harvey & Beaman, 2007; Maylor et al., 1999).

In sum, it appears that the IME is real and robust. Trivial explanations of the effect, like a concurrent articulation-like impact of late-list items on early items in a vocalized list, or visual advantages tied specifically to a split-list experimental design, are too restricted to account for all instances of the effect. Nevertheless, an explanation is clearly needed for how pre-recency performance on visual lists can match recency performance on auditory lists. Turning to the three major verbal STM accounts that have been the subject of this thesis, it seems that the interference-based approach (Nairne, 1990; Neath, 2000) is at a loss for an explanation for the IME. From the interference-based perspective, auditory items are represented in modality-dependent features that are not prone to interference from internal processes, and visual items are represented in modality independent internal features which can be interfered with by processes like inner speech. This explains the traditional modality effect: visual items in recency are interfered with by internal processes, while the memory trace for auditory items remains largely unaffected (Nairne, 1990). Given these premises, it is difficult to conceive, however, how a mid-list visual advantage could arise. One might suggest that visual lists are encoded in terms of modality-dependent interference-resistant visual features up until mid-list. This suggestion is echoed by store-based explanations that have been proposed to account for the IME. Thus it has been argued that participants sometimes opportunistically recruit additional visual codes, to be stored presumably in the visuo-spatial sketchpad (Baddeley and Hitch, 1974; Baddeley, 2000), to assist with visual list maintenance (see Baddeley & Larsen, 2007). This would explain how performance on visual items in pre-recency can match high performance in recency on auditory items, even though according to the store-based perspective these have direct access to an otherwise modality-neutral phonological store (Baddeley, 2003), or a dedicated Precategorical Acoustic Store (Crowder & Morton, 1969).

However, the notion that additional visual codes can be recruited to improve performance begs the question of why such a strategy does not produce a visual-list advantage throughout the list. While it is conceivable that there is an additional store or code that improves auditory performance at recency and an additional visual code that improves pre-recency performance, it is unclear why these codes should have differential effects on different portions of the list. Arguably, such an explanation of the effects in terms of two different mechanisms hardly goes beyond a redescription of the effects. Nevertheless, given that the additional recruitment of visual codes is currently the predominant explanation for the IME, it is necessary to subject it to careful scrutiny, before alternative explanations for the effect can be explored.

Experiment 5

Experiment 5 assesses the suggestion that the IME occurs because of a reliance on visual codes (Baddeley & Larsen, 2007; Beaman, 2002). If such codes can indeed be recruited strategically to improve performance, then it seems reasonable to expect that they will be recruited whenever they are available. This experiment tested this by contrasting forward serial recall of three types of list: Auditory lists, visual lists that were silently-read (‘visual-silent’) and visual lists that had to be vocalized (‘visual-vocalized’). On the visual-code recruitment account, pre-recency should be high in both the visual-silent and the visual-vocalized conditions, as both lists contain the same visual information. Conversely, if performance at pre-recency turns out to be inferior for both auditory and visual-vocalized lists when compared to visual-silent lists, then this would cast doubt on the visual-code recruitment account. Previous comparisons between visual-silent and visual-vocalized (e.g., N. R. Ellis, 1969), and visual-silent and auditory (e.g., Maylor et al., 1999) lists have both revealed a visual pre-recency advantage. Yet, it seems that the shapes of the serial position curves in recency and pre-recency have never been explicitly compared on auditory, visual-silent and visual-vocalized lists in the same forward serial recall experiment. Experiment 5 was designed to redress this shortcoming. Note that regardless of performance in pre-recency, previous studies suggest that both auditory and visual-vocalized lists should produce a stronger recency effect than visual-silent lists (e.g., Conrad & Hull, 1968; Maylor et al., 1999).

Method

Participants

Thirty-three Cardiff University Psychology undergraduates, native English speakers, (28 female) aged 18 – 29 years (Mean: 20.27 years) participated in exchange for course credit. All had normal or corrected to normal vision and hearing.

Materials

The program used in Experiment 5 was similar to the program used in Experiment 1, Chapter 2. The dissimilar letters that were used in Experiment 1 (H, Q, L, R, K, X, Y) were presented auditorily or visually. Letters were presented at the same pace as in that previous experiment, and for visual presentation, the same letter size and font was used. For auditory lists the items were recorded in a female voice with a 16-bit resolution, at a sampling rate of 48 kHz, and compressed digitally to 250 ms using Audacity 1.3.12 (Beta) software (http://audacity.sourceforge.net), without altering acoustic features such as pitch. Each trial started with a blank screen lasting 1 s. The order of the letters was random on each trial.

Design

This experiment had a 3 (presentation type) x 7 (serial position) within-participant design. Participants encountered in a random order three blocks of 30 trials—corresponding to the 3 presentation type conditions—each preceded by two practice trials. In the Auditory condition participants were presented with the to-be-remembered items through headphones, in the Visual-vocalized condition items were presented on the screen and participants were instructed to read them out loud, and in the Visual-silent condition they read the items silently. The dependent variable was the accuracy with which participants recalled each item in its correct serial position. At the end of each trial, seven buttons featuring the to-be-remembered letters appeared on screen. Participants were to click on the letters using the mouse with their dominant hand in the order in which they occurred in the just-presented list. Each ‘button’ could only be clicked once, and all buttons had to be clicked before the program would proceed to the next trial.

Procedure

The experiment was conducted in a sound-attenuated booth. At the beginning of the experiment the specific requirements of each condition were explained to the participants. It was emphasized that their main task was to remember the order of the letters on each trial. Participants were reminded of the requirements of a specific condition at the beginning of each trial block. With their permission, participants were monitored via an audio link to ensure compliance with the instructions relating to the Visual-vocalized condition. The overall experiment was about 45 min long.

Results

Figure 8 shows average recall performance across serial positions in each condition. Initially, it reveals that the standard modality effect was replicated in this experiment: Recall of the final couple of auditory as well as visual-vocalized items exceeded that for the visual-silent items. Furthermore, auditory and visual-vocalized items were recalled poorer than visual-silent items at pre-recency. Thus Figure 8 depicts the IME. These observations are supported by the statistical analysis: A 7 (serial position) by 3 (presentation type) within-participant ANOVA revealed a significant effect of serial position, F (6, 192) = 50.45, MSE = 0.02, $p < .05$, $\eta_p^2 = .61$, as is expected from normal serial position curves. There were also significant differences across presentation type, F (2, 64) = 5.21, MSE = 0.06, $p < .05$, $\eta_p^2 = .14$. However, a significant presentation type and serial position interaction, F (12, 384) = 22.44, MSE = 0.01 $p < .05$, $\eta_p^2 = .41$, indicates that these differences were not consistent across serial position.

Additional pair-wise simple effect comparisons of levels of presentation type show that, while Figure 8 clearly depicts superior performance for auditory items in recency, overall, there is no significant difference between auditory and visual-silent performance, F (1, 32) = 0.47, MSE = 0.06, $p = .5$, $\eta_p^2 = .01$. There is, however, a significant presentation type by serial position interaction when considering these two levels of presentation type, F (6, 192) = 22.9, MSE = 0.01, $p < .05$, $\eta_p^2 = .42$. Similarly, although Figure 8 reveals that recall of visual-vocalized lists is also higher in recency when compared to visual-silent lists, statistical analysis shows that overall there is a marginally significant superiority for visual-silent lists, F (1, 32) = 3.81, MSE = 0.09, $p = .06$, $\eta_p^2 = .1$, and a presentation type and serial position interaction when only these two levels of presentation type are included in the analysis, F (6,

192) = 32.19, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .5$. Finally, auditory performance is superior to visual-vocalized performance, *F* (1, 32) = 17.62, *MSE* = 0.03, $p < .05$, $\eta_p^2 = .36$. This does not, however, apply to all serial positions as is apparent in Figure 8 and evidenced by the serial position and presentation type interaction, *F* (6, 192) = 9.92, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .24$.

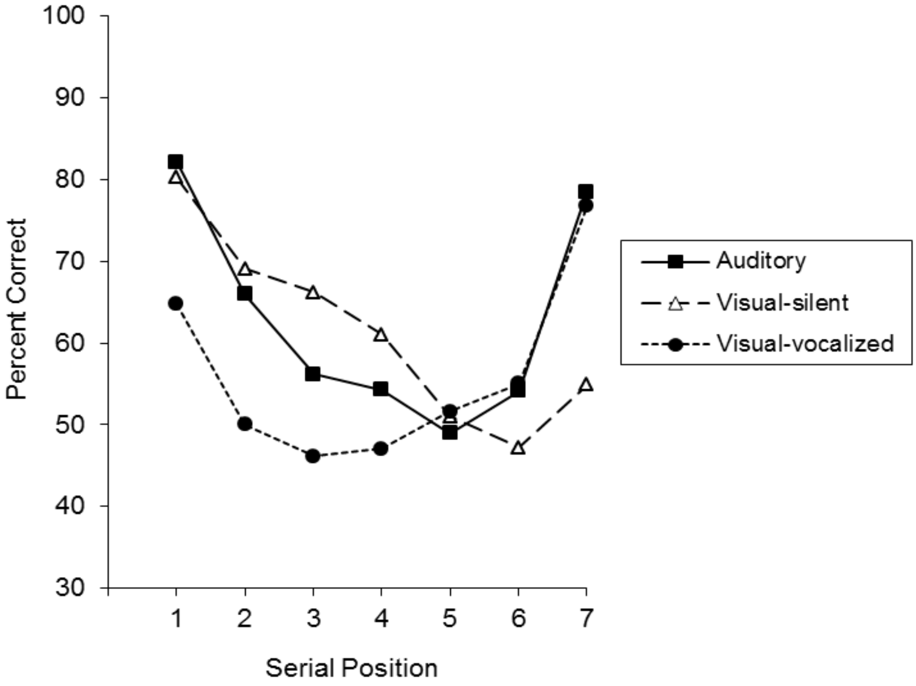


Figure 8. Percentage of correctly recalled items in each serial position, as a function of presentation type.

Discussion

The present results reveal a clear IME: Mid-list performance on visual-silent lists was higher than on visual-vocalized or on auditory lists. Indeed this pre-recency advantage of visual-silent lists matched the high recency performance of auditory and visual-vocalized lists. These data confirm that the split-list report method is certainly not a precondition for the IME (Beaman, 2002), although, it remains possible that having to restructure the list accentuates the effect (Beaman, 2002, Cowan et al., 2004). Moreover, the current results challenge the suggestion that recall for visual-silent lists in pre-recency is superior to auditory due to the recruitment of additional visual codes (Baddeley & Larsen, 2007; Beaman, 2002). In the visual-vocalized condition, visual information was also available, and could therefore presumably have been recruited to increase performance in pre-recency. Instead, pre-recency performance on visual-vocalized lists was in fact poorer than in any other condition, undermining the idea that additional visual codes can be recruited at will to boost performance.

However, defenders of the visual code-recruitment account could counter that visual codes were recruited after all in the visual-vocalized condition: Performance in pre-recency in that condition may not have been comparable to that in the visual-silent condition because the need to articulate each item created a concurrent articulation-like effect (cf. N. R. Ellis, 1969; Penney, 1975). That is, any performance boost that recruitment of visual codes to maintain items in pre-recency offered was superseded by the damaging impact of having to vocalize the TBR list items. The observation that recall of items in recency in the visual-vocalized condition was as high as in the auditory condition could then be explained with the acoustic trace of the last item being preserved in an additional store such as the Precategorical Acoustic Store (Crowder & Morton, 1969). Alternatively, the last acoustic item could have been protected from interference because it was maintained in a special acoustic code (Nairne, 1990; Penney, 1989). Experiment 6 addressed this counterargument.

Experiment 6

It has been shown that a concurrent verbal utterance is less disruptive if it is uttered silently than if it is uttered out loud (Macken & Jones, 1995). Hence, if in the visual-vocalized condition of Experiment 5 the beneficial effect of recruiting visual codes was nullified at pre-recency by having to articulate the items out loud, then this disruptive effect of vocalizing should be diminished if the TBR items are silently mouthed, and the IME should transpire. Experiment 6 therefore added to the conditions of Experiment 5 a condition requiring the silent mouthing of visually-presented lists (Visual-mouthed). If the visual code hypothesis is correct, then pre-recency performance in the visual-mouthed condition should be higher than in the visual-vocalized condition, as the supposedly recruited visual codes would be disrupted less by the mouthing. Note that the present manipulation not only contributes to the debate on the existence of a hardwired memory benefit to auditory information by helping to discern the nature of the IME, but has also a more immediate relevance for the central argument of this chapter: Verbal STM accounts that propose a hardwired recall benefit for lists with an acoustic component (e.g., Crowder & Morton, 1969) predict that in the visual-mouthed condition, performance in recency should be diminished compared to that for visual-vocalized and auditory lists.

Method

Participants

Twenty-two Cardiff University Psychology undergraduates, native English speakers, (19 female) aged 18 - 26 years (Mean: 19.48 years) participated in exchange for course credit. All had normal or corrected to normal vision and hearing.

Materials, Design and Procedure

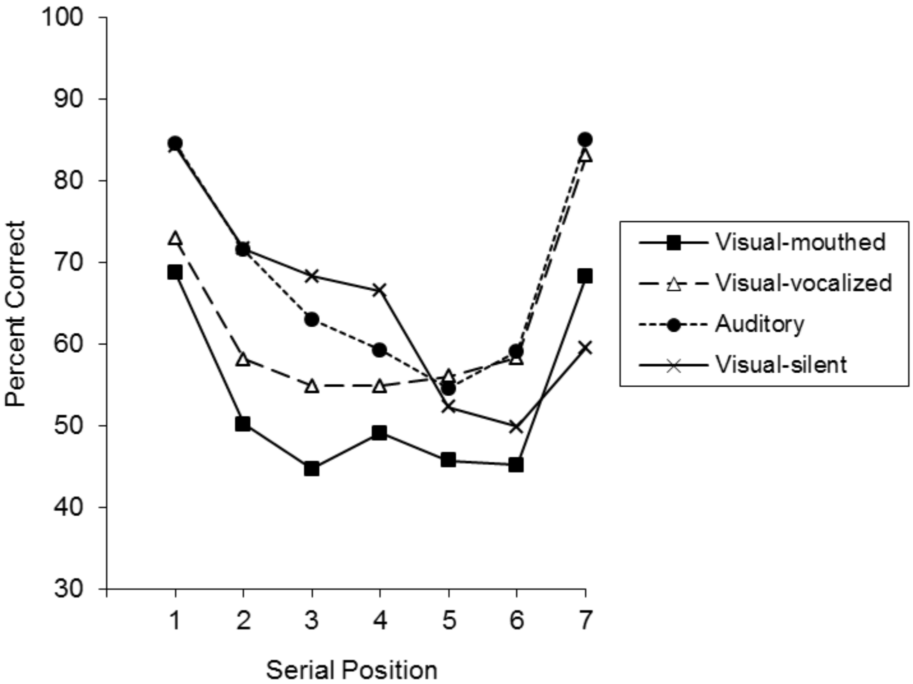
The only difference between the current study and Experiment 5 was that participants were presented with an additional block of 30 visually presented lists, preceded by two practice trials. Participants had to silently mouth each item as it appeared on the screen. The order of the four condition blocks was random. Compliance with instructions relating to the visual-vocalized and visual-mouthed conditions was ensured, respectively, via a sound and a video link. The experiment took about 60 min to complete.

Results

Figure 9 depicts a pattern of performance that is broadly consistent with Experiment 5 in those conditions shared between that experiment and this. Moreover, performance for the newly added visual-mouthed lists appears highly similar to performance on visual-vocalized lists, except for being lower overall. Generally, Figure 9 demonstrates the classical pattern of differences across serial positions which the 4 (presentation type) by 7 (serial position) ANOVA revealed as being significant, F (6, 126) = 41.48, MSE = 0.02, $p < .05$, $\eta_p^2 = .66$. The analysis furthermore showed that auditory performance was highest, followed by visual-silent, then visual-vocalized and then visual-mouthed performance with significant differences between the presentation types, F (3, 63) = 13.05, MSE = 0.05, $p < .05$, $\eta_p^2 = .38$. Yet there also was a significant interaction between presentation type and serial position, F (18, 178) = 11.61, MSE = 0.01, $p < .05$, $\eta_p^2 = .36$, indicating that differences between conditions were not consistent across the serial position curve.

Further pair-wise simple effects comparisons revealed no significant difference between auditory and visual-silent lists, F (1, 21) = 2.62, MSE = 0.04, p = .12, $\eta_p^2 = .11$, but a significant serial position by presentation type interaction, F (6, 126) = 18.99, MSE = 0.01, $p < .05$, $\eta_p^2 = .48$. Similarly, visual-silent performance was overall equal to visual-vocalized performance, F (1, 21) = 0.47, MSE = 0.07, p = .5,

$\eta_p^2 = .02$, but this was again qualified by a significant presentation type by serial position interaction, *F* (6, 126) = 23.23, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .53$. Auditory performance was marginally significantly higher than visual-vocalized performance, *F* (1, 21) = 4.14, *MSE* = 0.06, $p = .055$, $\eta_p^2 = .17$, and the interaction between presentation type and serial position was significant, *F* (6, 126) = 4.52, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .18$. Visual-mouthed performance was significantly lower than visual-vocalized, *F* (1, 21) = 18.18, *MSE* = 0.04, $p < .05$, $\eta_p^2 = .46$, auditory, *F* (1, 21) = 47.28, *MSE* = 0.04, $p < .05$, $\eta_p^2 = .69$, and visual-silent, *F* (1, 21) = 18.13, *MSE* = 0.06, $p < .05$, $\eta_p^2 = .46$, performance. For each of these comparisons there was a significant interaction between presentation type and serial position, *F* (6, 126) = 2.2, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .1$, *F* (6, 126) = 2.75, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .12$ and *F* (6, 126) = 14.39, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .41$, respectively.



<u>Figure 9</u>: Percentage of correctly recalled items in each serial position, as a function of presentation type.

Although performance on visual-mouthed lists appears to be generally inferior, an inspection of Figure 9 suggests that mouthing the list boosts recency to the same extent as listening to items or vocalizing visual items. To corroborate this, a further ANOVA was carried out on the recency portion of the lists. Recency was defined as the difference between performance on the last item in a list and the average performance on the remaining items (cf. Greene & Crowder, 1984). A significant main effect of presentation type was found for these recency scores, F (3, 63) = 21.75, MSE = 0.02, $p < .05$, $\eta_p^2 = .51$. However, additional pairwise comparisons revealed that the only recency score that is significantly different from the others is visual-silent, t (21) = 5.15, $p < .001$, t (21) = 5.7, $p < .001$, and t (21) = 8.12, $p < .001$, when compared to visual-mouthed, visual-vocalized and auditory recency, respectively. In other words, recency was equally strong for auditory, visual-vocalized and visual-mouthed lists, F (2, 42) = 1.43, MSE = 0.02, $p = .25$, $\eta_p^2 = .06$).

Discussion

As was the case in Experiment 5, Experiment 6 showed that pre-recency performance on visual-silent lists was higher than on any other presentation type in this experiment. Thus the IME was once more demonstrated. Further, as expected, performance on auditory and visual-vocalized lists in the present experiment was high in recency producing a U-shaped performance pattern overall. The same pattern of performance was also observed for the visual-mouthed lists, even though these lists were devoid of an auditory component.

There are several ways in which the present findings challenge the notion of a hardwired memory advantage for auditory material as propagated by the majority of verbal STM theories. First, they demonstrate that one of the main problems for that notion, namely the IME—the observation of a pre-recency visual advantage that matches the auditory advantage in recency—is robust. Moreover it cannot be explained away, as has been attempted at least from the store-based perspective (c.f. Baddeley & Larsen, 2007), with participants opportunistically encoding TBR visual items in a visual code to improve pre-recency performance. In both the visual-mouthed and the visual-vocalized condition of the present experiment, additional visual information was available, yet in both conditions performance was worse at pre-recency than in the auditory condition. It could be objected that this is because the performance boost in pre-recency gained from the recruitment of visual codes was superseded by a concurrent articulation-like effect of the respective oral activity (vocalizing or mouthing). However, this objection would carry with it the prediction that there is a greater chance of a pre-recency advantage transpiring with visual-mouthed list. This is because mouthed concurrent articulation is significantly less disruptive than vocalized concurrent articulation (Macken & Jones, 1995). Contrary to this prediction in the present experiment pre-recency performance in the visual-mouthed condition was lower than in the visual-vocalized condition, making the idea that additional visual codes are recruited to improve pre-recency performance even less likely.

The present results also raise the question of how relative recency performance can be equal on lists that consist of auditory input and lists that do not if there are additional stores (Crowder & Morton, 1969), codes (Penney, 1989) or item feature representations (Nairne, 1990) that selectively benefit auditory recency performance. Of course, one could argue that there are yet further codes (Penney, 1989) in which the mouthed list is maintained in an auditory-like fashion. However, having discarded additional visual codes as an explanation for the visual advantage in pre-recency, one has to wonder how useful it is in general to assume additional dedicated mechanisms or storage codes to explain superior performance on one task or modality or the other.

An alternative, more parsimonious, approach would be to appeal to a single construct that might explain the similar shapes of the auditory, vocalized, and mouthed serial position curves. According to the perceptual-gestural account (Hughes et al., 2009, 2011; Jones et al., 2004, 2006) such a construct might be perceptual organization. In the auditory domain, the perceptual system is known to organize sound into discrete auditory objects or streams based on physical characteristics of the different auditory inputs (Bregman, 1990), according to Gestalt principles similar to those operating in the visual domain (Koffka, 1935). For example, if an auditorily presented list is followed by silence, then this creates a clear figure-ground contrast that defines the list boundary. This boundary, in turn, serves as an edge for the temporally-extended acoustic object, and thus acts as an anchor that serves to disambiguate the order of items at the end of the auditory list (see Nicholls & Jones 2002).

While such an instantiation of the perceptual-gestural view might also seem to suggest that high performance in recency is reserved for lists with an acoustic component, it is easy to see how salient list-edges might also be present outside of the acoustic domain. Indeed, this idea is at the core of the remaining experiments reported in this chapter. For example, if each list item has to be accompanied (due to instructions) by an articulatory gesture—as in some of the conditions of Experiments 5 and 6—this list-processing-relevant activity may in effect transform the succession of TBR items into a discrete temporally-extended object, the beginning and end of which is defined by the onset and cessation of that activity (see also Macken & Jones, 1995, for a similar argument). The figure-ground contrast between the presence and absence of the salient activity can arguably result in as strong an anchor for item order as the silence at the end of an acoustic list. Moreover, if the processing of the visual-silent list is particularly engaging such as, for example, when the letter list is not presented as a sequence of alpha numeric characters, but instead a sequence of visually-presented silent gestures that the participant has to reconstruct into letter representations through lip-reading, then again the beginning and end of the (participant's) list-processing activity will create clear boundaries, hence accounting for the strong "auditory-like" recency found for lip-read lists (Campbell & Dodd, 1980; Greene & Crowder, 1984; de Gelder & Vroomen, 1992).

The perceptual-gestural account seems to offer a promising framework, therefore, for developing a unified explanation for high recency performance for auditory, visual-vocalized, visual-mouthed, and even lip-read lists, without having to invoke additional stores or codes. Instead, the simple assumption that the list processing activity/cessation creates a figure-ground contrast that acts as an anchor for item order recall seems sufficient to explain high recency performance for a range of conditions. However, it is still unclear why pre-recency performance on visual-silent lists should be higher than on auditory lists and indeed why this pre-recency performance should match the boost in performance that the auditory list receives from the order disambiguating silence at the end of the list. That is, how does the visual superiority in pre-recency (the IME) arise given that there is no apparent anchor in the middle of the visual list?

One possibility is that visual pre-recency performance is superior because of a greater flexibility in subjectively restructuring a visual compared to an auditory TBR list. There is some indication that the IME is more apparent when the TBR list has to be restructured, such as when lists have to be recalled in a split-list fashion, recalling the later list items first (Beaman, 2002; Cowan et al., 2004). Subjective restructuring also has an important role in the perceptual-gestural account. In typical serial recall tasks, such as the ones described in this chapter, the main challenge lies in the maintenance of order of a sequence of verbal items that are unconstrained by grammar or syntax. According to the perceptual-gestural view, there are several ways to meet this challenge. For example it is possible, utilizing motor planning skills, to subjectively impose a prosodic rhythm onto the sequence (Hughes et al., 2009), thus grouping the TBR sequence into smaller chunks, which has a clear benefit on performance (see e.g. Frankish, 1989). Yet in order to use motor planning skills to upload the TBR list onto a subjective prosodic rhythm that perhaps would even impose grouping constraints, the list needs to be unconstrained perceptually, like visual-silent lists. Any perceived structure might be at odds with the subjective motor planning strategy. An example of this would be lists that are perceived automatically as temporally-extended objects, like auditory lists. If the TBR list is encoded as an object then the objecthood itself generates order cues for list items, particularly at the object boundaries, that is, the beginning and end of the list. Yet, since the very idea of an “object” denotes a cohesive and bound entity that is rigid and immutable (Spelke, 1990), it seems probable that the list-object will be resilient to subjective motor planning-based strategies of imposing order. This could explain the traditional and the inverted modality effects: If an auditory TBR list is encoded as an object, memory will be particularly high at the list boundaries, but memory at mid-list will be higher for lists that, due to their perceptually unconstrained nature, can be easily restructured to fit a subjective motor plan. Experiment 7 addressed this hypothesis by testing whether visual lists lend themselves better to restructuring than auditory lists.

Experiment 7

If the visual pre-recency advantage is predicated on the ability to subjectively restructure the TBR list so as to group it or make it fit an ideal subjective articulatory plan, then a requirement to restructure a TBR list for recall should have less impact on visual than on auditory lists. If the presumed rigidity of the auditory perceptual object obstructs fitting it into a subjective prosodic rhythm, then arguably, that same rigidity should make it yet more difficult to subjectively restructure the TBR list. In order to test this, participants were presented with auditory and visual lists, which either needed to be recalled in a forwards serial manner, or in a forwards serial manner but with the items in odd serial positions being recalled first, as a group, followed by the items in the even serial positions, as another group. The requirement to restructure the TBR list was expected to have a greater negative impact on recall of auditory lists.

Method

Participants

Seventy-two Cardiff University students, native English speakers, (12 male) aged 18 – 32 years (Mean: 20.08 years) participated in exchange for course credit. All had normal or corrected to normal vision and hearing. Thirty participants were assigned to the Restructure group (aged 19-32 years, Mean: 20.96 years, 8 male). Forty-two participants were in the Forwards recall group (aged 19-24 years, Mean: 19.21 years, 4 male).

Materials

The same visual and auditory letter stimuli as in the visual-silent and auditory conditions in Experiment 5 and 6 were used in this experiment.

Design

This experiment had a mixed design with the two within-participant variables being modality (auditory vs. visual) and serial position at presentation, and the between-participants variable being recall-type (forwards vs. restructured). Recall-type was manipulated between participants in order to avoid a potential carry-over effect from the Restructured condition. All participants received a block of 30 auditory lists and a block of 30 visual lists, each preceded by two practice trials. Participants in the Forwards group had to reproduce the TBR items in their presented order. In the Restructured group, participants had to recall the order of all the items in the odd serial positions first, and then recall the order of the items in the even serial positions. For example, given the TBR list “H, K, L, Q, R, X, Y”, participants would have to reproduce: “H, L, R, Y, K, Q, X”. The dependent variable was the accuracy with which the correct item was placed in each serial position. As in previous experiments, participants clicked on buttons on the screen that corresponded to the list letters, in the order required by the condition. Note that this procedure made it impossible to circumvent the instruction to restructure in the Restructured condition (e.g., with written recall, a participant could write out the list in forward serial order, placing the first four items in odd output positions, and then ‘fill in the gaps’ with the last three items).

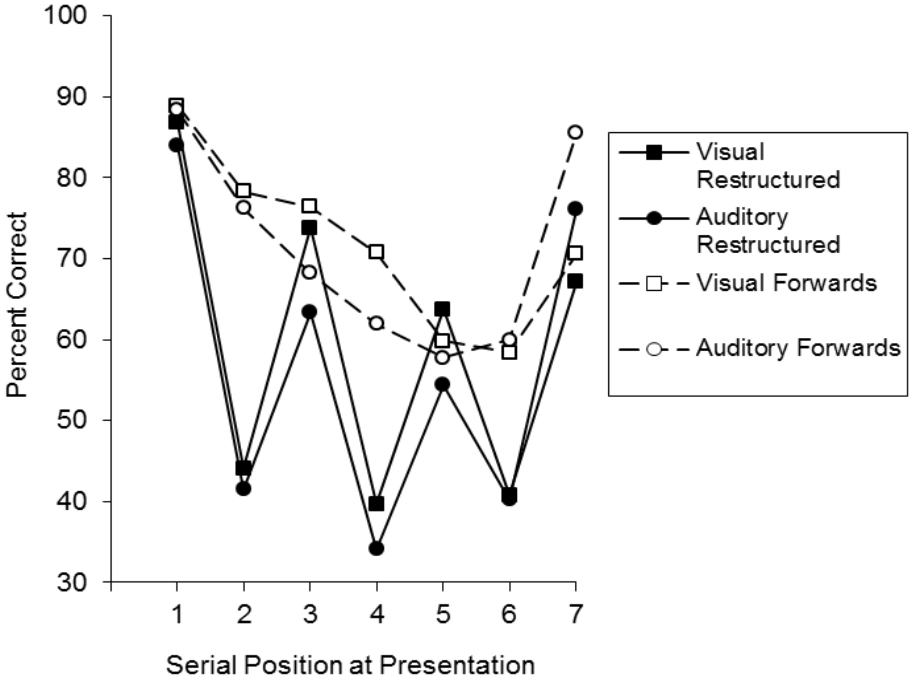
Procedure

The experiment was conducted in a sound-attenuated booth. At the beginning of the experiment, the specific requirements of the condition to which they had been assigned were explained to the participants. In the Forwards group participants were instructed to simply recall the order of the TBR list. In the Restructured group participants were asked to “mentally unzip” the TBR list and then reproduce the order of the odd items first. It was suggested that participants imagine the odd and even positioned TBR items as being in two different groups, the odd and the even group. The overall experiment took about 30 min.

Results

The average performance on each serial position in each condition was calculated and is depicted in Figure 10. Serial position here is defined as the serial position in which an item was presented, and not in which an item was recalled. Note that in the Restructured group these are indeed two different concepts, as, for example, items presented in the second, fourth and sixth serial position were actually output in the fifth, sixth and seventh serial positions, respectively.

A 2 (recall-type) by 2 (modality) by 7 (serial position at presentation) ANOVA revealed, as usual, a significant effect of serial position, *F* (6, 420) = 182.54, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .72$. More importantly, there was no significant difference between the auditory and the visual modalities, *F* (1, 70) = 1.33, *MSE* = 0.07, $p = .25$, $\eta_p^2 = .02$, but a significant interaction of serial position and modality, *F* (6, 420) = 17.85, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .20$, demonstrating that the auditory superiority in recency was once more matched by a visual mid-list superiority. However, there was no significant recall-type by modality interaction, nor recall-type by modality by serial position interaction, *F* (1, 70) = 0.53, *MSE* = 0.07, $p = .46$, $\eta_p^2 = .01$, and *F* (6, 420) = 1.09, *MSE* = 0.01, $p = .37$, $\eta_p^2 = .02$, respectively. Thus, contrary to the hypothesis, having to restructure the list did not impair recall of auditory lists any more than visual. Overall performance was, however, significantly impaired by the requirement to restructure the list, *F* (1, 70) = 20, *MSE* = 0.23, $p < .05$, $\eta_p^2 = .22$. Figure 10 indicates that restructuring the list particularly impaired performance on the second, fourth and sixth serial positions, that is, the items that were to be recalled last. This is supported by the significant recall-type by serial position interaction, *F* (6, 420) = 48.95, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .41$. An additional comparison of performance across each condition in the second, fourth and sixth serial position reveals a significant effect of recall-type, *F* (1, 70) = 71.41, *MSE* = 0.11, $p < .05$, $\eta_p^2 = .51$. In contrast, when performance on the first, third, fifth and seventh serial positions was considered, there was no effect of recall-type, *F* (1, 70) = 1.06, *MSE* = 0.14, $p = .31$, $\eta_p^2 = .02$, nor a recall-type by serial position interaction, *F* (3, 210) = 2.05, *MSE* = 0.01, $p = .11$, $\eta_p^2 = .03$.



<u>Figure 10</u>: Average recall performance in the Forwards and Restructured conditions as a function of serial position.

Discussion

The present experiment once more revealed an IME: Strong auditory performance in recency matched by superior visual pre-recency performance.

Whereas restructuring the list impaired overall performance, this was the case independently of modality. These results cast doubt on the hypothesis that visual performance in pre-recency is improved by the ability to flexibly fit the visual TBR list onto an articulatory plan, or subjectively group it. The notion that an auditory list is processed as a temporally-extended perceptual object suggests a certain internal rigidity of the mental representation of the auditory list (Spelke, 1990). However, the current experiment suggests that an auditory list is no more internally cohesive than a visual list. Thus, even if fitting a TBR sequence onto a subjective prosodic rhythm were important for maintaining the order of the sequence, as postulated by the perceptual-gestural account (Hughes et al., 2009, 2011), the current experiment does not offer evidence that this fitting process should be easier for visual silent-list than for auditorily lists. The notion of greater flexibility of restructuring a visual list does not therefore seem to be an adequate explanation for the IME.

Another motor planning skill that, according to the perceptual-gestural account (Hughes et al., 2009), can be drawn upon when assembling a list of unconstrained verbal items into a recallable sequence is co-articulation. In a list that is unconstrained by grammar or semantics it might be possible to impose order by adjusting the articulation of one item to allow a smoother articulatory transition to the next (Hughes et al., 2009; A. Murray & Jones, 2002; Woodward, Macken, & Jones, 2008). Visual-silent lists may afford such a strategy more than auditory, visual-vocalized or visual-mouthed lists. Hearing, vocalizing or even mouthing a TBR item might activate a certain schema of how to articulate the item. Deviating from that schema strategically, in order to shape the item-end through co-articulation as a cue for the next item to promote serial recall might thus be more difficult. Indeed, in the auditory domain there is much evidence for a tendency to articulatorily imitate even very subtle properties of a verbal utterance (Goldinger, 1998; Pickering & Garrod, 2007; Street & Cappella, 1989). In contrast, with visual-silent lists participants are free to modify their articulation of the TBR items in order to facilitate co-articulatory transitions between items. Thus, particularly in parts of the serial position curve where item memory is not anchored at perceptual object edges, i.e., in mid-list, visual-silent list performance could benefit. In order to establish whether articulatory factors underpin the visual pre-recency advantage Experiment 8 examines whether it is reduced when co-articulatory planning of the list is impeded through concurrent articulation.

Experiment 8

If articulatory processes play a role in the IME, then the effect should be reduced under concurrent articulation. In the present experiment performance on auditory and visual TBR lists with and without concurrent articulation was compared. It should be noted that a similar experiment was conducted previously (Jones et al., 2004, Experiment 2). The results of that study indicated the presence of an IME in the baseline condition which seemed to be reduced under concurrent articulation, suggesting that articulatory fluency contributes to the IME. Yet the interplay between concurrent articulation and performance in pre-recency and recency across modalities was not the subject of that study and hence was not commented upon or submitted to statistical analysis, a shortcoming that was addressed in the present experiment.

Method

Participants

Twenty-two Cardiff University students, native English speakers, (5 male) aged 18 – 32 years (Mean: 21.05 years) participated in exchange for course credit. All had normal or corrected to normal vision and hearing.

Materials

The same visual and auditory letter stimuli as in the visual-silent and auditory conditions in Experiment 5 and 6 were used in this experiment.

Design

This experiment had a within participant design with the independent variables being modality (auditory vs. visual), serial position and concurrent articulation (present vs. absent). Trials were blocked into four blocks, one for each combination of the modality and the concurrent articulation variables. Each block contained 24 trials, preceded by two practice trials. On concurrent articulation trials participants were required to whisper the digits 8, 9 and 10 at a rate of 3 items per second, during list presentation. The dependent variable was the accuracy with which the correct item was placed in each serial position.

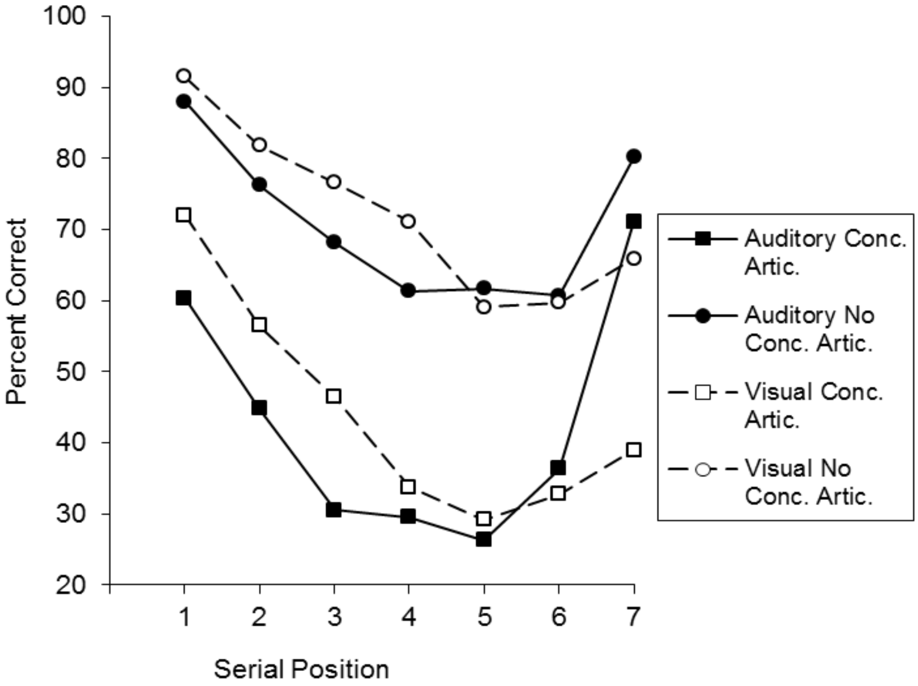
Procedure

The experiment was conducted in a sound attenuated booth. At the beginning of the experiment, the specific requirements of each condition were explained to the participants. Compliance with the concurrent articulation instruction was monitored via an audio link. The overall experiment took about 40 min to complete.

Results

Average performance in the presence and absence of concurrent articulation was calculated and is depicted in Figure 11. Inspection of this figure suggests that the IME was replicated and, more central to the aim of the present experiment, is still evident under concurrent articulation. A 2 (modality) by 2 (concurrent articulation) by 7 (serial position) ANOVA revealed the typical significant effect of serial position, F (6, 126) = 63.07, MSE = 0.02, $p < .05$, $\eta_p^2 = .75$. There was also as significant reduction in performance due to concurrent articulation, F (1, 21) = 119.58, MSE = 0.1, $p < .05$, $\eta_p^2 = .85$, but concurrent articulation did not affect one modality any more than the other, F (1, 21) = 0.004, MSE = 0.03, $p = .95$, $\eta_p^2 = .0002$. There was no significant effect of modality, F (1, 21) = 0.7, MSE = 0.05, $p = .41$, $\eta_p^2 = .03$, but a significant interaction between modality and serial position was observed, F (6, 126) = 22.94, MSE = 0.01, $p < .05$, $\eta_p^2 = .52$, and Figure 11 confirms that both in the presence and in the absence of concurrent articulation high auditory performance in pre-recency was matched by high visual performance in early to mid-list positions.

A significant interaction between modality, serial position and concurrent articulation was also observed, F (6, 126) = 3.61, MSE = 0.01, $p < .05$, $\eta_p^2 = .15$. Inspection of Figure 11 suggests that this might have primarily to do with recall of the last auditory item being relatively unaffected by concurrent articulation. Indeed, an additional comparison of recency of auditory lists with and without concurrent articulation, with recency being defined in the same way as it was in Experiment 6 (also cf. Greene & Crowder, 1984), reveals that auditory recency is significantly more pronounced in the presence of concurrent articulation, t (21) = 5.57, $p < .001$. However, since there is no significant overall effect of modality, the accentuation of auditory recency under concurrent articulation appears to have been matched by an accentuation of the pre-recency recall advantage for visual lists.



<u>Figure 11</u>: Average performance on visually and auditorily presented verbal lists with and without concurrent articulation.

Discussion

Once more, an IME was observed. Moreover, the size of the effect matched the size of the standard modality effect in recency, even when articulation was suppressed. The observation of the IME under concurrent articulation makes it unlikely that the visual pre-recency advantage is dependent upon the ability to rehearse visual TBR items more effectively than auditory.

According to the perceptual-gestural account, in order to maintain the order of a sequence of verbal items that is unconstrained by semantics, grammar or syntax, articulatory motor planning skills can be utilized. These are used to either fit the items onto a subjective prosodic rhythm or to modify the articulation of items to facilitate co-articulatory transitions between them (Hughes et al., 2009). Experiment 7 and the current experiment have addressed several explanations as to how utilizing motor planning skills might be more efficient given visual-silent lists and how this might account for the IME. Yet, the present experiment in particular makes it seem unlikely that the IME arises because of a greater ease in utilizing articulatory motor planning processes in the visual domain. If this were the case, then the IME should be reduced under concurrent articulation. Instead, it was exacerbated. On the one hand, this clearly excludes gestural processes as the origin for the effect. On the other hand, looking towards perceptual processes for an explanation of the IME also presents a conundrum: If auditory recency performance is improved by the order-disambiguating boundary that is provided by the silence at the end of a TBR list, how can performance on visual-silent lists in pre-recency match this improvement without an obvious mid-list boundary?

One possible difference between the various presentation types that might be the key to this paradox is the extent to which they constrain the obligatory encoding of the list in its entirety. This pertains to the idea of objects as cohesive entities (Spelke, 1990): If an item list is encoded, that is processed beyond a mere sensory stage by the neurocognitive system, as a temporally-extended object, as seems to be the case with auditory visual-vocalized and visual-mouthed lists, then it stands to reason that all the elements of the object will also be encoded. On the other hand, if the TBR item list is perceived as a collection of discrete events, like with visual-silent lists, then there is no reason to assume that all the elements of the list will be encoded. Furthermore, if encoding only a subset of the presented TBR items would somehow benefit recall, then a presentation type that affords the freedom to selectively avoid encoding or ‘ignore’ some of the items, i.e. visual-silent presentation, should have a recall advantage over a presentation type in which each TBR item is encoded obligatorily. Naturally, this advantage would only extend to the items that are not being ignored, that is items whose processing is not wilfully prevented beyond a perceptual stage.

In order to demonstrate that the first of these assumptions, namely that TBR lists that are represented as objects are indeed obligatorily processed in their entirety, a possible manipulation is to append an irrelevant, to-be ignored item, a suffix, at the end of a TBR list. If even a to-be-ignored suffix is obligatorily encoded as part of the TBR list, then all the TBR list items are likely to be subject to obligatory encoding, too. In point of fact, the detrimental effect of a suffix has been repeatedly demonstrated with auditory lists (Greene & Crowder, 1984; Jones et al., 2004, 2006; Nicholls & Jones, 2002) as well as vocalized and mouthed lists (Greene & Crowder, 1984). Specifically, it was found that if a suffix that shares perceptual properties with the TBR list items is appended at the end of the list then performance in recency is reduced. Moreover, it was shown that if the suffix is “captured”, meaning that it is made to be perceived as part of an additional to-be-ignored stream of acoustic events, then the negative impact of the suffix on the TBR list diminishes (Nicholls & Jones, 2002). Thus, it has been suggested that the suffix effect occurs because the suffix is obligatorily encoded as part of the TBR list which displaces the perceptual boundary at which the last TBR list item would otherwise be anchored. If, however, even a to-be-ignored suffix is obligatorily encoded as part of the TBR list then it follows that every item in the TBR list will be obligatorily encoded, too. In contrast, previous research using visual-silent lists has constantly failed to find a suffix effect, even if the suffix shared (visual) perceptual properties with the other items of the TBR list, and trivial explanations like gaze aversion from the to-be-ignored stimulus have been ruled out (Greene, 1987). Thus, even if the visual suffix is made to seem as part of the

TBR list, there is little difficulty in ignoring it. This suggests that other visual list items might also be strategically ignored.

While past research on the suffix effect makes it seem likely that in lists perceived as temporally-extended objects every item is obligatorily encoded, Experiment 9 offers a novel demonstration of this. Although originally designed to investigate the effects of a suffix on the IME, the data obtained in Experiment 9 and the way in which it was analyzed reveal a clear co-dependence between recall of the early and late items in auditory lists and a clear lack of such a dependence for visual-silent lists. Because these data demonstrate in an unprecedented way how temporally-extended object lists are encoded in their entirety and lists not encoded as objects are not, the experiment was included here to bolster the argument that the IME arises because participants can choose to ignore a part of a visually-presented list.

Experiment 9

The present experiment illustrates in a novel way that temporally-extended object lists are obligatorily processed in their entirety. Initially, performance on auditory and visual lists was compared with and without the addition of an auditory suffix. The suffix shared perceptual properties like voice and presentation duration with the rest of the TBR list items in the auditory condition; however, ignoring the suffix was an explicit experimental requirement. If participants would not be able to ignore the suffix in the auditory condition then, assumingly, they should also not be able to ignore any of the TBR list items either.

Previous studies conclude that the auditory suffix is incorporated into the TBR list because the impact of the suffix is moderated by the efficiency with which the suffix is captured into a different stream (Nicholls & Jones, 2002). The present study extended on this conclusion by supplementing the traditional analysis of variance of the serial memory performance data with a cluster analysis. This measurement, not previously used in serial recall studies, would identify any co-dependencies between recall responses at different serial positions of the auditory and the visual lists: If correct or incorrect responding at one serial position determines strongly the likelihood of correct or incorrect responding at another, the cluster analysis would show those serial positions as clustering closely together. The notion of auditory objecthood suggests that there should be some co-dependence between recall of early and late list items, the edges of the auditory object. Finding such co-dependencies between early and late list items would already suggest that there is some holistic processing of an auditory list. Moreover, obligatory incorporation of the suffix into the auditory list would mean that the suffix should displace the auditory object boundary in recency, and itself become the last item. Thus any co-dependence between early and late list items should be transformed into a co-dependence between the early items and the suffix. Empirically, the co-dependence between the early and late list TBR items should diminish.

In contrast, on lists not presumed to be objects, i.e. visual-silent lists, a co-dependence between early and late list items is not expected. Instead, in line with the idea that participants would ignore a portion of the visual list, it is expected that co-dependence between recalled items should diminish roughly as a function of their distance to each other. This is because with two items presented far apart, there is a greater chance for one of the items to end up in the ignored item category, and hence outside of the recall cluster of the other item.

Method

Participants

Twenty-seven Cardiff University Psychology undergraduates, native English speakers, (26 female) aged 18 - 20 (Mean: 19.04) participated in exchange for course credit.

Materials

The same letters as in the preceding experiments were used. Their visual presentation remained the same. However, for the auditory presentation the items were re-recorded in a male voice, using otherwise the same source stimuli as before. In addition, the word “go” was recorded in the same voice, to be used as an auditory suffix. Its presentation duration matched that of the other auditory items (250 ms). It was presented on suffix trials 750 ms after the last letter, in keeping with the inter-stimulus intervals between the TBR items. Its offset was immediately followed by the response screen. To equate the overall trial duration, on trials without a suffix there was a 1 s pause between the offset of the last list item and the appearance of the response screen.

Design

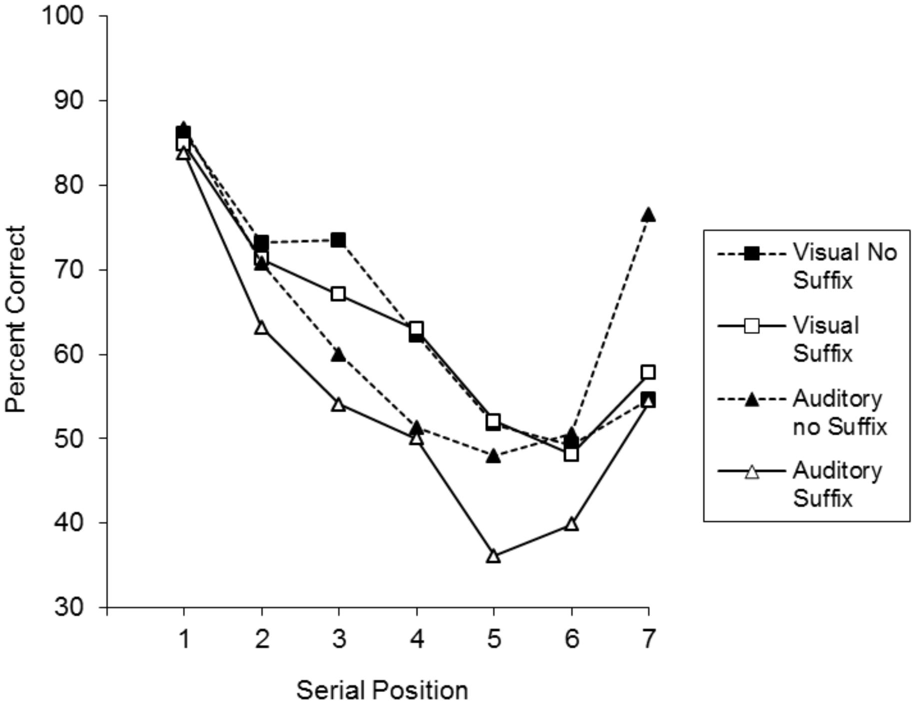
This experiment was a 2 (modality) by 2 (suffix present or absent) by 7 (serial position) within-participant design. There were two randomly ordered 40 trial-blocks, one for each modality. In each block, on half the trials the to-be remembered items were followed by a suffix. Suffix and non-suffix trials were presented in a quasi-random order with the constraint that no more than two trials of the same type could follow each other.

Procedure

The participants were placed in a sound attenuated booth. They were warned that on some trials the to-be-remembered list would be followed by the word “go” which they should try to ignore and not recall. Reminders of these instructions were presented at the beginning of each modality block. The experiment lasted about 40 min.

Results

Figure 12 once more reveals the IME. Overall the average performance on visual items was significantly higher than auditory performance, $F\ (1, 26) = 9.43$, $MSE = 0.05$, $p < .05$, $\eta_p^2 = .27$. The significant interaction between modality and suffix, $F\ (1, 26) = 17.56$, $MSE = 0.02$, $p < .05$, $\eta_p^2 = .4$, suggests, however, that this is not always the case. Inspection of Figure 12 and further pair-wise comparisons show that when the suffix is absent, there is no significant difference between auditory and visual performance, $F\ (1, 26) = 0.23$, $MSE = 0.04$, $p = .63$, $\eta_p^2 = .01$, but there is a significant interaction between modality and serial position, $F\ (6, 156) = 15.23$, $MSE = 0.01$, $p < .05$, $\eta_p^2 = .37$, reflecting the IME. The interaction between modality and serial position is still evident in the suffix-present condition, $F\ (6, 156) = 3.71$, $MSE = 0.01$, $p < .05$, $\eta_p^2 = .13$, but visual performance is now overall superior, $F\ (1, 26) = 27.21$, $MSE = 0.03$, $p < .05$, $\eta_p^2 = .51$.



<u>Figure 12</u>: Percentage of correctly recalled items as a function of serial position, modality and presence/absence of a suffix.

Returning to the overall analysis, it is evident that suffix presence significantly affected performance, *F* (1, 26) = 22.97, *MSE* = 0.02, $p < .05$, $\eta_p^2 = .47$, but the effect was not unitary across modality or serial position, *F* (6, 156) = 8.65, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .25$. Indeed, as depicted in Figure 12, visual performance was not affected by an auditory suffix at all. This is confirmed by the lack of a significant difference between visual lists with and without suffix, as revealed by a comparison of these two conditions, *F* (1, 26) = 0.43, *MSE* = 0.02, $p = .52$, $\eta_p^2 = .02$, and a lack of interaction between suffix and serial position, *F* (6, 156) = 1.55, *MSE* = 0.01, $p = .16$, $\eta_p^2 = .06$. In contrast, a comparison of the auditory suffix-present and suffix-absent conditions shows that the presence of the suffix did reduce performance on the final items in the auditory list. This is evidenced by the significant difference between auditory performance with and without suffix, *F* (1, 26) = 40.95, *MSE* = 0.02, $p < .05$, $\eta_p^2 = .61$

and the significant interaction of the effect of suffix with serial position, *F* (6, 156) = 9.69, *MSE* = 0.01, $p < .05$, $\eta_p^2 = .27$. Using the same recency measure as in Experiment 6, it was further established that the suffix significantly reduced auditory recency, *t* (26) = 6.59, $p < .001$.[5]

Cluster Analysis

The Squared Euclidean distance between the response accuracies of each participant on each serial position was used to construct hierarchical clusters of interrelating responses for each of the four conditions in Experiment 9. Figures 13a-d depict the dendrograms that this analysis produced. If there was a high likelihood for a participant to respond correctly on two serial positions on the same trial (henceforth: “co-recalled”), these serial positions would cluster closely together. The more unlikely it was that two items were co-recalled, the further apart would the serial positions of these items be on the dendrograms and the later they would join in the same cluster. For example, if correct and incorrect responses on the first and second serial positions would constantly coincide, whereas the accuracy of responses on the fourth serial position would only coincide with the correctness of the responses at the first and second serial position occasionally, then the first and second items would closely cluster together and the fourth serial position would be added to the cluster much later (see e.g. Figure 13a).

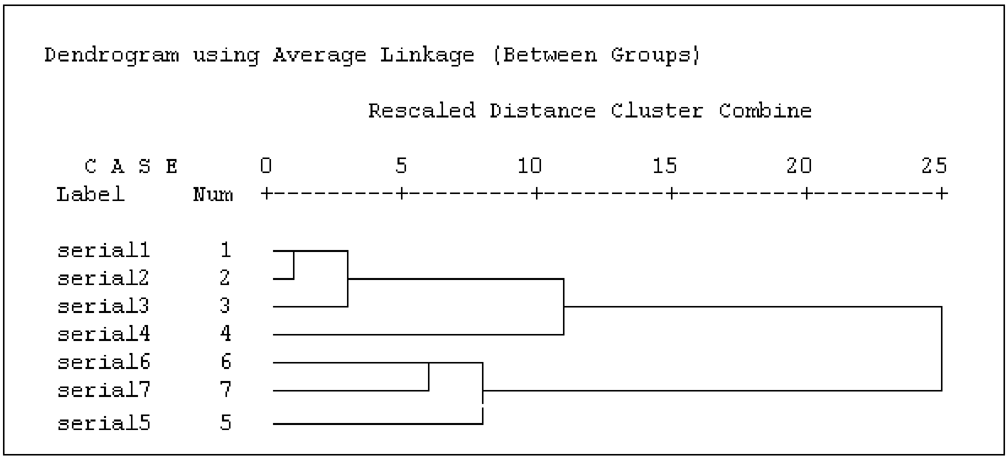
Figures 13a and 13b indicate that in the visual condition, regardless of whether an auditory suffix was present or not, the likelihood that participants co-recalled items on any trial declined as a function of the serial position: The first four items from a

---

[5] It should be noted that the original purpose of Experiment 9, namely to establish that the IME is not simply an artifact of the high auditory end-list performance, was also accomplished here, as the visual pre-recency advantage persisted even when auditory recency was reduced by a suffix.

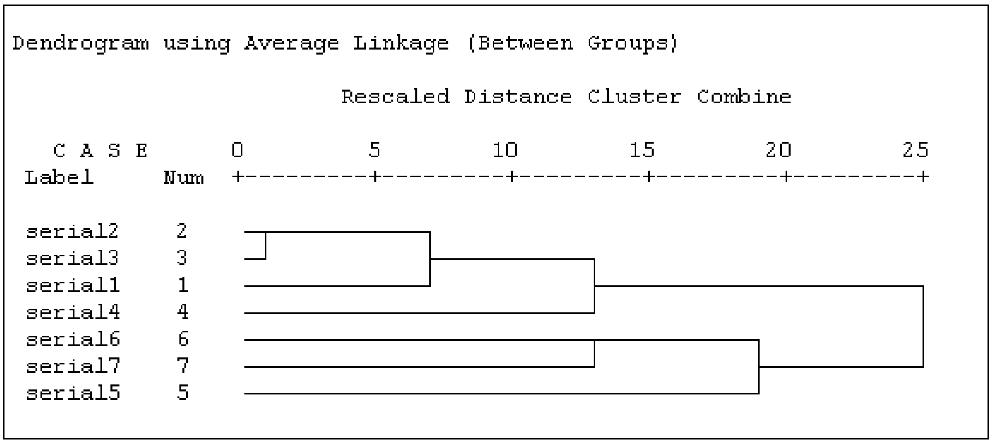
single recall cluster to which the last three items are added much later. Of particular interest is the fact that the same pattern only emerged for auditory lists in the suffix present condition (Figure 13d). When the suffix was absent (Figure, 13c) participants were more likely to co-recall the first three and the last item, in an overarching recall cluster. Thus, as would be expected from a cohesive object, there seems to be a codependence between the early and the late items in the auditory list. Moreover, a suffix seems to reduce the frequency with which the last item is co-recalled with the early items, presumably by displacing the last item from the list-object boundary.

<u>13a</u>: Visual lists, no suffix:



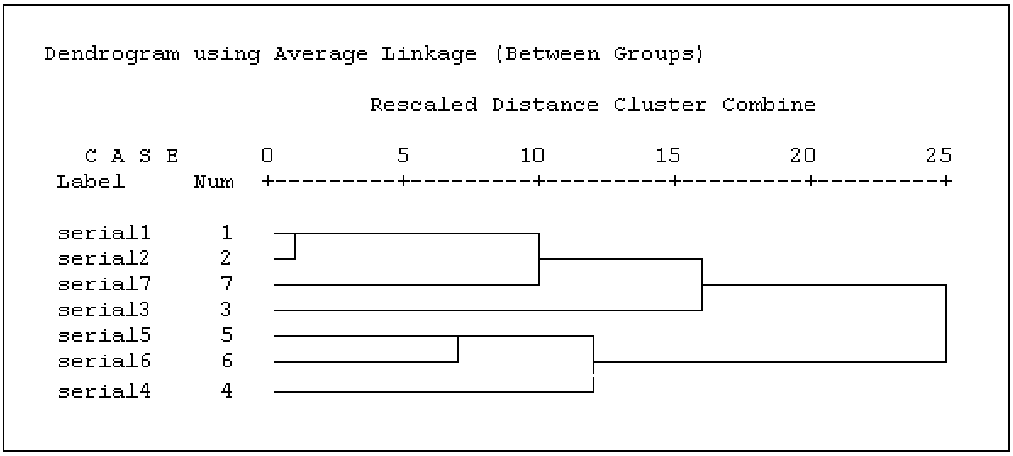
```
Dendrogram using Average Linkage (Between Groups)

                    Rescaled Distance Cluster Combine

   C A S E      0         5        10        15        20        25
  Label     Num  +---------+---------+---------+---------+---------+

  serial1     1
  serial2     2
  serial3     3
  serial4     4
  serial6     6
  serial7     7
  serial5     5
```

<u>13b</u>: Visual lists, with suffix

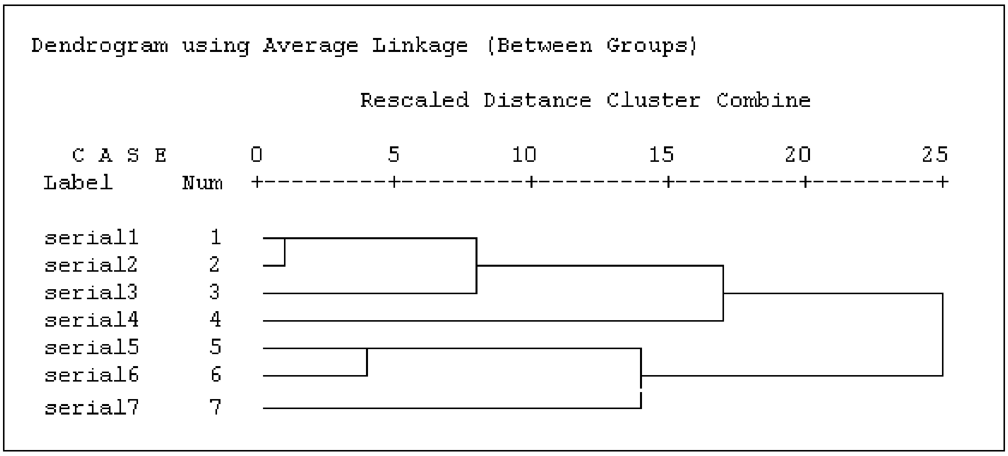


```
Dendrogram using Average Linkage (Between Groups)

                    Rescaled Distance Cluster Combine

   C A S E      0         5        10        15        20        25
  Label     Num  +---------+---------+---------+---------+---------+

  serial2     2
  serial3     3
  serial1     1
  serial4     4
  serial6     6
  serial7     7
  serial5     5
```

<u>13c</u>: Auditory lists, no suffix



<u>13d</u>: Auditory lists, with suffix



<u>Figures 13a-d</u>: Dendrograms depicting the clusters of correct responses on each serial position for each condition in Experiment 9.

Discussion

The IME was once more replicated in this experiment. The superior performance on auditory lists in recency was accompanied by superior performance on visual lists in pre-recency. Moreover, in line with previous research (Greene & Crowder, 1984; Jones et al., 2004, 2006; Nicholls & Jones, 2002), the auditory suffix reduced auditory list performance in recency. If the suffix cannot be ignored, however, it would follow that all the TBR items are also encoded.

The cluster analysis corroborates this conclusion. In the absence of the suffix, early and late list auditory items have a high likelihood of being co-recalled. This suggests that some obligatory holistic processing of the entire auditory list takes place, whereby whenever the early-list items are recalled well the last item is recalled well, too. This supports the notion of auditory list objecthood postulated by the perceptual-gestural account. The cluster analysis also shows that the suffix displaces the last TBR list item from the auditory list object edge. In conjunction, the cluster analysis, in addition to the more traditional descriptive and inferential statistical analysis of the data, make it doubtful that any TBR items from the auditory list can be ignored. In contrast, cluster analysis of the visual lists reveals that there is a consistent trend to co-recall the first four visual items, with the late list items falling outside of this recall cluster. This, together with the comparatively low visual performance in recency, is in line with the idea that the visual list is not obligatorily encoded in its entirety.

In order to explain how the ability to strategically ignore portions of the visually presented TBR list might result in an encoding-based visual pre-recency advantage it is necessary to assume that the STM process is capacity limited and that overburdening the process beyond its capacity has a negative effect on recall. The idea of a capacity limit on STM is widely accepted (see Baddeley, 2003; Cowan, 2001; Miller, 1956). From the perceptual-gestural perspective (Hughes et al., 2009, 2011; Jones et al., 2004, 2006) in order to maintain the order of a sequence of dissociated verbal tokens articulatory motor skills can be utilized to impose order, for example by modifying item articulation to facilitate co-articulatory transitions. Yet, clearly, there must be a limit on the number of such co-articulatory modifications, and their complexity, that can be planned and maintained at any one time. When this number is exceeded, e.g., as a result of encoding too many items, then accurate recall of order can no longer be guaranteed.

To speculate further, suppose that the average experimental participant has the ability to plan sufficient co-articulatory modifications to accurately maintain the order of four items. Previous research (e.g., Cowan, 2001) as well as the cluster analyses of visual lists in the present experiment confirms that this is a representative number of items that can be remembered by the average participant. Having to maintain a sequence of five or more items will pose a challenge for the average participant. If the surplus items are obligatorily encoded then the number of item transitions would exceed the number of co-articulatory item modifications that can be planned. In order to compensate, less sophisticated co-articulatory modifications could be implemented, which would however be of inferior quality. Hence confusions between individual items would be more likely. The present experiment, in conjunction with past studies on the suffix effect (Greene & Crowder, 1984; Nicholls & Jones, 2002), suggests that temporally-extended object lists, like auditory, visual-vocalized and visual-mouthed lists, are obligatorily encoded in their entirety. Given the average capacity of four items, the seven item TBR lists used in the experiments in this chapter must have frequently exceeded the participants' capacity to plan co-articulatory modifications to maintain the order of all items in the lists, resulting in a loss in recall performance. The disadvantage of obligatorily encoding the entire list would be, however, offset by the perception of clear list boundaries serving as strong order cues at the beginning and end of the lists. Thus, the reason for the strong recency on auditory, visual-vocalized, and visual-mouthed lists, namely objecthood, is also the reason for a low mid-list performance on those lists.

In contrast, with visual-silent lists (Greene, 1987) the average participant seems to be able to surmount the challenge of perceiving surplus items by simply ignoring them. Thus a participant with the ability to plan co-articulatory transitions between four items would most likely encode the first four items and ignore the rest. In such an instance only the first four items would be recalled very well, and the remaining items would be mostly forgotten. This number is not a constant, however: First, it would vary due to individual differences in articulatory planning aptitude. Further, it is doubtful that participants have a sufficient level of introspection to precisely determine how many items they are capable of accurately maintaining. Thus, should participants choose to strategically ignore a number of items it is likely that they will either ignore too few or too many items, and in either case their performance will not be optimal. This may be why performance on the first four items is not at ceiling, but high enough to match the auditory advantage in recency.

General Discussion

The aim of this chapter was to scrutinize traditional verbal STM accounts by focusing on the predictions these accounts make in regards to serial recall performance differences across modalities. Thus, store-based (e.g., Baddeley, 2003; Crowder & Morton, 1969; Henson, 1998) interference-based (Nairne, 1990; Neath, 2000), as well as some independent accounts that are not immediately associated with either the store or the interference-based tradition (Glenberg & Swanson, 1986; Penney, 1989) argue that there are hardwired entities, like a special store (Crowder & Morton, 1969), or processes, like encoding of auditory items in terms of physical features as opposed to features that arise from internal cognitive processes (Neath, 2000), which ensure a recall advantage for verbal information that is presented auditorily. As evidence, these theories point towards the modality effect, the robust finding that verbal lists presented auditorily are recalled better in recency than verbal lists that are presented visually. In the present chapter, however, two challenges to the notion of a hardwired advantage for auditory items were presented. First, there are lists without an auditory component, like visual-mouthed lists, on which recall performance in recency is as high as on auditory lists. Second, the auditory recency advantage is often matched by a visual advantage in pre-recency. The present experiments sought to demonstrate the robustness of this phenomenon, the IME, as well as determining whether its presence could be reconciled with any of the traditional STM accounts. In Experiment 5 and 6 it was established that store or interference based explanations cannot readily accommodate the inverted modality effect. These two experiments demonstrated, against the explicit prediction of several STM models (Baddeley and Larsen, 2007; Penney, 1989) that the IME is not observed because additional visual codes are utilized by the participants to improve pre-recency performance on visual-silent lists. Performance on visual lists that had to be vocalized or mouthed resembled performance on auditory lists. Participants did not utilize the available visual information to match pre-recency performance on visual-vocalized or visual-mouthed lists to performance on visual-silent lists. Indeed, it was concluded that the only way store or interference based accounts can accommodate the finding of high recency with visual-mouthed lists and high pre-recency with visual-silent lists is to invoke a different code, or storage device for each presentation type.

Instead of arguing that auditory and visual-vocalized lists are stored in an acoustic code that promotes high recall in recency, that visual-mouthed list are stored in an acoustic-like code that also promotes high recall in recency (Penney, 1989), and that visual-silent lists are stored in yet another code that promotes high recall in pre-recency, a more unitary explanation was sought. Based on the perceptual-gestural view of STM (Hughes et al., 2009, 2011; Jones et al., 2004, 2006), it was suggested that whenever there is a high degree of perceptual cohesiveness between the TBR list items the list will be perceived as a temporally-extended object, so that cessation of list processing activity will represent a clear object edge that disambiguates item order at that edge. Note that in order to be cohesive the TBR list does not need to be auditory but list processing needs to be salient, which is, for example, the case when each list item is mouthed or lip-read.

It was initially unclear how this account might explain the pre-recency advantage on visual-silent lists matching the recency advantage that auditory visual-vocalized and visual-mouthed lists had because of their objecthood. The initial idea was that visual-silent lists are somehow grafted more easily onto a subjective prosodic rhythm, or that visual items lend themselves particularly well to co-articulatory modifications, both articulatory planning processes which according to the perceptual-gestural account facilitate order recall. These hypotheses were however found to be inadequate. First, Experiment 7 failed to confirm that visual-silent lists are uploaded more easily onto a subjective structure than auditory lists. The experimental requirement to recall the TBR list in a restructured form had as much impact on visual-silent lists as it had on auditory lists. If the visual lists cannot be restructured any easier than auditory lists as per experimental instructions then there is no reason to assume that visual lists can be restructured any easier than the auditory lists to fit a subjective prosodic rhythm which would promote recall. Subsequently, Experiment 8 raised doubts that the visual pre-recency advantage has anything to do with articulatory motor planning processes benefitting visual-silent list recall. When articulatory planning was constrained with concurrent articulation the IME remained.

Moreover, the traditional modality effect and the IME were equally accentuated. This suggested that the locus of the IME is probably also perceptual.

Experiment 9 hinted at a potential explanation for the IME. It suggested that auditory lists are encoded in their entirety. It was shown that a suffix at the end of the auditory TBR list is obligatorily encoded as part of the list. If a to-be-ignored item at the end of the TBR list is encoded as part of the list, it would seem reasonable to suppose that all the TBR auditory list items are also obligatorily encoded. Previous research has also demonstrated similar suffix effects with vocalized and mouthed lists (Greene & Crowder, 1984), hence it is likely that these lists are also obligatorily encoded in their entirety. In contrast to this, it was shown that having to ignore a suffix at the end of a visual-silent list has little effect (Greene, 1987). This implies that there is no obligatory encoding of the entire visual list; visual items at the end of the list can be strategically ignored.

The differences in how the lists from the various presentation types are processed, with every item being encoded in lists perceived as temporally-extended objects and items being potentially ignored in lists that are not perceived as objects, could be the key to explaining the IME, in accordance with the perceptual-gestural account of STM. From this standpoint, it is necessary to assume that the processes supporting serial recall are capacity limited and that there is a detriment to recall if the number of encoded items exceeds this capacity. According to the perceptual-gestural account (Hughes et al., 2009), the order of an unstructured verbal list can be maintained by utilizing articulatory planning skills to facilitate co-articulatory transitions between the items. Arguably there must be an upper limit on the number and complexity of the co-articulatory modifications that can be planned at any given time. Moreover, the likelihood of exceeding that limit would increase as a function of, amongst other factors, the number of to-be-planned co-articulatory inter-item transitions, and hence as a function of the number of encoded items. On temporally-extended object lists, where all items seem to be obligatorily encoded, the likelihood of encoding more items than can be accurately maintained would thus be quite high. In order to maintain the surplus items less complex co-articulatory modifications could be devised, which however would probably be less efficient cues for order, so that the likelihood for item confusions would increase. The resulting drop in performance would be compensated, however, on object lists in primacy and recency. Here the object edges provide perceptual cues, which serve as positional anchors for list items. Note that these cues are independent of articulatory planning processes.

On visual-silent list there are no immediately apparent object cues, and perceptual organisation processes do not seem to play a role in facilitating recall of visual verbal material. However, since it seems possible to selectively ignore perceived items on visual-silent lists (Greene, 1987) it seems easier to regulate how many items are encoded from a TBR list. Thus with non-object lists it is possible to only encode as many items as can be accurately maintained in order using motor planning skills. Indeed, if the amount and complexity of co-articulatory modifications that can be planned is constrained, for example by concurrent articulation then this can be accommodated by simply encoding fewer items from the visual-silent list. This is why in Experiment 8 the IME was as much accentuated as the traditional modality effect. Whereas, on the auditory list the complexity and quality of co-articulatory transitions between list items declined and recall performance became more defined by the perceptual cues at the list edges, in the visual list recall efforts were simply restricted to fewer early items, which accentuated the early-list visual advantage. It should also be noted that the present explanation of the IME, emphasizing presentation type-based differences in obligatory encoding of the entire list, is not at odds with the results from Experiment 7. Thus it could be argued that Experiment 7 shows that auditory and visual items can be ignored equally successfully, given that for either presentation type there was a similar decline in performance as a result of the restructuring manipulation. However, it seems unlikely that the instruction to mentally restructure the TBR lists was interpreted as an instruction to ignore a group of items. Instead, the restructuring probably had its impact at a later stage, after the encoding differences between the presentation types would have contributed to the re-emergence of the IME in that experiment.

Finally, the ability to strategically ignore a portion of the visual TBR list may also explain why the IME is not always evident in serial memory experiments (cf. Penney, 1975). Should the majority of experimental participants interpret an instruction to recall the entire TBR list as an instruction to encode the entire list and reproduce it as best as they can, as opposed to an instruction to perform as well as possible on the recall task, then those participants may encode the entire TBR visual list and performance on both visual and auditory lists will be equal, except at the list-edges where performance on auditory lists would receive a boost from perceptual factors. Indeed, appropriate manipulations of experimental instructions might be a simple way to address some of the speculations offered in this chapter in regards to the origin of the IME.

In conclusion, both the presence of a modality effect in lists without an auditory component and an IME pose considerable challenges to the store-based, interference-based, and the perceptual-gestural accounts of verbal STM. Yet, in contrast to the store- and interference based accounts, the perceptual-gestural can be plausibly expanded to accommodate both findings, by allowing for any list that might be perceived as an object to be well recalled at the list edges, and by suggesting that the gestural aspect of the verbal STM process is capacity limited and that overburdening the process beyond its capacity has a negative effect on average recall performance.

CHAPTER 5: DISCUSSION

At the beginning of this thesis the "dragon" metaphor was introduced. The purpose of the metaphor was to illustrate how a well-defined single construct can easily explain a wide range of data, whilst still being improbable. At the same time, a belief in dragons constitutes a prime example of how inappropriate it is to believe in something that is nowhere to be found. The aim of this thesis was to show that the dragon metaphor is highly relevant for verbal STM research. Thus, in the debate about the nature of verbal STM, some (e.g., Baddeley, 2003) propose that it is accomplished by a dedicated mechanism, comprising a bespoke store which serves as a passive repository of decay-prone phonological memory traces, and an active articulatory control process to refresh these traces. Yet although the simple and elegant concept of a bespoke store has considerable explanatory power, it seems doubtful that a system dedicated exclusively to retaining phonological information for a brief period of time would have evolved. Indeed, the notion of the passive short-term store, because of its elegance and simultaneous unlikelihood, is rather reminiscent of the notion that travellers disappear on long journeys because of dragons.

Other theorists (e.g., Nairne, 1990), rejecting the idea of a passive store in which information decays and an active refreshing mechanism, have argued that verbal STM is an entirely passive process that is governed by retroactive interference. Yet, the processes that prominent interference-based models propose have not been identified with any neurological equivalent; they lurk somewhere within the neural architecture but only like a dragon behind a hill.

In the present thesis, the attempt was made to advocate yet another alternative approach, considering STM as an emergent property of general receptive and productive mechanisms, with speech-related productive mechanisms being employed particularly frequently on verbal STM task due to the linguistic nature of the TBR material. This allows a clear mapping of processes associated with verbal STM onto brain regions associated with perception, (speech-) planning and production. At the same time, there is no necessity to invoke yet another dedicated system, thus violating evolutionary parsimony. Thus, the perceptual-gestural approach to verbal STM could hold the key to solving the "dragon" problem. To confirm this, the empirical work of this thesis focussed on demonstrating that the perceptual-gestural account offers an equal or better explanation of a range of data obtained from behavioural and neuroscientific studies, than either the store-based or the interference-based approach.

Summary of Empirical Chapters

Chapter 2: The impact of non-verbal concurrent tasks on verbal STM

In the first empirical chapter of this thesis the predictions of the store-based Phonological Loop model (Baddeley, 2003), the interference-based Feature Model (Nairne, 1990) and the perceptual-gestural account (Jones et al., 2004) were examined in regards to the role articulatory processes play in short-term retention of verbal information. Both the Phonological Loop model and the perceptual-gestural account argue that the articulatory process is critical in order to, respectively, refresh decaying traces of information inside a bespoke store or to enable the assembly of an unconstrained list of verbal events into a coherent sequence of articulatory gestures. The Feature Model, on the other hand, argues that articulatory action is unnecessary to maintain TBR information. On this model, the commonly observed reduction in verbal STM performance as a result of concurrent articulation (e.g., Baddeley, 1986) is attributed to the irrelevant verbal features of the concurrently repeated items interfering with the TBR item traces. The simultaneous constraint that concurrent articulation puts on articulatory fluency does not, according to the Feature Model, immediately affect the STM process.

In order to show that a constraint on the articulatory process can disrupt STM even if the confound of interfering verbal representations is removed, concurrent articulation was substituted with a non-verbal complex oral motor action, namely chewing gum. Experiment 1 revealed that, in line with the predictions of the Phonological Loop model and the perceptual-gestural account, but against the predictions of the Feature Model, chewing gum, like concurrent articulation, has a detrimental effect on verbal STM. However, unlike concurrent articulation, chewing gum did not attenuate the canonical PSE. In order to learn more about the intricacies of the effect that chewing has on STM, two further experiments were conducted. Experiment 2 compared the effects of chewing on a probed order and a missing item task. These tasks are considered to be well matched except that the probed order task requires order recall and has been shown to be particularly sensitive to disruption by concurrent articulation (Klapp et al., 1983). Yet, chewing gum had a similar effect on both tasks indicating that its effects cannot be considered equivalent to the effect of concurrent articulation. Indeed, Experiment 3 showed that the effects of chewing rather resemble the effects of simple tapping, a concurrent motor task which, at first glance, seems far removed from the articulatory process. These results seemed initially to resonate with the Feature Model, which claims that irrelevant tasks that deplete a central attentional resource can have a detrimental effect on memory, without introducing additional features (Neath, 2000). However, since that model does not sufficiently specify when or why some concurrent tasks affect STM by introducing interfering features and others by depleting an attentional resource, the

Feature model was not deemed to provide a satisfactory account of the effects of chewing gum and tapping on verbal STM. Instead, the search for an explanation turned back towards the perceptual-gestural and Phonological Loop accounts, both of which acknowledge that peripheral motoric constraints on the articulatory process can reduce overall performance on verbal STM tasks (cf. Baddeley & Wilson, 1985; Macken & Jones, 1995). It was thus concluded that tapping and chewing represent such constraints.

Chapter 3: Theta Burst Stimulation of Broca's area modulates verbal STM

The empirical evaluation of the most prominent instantiations of the interference-based, the store-based and the perceptual-gestural approaches to verbal STM continued in Chapter 3, which reprised the scrutiny of the various predictions the respective accounts make about the role of articulatory processes in verbal STM. Thus, as discussed in Chapter 2, the interference-based approach (Nairne, 1990; Neath, 2000) does not consider an unconstrained articulatory process to be necessary for STM. The negative effects of concurrent tasks which seemingly disrupt the STM process by constraining articulatory fluency, like concurrent articulation (Baddeley, 1986) or chewing gum (Chapter 2), are explained either with reference to interfering verbal features or the depletion of a central attentional resource (cf. Neath, 2000; but see also Lewandowsky, Geiger, & Oberauer, 2008).

The store-based Phonological Loop model (Baddeley, 2003), on the other hand, regarding articulatory processes as vital, has identified the location of the articulatory control process that serves to refresh decaying phonological representations of TBR information in the phonological store with BA 44, Broca's area. The model thus clearly predicts that a lesion of Broca's area should result in a significant reduction in verbal STM performance. This is because in the absence of the articulatory control process, information in the phonological store would decay rapidly and be forgotten. Moreover, visual information, which according to the Phonological Loop model gains access to the phonological store via the articulatory control process would not be stored at all, or only to a very limited degree. Note that according to the model, a side-effect of the disruption of access of visual material to the phonological store should be a reduction of the impact on recall of inter-item phonological similarity.

Like the Phonological Loop model, the perceptual-gestural account (Hughes et al., 2009, 2011; Jones et al., 2004, 2006), in regarding verbal STM an emergent property of perception and speech-planning, considers Broca's area to be involved in the STM process, given that the area is commonly associated with speech-planning (e.g., Davis et al., 2008). Yet the account does not specify Broca's area as the sole seat of the speech-plan assembly process. It is rather in line with the account to consider Broca's area as one important node in a network of brain areas responsible for speech production, damage to which should impact upon but not necessarily disrupt the verbal STM process.

In order to address the predictions of the models, Broca's area was carefully localized in a sample of volunteers, first structurally on their MRI scan then functionally, by applying repetitive TMS to various locations within the identified region until a speech-planning arrest hotspot was found. There, a theta burst of TMS was administered. Participants were subsequently given a visual-verbal STM task. A careful analysis of performance of each participant on the STM task revealed that TBS of Broca's area appears to reduce the PSE without affecting average performance. This observation is incompatible with both the interference-based account and the store-based account. Thus, it is not clear from an interference-based perspective how a virtual lesion of a brain area associated with speech should have either introduced interfering verbal features or depleted a central attentional resource. Moreover, from the store-based perspective of the Phonological Loop model, the primary effect of inhibiting Broca's area should have been a reduction in average performance and not an isolated reduction of the PSE. Only the perceptual-gestural approach can accommodate the selective effect that temporarily lesioning Broca's area had on verbal STM. This is because the approach, while suggesting that Broca's area is important for the verbal STM process due to its involvement in speech planning, does not postulate that the area is the seat of the process (for further discussion of this point, see section "Towards a neural model of verbal STM" below). This allows for the occurrence of peculiar behavioural effects like a PSE reduction in the absence of a reduction of overall performance arising from a stimulation of Broca's area. Nevertheless, it was concluded that more research is needed to identify the precise role of Broca's area in verbal STM, and how inhibition of the area could reduce recall of dissimilar verbal items, while simultaneously improving recall of similar items.

Chapter 4: A new approach to modality effects in verbal serial recall: Meeting the challenge of explaining a visual mid-list advantage

In the third empirical chapter of this thesis the emphasis shifted towards evaluating the store-based, interference-based and perceptual-gestural approaches to verbal STM based on their predictions about the impact of differences in presentation modality on the ability to maintain the order in a verbal sequence. Thus, all three approaches seem to predict a physiologically hardwired memory advantage, manifesting itself in particular in recency, for auditorily over visually presented lists. Nevertheless, the accounts differ in the reasons they give for this advantage. From the store-based perspective, the recency advantage is either considered a consequence of the preferential access of acoustic material to the bespoke phonological short-term storage mechanism (Baddeley, 2003), or an additional low capacity acoustic storage mechanism (Crowder & Morton, 1969). From the interference-based perspective it is suggested that auditory information is represented to a greater extent in terms of modality-dependent features, which are less prone to interference from internal activity, so that auditory memory traces in recency are relatively well preserved (Nairne, 1990). Finally, from the perceptual-gestural perspective, auditory lists are encoded as temporally-extended objects, so that the silence at the end of the auditory list constitutes an order-disambiguating boundary (Jones et al., 2004, 2006).

The advantage at recency for auditory compared to visual serial recall has been dubbed the modality effect. Yet, the assumption that this auditory advantage is hardwired faces two strong challenges: First, high recency performance can be observed in the absence of acoustic input such as with visual-mouthed or lip-read lists (Greene & Crowder, 1984). Moreover, a visual mid-list advantage matching the auditory advantage in recency, the IME (Beaman, 2002), is also frequently observed. As discussed in Chapter 4, these two challenges have not sufficiently been addressed by previous research, with the IME in particular being rarely commented upon, and are commonly dismissed as fluctuations in the data resulting from participants recruiting additional visual codes (Penney, 1975, 1989) or the use of additional STM mechanisms that are primarily dedicated to the retention of visuo-spatial information (Baddeley & Larsen, 2007). These explanations appeared inadequate, however, in light of Experiment 6 in Chapter 4. In that experiment performance on visual-silent lists was compared to performance on auditory, visual-vocalized and visual-mouthed lists. Although visual codes were present in both conditions, and acoustic codes were not present on either, performance on visual-silent and visual-mouthed lists differed significantly, with early list recall being high on visual-silent list, and a distinct recency advantage being present on visual-mouthed lists. Given these results one could of course still maintain that participants simply recruit an additional store or code for each instance of high performance. Such an explanation would, however, hardly go beyond a re-description of the data. Indeed the ease with which such a "simple" explanation can be dismissed as not being parsimonious raises questions about the general validity of ever invoking dedicated mechanisms to explain a performance pattern.

In order to reconcile the existence of the IME and of high recency in the absence of auditory input with verbal STM theory, Chapter 4 turned to the perceptual-gestural account. It was suggested that perhaps in order to be perceived as temporally-extended objects, verbal lists do not necessarily need to contain an auditory component. Engaging in a salient activity to process a TBR list, like gesturing the TBR list items or reading them from a person's lip movements might be sufficient. Thus, cessation of the salient activity would still constitute an order disambiguating boundary, explaining high recency in the absence of an auditory input. Moreover, Chapter 4 explained the IME by proposing that if a list is organized by perceptual processes as an object, then it is encoded in its entirety. This assumption was supported by Experiment 9 in which even a suffix is obligatorily encoded as part of the auditory list-object. While obligatory encoding of an entire list as an object might benefit recall at the list-object edges, it is also in line with the perceptual-gestural account to argue that it will likely impact performance at mid-list. This is because, according to the account, the primary challenge with reproducing the order of a grammatically and semantically unconstrained list of familiar items lies with imposing appropriate transitional markers between items (Hughes et al., 2009). Thus, utilizing speech planning skills, articulation of the TBR items could be modified in a fashion that would facilitate co-articulatory transitions between the items. Arguably, however, only a limited number of co-articulatory modifications can be planned at any one time. Thus the quality of each modification, and hence its effectiveness as an order cue, would likely decline as a function of the quantity of to-be planned modifications. Thus longer lists, if encoded in their entirety, would be held together with poorer co-articulatory modifications and hence there would be a greater chance for inter-item confusions. On the other hand, if a portion of the TBR list could be strategically ignored, as seems possible in visual lists that are not perceived as objects (Greene, 1987), then the co-articulatory transitions between the encoded list items would be of high quality and hence errors would be rare. Thus, it was concluded in Chapter 4 that the visual mid-list advantage results from an ability to strategically ignore the latter portion of the visual list if it is not obligatorily organized as an object by the perceptual system. This perceptual-gestural explanation for the presentation type-based differences in verbal STM performance appears more adequate and parsimonious than store- or interference-based proposals of additionally recruited stores or codes.

Additional Challenges for Verbal STM Theory

The problem of serial behaviour

The empirical findings just summarised suggest that the perceptual-gestural approach to verbal STM is more capable of accommodating the effects of various types of non-verbal constraints on articulation and presentation-type based differences than either an interference or a store-based approach. However, in particular the explanation that the perceptual-gestural perspective offers for the presentation type-based differences reveals an important issue with the perceptual-gestural view that, while being somewhat outside the scope of this thesis, needs to be addressed here. A potential challenge is that by assuming that order in a TBR verbal sequence is maintained primarily via co-articulatory links between items (Hughes et al., 2009), the perceptual-gestural account seems to subscribe to an associative chaining account of serial order maintenance. Associative chaining generally refers to a way of conceptualizing serial behaviour, whereby each action is thought to trigger a subsequent action (Lashley, 1951). In STM research, theories that postulate that items serve as cues for subsequent items are referred to as associative chaining models (Henson, 1998). Associative chaining has, however, fallen out of favour in verbal STM research in recent years. Instead, positional coding is considered to offer more accurate predictions about patterns of serial recall performance (cf. Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Henson, 1998). In contrast to associative chaining models, positional coding models argue that items are associated with a certain context, for example the distance from the start or the end of a TBR list (Henson, 1998) or the state of a set of neural oscillators (Burgess & Hitch, 1999).

A prominent objection against associative chaining is that the execution of successive actions is often too quick for each action to serve as a trigger for the next action. For example, the speed with which a piano player presses successive keys is too great to assume that feedback about each button press travels back to the brain and there triggers the execution of the next keypress (Lashley, 1951). However, this does not exclude the possibility that sequences of actions are *planned* in a manner such that the execution of the first action triggers a cascade of successive actions, as the perceptual-gestural account suggests. Yet there are other, more substantial, challenges to the associative chaining idea as envisaged by the perceptual-gestural view. For example, recalling the serial order of mixed lists containing, alternatingly, phonologically dissimilar and similar items yields a saw-tooth pattern of performance with dissimilar items being recalled well and similar items being recalled poorly. Importantly, average recall performance on dissimilar items in mixed lists is as high as for such items when presented in a ‘pure’ dissimilar list (Baddeley, 1968; Henson, Norris, Page, & Baddeley, 1996). This finding is thought to constitute a strong challenge for associative chaining theories: It is argued that similar items don’t provide good discriminative cues for their successors. Similarly it could be suggested that the low performance on phonologically similar items in mixed lists indicates that those items are frequently forgotten. An item that is forgotten, however, cannot serve as a cue for its successor. This means that on mixed lists, forgetting a similar item should lead to forgetting its successive dissimilar item. Recall performance on dissimilar items in mixed lists should therefore be poorer on average than on a matched pure dissimilar list. Since this is not the case, and mixed-list dissimilar items are recalled just as well as items on pure dissimilar lists, it is argued that the dissimilar items must be associated with a certain position in the list as opposed to being chained to their preceding items (Henson et al., 1996). In regards to the perceptual-gestural approach to serial order, it could thus be criticized that the assumption that serial order is maintained by introducing co-articulatory cues linking successive items implies that forgetting a single item should lead to forgetting of its successor, or that the co-articulatory cues following similar items should not link to successive items as well as co-articulatory cues following dissimilar items. The perceptual-gestural account thus seems to incorrectly predict that performance on dissimilar items on mixed lists should be lower than performance on dissimilar items on pure lists.

The second strong challenge to associative chaining is the prevalence of fill-in errors over infill errors on serial recall tasks. For example, if it is assumed that the items “ABC” have to be reproduced in order, and the second item “B” is reproduced as the first item, then a fill-in error would follow if the first item “A” would be recalled in the second serial position, i.e. the response would be “BAC”. An infill error would be observed if the misplaced item “B” would be followed by its original successor, namely “C”, so that “BCA” would be the response given. While the perceptual-gestural assumption of co-articulatory cueing between items seems to predict a great inter-item cohesiveness, and hence a greater frequency of infill errors, research seems to suggest that fill-in errors in fact occur twice as often as infill errors (Henson, 1998; Surprenant, Kelley, Farley, & Neath, 2005).

While these issues pose considerable challenges to the perceptual-gestural view, they are not insurmountable. Consider the finding that recall of phonologically dissimilar items in mixed lists is as high as on pure dissimilar item lists (Henson et al., 1996). The perceptual-gestural view could respond to this challenge by suggesting that on mixed lists phonologically similar and dissimilar items are processed as belonging to separate streams. This argument applies particularly in the auditory domain. Here it has been shown that if the acoustic properties of items in a TBR list are alternated, so that for example the first item is presented in a male voice, the second in a female, the third in a male voice again, and so on, then the TBR list is perceived as two separate lists (Bregman, 1990; Hughes et al., 2009). Yet, comparable performance between dissimilar items in mixed and pure dissimilar lists has also been observed when lists were presented visually (Henson et al., 1996). Nevertheless, this does not exclude some sort of post-perceptual, modality-independent segregation of the mixed list into similar and dissimilar items. Such a speculation is supported by the finding that on mixed lists transpositions of items tend to take place primarily between items from the same similarity-category, i.e., similar items tend to swap places with other similar items rather than with their direct neighbours, which is different from pure dissimilar lists where transpositions are most frequent between direct neighbours (Henson et al., 1996). Arguably, this could indicate that dissimilar and similar items in mixed lists are maintained as two separate lists, that is a well-remembered dissimilar item list and a badly remembered similar item list. Indeed, Experiment 7 of the present thesis (Chapter 4) might have serendipitously illustrated the effects of comparable double-list maintenance. In that experiment participants were instructed to mentally segregate items in odd serial positions from items in even serial positions and then reproduce both groups in serial order, starting with the items in odd positions. With visually as well as with auditorily presented lists, items in odd positions were recalled just as well as items in the same serial positions on lists that did not need mental restructuring, yet items in even positions on restructured lists were recalled much worse than their counterparts on non-restructured lists. If the serial order of items is maintained though positional coding, it is difficult to see how the odd-positioned items were recalled so well, even though their serial position at encoding did not match their serial position at retrieval. It seems that, if the order of the odd-positioned items was maintained by associating each item at encoding with a certain contextual cue, like distance from the beginning and end of the list (Henson, 1998), then at retrival that contextual cue should point to the item position at encoding and not the new position in which the item had to be placed as per task demand. On the other hand, if participants in that study used their vocal-articulatory planning skills to generate two separate coherent lists of items focussing their maintenance efforts on the item list that was supposed to be recalled first, i.e., the odd item group, then it would make sense that recall of items from that group remained high.

Further, there are several objections that can be raised from a perceptual-gestural perspective in regards to the prevalence of fill-in errors over infill errors. First it should be noted that the evidence for order errors being usually followed by fill-in errors has been contested in a recent study showing that erroneously recalled items are most frequently followed by their original successor (Solway, Murdock, & Kahana, 2012). Secondly, given the experimental requirement to recall each item in the position it was presented, any infill error would necessarily lead to more errors than a fill in error. To return to the example given above, if ABC is presented and an order error followed by a fill-in error is made (i.e. BAC) then only two errors are made. If however the order error is followed by an infill error (i.e. BCA) then three errors are made. The supposed prevalence of fill-in over infill errors could therefore be as much a demonstration of participants' ability to minimize errors as a demonstration of a fundamental process of serial behaviour. If it is assumed that participants are capable of monitoring their own performance to the extent that they strategically commit fill-in rather than infill errors to keep errors low, the prevalence of fill-in errors poses no challenge for associative chaining-based accounts of STM.

Finally, from the perspective of the perceptual-gestural account, the employment of co-articulatory modifications to form associative links between items is not the only means by which item order can be maintained. Thus, proponents of the perceptual-gestural account suggest that prosody might play an important role in the maintenance of item order (Hughes et al., 2009). Associating items with a prosodic rhythm could therefore constitute a way of positional coding that would be defensible from a perceptual-gestural standpoint. In sum, while some challenges to associative chaining could be considered challenges to the perceptual-gestural account, it appears capable of accommodating or circumventing these challenges, even though more research is required to discern the details.

Endorsement of principles rather than details

The problem of serial order reveals that the empirical work conducted within the framework of the present thesis does not suffice to confirm all the assumptions of the perceptual-gestural account. Moreover, at several points throughout the thesis it was pointed out how the perceptual-gestural view would need to be modified to accommodate the empirical data, for example, by expanding the concept of temporally-extended objects to include visual-mouthed and lip-read lists. It is hence probably inaccurate, at least from the perspective of this thesis, to consider the perceptual-gestural approach as the final definitive word on verbal STM. However, the studies reported here do point towards the general principles of the perceptual-gestural approach, like the rejection of bespoke memory buffers, and the conceptualization of memory as a fully active process, as a more promising basis for future STM research, than either the store-based, or the interference based approach.

Yet, perhaps, it could be objected that by contrasting the perceptual-gestural view primarily against the Feature Model and the Phonological Loop model the thesis does not capture the breadth of the interference-based or the store-based perspectives and is thus premature in challenging them as adequate approaches to verbal STM. However, while this thesis, admittedly, has not looked in any detail at how its empirical findings might apply to other models of verbal STM, the majority of challenges that the present thesis poses for the Phonological Loop and Feature models are germane to all forms of the store or interference-based approach. Thus, any theory that insists on retroactive interference of content being the sole source of forgetting (e.g., Lewandowsky & Farrell, 2008; Nairne, 1990) is at odds with studies that observe a reduction or modification of memory performance that cannot be attributed to interfering content. Studies of patients with speech-lesion-based verbal STM impairments (Waters et al., 1992), as well as studies in which recall output is modified as a consequence of TMS (Chapter 3), clearly show that interfering content is not a prerequisite for forgetting. Instead, these studies point clearly towards STM being an active process. Indeed, they suggest that cognitive studies that seem to demonstrate forgetting by content-dependent interference (e.g., Lewandowsky et al., 2008), are in fact simply demonstrating interference with the active process of STM.

Of course, an alternative explanation for an observed brain lesion-based reduction of memory could be that the damaged region constitutes some sort of passive storage device, like a phonological store (Baddeley, 2003). Such an assumption, however, as argued at several points throughout this thesis, gives rise to a number of problems. First, theories assuming that damage to selective brain regions causes information storage deficits because the region constitutes a passive information buffer are constantly challenged by findings implicating the respective region in some active process. Indeed, the active function ascribed to the region is usually much more in line with the specific disorders associated with damage to the region than any passive storage role that is ascribed to it. For example, while memory impairment following damage to BA 40 could be interpreted as demonstration of BA 40 containing a passive STM store (Baddeley, 2003; Vallar & Baddeley, 1984) much more evidence is accommodated by considering the region as actively involved in the integration of perception and oral motor action (Hickok, 2009). This accounts for the selective verbal memory impairments associated with damage to the region—perceived verbal information does not get through to vocal-articulatory planning and production mechanisms to be maintained—as well as the area’s implication in linguistic processes—the region being frequently impaired in patients with Wernicke’s aphasia (Buchsbaum & D’ Esposito, 2008). Another problem is that once a dedicated storage mechanism is invoked to explain information buffering in one cognitive domain, it would have to be assumed that in other cognitive domains information buffering is accomplished by dedicated buffers, too. Yet, the evolutionary viability and hence existence of a dedicated store for every buffering activity of the nervous system seems implausible (cf. also D’Esposito, 2007), in particular since evidence indicates that a small number of neurons in a petri dish are capable of maintaining information—keeping an electric stimulus active for up to several seconds—without the presence of any dedicated buffer structure (Vishwanathan, Bi, & Zeringue, 2011). The advantage of using the perceptual-gestural approach as a basis for future research is therefore that it discourages a search for bespoke entities that accomplish a certain cognitive process, but instead encourages an understanding of how complex cognitive processes like verbal STM can emerge from rudimentary processes like perception and preparation for action.

Overreliance on neuroscientific evidence

Perhaps another criticism that could be raised at this point against the current argument might be that it relies too heavily on neuroscientific evidence, and that such evidence should not play a role in determining the validity of cognitive theories. This objection might be justified if the only goal of cognitive science was to predict behaviour. Then, indeed, purely cognitive models like the Feature Model could be considered valid if they predicted behaviour with adequate accuracy. However, arguably, the goal of science, cognitive or otherwise, is to explain as well as to predict. In order to explain behaviour it is necessary to understand that behaviour originates within the biological confines of a living organism (cf. Glenberg, 1997). Hence, the accuracy with which a cognitive model can predict behaviour cannot be the only criterion on which the merits of a model are determined: If the assumptions of the model do not appear reasonable within a biophysical world, the model cannot be argued to explain behaviour. Even if the data that a model predicts correlate with behavioural performance it does not mean that the mechanisms the model postulates are the mechanisms underpinning the behaviour. In this spirit, the final section of this chapter proposes a rudimentary neural model of verbal STM, incorporating the principles of the perceptual-gestural approach, to be used as a basis for future research.

Towards a neural model of verbal STM

Towards the end of this thesis, it seems appropriate to speculate how a neural model of verbal STM could be conceptualized based on the arguments put forward so far. It is clear that it would need to be assumed that such a system would maintain information actively, and that the maintenance would not be accomplished by a single dedicated mechanism, but instead by a number of brain regions, whose primary function would be associated with perception and (speech) planning. These regions would play a role in the verbal STM process to the extent to which their recruitment might be appropriate to accomplish a given verbal STM task.

The neural path of any TBR information begins with perceptual processes. This thesis touched upon two routes via which TBR verbal information can reach the brain, namely the auditory and the visual route. In the auditory domain, perceptual sequential-organization processes appear to define verbal STM performance, being manifest in effects like the talker variability effect (Hughes et al., 2009, 2011) and the high recency performance on auditory TBR lists (Nicholls & Jones, 2002; see also Chapter 4). It therefore stands to reason that the first component of the auditory verbal STM process is the auditory cortex, where the acoustic input is categorized and segregated into streams of information (cf. Rauschecker & Scott, 2009). It is probably here that an auditory list consisting of items with great perceptual coherence is organized into a temporally extended object, as envisaged in Chapter 4. The primary visual cortex, on the other hand, plays little role in the visual verbal STM process. This is not to say that the visual cortex is not essential for identifying visual information in the first place, but it seems doubtful that this region is involved in verbal information buffering beyond this function. Only when the visual input is identified as verbal, which seems to happen in the anterior fusiform gyrus (Nobre, Allison, & McCarthy, 1994), does it become relevant for the verbal STM process. The next main node in the neural network recruited for verbal STM is the left planum temporale where the Sylvian-parietal-temporal region seems to be of particular importance. This region seems to be involved in integrating auditory perception with oral motor action (Hickok, 2009) but has also been reported to be active during reading (Buchsbaum, Olsen, Koch, Kohn, Kippenhan, & Berman, 2005). Thus, it seems reasonable to conclude that this region plays a crucial role in the upload of the perceived visual and auditory verbal information into an articulatory motor-plan.

With the Sylvian-parietal-temporal region linking perception to oral-motor action, it is the function of motor action regions that has to be considered next. Within the perceptual-gestural framework, oral motor planning regions are particularly involved in imposing coherence onto unconstrained verbal items, to enable maintenance of their order. As was demonstrated in Chapter 3, Broca's area seems to be among the regions likely to be involved in this process. Further, there is some evidence that Broca's area is involved in the generation of co-articulatory modifications between TBR items (e.g., Katz, 2000), which according to the perceptual-gestural view is a crucial skill for maintaining verbal list order over the short term. Indeed, one could speculate that TBS of the area reduces performance on dissimilar items while simultaneously improving performance on similar items (Chapter 3) because the stimulation somehow affects the nature of the co-articulatory transitions between successive items.

As has been emphasized at several points throughout the thesis, however, Broca's area cannot be considered the only area involved in the assembly of the motor plan for verbal list maintenance. Indeed, speech motor-planning deficits have been observed following damage to a vast number of brain regions, which according to the perceptual-gestural view implicates them in the verbal STM process. Thus, the left superior precentral gyrus of the insula has been shown to be involved in articulation of complex multi-syllabic words (Baldo, Wilkins, Ogar, Willock, & Dronkers, 2011), with damage to the region contributing to speech-planning impairments (Dronkers, 1996). Further regions that have been associated with speech-planning are the thalamus, where haemorrhages have been reported to cause speech-planning deficits (Ozeren, Koc, Demirkiran, Sönmezler, & Kibar, 2006), and the superior cerebellum and the supplementary motor area, which have been found active in fMRI studies of articulatory motor-planning (Liegeois & Morgan, 2012). In line with the view that these regions are involved in the verbal STM process, all three regions have also been found to be active during verbal STM tasks (cf. Awh, Jonides, Smith, Schumacher, Koeppe, & Katz, 1996).

Finally, there is a number of adjacent regions like the inferior cerebellum, the primary motor cortex and parts of the thalamus that are involved in speech production as opposed to speech planning (Liegeois & Morgan, 2012). As was discussed in Chapter 2, peripheral motoric impediments like chewing gum, tapping or anarthria (Baddeley & Willson, 1985) do significantly reduce verbal STM performance. It seems therefore reasonable to argue that even regions involved only in speech production play a role in the verbal STM process. On the other hand, Chapter 2 also demonstrated that peripheral motoric impediments do not specifically impact upon the PSE or on STM tasks requiring serial memory. Thus, if a brain region is indeed involved only in speech production, which seems to apply to the inferior cerebellum and the primary motor cortex, it is probably most accurate to regard it as a region to which the maintenance of the assembled speech motor-plan can be outsourced for optimal performance but which does not otherwise determine the nature of the STM process.

In order to test the accuracy of this neurological model, TMS seems once more an appropriate technique. For example, TMS could be used to confirm the role of the Sylvian parietal temporal region in this model (cf. also Buchsbaum & D'Esposito, 2008; Hickok, 2009). As it stands, current knowledge about the area suggests that its disruption should result in the survival of auditory recency but reduce overall recall performance on all types of verbal lists, because their maintenance could not be passed on to regions specialized in articulatory planning and production. It is not entirely clear whether inhibition or impairment of the Sylvian-parietal-temporal region should affect high recency on visual-vocalized, visual-mouthed, and lip-read lists. Yet, it stands to reason that since the objecthood of these lists seems to be primarily determined by a salient oral motor-processing activity being associated with the visual-verbal list presentation (Chapter 4)—a link that is evidently supported by the activity in the Sylvian-parietal-temporal region (Buchsbaum et al., 2005)—the likelihood of perceiving these lists as objects, and hence performance on these lists in recency, should diminish if the Sylvian-parietal-temporal region is inhibited or damaged.

Another series of TMS studies could follow up on the observation that while concurrent articulation reduced both the PSE and overall performance, inhibition of Broca’s area only affected the size of the PSE (Chapter 3). This suggests that the impact that concurrent articulation has on STM involves regions beyond Broca’s area. Consequently, inhibiting the other speech-planning regions that have been mentioned here, like the supplementary motor area or the left superior precentral gyrus of the insula, should impact memory in ways complementary to the effects of inhibiting Broca’s area. Here it needs to be pointed out, however, that whereas using TMS to investigate the motor cortex should not pose a problem, stimulation of deeper regions like the insula might have to wait for full approval of the deep TMS (Harel et al., 2010) or transcranial pulsed ultrasound (Tufail et al., 2010) techniques that are currently being developed for non-invasive manipulation of subcortical brain regions.

Finally, inhibition of speech production areas, like the primary motor cortex, using TMS should be considered, too, as this might permit comparisons to the impact of concurrent motoric tasks like chewing or tapping on STM performance. Indeed, given that TBS of Broca’s area produced a reduction of the PSE in the absence of a reduction of overall performance, and chewing produced a reduction of overall performance in the absence of a reduction of the PSE on comparable verbal STM tasks, it seems reasonable to propose that the interplay of Broca’s area and some speech production region that is responsible for speaking, chewing and tapping is sufficient to accomplish some basic assembly of TBR verbal information into a coherent motor-plan and the subsequent maintenance thereof.

If confirmed, the proposed neurological model would demonstrate that the process of STM can be fully accounted for by the function of general receptive and productive mechanisms. This could expand upon current understanding of neurological disorders, linking language related disorders and verbal STM deficits closer together (e.g., Jodzio & Taraszkiewicz, 1999). Yet more importantly, one could begin to understand how the interaction of mechanisms designed to accomplish relatively simple processes like perception and action could give rise to complex, seemingly metaphysical, entities like memory and cognition. Indeed, if the admittedly rudimentary description of the interplay of regions responsible for perception and oral motor-action provided here should hold the key to verbal STM then perhaps that key could be also used to unlock the metaphorical riddle box that memory has proved to be ever since research into it began. Raising awareness not to take the metaphor of the riddle box of memory literally is probably the greatest contribution this thesis can offer to that endeavour.

BIBLIOGRAPHY

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,* 50-68.

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

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., Marshall, J. C., Shah, N. J., Fink, G. R., & Zilles, K. (2004). Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space - the roles of brodmann areas 44 and 45. *NeuroImage, 22,* 42-56.

Anderson, J. A. (1995). *An introduction to neural networks*. Cambridge, MA: MIT Press.

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

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

Aziz-Zadeh, L., Cattaneo, L., Rochat, M., & Rizzolatti, G. (2005). Covert speech arrest induced by rTMS over both motor and nonmotor left hemisphere frontal sites. *Journal of Cognitive Neuroscience, 17*, 928-938.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *The Quarterly Journal of Experimental Psychology, 20,* 249-263.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D. (1992). Is working memory working? The fifteenth Bartlett lecture. *Quarterly Journal of Experimental Psychology, 44A*, 1–31.

Baddeley, A. D. (1997). *Human memory: Theory and Practice (Revised Edition).* Hove: Psychology Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-432.

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

Baddeley, A. D. (2007). *Working memory, thought and action.* Oxford, England: Oxford University Press.

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

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp 47-89). New York: Academic Press.

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

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

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*, 575-589.

Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language, 24,* 490-502.

Baker, J. R., Bezance, J. B., Zellaby, E., & Aggleton, J. P. (2004). Chewing gum can produce context-dependent effects upon memory. *Appetite, 43,* 207–210.

Baldo, J. V., Wilkins, D. P., Ogar, J., Willock, S., & Dronkers, N. F. (2011). Role of the precentral gyrus of the insula in complex articulation. *Cortex, 47,* 800-807.

Basso, A., Spinnler, H., Vallar, G., & Zanobio, M. E. (1982). Left hemisphere damage and selective impairment of auditory verbal short-term memory. A case study. *Neuropsychologia*, *20*, 263-274.

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

Beaman, C. P., & Jones, D. M. (1997). The role of serial order in the irrelevant speech effect: Tests of the changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 459-471.

Beaman, C. P., Neath, I., & Surprenant, A. M. (2007). In D. S. McNamara & J. G. Trafton (Eds.), Phonological similarity effects without a phonological store: An individual differences model. *Proceedings of the 29th Annual Conference of the Cognitive Science Society.* (pp 89-94). Austin, TX: Cognitive Science Society.

Bishop, D. V. M., & Robson, J. (1989). Unimpaired short-term memory and rhyme judgement in congenitally speechless individuals: implications for the notion of "articulatory coding". *Quarterly Journal of Experimental Psychology*, *41*, 123-140.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organisation of sound*. Cambridge, MA: MIT Press.

Buchsbaum, B. R., & D’Esposito, M. (2008). The search for the phonological store:

From loop to convolution. *Journal of Cognitive Neuroscience, 20*, 762–778.

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

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

Buschke, H. (1963). Relative retention in immediate memory determined by the missing scan method. *Nature, 200,* 1129-1130.

Campbell, R., & Dodd, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology, 32*, 85-99.

Carello, C., LeVasseur, V. M., & Schmidt, R.C. (2002). Movement sequencing and phonological fluency in (putatively) nonimpaired readers. *Psychological Science, 13,* 375-379.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

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

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-114.

Cowan, N., Saults, J. S., & Brown, G. D. A. (2004). On the auditory modality superiority effect in serial recall: separating input and output factors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 639–644.

Crowder, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology, 1*, 157-178.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS).

Perception and Psychophysics, 5, 365-373.

Davis, C., Kleinman, J. T., Newhart, M., Gingis, L., Pawlak, M., & Hillis, A. E. (2008). Speech and language functions that require a functioning Broca's area. *Brain and Language, 105,* 50-58.

de Gelder, B., & Vroomen, J. (1992). Abstract versus modality-specific memory representations in processing auditory and visual speech. *Memory & Cognition, 20*, 533–538.

Della Sala, S., & Logie, R. H. (1997). Impairments of methodology and theory in cognitive neuropsychology: A case for rehabilitation? *Neuropsychological Rehabilitation, 7*, 367-385.

D’Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 362,* 761–772.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature, 384*, 159-161.

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,* 624–634.

Ellis, N. R. (1969). Evidence for two storage processes in short-term memory. *Journal of Experimental Psychology, 50,* 390-391.

Epstein, C. M., Lah, J. J., Meador, K., Weissman, J. D., Gaitan, L. E., & Dihenia, B. (1996). Optimum stimulus parameters for lateralized suppression of speech with magnetic brain stimulation. *Neurology, 47*, 1590-1593.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59-79.

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15,* 469–479.

Freeman, G. L. (1940). Dr. Hollingworth on chewing as a technique of relaxation. *Psychological Review, 47,* 491-493.

Friedman, L., Kenny, J. T., Wise, A. L., Wu, D., Stuve, T. A., Miller, D. A., Jesberger, J. A., & Lewin, J. S. (1998). Brain activation during silent word generation evaluated with functional MRI. *Brain and Language, 64*, 231-256.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20,* 1-55.

Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12*, 3-15.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279.

Greene, R. L. (1987). Stimulus suffixes and visual presentation. *Memory* & *Cognition, 5*, 497-503.

Greene, R. L., & Crowder, R. G. (1984). Modality and suffix effects in the absence of auditory stimulation. *Journal of Verbal Learning and Verbal Behavior, 23,* 371-382.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences, 12,* 474-480.

Guerard, K., Jalbert, A., Neath, I., Surprenant, A. M., & Bireta, T. J. (2009). Irrelevant tapping and the acoustic confusion effect: The effect of spatial complexity. *Experimental Psychology, 56,* 367-374.

Harel, E.V., Zangen, A., Roth, Y., Reti, I. M., Braw, Y., & Levkovitz, Y. (2011). H-coil repetitive Transcranial Magnetic Stimulation for treatment of Bipolar Depression: An Add-on, Safety and Feasibility study. *World Journal of Biological Psychiatry, 12,* 119-126.

Harvey, A. J., & Beaman, C. P. (2007). Input and Output modality effects in immediate serial recall. *Memory, 15,* 693-700.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36,* 73-137.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology, 49,* 80-115.

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

Huang, Y., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron, 45,* 201-206.

Hubl, D., Nyffeler, T., Wurtz, P., Chaves, S., Pflugshaupt, T., Lüthi, M., von Wartburg, R., Wiest, R., Dierks, T., Strik, W. K., Hess, C. W., & Müri, R. M. (2008). Time course of blood oxygenation level-dependent signal response after theta burst transcranial magnetic stimulation of the frontal eye field. *Neuroscience, 151*, 921-928.

Hughes, R. W., & Jones, D. M. (2005). The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 316-327.

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,* 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,* 1435–1447.

James, W. (1890). *The principles of psychology.* New York: Holt, Rinehart & Winston.

Jodzio, K., & Taraszkiewicz, W. (1999). Short-term memory impairment: Evidence from aphasia. *Psychology of Language and Communication, 3*, 39-48.

Johnson, A., & Miles, C. (2007). Evidence against memorial facilitation and context-dependent memory effects through the chewing of gum. *Appetite, 48,* 394-396.

Johnson, A. J., & Miles, C. (2008). Chewing gum and context-dependent memory: The independent roles of chewing gum and mint flavour. *British Journal of Psychology, 99,* 293–306.

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,* 265-281.

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

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*, 656-674.

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin and Review, 7,* 550-558.

Katz, W. F. (2000). Anticipatory coarticulation and aphasia: Implications for phonetic theories. *Journal of Phonetics, 28,* 313-334.

Kennedy, D. O., Scholey, A. B. & Wesnes, K. (2000). The dose dependent cognitive effects of acute administration of Ginkgo biloba to healthy young volunteers. *Psychopharmacology, 151,* 416-423.

Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve working memory of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General, 112,* 240-264.

Kirschen, M. P., Davis-Ratner, M. S., Jerde, T. E., Schraedley-Desmond, P., & Desmond, J. E. (2006). Enhancement of phonological memory following transcranial magnetic stimulation (TMS). *Behavioural Neurology, 17,* 187-194.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace & World.

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

LeCompte, D. C. (1996). Irrelevant speech, serial rehearsal and temporal distinctiveness: A new approach to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1154-1165.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory and Cognition, 24*, 305-321.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate serial recall. *Quarterly Journal of Experimental Psychology, 53,* 626–646.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *The Psychology of Learning and Motivation, 49,* 1-48.

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

Liégeois, F. J., & Morgan, A. T. (2012). Neural bases of childhood speech disorders: lateralization and plasticity for speech functions during development. *Neuroscience and Biobehavioral Reviews, 36*, 439-458.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 436-448.

Maylor, E. A., Vousden, J. I., & Brown, G. D. A. (1999). Adult age differences in short-term memory for serial order: *Data and a model. Psychology and Aging, 14*, 572-594.

Metcalfe, J., Glavanov, D., & Murdock, M. (1981). Spatial and temporal processing in the auditory and visual modalities. *Memory and Cognition, 9,* 351-359.

Miles, C., & Johnson, A.J. (2007). Chewing gum and context-dependent memory effects: A re-examination. *Appetite, 48,* 154-158.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Murdock, B. B., Jr. (1968). Serial order effects in short-term memory. *Journal of Experimental Psychology, 76*, 1–15.

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,* 594-598.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78,* 679-684.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18,* 251–269.

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin and Review, 7,* 403–423.

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.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature, 372,* 260-263.

Nyffeler, T., Cazzoli, D., Wurtz, P., Lüthi, M., Von Wartburg, R., Chaves, S., Déruaz, A., Hess, C. W., & Müri, R. M. (2008). Neglect-like visual exploration behaviour after theta burst transcranial magnetic stimulation of the right posterior parietal cortex. *European Journal of Neuroscience, 27,* 1809-1813.

Nyffeler, T., Wurtz, P., Lüscher, H., Hess, C. W., Senn, W., Pflugshaupt, T., Von Wartburg, R., Lüthi, M., & Müri, R. M. (2006). Repetitive TMS over the human oculomotor cortex: Comparison of 1-hz and theta burst stimulation. *Neuroscience Letters, 409,* 57-60.

Ogar, J., Slama, H., Dronkers, N., Amici, S., & Gorno-Tempini, M. L. (2005). Apraxia of speech: An overview. *Neurocase, 11,* 427-432.

Overman, A.A., Sun, J., Golding, A.C., & Prevost, D. (2009). Chewing gum does not induce context-dependent memory when flavor is held constant. *Appetite, 53,* 253-255.

Ozeren, A., Koc, F., Demirkiran, M., Sonmezler, D., & Kibar, M. (2006). Global aphasia due to left thalamic hemorrhage. *Neurology India, 54*, 415-417.

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

Penney, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin*, *82*, 68-84.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17,* 398-422.

Penney, C. G., & Blackwood P. A. (1989). Recall mode and recency in immediate serial recall: Computer users beware! *Bulletin of the Psychonomtc Soctety, 27,* 545-547.

Pickering, M. J., & Garrod S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*, 105-110.

Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*, 718–724.

Romero, L., Walsh, V., & Papagno, C. (2006). The neural correlates of phonological short-term memory: A repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience, 18*, 1147-1155.

Saito, S. (1994). What effect can rhythmic finger tapping have on the phonological similarity effect? *Memory & Cognition, 22*, 181–187.

Sakamoto, K., Nakata, H., & Kakigi, R. (2009). The effect of mastication on human cognitive processing: A study using event-related potentials. *Clinical Neurophysiology, 120,* 41-50.

Shallice, T., & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia, 15*, 729–735.

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition, 40,* 177–190.

Spelke, E. S., (1990). Principles of object perception. *Cognitive Science, 14,* 29-56.

Stephens, R., & Tunney, R. J. (2004). Role of glucose in chewing gum-related facilitation of cognitive function. *Appetite, 43,* 211–213.

Stewart, L., Walsh, V., Frith, U., & Rothwell, J. C. (2001). TMS produces two dissociable types of speech disruption. *NeuroImage, 13,* 472-478.

Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., & Mattingley, J. B. (2005). Simple metric for scaling motor threshold based on scalp-cortex distance: Application to studies using transcranial magnetic stimulation. *Journal of Neurophysiology, 94,* 4520-4527.

Street, R. L. Jr., & Cappella, J. N. (1989). Social and linguistic factors influencing adaptation in children’s speech. *Journal of Psycholinguistic Research, 18,* 497-519.

Surprenant, A. M., Kelley, M. R., Farley, L. A., & Neath, I. (2005). Fill-in and infill errors in order memory. *Memory, 13,* 267–273.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General, 138*, 236–257.

Tremblay, S., Parmentier, F. B. R., Guérard, K., Nicholls, A. P., & Jones, D. M. (2006). A spatial modality effect in serial memory. *Journal of Experimental*

Psychology: Learning, Memory & Cognition, 32, 1208-1215.

Tucha, O., Mecklinger, L., Maier, K., Hammerl, M. & Lange, K. W. (2004). Chewing gum differentially affects aspects of attention in healthy subjects. *Appetite, 42,* 327-329.

Tufail, Y., Matyushov, A., Baldwin, N., Tauchmann, M. L., Georges, J., Yoshihiro, A., Helms Tillery, S. I., & Tyler, W. J. (2010). Transcranial pulsed ultrasound stimulates intact brain circuits. *Neuron, 66*, 681-694.

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken Expec tations: Violation of expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 164-177.

Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for Cognitive Neuropsychology. *Cognitive Neuropsychology, 23,* 135-155.

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior, 23*, 151-161.

Verbruggen, F., Aron, A. R., Stevens, M. A., & Chambers, C. D. (2010). Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 13966-139671.

Vishwanathan, A., Bi, G. Q., & Zeringue, H. C. (2011). Ring-shaped neuronal networks: a platform to study persistent activity. *Lab Chip, 11,* 1081-1088.

Waters, G. S., Rochon, E., & Caplan, D. (1992). Role of high level speech planning in rehearsal: Evidence from patients with apraxia of speech. *Journal of Memory and Language, 31,* 54-73.

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience, 16*, 978-987.

Wilkinson, L., Scholey, A., & Wesnes, K. (2002). Chewing gum selectively improves aspects of memory in healthy volunteers. *Appetite, 38,* 235-236.

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 & Language, 58*, 48-65.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin, 133,* 273-293.

APPENDIX

TMS Pilot Experiment

This pilot experiment was conducted to validate the behavioural task to be used in the TMS study described in Chapter 3. Specifically the aim was to establish whether the phonological similarity effect can be observed with the probed order task, and whether the effect can be attenuated with concurrent articulation.

Method

Participants

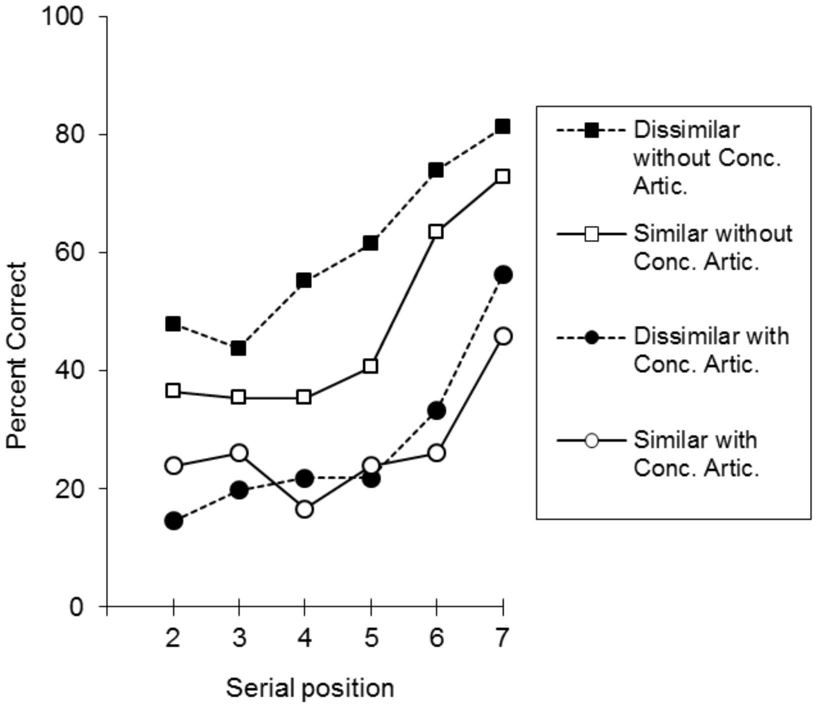
The participants were 24 volunteers from Cardiff University (2 male, aged 18-24, mean: 19.17). All participants had normal or corrected-to-normal vision and were native English speakers. None of these participants took part in the later TMS study.

Materials, Design and Procedure

There was a close match between the present study and the later conducted TMS experiment. The stimuli used in this experiment, their presentation rate and the response recording were the same as in the later experiment. However, the interferer variable had only 2 levels (concurrent articulation present and absent), although instructions about how to engage in suppression matched those from the main experiment. Another deviation from the later TMS experiment was the trial number. There were overall 96 trials, so that each serial position (except the first) would be probed 4 times in each condition. The trials were blocked by concurrent articulation, with 24 phonologically similar and 24 phonologically dissimilar trials being presented in quasi-random order in the concurrent articulation present and absent blocks. There was no pause between the blocks. The entire experiment was conducted in a single session and took about 30 minutes.

Results and Discussion

The average performance of each participant on each probed serial position is depicted in Figure 13.



<u>Figure 13</u>: The average performance accuracy on the probed order task in the presence and absence of concurrent articulation.

A 2 (similarity) by 2 (suppression) by 6 (serial position) within-participant ANOVA revealed a significant effect of serial position, *F* (5, 115) = 24.95, *MSE* = 0.07, $p < .05$, $\eta_p^2 = .52$. There was also a main effect of similarity, *F* (1, 23) = 10.28, *MSE* = 0.07, $p < .05$, $\eta_p^2 = .31$, indicating that overall performance on dissimilar lists was better than on similar lists. There was furthermore a main effect of concurrent articulation, *F* (1, 23) = 71.15, *MSE* = 0.14, $p < .05$, $\eta_p^2 = .76$, indicating that performance under concurrent articulation was reduced. Concurrent articulation significantly interacted with phonological similarity, *F* (1, 23) = 12.23, *MSE* = 0.05, $p < .05$, $\eta_p^2 = .35$. This, in conjunction with Figure 13 indicates that the phonological similarity effect was attenuated under concurrent articulation.

These results confirm that the phonological similarity effect can be observed with probed order recall. Furthermore, concurrent articulation reduces the effect in this paradigm. It thus stands to reason that TMS of an articulatory planning area will also reduce the phonological similarity effect in this paradigm, a hypothesis addressed in the main experiment.