Faculté de Psychologie, Logopédie, et Sciences de l'Éducation

The nature of serial order coding in verbal working memory

Thèse présentée par

Hong XIAO

En vue de l'obtention du grade académique de
Doctorat en Sciences Psychologiques

Sous la supervision de
Steve MAJERUS et Lucie ATTOUT

Année académique 2025-2026

PsyNCog
Psychology & Neuroscience of Cognition

Summary

This PhD thesis investigates how serial order is coded in verbal working memory (WM) from linguistic, spatial, and temporal perspectives. Three experimental studies employed the immediate serial recall task to test the contribution of syntactic legality (Study 1), spatial manipulation (Study 2), and temporal regularity (Study 3) to serial order performance. Results indicate a robust effect of syntactic knowledge, a context-dependent effect of spatial factors, and no reliable effect of temporal regularity under the conditions tested. Overall, the findings support accounts of WM as a dynamic, attention-shaped system grounded in long-term memory, and point to serial order as a flexible, multi-source representation rather than a fixed code or a dedicated buffer.

Acknowledgement

My PhD journey in Belgium has passed in the blink of an eye. I sincerely thank everyone who has helped and supported me along the way. Without your guidance and encouragement, I would not have been able to complete this doctoral dissertation on my own.

First and foremost, I would like to express my deepest gratitude to my supervisor, **Steve Majerus**. Thank you for accepting my PhD application and for giving me the opportunity to complete my doctorate under your supervision. You are an exceptional scholar, and your passion and dedication to research have profoundly influenced me. Whenever I encountered difficulties in my academic work, you were always the first to offer help and clarification. At the same time, you trusted me and encouraged me to develop independence and intellectual freedom in my research, which helped me build professionalism and confidence as a researcher. I am also very grateful for your kindness. As an international student who does not speak French, I deeply appreciated your patience in communicating with me. At the end of every meeting, your smile, encouragement, and affirmation meant a great deal to me and gave me tremendous motivation. I truly value the chance to work with such a dedicated scholar and supportive mentor.

I would also like to thank my co-supervisor, **Lucie Attout**, for her continuous support and assistance throughout my PhD. Thank you for our discussions and meetings, and for offering many valuable suggestions that greatly strengthened my work.

I am grateful to the members of my thesis committee, **Christel** and **Martine**, for following my progress over the past four years and for supporting the continuation of my doctoral research.

I also warmly thank **Prof. Guida** and **Prof. Fias** for agreeing to serve as external reviewers of my thesis, for reading my dissertation, and for traveling to attend my defense.

I would like to thank **Judith** and **Friederike** for their support and help with my research. It was a great pleasure to work with you, and I found our collaboration both enjoyable and inspiring.

I am grateful to the Chinese Scholarship Council (CSC) for providing the platform and financial support that made it possible for me to pursue my PhD abroad.

My sincere thanks go to all the students who participated in my studies. I would like to thank **Louis** for working with me during the first year of my PhD. Thank you all for contributing your time and effort to help collect the experimental data. I also thank all participants in my experiments – your participation and the data you provided made this thesis possible.

I would like to thank my colleagues – though I would rather call you my friends – for the friendship you have given me. You can finally take a deep breath now: no more speaking English all the time! Thank you for all the lunches we shared, and for the many conversations we had about research, politics, and entertainment. You are all outstanding scholars. Nathan, Arya, Charlène, Charline, David, Iden, Bastien, Chloé, Emma, Olivier, Arnaud, Paradise, Wenrui, thank you for your friendship. **Coline**, thank you for your kindness and help when I first arrived. **Pauline**, **Tania**, and **Claudia**, thank you for always treating me with such gentleness. **Robin** and **Benjamin**, you are among the most interesting people I have ever met.

I also want to thank my Chinese friends. Being alone in a foreign country, it is truly precious to be able to speak in our familiar dialects and share a sense of home. **Wen Xiang**, thank you for being my roommate and for tolerating my little habits. Thank you to the kind and brave **Jiayu**, who always brought me handmade buns and authentic food. Thank you, **Xiaobo**, my talkative online friend – although you often complain about work, you always listened to me and comforted me when I was struggling. Thank you, **Peng Wen**, **Yuxia**, and **An Ning**, for every holiday and every traditional Chinese festival we spent together.

To **my parents**, thank you for supporting my choices, for allowing me to pursue what I truly want, and for accepting, protecting, and caring for me. You have given me the drive to move forward, and you are also the harbor where I can always rest. I also thank my other **family members** for their care and concern over the years, and for always being there for my parents.

Thank you to **Jean** for giving me another home in Belgium, and for introducing me to a family that is so kind and wonderful. Thank you, **Alain** and **Monique**, for coming to attend my PhD defense.

谢谢这一路上所有的人和事，我不知道未来的路会是什么样，也不知道我们是否还会再次相聚，但是一切都会有痕迹，这些痕迹会成为涓涓细流滋养我，成为我的一部分。谢谢你们出现在我的生命中。

Enfin, je voudrais remercier tout ce que j’ai rencontré sur ce chemin, les personnes comme les moments. Je ne sais pas de quoi demain sera fait, ni si nos routes se croiseront encore, mais rien ne passe sans laisse rune empreinte. Ces empreintes, peu à peu, deviennent un mince filet d’eau – silencieux et fidèle – qui me nourrit, me façonne, et finit par se mêler à moi. Merci d’avoir traversé ma vie.

Finally, I would like to thank everything and everyone I have encountered on this path. I do not know what the future will look like, nor whether we will meet again, but everything leaves traces. These traces become a gentle stream that nourishes me and becomes part of who I am. Thank you for being part of my life.

肖红 Hong Xiao

2025.12.30 in Liège

…Mais il n’y a pas de limites pour aimer, et que m’importe de mal étreindre si je peux tout embrasser.

— Albert Camus, *L’Envers et l’Endroit*, Amour de vivre.

Table of Contents

General Introduction

Whenever someone asks me what my study is about, I usually give the same example: remembering a phone number. It's not enough to recall the digits themselves—you also need to keep them in the right order. If you swap a 3 and a 7, you'll end up calling the wrong person. My research looks at how we keep track of this order.

What makes this tricky to explain is that remembering order feels so natural. We just do it, without thinking. For a long time, researchers also treated order as if it simply "came along for the ride" when we remembered items. Only more recently did studies start to separate the two and show that remembering what something is and remembering when it came are not the same thing. Evidence from both behavioral and neuroscience studies suggests that order memory relies on its own mechanisms.

This thesis explores how people code serial order in verbal working memory. Since the twentieth century, researchers have proposed different ways this coding might work. Some studies find strong evidence for specialized, temporal, or linguistic order codes, but others fail to observe evidence for these codes when tasks feel more natural. In other words, direct proof of how order is coded is still missing.

Theoretical Introduction

Chapter 1

Verbal working memory and the problem of serial order

This chapter provides a stepwise introduction to serial order in verbal working memory (WM). We begin by introducing the classic WM model of Baddeley and Hitch and proposing the term serial order memory. Next, we define serial order information and distinguish it from item-level information, highlighting why order is a fundamental aspect of memory beyond simple item recall. Then, we review the classic methodologies used to study serial order memory, with a focus on the immediate serial recall task, which serves as the principal paradigm in this thesis. We then summarize the behavioral benchmarks derived from serial order tasks, outlining the characteristic patterns and errors that inform theoretical accounts.

Verbal working memory and serial order

In a seminal work, Lashley (1951) emphasized the fundamental cognitive ability to process serial order and proposed the problem of serial order in behavior. From then on, the ability to code and maintain information about the sequential order of events in WM became a crucial research topic. Serial order processing is considered particularly important for advanced cognitive functions, such as vocabulary acquisition and language production (e.g., Burgess & Hitch, 1999). Besides, in the nonverbal domain, memory for serial order is thought to be crucial for acquiring many motor skills and social behaviors, which are often learned by observing and imitating sequences of actions performed by others (Agam et al., 2005, 2010; Baddeley, 2007).

The original multicomponent model of Baddeley and Hitch (1974) proposed three components in WM: (1) *the phonological loop*, which serves to hold and keep verbal information active in memory; (2) *the visuospatial sketchpad*, for retaining visual and spatial information; and (3) *the central* *executive*, which serves to manipulate data in memory and control the system (see Figure 1.1).

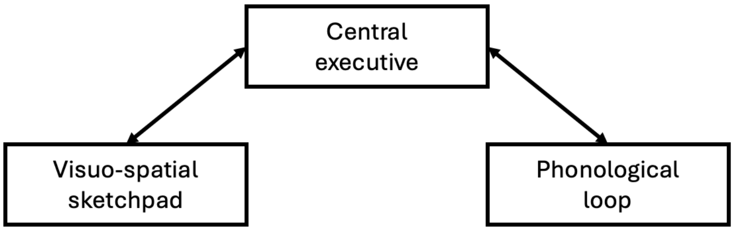

Figure 1.1. The initial working memory model, adapted from Baddeley and Hitch (1974).

The phonological loop is a temporary buffer that holds verbal information and allows for more prolonged maintenance via the articulatory rehearsal mechanism. Meanwhile, the phonological code has been shown to explain the wealth of serial recall data at a qualitative level, including the effects of phonological similarity, word length, articulatory suppression, subvocal rehearsal, and presentation modality. However, it does not make an explicit distinction between item and serial order information, as both types of information are considered to be represented via phonological codes. Furthermore, the representation of serial order information only via phonological mechanisms cannot account for the complexity of serial order phenomena that I will describe later. The question of representing serial order information remains a major theoretical challenge for this as well as other models of working memory (Majerus, 2019).

Serial order memory vs. item memory

Before examining the mechanisms underlying the representation of serial order information, it is important to clarify its distinction from item information (e.g., Healy, 1974; Murdock, 1976; Nairne & Kelley, 2004). Item-level information typically reflects the verbal units that constitute the memoranda, whereas serial order information refers to the temporal succession of these units within a list of stimuli (e.g., Lee & Estes, 1981; Nairne & Kelley, 2004). For example, for the consonant list "*B-S-J-R-A*", the item-level information is the individual consonants (e.g., there was a consonant "*B*" in the list). In contrast, the order-level information is the order of consonants (e.g., "*J*" is followed by "*R*", and "*R*" is followed by "*A*").

Based on Majerus (2019) (see Table 1.1), item information typically involves language representations of individual words but can also represent sub-word level information, such as syllables and phonemes in the case of nonword memoranda (e.g., Lee & Estes, 1981; Nairne & Kelly, 2004). More recently, evidence has also pointed to the contribution of supralexical information, including morphosyntactic representations (Guitard et al., 2025; Poirier et al., 2015). Serial order information, in contrast, concerns the unpredictable and arbitrary temporal order of occurrence of each word in the list or each phoneme in the novel non-word string (Gathercole, 1995; Gathercole et al., 1999; Majerus et al., 2004, 2012). These distinctions are elaborated in the following chapter.

Table 1.1. The different levels of item and serial order information encoded in verbal WM tasks by Majerus (2019).

Item level	Long-term serial order level	Short-term serial order level
Word strings	Serial position of phonemes within words	Serial position of words in a list
Nonword strings	Serial position of phonemes for nonwords with highly familiar phoneme co-occurrences	Serial position of nonword strings in a list
		Serial position of phonemes for nonwords with less familiar or novel phoneme co-occurrences

Behavioral Evidence

Several behavioral studies support a distinction of processes involved in WM for item and serial order information. Early work by Bjork and Healy (1974) used four-consonant lists with three retention intervals and demonstrated that order information decayed more rapidly than item information as retention interval increased. Sternberg (1967) presented lists of items followed by a probe digit and asked participants either to judge whether the probed item had happened in the list (item recognition) or to report the item adjacent to the probe in the list (order recognition). Reaction time was faster for item recognition than for order recognition. Converging evidence comes from studies comparing the temporal dynamics (speed/accuracy) of an item probe task (e.g., judging which of "A-B" appeared in the list) with an order probe task (e.g., judging which of "A-B" occurred more recently), which have consistently indicated slower and more effortful processes in order memory than for item memory (Matthews & Henderson, 1970; McElree & Dosher, 1993). Similarly, presenting items in small temporal groups enhances serial order recognition but not item recognition (Henson et al., 2003). Rhythmic concurrent tasks, such as finger tapping, interfere with order memory but not item memory (Engelkamp & Dehn, 2000; Henson et al., 2003). Further evidence comes from studies on intra-list phonemic similarity, which impairs serial order recall but facilitates item recall (Crowder, 1979; Lian et al., 2004).

Other studies have demonstrated that item recall is particularly sensitive to the linguistic properties and associated long-term knowledge of memoranda. Item WM performance is enhanced for word lists compared to nonword lists, a phenomenon known as the lexicality effect in verbal WM (Brener, 1940; Gathercole et al., 2001; Hulme et al., 1991, 1995; Kowialiewski & Majerus, 2018; Majerus & Van der Linden, 2003; Patterson et al., 1994; Saint-Aubin & Poirier, 2000). Recall performance also improves for words with richer or more accessible lexical and semantic representations, such as high-frequency words versus low-frequency words (Gregg et al., 1989; Hulme et al., 1997; Kowialiewski & Majerus, 2018; Poirier & Saint-Aubin, 1996; Roodenrys et al., 1994; Watkins & Watkins, 1977), or concrete words versus abstract words (Bourassa & Besner, 1994; Hulme et al., 1997, 2003; Kowialiewski & Majerus, 2018; Majerus & Van der Linden, 2003; Walker & Hulme, 1999). These results have been observed based on basic item vs. order recall measures during a verbal WM task, or based on more complex process-dissociation procedures (Nairne & Kelley, 2004).

Moreover, if there is any effect of linguistic knowledge on serial order memory, then it is a detrimental effect observed for recall of semantically similar vs. dissimilar word lists (Guérard & Saint-Aubin, 2012; Neale & Tehan, 2007; Murdock, 1976; Saint-Aubin & Poirier, 1999a, 1999b). For instance, Saint-Aubin et al. (2005) reported that semantic relatedness impaired serial order memory, even when the number of recalled items was equated across semantically related and unrelated lists. Similar effects were observed by Tse (2009) and Tse et al. (2011), who demonstrated that both categorical and associative forms of semantic relatedness disrupt order memory, independently of item recall performance. More recently, Guitard et al. (2025) showed that participants produced more migration errors when recalling semantically related lists compared to unrelated lists. At the same time, it should be noticed that other studies have observed different findings, with no impact of semantic similarity on serial order recall (Nairne & Kelley, 2004), or even with improved serial order recall for semantically related word lists compared to unrelated ones, but only when the similar items were presented in a grouped manner (Kowialiewski et al., 2024).

In sum, these behavioral findings suggest that item and serial order information are dissociable and that an effect of linguistic knowledge on WM abilities is most reliably observed for item WM.

Neuroimaging and neuropsychological evidence

The dissociation between item and serial order memory is also supported by neuroimaging studies. The fronto-parietal cortex has been shown to play a central role in serial order memory. This network includes the dorsolateral prefrontal cortex (*DLPFC*), which may play a role in explicit temporal context encoding or retrieval (Barbey et al., 2013; Miller & Cohen, 2001), and more importantly, the bilateral intraparietal sulci (*IPS*; Attout et al., 2022; Henson et al., 2000; Majerus et al., 2006, 2010; Marshuetz et al., 2000, 2006). More specifically, studies controlling for task difficulty and attentional requirements associated with item vs. serial order probe recognition tasks revealed an involvement of portions of the IPS in serial order WM conditions (Attout et al., 2014, 2019, 2022; Cristoforetti et al., 2022; Majerus, Poncelet, Van der Linden, et al., 2006; Majerus et al., 2010), especially the right posterior IPS (Attout et al., 2022; Cristoforetti et al., 2022; Fias et al., 2007). Interestingly, IPS involvement has also been observed for serial order probe recognition tasks involving sequences of visual information, such as unfamiliar faces (Majerus et al., 2007, 2010). In addition, some studies have highlighted a role of the hippocampus in serial order memory (Cristoforetti et al., 2022; Fortin et al., 2002; Roberts et al., 2018), although this is not always observed in verbal serial order WM tasks (Attout et al., 2019, 2020; Majerus, Poncelet, Van der

Linden, et al., 2006; Majerus et al., 2010). On the other hand, verbal item WM has been shown to involve dorsal (phonological) and ventral (semantic) pathways of the language network, and the contents of items maintained in WM are decodable based on multivariate neural patterns in these language pathways.

Evidence from neuropsychological studies further supports the dissociations between item and serial order WM, with verbal item WM deficits being more dependent on underlying language impairment (Majerus, 2008; Majerus et al., 2015). For example, adult patients with progressive semantic impairment exhibit poor item recall in word serial recall tasks, whereas serial order recall remains intact (Majerus, Norris, et al., 2007). More generally, patients with aphasia can show selective verbal item or serial order WM deficits (Majerus et al., 2009, 2015).

In sum, current behavioral, neuroimaging, and neuropsychological studies suggest that encoding and retention of serial order information in verbal WM is coded with a specific cognitive mechanism related to the attention system, as well as some regions related to the language system, sharing with item information.

Immediate serial recall task

Memory for serial order is particularly crucial in the verbal domain. For instance, vocabulary acquisition depends on remembering the phonemes that constitute new words in the correct order (Baddeley et al., 1998; Gathercole & Baddeley, 1990; Leclercq & Majerus, 2010; Majerus, Poncelet, Elsen, et al., 2006), or meaningful sentences require words in an order obeying the correct grammar (e.g., Bonhage et al., 2014; Scheerer-Neumann, 1981). This reliance on order is further reflected in the tendency that individuals recall information in a forward order automatically, even without explicit instructions, suggesting that forward-order recall may reflect a general organizational principle in verbal WM (Bhatarah et al., 2006, 2008, 2009; Grenfell-Essam & Ward, 2012; Howard & Kahana, 1999; Kahana, 1996; Ward et al., 2010). In these studies, memory for serial order has typically been examined using an *immediate serial recall (ISR)* task, in which participants are asked to recall short sequences of familiar verbal items (e.g., letters, digits, or words) in the temporal order presented.

Another task, the immediate serial recognition task, has also been widely used, in which, after the presentation of a memory list, a probe list is presented, and participants need to determine whether the order of these items is the same or different from the memory list. Compared to the immediate serial recognition task, the ISR task provides a more direct window into how verbal WM maintains both item and serial order information, and analysis of recall performance on serial position enables researchers to examine the cognitive processes underlying serial order memory (e.g., Henson, 1998; Lewandowsky & Farrell, 2008). The serial recall curve in ISR tasks exhibits characteristic serial position effects, with enhanced recall for items at the beginning of the list (the primacy effect) and at the end (the recency effect). For the ISR task, serial order recall performance can be calculated with a specific score, which is the proportion of the sum of items recalled at the correct position divided by the sum of items recalled at whatever position. More importantly, item and serial order recall performance can be studied via the specific errors in ISR. For example, omissions are typically used to measure item recall performance. Transposition errors, which are items recalled in the wrong order, in particular, provide insights into the mechanisms of serial order representation.

Behavioral benchmarks in serial order memory

Empirical research has identified a set of robust regularities—referred to as benchmarks—that characterize serial order memory. These benchmarks, primarily derived from the ISR task, provide critical constraints for theoretical models of verbal WM, as well as serial order representation and maintenance (Oberauer et al., 2018).

Primacy and Recency Effects

The pattern between the accuracy of retrieval for items and their positions in the list is known as the serial position curve. For verbal lists, the recall accuracy is shown to be better for those items presented at the start of the list, called the primacy effect, and at the end of the list, called the recency effect. Besides, this serial position curve is also found in nonverbal stimulus sequences, such as visuospatial locations (Avons, 2007; Farrand et al., 2001; Guérard & Tremblay, 2008; Jones et al., 1995; Smyth & Scholey, 1996; Tremblay et al., 2006), auditory–spatial locations (Parmentier & Jones, 2000;

Tremblay et al., 2006), and unfamiliar faces (Smyth et al., 2005; Ward et al., 2005). Because of their universality, primacy and recency effects become one of the most important benchmarks in serial order memory. Additionally, studies have also shown that the recall timing latency to initiate the first item in a verbal sequence is far longer than for the rest items, and the recall latencies for these subsequent serial positions follow an inverted U-shaped profile (Anderson et al., 1998; Farrell, 2008; Farrell et al., 2011; Farrell & Lewandowsky, 2004; Haberlandt et al., 2005; Maybery et al., 2002; Parmentier & Maybery, 2008; Thomas et al., 2003).

Transposition Errors

In ISR tasks, recall errors often involve the confusion between the target item and other items from the same memory set (Aaronson, 1968; Guérard & Tremblay, 2008; Henson, 1996; Smyth et al., 2005). These errors are called transposition errors. Most transpositions happen in the adjacent items (e.g., one position before or after the target item) by switching the positions of these two items. This tendency for transpositions around their correct positions is known as the *locality constraint* (Henson, 1996; Hurlstone & Hitch, 2015; Lee & Estes, 1977; Nairne et al., 1991; Smyth et al., 2005). More specifically, transpositions with negative distance values, in which items are recalled before their correct positions, are known as *anticipation* errors. Transpositions with positive distance values, in which items are recalled after their correct positions, are known as *postponement* errors. Additionally, when an item is skipped during recall—for example, in the memory list "*A–B–C–D–E–F–G*", if letter *B* is omitted and the recall sequence begins as "*A–C…*", the next recalled item may either be the skipped item ("*A–C–B…*") or the subsequent item in the original list ("*A–C–D…*"). The former case is referred to as a *fill-in* error, whereas the latter is termed an *infill* error (Surprenant et al., 2005). Items recalled in their correct positions are represented by a distance value of zero. Transpositions are typically measured in terms of transposition gradients that plot the probability of transpositions as a function of distance.

Temporal Grouping Effects

Memory performance improves when stimuli are grouped by short inter-stimulus intervals (*ISIs*) and separated by longer between-group intervals (Frankish, 1989; Hartley et al., 2016; Hitch, 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b), and this advantage is accompanied by systematic changes in order errors (Ryan, 1969a, 1969b). Specifically, within-group primacy and recency effects emerge in addition to the main primacy and recency effects that characterize the entire lists (Frankish, 1989; Hartley et al., 2016; Hitch et al., 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b). Temporal grouping also influences transposition errors, giving rise to interposition errors (Hartley et al., 2016; Henson, 1999). For example, the memoranda "*A-B-C, D-E-F*" would be recalled as "*A-E-C, D-B-F*". Serial order errors tend to involve within-group items rather than items stemming from different groups (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Hartley et al., 2016; Henson, 1999; Ng & Maybery, 2002, 2005; Ryan, 1969a, 1969b). Temporal grouping effects align with models of serial order memory that posit a hierarchical temporal structure, with order encoded simultaneously at multiple levels, one tracking item position within groups and another encoding group position at the whole-list level (Brown et al., 2000; Hartley et al., 2016).

Chapter summary

In this chapter, we introduced the concept of serial order memory and its distinction from item memory in verbal WM tasks. Based on evidence from behavioral, neuropsychological, and neuroimaging data, serial order and item information in verbal WM are considered to be represented by distinct cognitive and neural processes. We also presented the ISR task and several associated benchmark effects that characterize serial order memory. In the next chapter, we will introduce the main theoretical models of serial order memory.

Chapter 2

Theoretical models of serial order WM

Different theoretical models have been proposed to account for serial order memory and associated benchmark effects. In this chapter, we will introduce two major families of models, the associative vs. positional accounts of serial order WM (see Figure 1.3).

Associative Chaining Models

One of the earliest theoretical approaches to serial order is the associative chaining account (Ebbinghaus, 1964; Lashley, 1951a; see Figure 1.2 (A)). The idea of this account is that serial order is maintained through the associations between successive items, with the former recalled item serving as a retrieval cue for the latter. Thus, serial order recall is achieved by traversing these associations. For example, a digit sequence such as "*3-6-2-7-1*" could be represented as "*3* precedes *6*, *6* precedes *2*," and so on. Associative chaining provides an intuitively appealing explanation for certain empirical findings, such as the serial position effect, the phonological similarity effects, the Hebb repetition effect, and so on (e.g., Ebbinghaus, 1964; Lewandowsky & Murdock, 1989; Murdock, 1993, 1995).

However, this associative chaining account faces some objections. One of the most important objections is, when order recall fails at one position of the list, chaining models consider that the item-item association would be broken, and it should not be possible anymore to recall any subsequent item in the list. This assumption is contradicted by behavioral evidence, among which is most critically the recency effect in serial position curves. The theory of the distributed associative memory (*TODAM*) model (Lewandowsky & Murdock, 1989) proposes to solve this shortcoming in the following manner. The central idea of TODAM is that items are represented as vectors composed of random elements, and order information is encoded by combining the vectors of contiguous items. Serial order is represented in memory by adding both the item and associative vector representations to a common memory vector. During serial recall, retrieval begins with the first item, which then serves as a cue for the second item; the recall of these two items jointly cues the third, and the process continues in this manner. For example, in the list "*A-B-C-D-E*", "*A*" may cue retrieval of "*B*", the pair "*A-B*" may cue "*C*", and so on; see Figure 1.2 (B). Because representations in TODAM are distributed, the retrieval cue is not an exact replica of an item, as in classical chaining models, but rather a blurred approximation. Moreover, noise is introduced during both encoding and retrieval. When an item is retrieved successfully, the noisy output vector must be deblurred before the item can serve as an effective cue for the next response. If the system fails to fully deblur the noisy vector, the associative chain is not entirely disrupted—the blurred vector can still function as a cue, often leading to the correct retrieval of the next item. However, TODAM faces a major limitation in accounting for transposition errors involving nonadjacent items. In other words, like other chaining-based approaches, it struggles to explain positional exchange errors in which two items that are far apart in the sequence swap their positions.

Because of these limitations, associative chaining models have largely been abandoned in favor of positional marker models, which will be presented in the next section (Burgess & Hitch, 2006a; Farrell & Lewandowsky, 2002; Henson, 1996; see Figure 1.2 (C)). However, recent research suggests that both chaining and positional coding mechanisms may coexist, working in tandem to maintain serial order (Burgess & Hitch, 1992; Caplan et al., 2022; Logan & Cox, 2021; Osth & Dennis, 2015; Osth & Hurlstone, 2023). This integrative hypothesis is supported by a recent study by Logan et al. (2024), who found that positional marker accounts can simulate behavioral performance patterns in serial recall tasks but less so in cued recognition tasks, while the reverse was observed for stimulations run based on chaining accounts. This integrated view acknowledges the associative chaining while emphasizing the need for a more robust positional representation to explain error patterns and distance effects.

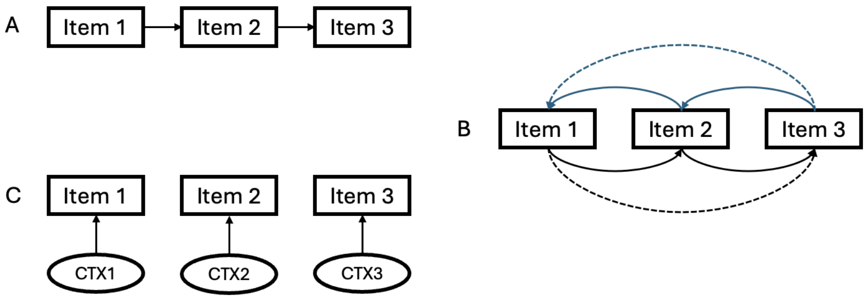

Figure 1.2. Mechanisms for the representation of serial order: (A) simple chaining, (B) compound chaining, and (C) positional marker, adapted from Hurlstone (2021).

Positional Marker Models

In contrast to associative chaining models, position marker models propose that serial order is represented via the association of items with some independent and varying contextual representations of their position within a sequence of items, such as time signals or positional information relative to the beginning or the end of the list (e.g., Henson, 1998; Brown et al., 2000; Hartley et al., 2016; see Figure 1.3). At recall, the context signal is reset and re-evolves along its original path, with the sequence item being activated according to the degree of similarity between the current state of the context signal and the state to which each item was associated.

These computational context-based models can account for many of the behavioral phenomena that characterize serial order memory, such as the temporal grouping effect or the locality constraint on transpositions (Henson, 1998; Lewandowsky & Farrell, 2008; Oberauer et al., 2018). Moreover, neuroimaging studies showing distinct brain areas involved in coding serial order and item information support these item-position association models, as presented in the previous chapter, by showing that serial order memory recruits specific fronto-parietal cortices (see Chapter 1), as compared to item memory, which recruits to a larger extent fronto-temporal cortices.

At the same time, positional models of serial order memory are not uniform and differ in the type of temporal codes they assume to be used for coding serial order. We present here an overview of this type of model before moving to the more fundamental question of the nature of the codes used for representing serial order information.

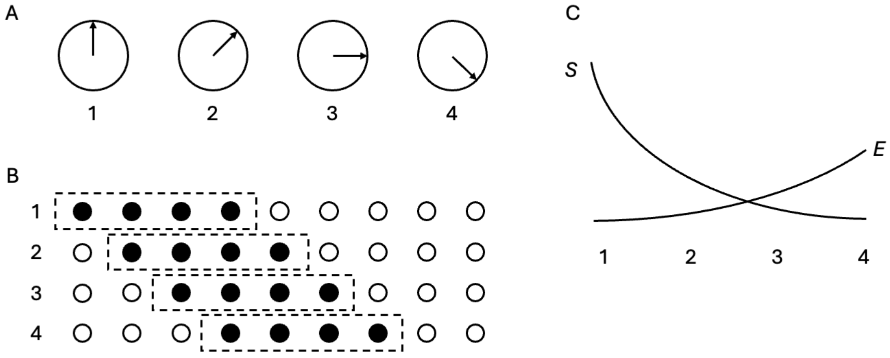

Figure 1.3. Illustrations of positional coding in (A) the temporal terms, (B) the absolute terms, and (C) the relative terms, adapted from Henson (1999).

Temporal coding models

The oscillator-based associative recall (*OSCAR*) model (Brown et al., 2000) is one of the models relying on a temporal coding scheme. It is supposed that serial order is linked with the different states of a time-varying context signal driven by sets of temporal oscillators operating at various frequencies. This mechanism can be likened to an analog clock. Such clocks have several hands (hour, minute, second), all oscillating at different frequencies. For example, the second hand is fine-grained, but it repeats itself every minute. Thus, the minute and hour hands are necessary for discrimination (e.g., from 8:00 to 12:00). These oscillatory codes are then associated with each item. Confusions or loss of positional information can occur when the oscillatory codes are similar to one another (e.g., 8:01 to 8:02). These temporal oscillators are dynamic, internal learning-context signals that could store and retrieve the sequences in a multi-layer. An updated temporal coding scheme is the SIMPLE model (Brown et al., 2007), which proposed that order is coded by temporal distance. It predicts that the temporal distinctiveness item would perform better in recall performance, and this temporal distinctiveness is determined by the time distance between the target item and other items (more information in Chapter 3 about *temporal distinctiveness effects*).

Although OSCAR proposed an explicit mechanism for serial order coding, one limitation of this time-based model is that the timing signal is based on free-running oscillators that are not influenced by bottom-up properties of the input sequence. The bottom-up entrainment of oscillators is not implemented in this model, as Brown et al. (2000) were careful to note. Recently, Hartley et al. (2016) proposed an alternative model based on the bottom-up hypothesis, known as the bottom-up multi-scale population oscillator (*BUMP*) mechanism. Unlike the OSCAR model, it emphasizes the role of the context itself rather than relying on top-down expectations. In their study, participants were presented with different temporal grouping sequences (e.g., 2-2-5, 2-3-4, or 3-3-3) in either a block or random order. They found that expected and unexpected grouping sequences presented similar serial recall and error patterns. In this case, they assumed that a bottom-up analysis of the speech input encodes the serial order of items in a sequence. With the simulation, they supposed that this analysis is achieved by means of a set of filters that respond to local changes in the magnitude and phase of its amplitude envelope at different frequencies. More specifically, they proposed that in the auditory-verbal temporal grouping material, the grouping items with more low-frequency oscillators (at or near the list presentation rate) will be more prone to local transpositions and correspondingly less prone to interposition errors. On the other hand, at higher frequencies, the oscillators (corresponding to the temporal scale of groups encountered in the input) will tend to resonate between the groups (especially for regularly grouped lists) and induce interposition errors.

Absolute coding models

Coding of absolute position assumes that items are associated with signals representing specific ordinal positions (first, second, third, etc.) (Burgess & Hitch, 1992; Lewandowsky & Farrell, 2008). These models propose that each item is linked to a distinct positional signal that uniquely defines its place within a sequence. For instance, when memorizing the list "*A–B–C–D*", such models posit that *A* is associated with position 1 and *B* with position 2, rather than encoding *B* as following *A* or *C* as following *B*, as in chaining models. Because each item's serial position is represented independently, absolute coding models can explain recall performance and associated serial position curves when certain items are omitted. Transposition errors are thought to result from similarities between adjacent positional signals.

The absolute coding approach was further developed within a framework that integrates WM and long-term memory (*LTM*) processes (Burgess & Hitch, 1999, 2005). Central to this integration is the Hebb repetition effect, which demonstrates the contribution of long-term learning to WM, as evidenced by improved recall for sequences that are periodically repeated (Hebb, 1961). Building on this finding, and subsequent work by Chiara Fastame et al. (2005), Burgess and Hitch (1999, 2005) proposed a computational model in which serial order in verbal WM emerges from interactions between short-term and long-term learning mechanisms (see Figure 1.4). The model comprises three layers: item layers representing lexical and phonological information, and a context or timing signal layer. Item–context associations are key to explaining immediate serial recall. This revised architecture assumes that each verbal item is associated with a gradually changing context signal that provides an absolute representation of serial position. During encoding, item-to-context associations between each presented item and its current context state are learned by updating connection weights between these layers. During recall, the active context inhibits items not associated with it and activates those that are, via context-to-item associations, while a competitive queuing mechanism determines the next recalled item. Connection weights updated in a fast manner will also quickly diminish, thereby simulating a typical working memory context. If the same item-context associations occur repeatedly, connection weights will also be updated in a slow but more stable manner, thereby simulating long-term learning as in the Hebb effect.

This dual-memory framework accounts for a wide range of empirical phenomena, including the serial position curve, transposition gradients, fill-in errors, and the transition from novel to familiar sequence learning.

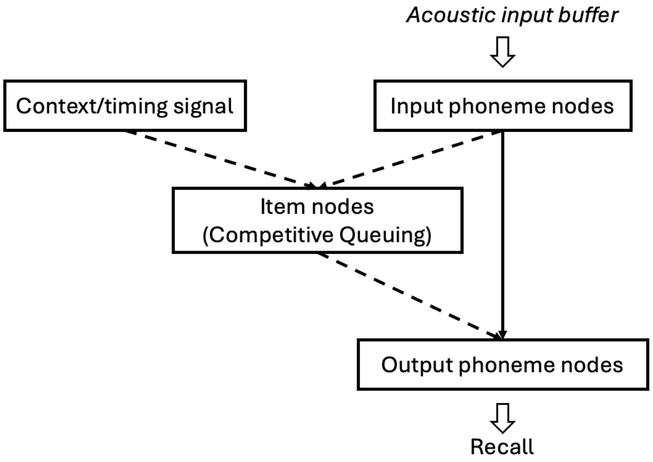

Figure 1.4. Outline adapted from the Burgess and Hitch (1999, 2006) model. The dashed line represents the modifiable all-to-all connections. The solid line represents the pre-wired one-to-one connections.

Relative coding models

Relative coding models, such as the Start-End Model (*SEM*) by Henson (1998), assume that items are coded relative to both start and end markers of a sequence. The start marker is strongest at the beginning of the sequence and decreases in strength towards the end, while the end marker is strongest at the end of the sequence and decreases in strength towards the beginning. The relative strengths of the start and end markers, therefore, provide an approximate two-dimensional code for each position within the sequence. Associating items with these marker strength codes, their position relative to both the start and end of the sequence. Simulations confirm its ability to capture the main phenomena and errors in serial recall, such as the effects of primacy, recency, transpositions, and, in particular, positional errors between groups and between trials.

Chapter summary

In this chapter, we introduced two main sets of theoretical models about serial order information in verbal WM. Associative chaining models represent serial order information via successive item-to-item bindings, while positional marker models associate items with specific contextual signals, often in the form of time-based signals. However, the exact nature of the codes used for representing serial order in verbal WM remains unclear. In the next chapter, we will see that, in addition to time-based coding of serial order information, spatial coding as well as linguistic coding have also been proposed.

Chapter 3

The question of serial order codes: temporal vs. spatial vs. linguistic

This chapter reviews empirical evidence for temporal coding of serial order information and then moves to a different line of studies that suggest that serial order codes may also have a spatial and/or a linguistic nature.

Temporal Coding of Serial Order Memory

In Chapter 2, we introduced the time-based models of serial order in verbal WM. What is the empirical evidence for time-based coding of serial order information?

First, there is a vast literature highlighting the general sensitivity of verbal WM performance to temporal factors. A major effect to be mentioned here is the *Temporal grouping effect,* already mentioned in Chapter 1. Generally, memory performance improves when stimuli are grouped by short ISIs and separated by longer between-group intervals (Frankish,1989; Hartley et al., 2016; Hitch, Burgess, Towse, & Culpin, 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b), and this advantage is accompanied by systematic changes in order errors (Ryan, 1969a, 1969b). Specifically, within-group primacy and recency effects emerge in addition to the main primacy and recency effects that characterize the entire lists (Frankish, 1989; Hartley et al., 2016; Hitch et al., 1996; Ng &Maybery, 2002; Ryan, 1969a, 1969b). Temporal grouping also influences transposition errors, giving rise to interposition errors, which serial order errors tend to involve within-group items rather than items stemming from different groups (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Hartley et al., 2016; Henson, 1999; Ng & Maybery, 2002, 2005; Ryan, 1969a, 1969b). Temporal grouping effects align with temporal models of serial order memory that posit a hierarchical temporal structure, with order encoded simultaneously at multiple levels, one tracking item position within groups and another encoding group positions at the whole-list level (Brown et al., 2000; Burgess & Hitch, 2006; Hartley et al., 2016).

A related effect is the *temporal distinctiveness effect*, in which items isolated by larger temporal gaps at encoding are recalled more accurately (Morin et al., 2010; Geiger & Lewandowsky, 2008; Lewandowsky et al., 2008). This temporal-isolation effect is general across many paradigms (e.g., free-recall tasks or recognition tasks) (Brown et al., 2000; Lewandowsky et al., 2004).

Second, a number of studies have manipulated temporal dimensions in a more explicit manner, by presenting temporal signals along with a memory list, with participants needing to overtly track these signals or not. Plancher et al. (2018) observed improved serial recall when a regular rhythm beat was presented during maintenance, compared to silent or irregular conditions. Fanuel et al. (2018) replicated this rhythm-related benefit during maintenance under both low and high cognitive load conditions. In contrast, when temporal regularity is introduced during the encoding phase, evidence suggests a more detrimental effect on serial order memory. Henson et al. (2003) instructed participants to tap during list presentation and found that paced tapping impaired performance on a list probe task, regardless of whether tapping followed a regular or complex rhythm. Hall and Gathercole (2011) similarly reported reduced WM performance when a regular-paced sound was presented during encoding and maintenance compared to a silent control condition. Parmentier and Beaman (2014) found that the presence of rhythmic distractors during encoding and maintenance phases impaired recall compared to silence, and that regular rhythms were even more disruptive than irregular ones. These studies reveal a contradictory pattern of results, particularly regarding the nature of the impact of an external regular rhythm.

One important element here could be the nature of rhythm itself, which, albeit regular, may not be isochronous with the encoding rhythm of the items. Only an isochronous rhythm presented during encoding may lead to facilitated encoding of memoranda and their serial order, or, at the least, not impair verbal WM recall performance. In line with this hypothesis, Pannell et al. (2023) found comparable recall performance between silence and regular single-tone presentation during encoding when the tone series was synchronized. Similarly, Attout et al. (2025) observed detrimental recall performance when the encoding of memoranda was associated with a random rhythm (in contrast to an isochronous rhythm), and the isochronous rhythm condition led to the same level of performance as the silent baseline condition.

A further issue is the use of concurrent tasks or dual processing paradigms in the above-mentioned studies, with participants encoding and maintaining memoranda while also following an external temporal signal. These dual task demands may lead to interactions due to the division of attentional resources during encoding, which may partly explain the detrimental effects of both regular and irregular rhythms. It should also be noted that Plancher et al. (2018) interpreted their findings as evidence for an attentional refreshing mechanism, the rhythmic cue during maintenance favoring the sequential refreshing of items.

To date, only one study has implemented a more direct approach to manipulate the temporal regularity of memory list presentation. Gorin (2020) compared recall performance for lists presented using a regular versus irregular temporal rhythm. Specifically, in the irregular condition, ISIs were arranged so that memory lists were presented either in a pseudo-random manner or in a completely random rhythm. Despite these manipulations, no reliable differences in serial order performance emerged between regular and irregular encoding rhythms. One possible explanation for this null effect is the influence of uncontrolled temporal grouping effects. In the irregular conditions, the alternation of relatively long vs. short ISIs may have allowed some items to benefit from spontaneous grouping, potentially attenuating the effect of rhythm irregularity (see also Hartley et al., 2016; as well as Ryan, 1969a, for related findings).

In sum, while there is a relative wealth of studies showing interactions between temporal dimensions and performance in serial order memory tasks, direct evidence for the coding of serial order information by temporal signals, for example, in the form of temporal oscillators, is still missing.

Spatial Coding in Serial Order Memory

While some positional models of serial order memory posit time-based signals for coding serial order information, there is also evidence for the intervention of spatial codes during serial order memory.

In a seminal study, van Dijck and Fias (2011) employed a dual-task paradigm in which participants judged digit parity for items maintained in WM. They found that parity judgements for items early in the memory list facilitated left-hand responses, whereas judgments for later items facilitated right-hand responses—a pattern suggesting the use of an internal left-to-right spatial template. van Dijck et al. (2013) further showed that reactivating start-of-list WM items induced a leftward attentional spatial bias in a concurrent dot detection task, while end-of-list items induced a rightward spatial bias.

These findings have been conceptualized within *the mental whiteboard hypothesis*, which proposes that sequential items in verbal WM are associated with spatial position markers along an internalized spatial axis (Abrahamse et al., 2014; 2017). These spatialization effects in verbal WM are also referred to as the *ordinal position effect* (Ginsburg et al., 2014) or the *Spatial–Positional Association of Response Codes* effect (*SPoARC*; Guida et al., 2016; Guida & Lavielle-Guida, 2014).

The mental whiteboard hypothesis proposes that serial order WM is grounded in the spatial attention system. More specifically, positional markers, which provide the multi-item WM with a serial context, should be interpreted as coordinates within an internal, spatially defined system. Information maintained via these spatial can then be inspected and retrieved via spatial attention mechanisms. For example, when we need to remember a series of items in a specific order, we would represent and maintain these items (e.g., from left to right) in an internal "mental whiteboard", like writing the items down on a physical whiteboard and organizing them on a horizontal line. In Western participants, these spatialization effects typically involve a left-to-right horizontal dimension, but cultural reading habits modulate this spatial mapping. For example, in the paradigm used by van Dijck and Fias (2011), reversed spatial mappings have been observed in right-to-left readers (Guida et al., 2018; Park et al., 2024; Shaki & Fischer, 2012). At the same time, top-down vertical spatialization effects have also been shown in verbal WM tasks (Hartmann et al., 2021). When the context does not cue a specific spatial coding direction or orientation, we assume that it spontaneously occurs in writing direction (Whorf, 1956; Maass & Russo, 2003; Spalek & Hammad, 2005; Bonato et al., 2012), and specially, reading direction, which may be the most important force in shaping mental representations of space (Zebian, 2005; Shaki & Fischer, 2008; Shaki et al., 2009; Bonato et al., 2012).

These behavioral effects are also supported by neuroimaging and eye-tracking studies. At the neural level, Cristoforetti et al. (2022) adapted the original paradigm of van Dijck and Fias (2011) to examine the multivariate neural codes associated with serial order coding (see also Attout et al., 2014; Attout et al., 2022; Rasoulzadeh et al., 2021; Tian & Fischer-Baum, 2025). They observed that the retrieval of items from early vs. late serial positions was associated with specific neural patterns in the intraparietal sulcus and the hippocampus, two regions that have also been associated with spatial representations and navigation. Furthermore, an eye-tracking study by Sahan et al. (2022) found that retrieving verbal items at a specific position in a sequence in WM was accompanied by horizontal eye movements, which the gaze diverted more to the left side of space when searching items from initial parts of the memorized sequence and more to the right side for later parts, supporting the idea that serial order information in verbal WM is spatially coded and governed by spatial attention. A similar effect was reported in a recent recall study, where gaze shifts in visual space correlated with spatial shifts of attention along a left-to-right mapping of serial positions (Schroth et al., 2025).

A further study by Guida, Abrahamse, et al. (2020), however, indicates that left-to-right spatial codes may not be a default coding format for serial order (see also Guida, Mosinski, et al., 2020). In their study, they manipulated the spatial presentation format (left-to-right, central, right-to-left) of consonant sequences in a recognition task. Participants were asked to indicate whether the probe was in the sequence or not by pushing either the "yes" or "no" button with their left or right hand. They observed the expected faster left-hand reaction on beginning items and faster right-hand reaction on later items in central and left-to-right presentation (encoding) formats, but this effect reversed in the right-to-left presentation condition. These findings led to an assumption of the principle of economy, whereby spatial codes are flexibly recruited to provide a functional advantage in serial recall.

While there is now a relatively large set of studies showing the involvement of spatial processes during serial order WM tasks, these studies do not yet directly show that serial order information is actually coded using spatial codes, or that these codes are essential for the maintenance of order information in WM In order to answer this question, we would need to know whether disturbing or facilitating the build-up of spatial positional codes has a measurable effect on serial order memory. A few studies provide useful information to this question, although it should be noted that not all of these studies were designed to examine this specific question. Hitch and Morton (1975) contrasted lists presented from left to right versus in identical positions on the display without observing any difference in recall performance. Battacchi et al. (1990) found a stronger recency effect when sequences were presented in spatially distributed formats (e.g., left-to-right or up-to-down) compared to a central, non-distributed format. However, LeCompte (1992) attempted to replicate this advantage but failed to observe any effect. Similarly, Li and Lewandowsky (1995) found no recall benefit when comparing centrally presented items versus items presented in random spatial positions. Similarly, using a dual-task paradigm, Attout et al. (2025) did not observe facilitated serial order recall performance when memoranda were presented in association with spatial cues that moved to the right with each successive presentation of the memory item. Beyond the inconsistency of the findings, a key problem of these studies is that they compare situations that provide an encoding structure versus no encoding structure, and hence, it is difficult to draw any conclusions regarding the specific impact of spatial variables or their absence on verbal WM recall performance. One study contrasted two spatially organized encoding conditions by comparing recall performance for left-to-right vs right-to-left presentation of the memoranda during encoding. Fischer-Baum and Benjamin (2014) reported an advantage for the canonical, left-to-right condition, but the advantage was small and restricted to the final list positions.

In sum, while there is robust evidence for the presence of spatialization effects in verbal WM tasks, direct evidence for the use of spatial codes for representing serial order information in WM remains scarce and contradictory.

Linguistic Coding of Serial Order Memory

Coming back with a chapter on linguistic support of serial order memory may appear paradoxical, given that in the previous chapters we have developed a vast literature suggesting linguistic support mainly characterizes the encoding and maintenance of item information. However, there is one aspect of linguistic knowledge that we have not yet considered, and this is syntactic knowledge. Syntactic knowledge precisely defines the order of words within a verbal sequence, as a function of their grammatical function. Hence, in natural WM situations involving the maintenance of sequences of grammatically (and semantically) connected words rather than of arbitrary word sequences, syntactic knowledge may also support the maintenance of information in WM, and this most specifically at the serial order level. Basic syntactic knowledge structures, such as the rule that a verb is usually followed by a noun or a noun is typically followed by an adjective in many languages (Bock, 1995; Garrett, 1980), restrict the possible positions of words within a sentence.

Direct evidence for a role of syntax on verbal WM, and specifically serial order memory, is very scarce. Some indirect and nonspecific evidence comes from *the sentence superiority effect (SSE).* A number of studies have shown that well-formed sentences are recalled more accurately than scrambled word lists (Allen et al., 2018; Baddeley et al., 2009; Brener, 1940; Cattell, 1886; Jefferies et al., 2004; T. Jones & Farrell, 2018; Marks & Miller, 1964; Miller & Isard, 1963; Miller & Selfridge, 1950; Poirier & Saint-Aubin, 1996; Savill et al., 2015, 2018). However, while sentences are formed on the basis of syntactic knowledge, they also form semantically coherent structures, which are missing in scrambled word lists. Any advantage in WM recall for sentences vs. scrambled word lists is therefore difficult to interpret with regard to the specific supportive role of syntactic knowledge. More recently, Jones and Farrell (2018) showed that recall accuracy for verbal sequences, measured by a reconstruction task, improves if they follow proper syntax, even when sequences are meaningless. Similar results were previously observed (Perham et al., 2009) for recalling lists of semantically unrelated adjective-noun pairs. While these two studies control for the impact of semantic knowledge and show a general role of syntactic knowledge on verbal WM performance, they do not yet provide specific evidence for an impact on serial order memory.

To date, only two studies have directly examined the impact of syntactic knowledge on item versus order memory. Schweppe and her colleagues (2022) presented lists of syntactically correct (adjective before noun) or incorrect (noun before adjective) word pairs to German-speaking participants for immediate serial recall, specifically determining item and order recall scores. While observing a robust impact on item recall, they observed no impact on order recall. A similar pattern was observed in a French-language study (Querella & Majerus, 2024). French has the specificity of allowing for both adjective anteposition (i.e., adjective before noun) and postposition (i.e., noun before adjective), but the specific type of position is determined in a probabilistic manner by the nature of the adjective. Despite these more flexible syntactic rules, Querella and Majerus still found no specific impact of adjective-noun syntactic regularity on order recall, the effects being restricted to item recall. The impact of syntax on item recall, but not order recall, is surprising given that, by nature, syntactic knowledge about the sequential organization of words should impact memory for serial order. However, the way the lists were constructed in these two studies (as well as in the study by Perham et al., 2009) might have contributed to this pattern of results. More specifically, for a given list, all pairs were either presented in adjective-noun order or noun-adjective order, and hence the ordering of adjectives and nouns became predictable. This is likely to have allowed for serial position prediction effects, which may have constrained the order in which the adjectives and nouns were output, irrespective of the syntactic legality of the sequences.

In sum, while the syntactic level of knowledge is the most likely linguistic variable to provide support to the retention and reproduction of serial order information in a verbal WM task context, direct evidence is still missing.

Chapter summary

In this chapter, we introduced three approaches to serial order coding in verbal WM. Temporal coding emphasizes the association between serial order and time-based signals, which may be governed either by top-down intrinsic oscillatory mechanisms or by bottom-up, context-dependent external signals. Spatial coding proposes that sequential memoranda are mapped onto a spatial axis in “mental space”, shaped by reading and writing habits (e.g., left-to-right in Western cultures) as well as by contextual information. Linguistic coding suggests that serial order is supported by language-based representations from long-term memory, particularly syntactic knowledge. However, it remains unclear whether any of these codes predominates in verbal serial order processing and, more generally, how they contribute to the maintenance of serial order information in verbal WM. In Experiment Part, we present three studies, each focusing on one coding approach, to clarify the nature of serial order coding in verbal WM from temporal, spatial, and linguistic perspectives.

Experiment Part

Goals and hypotheses

In the theoretical part of this thesis, we have discussed evidence for the temporal, but also for the spatial and potentially linguistic nature of codes used for representing and maintaining serial order information in a verbal WM task context. While the literature is generally supportive of an intervention of temporal, spatial, and linguistic/syntactic factors in verbal WM tasks, their exact role in actually coding serial order information still remains to be shown. This PhD thesis aims at examining this role by manipulating the ease with which presumed temporal, spatial, or linguistic codes can be set up and have a direct facilitatory or interfering impact on serial order WM performance.

First, in **Study 1**, we investigated the role of linguistic-syntactic coding of serial order information in verbal WM. As already mentioned, two studies previously showed that syntactic knowledge about adjective-noun order can have an impact on verbal WM performance (Schweppe et al., 2022; Querella & Majerus, 2024). These studies, however, were only able to demonstrate an impact on item recall, contrary to the hypotheses formulated in this thesis. We had mentioned that this finding may have been the result of the specific manner in which the memory lists were set up. For a given trial, all adjective-noun pairs were presented in the same order (either adjective anteposition or adjective postposition), which may have led to a predictability of syntactic parsing and prevented serial order errors from occurring. To avoid this potential bias, Study 1 of this thesis used adjective-noun pair lists mixing adjective anteposition and postposition, thereby minimizing within-list syntactic order predictability. In addition, experiments were conducted in both French and German languages to examine the generality of the effects under investigation, but also because the two previous studies showing higher item recall performance for syntactically legal adjective-noun pair lists had been conducted in French and German.

Study 2 examined the role of spatial coding for representing and maintaining serial order information in verbal WM. As mentioned in the theoretical part of this thesis, very few studies have directly examined the extent to which spatial coding provides real support to serial order WM. We currently cannot rule out the hypothesis that the spatialization effects that have been observed in verbal WM tasks are secondary effects of serial order coding rather than reflecting their fundamental basis. To examine this question, we conducted a large, multi-experiment study based on a verbal immediate serial recall task that was submitted to several spatial manipulations. The first experiment aimed to replicate the results observed by Fischer-Baum and Benjamin (2014) by having participants encode verbal memoranda either in a (for Westerners) canonical left-right or in a non-canonical right-left encoding direction. We furthermore used longer lists (seven items) to increase task sensitivity and thereby maximize the intervention of spatial codes. In the second experiment, items were presented centrally, which—based on the evidence from the spatialization effect literature (Guida, Abrahamse, et al., 2020)—should lead to the setup of a left-to-right mental spatial representation of the memory sequence. If this spatial structure is essential for serial order maintenance and recall, then participants should have difficulty if they are instructed to recall the items in a direction opposite to this "spontaneous" spatial structure (i.e., from right to left). The last experiment provided a particularly stringent test of the intervention of spatial dimensions during encoding and recall by manipulating both spatial encoding direction and the intervention of spatial cues during recall.

In **Study 3**, we examined the role of temporal manipulations on serial order WM performance, in order to determine in a more direct manner the extent to which serial order information is based on temporal codes. According to the time-based models of serial order WM (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Hartley et al., 2016), memory lists presented at a temporally irregular rhythm should be more difficult to encode than memory lists presented at a regular rhythm, given that time-based models presume the intervention of temporal oscillators based on a regular oscillation frequency. For a temporally irregular incoming stream of items, associating the sequence of items and their ordinal position with a specific oscillator should become more difficult, especially when the temporal irregularities are large. As we have already mentioned, few studies have directly manipulated the encoding rhythm of a memory sequence itself and its impact on serial order memory. Those who did observe null findings, but without being able to discard potential confounding factors such as spontaneous grouping of segments of items induced by the temporal irregularity (Gorin, 2020). Study 3 of this thesis was a multi-experiment study based on a design similar to Gorin (2020), by using an immediate serial recall task and by presenting the memory lists in a regular or irregular rhythm. We designed the ISIs to be either regular or irregular, while, critically, minimizing the potential for grouping effects in the irregular lists via a specific design of the ISI variability.

Study 1

Syntactic Knowledge Does Support Working Memory For Serial Order: A comparison across French and German languages

Hong Xiao, Judith Schweppe, Pauline Querella, Lucie Attout, Friederike Contier, and Steve Majerus

Published in

Memory & Cognition (2025), 1-15

Abstract. A large body of research demonstrates robust interactions between verbal working memory (WM) and phonological and lexico-semantic language knowledge, particularly for item recall. The role of syntactic knowledge, involving knowledge about word positioning in a verbal sequence, has been explored to a lesser extent but is of theoretical interest given that this type of knowledge may also support serial order recall in multi-item sequences. This hypothesis has not been supported so far, either by French or German language studies. The present research re-examines the impact of syntactic knowledge on WM for lists of adjective-noun pairs by controlling for short-term syntactic predictability of memoranda. Contrary to previous studies, across two experiments involving French and German, we observed enhanced order recall accuracy for adjective-noun pairs presented in legal syntactic order, as well as more order migration errors for pairs in illegal syntactic order. This study is the first to demonstrate that long-term syntactic structures do support recall of serial order information in WM, consistent with hybrid and full linguistic accounts of verbal WM.

Introduction

There is increasing consensus that verbal working memory (WM) interacts with long-term language representations (Ebbinghaus, 1964). Numerous studies have demonstrated that verbal WM performance is enhanced for word lists compared to nonword lists, a phenomenon known as the lexicality effect in verbal WM (Brener, 1940; Gathercole et al., 2001; Hulme et al., 1991, 1995; Kowialiewski & Majerus, 2018; Majerus & Van der Linden, 2003; Patterson et al., 1994; Saint-aubin & Poirier, 2000). Furthermore, recall performance improves for words with richer or more accessible lexical and semantic representations, such as high-frequency words versus low-frequency words (Gregg et al., 1989; Hulme et al., 1997; Kowialiewski & Majerus, 2018; Poirier & Saint-Aubin, 1996; Roodenrys et al., 1994; Watkins & Watkins, 1977), or concrete words versus abstract words (Bourassa & Besner, 1994; Hulme et al., 1997, 2003; Kowialiewski & Majerus, 2018; Majerus & Van der Linden, 2003; Walker & Hulme, 1999). These findings underscore the robust interactions between verbal WM and language knowledge, particularly regarding item memory at phonological and lexico-semantic levels. They have led to models assuming that verbal WM is determined by access to language knowledge at least for item-level aspects (encoding, retrieval, and/or completion of a word within a list of words to be memorized, e.g., Baddeley et al., 2017; Burgess & Hitch, 1999, 2006; Cowan, 1999; Gupta & MacWhinney, 1997; Majerus, 2013). Other theoretical accounts consider that linguistic knowledge should support both item and sequence-level representation of information in WM (e.g., Acheson & MacDonald, 2009; Majerus, 2019; Martin et al., 1994; Schwering & MacDonald, 2020). The present study focuses on the potential impact of syntactic knowledge on verbal WM. This type of knowledge is of particular interest given that it concerns the serial positioning of words within a sentence, and hence, according to linguistic accounts of verbal WM, should also support recall of serial order information in verbal WM (Perham et al., 2009; Querella & Majerus, 2024; Schweppe et al., 2022).

Regarding the distinction between item and serial order aspects in WM, a significant number of studies have revealed dissociations between these two aspects. Most empirical studies have shown that phonological and lexical-semantic knowledge support item recall, but not order recall (Gathercole et al., 2001; Hulme et al., 1991, 1997; Majerus & D'Argembeau, 2011; Nairne & Kelley, 2004; Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999; Saint-Aubin & Poirier, 2000; Walker & Hulme, 1999). This is also supported by neuropsychological studies revealing the possibility of selective item versus serial order impairment, with item WM impairment often being linked to language deficits (Majerus, 2008; Majerus et al., 2015). For example, patients with semantic impairment may show poor item recall in word serial recall tasks, while serial order recall remains intact (Majerus, Norris, et al., 2007). Additionally, neuroimaging studies comparing item and order recall tasks have found that WM for serial order recruits specific fronto-parietal networks, including the intraparietal sulcus, while item recall tasks engage fronto-temporal networks involved in linguistic processing, including the supramarginal gyrus (Henson et al., 2000; Majerus, Poncelet, Van der Linden, et al., 2006; Majerus et al., 2010; Marshuetz et al., 2000a, 2006). These studies suggest that encoding and retention of item information in WM engage access to phonological and semantic aspects of words, while encoding and retention of serial order information involve a more specific mechanism, whose nature still remains to be determined (Hartley et al., 2016; Majerus, 2019).

While semantic information is generally considered to exert a stronger influence on item memory than on serial order memory, a growing body of evidence suggests that semantic processing can nonetheless interact with serial order representations, albeit in a complex manner (Guérard & Saint-Aubin, 2012; Neale & Tehan, 2007; Murdock, 1976; Saint-Aubin & Poirier, 1999a, 1999b). For instance, Saint-Aubin et al. (2005) reported that semantic relatedness impaired serial order memory, even when the number of recalled items was equated across semantically related and unrelated lists. Similar effects were observed by Tse (2009) and Tse et al. (2011), who demonstrated that both categorical and associative forms of semantic relatedness disrupt order memory, independently of item recall performance. These findings align with the hypothesis that semantic proximity among items increases the likelihood of order errors, such as item migrations (Poirier et al., 2015; Saint-Aubin et al., 2023). Consistent with this view, Guitard et al. (2025) recently showed that participants produced more conditional order errors when recalling semantically related lists compared to unrelated lists, further supporting the notion that semantic relatedness exerts a destabilizing effect on serial order representations. Other studies have observed opposite findings, by observing improved serial order recall for semantically related vs. unrelated word lists, but only when the similar items were presented in a grouped manner (Kowialiewski et al., 2024).

Less ambiguous evidence in favor of the supporting influence of language knowledge on memory for serial order may be obtained when considering syntactic aspects of language knowledge. Syntactic knowledge, which governs the order of words in a sentence, should be most strongly expected to exert an impact on the retention of serial order information in verbal WM tasks. Basic syntactic knowledge structures (such as the rule that a verb is generally followed by a noun, or a noun is generally followed by an adjective in a large range of languages; Bock, 1995; Garrett, 1980) constrain the possible positions of words within a sentence. These positional rules should also be expected to support WM for verbal sequences. The impact of syntactic knowledge on serial order memory has not yet been extensively explored. Some evidence comes from studies that have investigated the sentence superiority effect (SSE; e.g., Bonhage et al., 2014; Scheerer, 1981). These studies showed that regular sentences are recalled more accurately than scrambled word lists (Allen et al., 2018; Baddeley et al., 2009; Brener, 1940; Cattell, 1886; Jefferies et al., 2004; Jones & Farrell, 2018; Marks & Miller, 1964; Miller & Isard, 1963; Miller & Selfridge, 1950; Poirier & Saint-Aubin, 1996; Savill et al., 2015, 2018). This pattern of findings can be accounted for within the conceptual regeneration hypothesis (Potter & Lombardi, 1990; 1998), according to which sentence recall is based on a conceptual representation of the sentence that is generated while processing the to-be-recalled sentence. At recall, this conceptual representation is regenerated, and lexical and syntactic priming of words and syntactic structures presented in the sentence are particularly available when regenerating the sentence. Based on this, lexical and syntactic representations in long-term memory are key to the immediate recall of sentences. However, although Potter and Lombardi (1998; see also Lombardi & Potter, 1992) already emphasize the importance of syntactic structure for ordering words in sentence recall, their studies do not focus on order memory. Another framework that can account for the sentence superiority effect is the Construction-Integration (CI) model of discourse comprehension (Kintsch, 1988), which highlights the dynamic interplay between working memory, prior knowledge, and textual input in the construction of meaning. According to this model, sentence processing involves the initial extraction of multiple levels of linguistic information—phonological, lexical, semantic, and syntactic—which are used to construct a preliminary mental representation. This representation is subsequently integrated with relevant long-term knowledge structures, enabling the formation of a coherent and contextually grounded mental model. In the context of immediate serial recall tasks, such integrative processes may facilitate both the comprehension and recall of verbal sequences by enriching item representations and establishing associative links that support retrieval. The results of these studies and associated models are, however, difficult to interpret in terms of a specific influence of syntactic knowledge on serial order memory, given that sentence recall not only involves access to syntactic structures, but it also involves semantic and conceptual knowledge. More recently, Jones and Farrell (2018) demonstrated that recall accuracy for verbal sequences is improved if they conform to legal syntax, even for meaningless sequences. Similar findings had already been observed (Perham et al., 2009) for recall of lists for semantically unrelated adjective-noun pairs, the pairs occurring in either legal or illegal syntactic order. However, both of these studies did not dissociate the contribution of syntactic structure to item versus serial order retention, as either no specific measure of item-level working memory performance was included or only overall serial recall performance was determined. Another study by Schwering and MacDonald (2023) demonstrated that specific lexico-semantic and syntactic constraints, such as word context and part of speech, influence verbal WM, particularly concerning item recall performance, with less clear evidence for an impact on serial order recall.

So far, only two studies have aimed more directly at examining the impact of syntactic knowledge on item vs. order memory. Schweppe and her colleagues (2022) presented lists of syntactically legal (adjective before noun) or illegal (noun before adjective) word pairs to German-speaking participants for immediate serial recall and specifically determined item and order recall scores. While observing a robust impact on item recall, they observed no impact on order recall. A similar pattern was observed in a French language study (Querella & Majerus, 2024). French has the specificity of allowing for both adjective anteposition (i.e., adjective before noun) and postposition (i.e., noun before adjective), but the specific type of position is determined in a probabilistic manner by the nature of the adjective. Despite these more flexible syntactic rules, Querella and Majerus still found no specific impact of adjective-noun syntactic legality (*) on order recall, the effects being restricted to item recall. This impact of syntactic legality on item recall, but not order recall, is surprising given that, by nature, syntactic knowledge about the sequential organization of words should impact memory for serial order. However, the manner in which the lists were constructed in these two studies (as well as in the study by Perham et al., 2009) might have contributed to this pattern of results. More specifically, for a given list, all pairs were either presented in adjective-noun order or noun-adjective order, and hence the ordering of adjectives and nouns became predictable. This is likely to have allowed for the creation of temporary syntactic frames during the syntactic parsing process of list encoding, which may have constrained the order in which the adjectives and nouns were output, irrespective of the syntactic legality of the sequences, hiding the potential impact of long-term syntactic knowledge about adjective-noun output order.

The present study re-examines the impact of syntactic knowledge about adjective-noun order on verbal WM performance by using lists mixing adjective anteposition and postposition. This should prevent the build-up of short-term syntactic predictability rules during list encoding and allow for a specific assessment of the impact of long-term syntactic knowledge on order recall for lists of adjective-noun pairs. Experiment 1 tested this hypothesis for adjective-noun pairs in French, which allows for a full crossing of adjective-noun position (ante- vs. post-position) with syntactic legality*. Experiment 2 used a cross-language design by adapting the lists used in Experiment 1 for an administration in both French and German.

Experiment 1

Experiment 1 was conducted in French and was based on a 2 x 2 within-subject design with the factors, syntactic legality* (lists consisting of syntactically legal* pairs vs. lists consisting of syntactically illegal* pairs) and adjective position (anteposition vs. postposition). Critically, a single trial involved both postposition and anteposition pairs to avoid the build-up of temporary syntactic predictability of adjective position within a list. The key

---

[*] When referring to French, we will append a '*' symbol to the term "legality" to acknowledge that French adjective-noun orderings cannot be categorized as exclusively legal or illegal, but legality is determined in a probabilistic manner.

dependent variables were serial recall scores, order recall scores, and within-pair order migration errors. In addition, we recorded item recall scores and omission errors.

Method

Transparency and Openness. We reported how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The data and materials for all experiments are available at https://osf.io/kuj6h/?view_only=b5a91544211941c3b7ff9e0cee455e76.

The design, hypotheses, and analysis plan for Experiment 1 were preregistered at https://osf.io/x7sgk.

Participants. Ninety-three participants (49 females and 44 males; *average age* = 21.39, *SD* = 1.90) were recruited (see *Scoring and Analysis Procedure* section for justification of sample size) from the Université de Liège staff and student community. All participants were native French speakers without any history of language, learning, neurological, or psychiatric disorder, and reported no current drug use (e.g., cannabis) or alcohol abuse. Participants who were not monolingual or spoke another language during half-time in their daily life were not eligible for inclusion in this study. Seventy-six participants were right-handed and 7 were left-handed. One participant was excluded because of misunderstanding the task procedure. The study had been approved by the ethics committee of the Faculty of Medicine of the University of Liège. Participants were informed that no financial compensation was provided. All participants have given their informed consent to participate in the study before testing.

Materials. To select adjective-noun pairs and determine their syntactic preference in terms of adjective position, 57 additional young adults from the University of Liège community were recruited for an online syntactic preference judgment test (26 males and 31 females, 18 to 28 years old, *average age* = 21.84, *SD* = 2.11). Participants were presented with 130 French adjective and noun pairs with minimized semantic plausibility: 36 pairs taken from Querella and Majerus (2024) and 94 additional newly created pairs. For each pair, the two possible orders (adjective before noun or reverse) were displayed, one on each side of a switch bar. Participants needed to slide the bar according to their preferred order. The switch bar was divided into a scale with five different levels (*1* = strongly prefer left pair, *2* = prefer left pair, *3* = no preference, *4* = prefer right pair, *5* = strongly prefer right pair). To control for the effect of inflectional morphology, all pairs involved masculine nouns that are not inflected.

Based on the judgment results, we kept the 48 pairs that showed the strongest judgment preference in either direction. We recoded the responses to 1 means strongly preferring postposition and 5 means strongly preferring anteposition. 24 pairs with preferred anteposition (*mean preference* = 4.24, *SD* = 0.49, *Minimum* = 3.46, *Maximum* = 4.83) and 24 with preferred postposition (*mean preference* = 1.44, *SD* = 0.23, *Minimum* = 1.21, *Maximum* = 1.93) were used to create 16 lists containing each three adjective-noun pairs and by mixing adjective anteposition and postposition in each list. These lists were called the legal* lists. The 16 illegal* lists were created by simply reversing the order of the noun and adjective within each pairing of the legal* lists. This procedure ensured that legal* and illegal* conditions were matched for phonological similarity and the lexico-semantic status of the words. In addition, the number of syllables was kept constant at 11 syllables per list.

The material was recorded by a French-native female speaker using a neutral voice. Each item was recorded separately and then combined to form adjective-noun/noun-adjective pairs. The full stimulus lists are presented in Appendix Tables A.7 and A.8, the judgment results for each pair are in Table A.9, and the characteristics descriptions of stimuli are in Table A.10.

Procedure. The main experiment was conducted via OpenSesame software (https://osdoc.cogsci.nl/4.0/). Subjects were pseudo-randomly assigned to one of two presentation conditions (condition A: block of legal* lists followed by block of illegal* lists; condition B: block of illegal* lists followed by legal* lists). The order of the lists in each block was the same for all participants. They were tested individually in silent testing rooms. Participants were instructed to carefully listen to each list of six items, presenting the items in three pairs via the specific timing of presentation, followed by immediate oral serial recall, asking the participants to recall all the words, from first to last word. If participants could not remember a word at a particular position of the list, they had to say "oublié" (forgotten) for that position. Before the presentation of the experimental lists, participants had to complete two practice trials to ensure that they had correctly understood the task instructions. In each trial, every word was presented for 1000ms. The inter-stimulus interval was 350 milliseconds between nouns and adjectives in the same pair. Between each adjective-noun or noun-adjective pair, the inter-stimulus interval was longer (1000ms) to clearly mark the boundaries between the different pairs. By default, stimuli were presented via the speakers of the presentation PC. Headphones were also available if preferred by participants. Experimenters ensured that participants could clearly and comfortably perceive the stimuli during practice trials. There was no time limit for the participants to respond. The experiment lasted about 20 minutes per participant. All responses were digitally recorded for transcription and scoring.

Scoring and Analysis Procedure. Two standard recall accuracy scores were determined: serial recall score (number of items correctly recalled in correct serial position) and item recall score (number of items correctly recalled independently of serial position). A specific order recall score was also computed by dividing the sum of the serial recall score by the sum of the item recall score, as preregistered.

The following error types were identified: Within-pair order migration errors (the adjective and noun for a specific pair exchange their position) and omission errors (items for which the participant said "oublié" (forgotten) or items not recalled).

All analyses were conducted using a Bayesian statistical approach (see, e.g., Dienes, 2011; Morey & Rouder, 2011) according to our preregistration. This approach has the advantage of relying on a model comparison rationale to select and quantify the strength of evidence associated with each model, and crucially, allows for testing the strength of evidence for and against an effect of interest (i.e., positive evidence for the null hypothesis). The Bayesian framework does not involve traditional p-values, thereby avoiding multiple testing problems such as alpha inflation (Wagenmakers et al., 2008). All analyses are based on Bayes factors (*BF*), which can be considered as a relative measure of statistical evidence (Morey et al., 2016). The *BF* represents the degree to which the observed data update the initial belief in favor of one hypothesis relative to another. The *BF* is the likelihood ratio of a given model, the best-fitting model being the one with the highest *BF*. BF_{01} indicates evidence in favour of the null hypothesis, while BF_{10} indicates evidence in favour of the alternative hypothesis. Although there are no fixed thresholds for *BF* values, we used the following categories for describing strength of evidence: a *BF* of at least 1 is considered to indicate anecdotal evidence, a *BF* of at least 3 is considered to indicate moderate evidence, a *BF* of at least 10 is considered to provide strong evidence, a *BF* of at least 30 is considered to provide very strong evidence, and a *BF* of at least 100 is considered to indicate decisive evidence (Jeffreys & Jeffreys, 1998).

The R platform and the brms package (Bürkner, 2017) were used for running Bayesian generalized linear mixed models. The dependent measures were the different scores and error types as defined above. Fixed factors included legality* (legal* vs. illegal*), adjective position (anteposition vs. postposition), and their interaction. Serial position was excluded from the analyses to reduce the model's complexity. The random intercept included the subject variable. To determine the random slopes to be included in the model, we first ran four models including the random intercept and different random slopes combinations (none, legality*, adjective position, interaction of legality* and adjective position). Random slopes were included in the subsequent models if the *BF* value for a given model was larger than 3, relative to the null model including only the intercept.

Sample size was monitored via a Bayesian sequential sampling approach (Schönbrodt & Wagenmakers, 2018) combined with an effect size stabilization procedure (Anderson et al., 2022). Previous work has demonstrated that relying on a fixed critical p-value as the sole criterion for terminating data collection can lead to inflation of the Type I error rate, thereby compromising the validity of the resulting p-values. When Bayesian methods are employed, the use of a threshold Bayes factor as a stopping rule has been shown to bias the evidential value of the data due to the influence of optional stopping. In contrast, adopting effect size stabilization as a stopping criterion does not introduce such bias and has been found to preserve the integrity of statistical inference under both frequentist and Bayesian frameworks. Beginning from the 20th participant, we computed, for each additional participant, the effect sizes for the effect of interest (syntactic legality* - signed *Cohen's d*) and continued sampling until the effect size, for both item and serial order recall scores, stabilized (a priori defined minimal absolute change of effect size <.05 over 5 consecutive analyses; Anderson et al., 2022). The effect sizes of the item recall score and serial recall score were both stable after the 35th participant. In addition, in order to guarantee similar statistical power between the present study and previous studies by

Schweppe and her colleagues (2022) and Querella and Majerus (2024), which involved larger samples, we continued recruitment until reaching a comparable sample size. Also note that Kowialiewski (2024) recently showed that the effect size stabilization procedure does not guarantee the detection of a true effect in the population if the stabilized effect size is small.

Besides, for the display of the data in the figures, we presented the data as proportions. The proportion of item recall scores and serial recall scores was determined by dividing the number of correctly recalled items (or of the items recalled in the correct serial position) by the total number of items. The proportion of within-pair order migration errors was calculated by dividing the number of within-pair order migration errors by the total number of order errors (= difference between item and serial recall scores). Similarly, the proportion of omission errors was obtained by dividing the number of omission errors by the total number of item errors (= difference between item recall score and total number of items in the task).

Results

Serial Recall Score. A first Bayesian generalized linear mixed model analysis was run on the serial recall score. The best-fitting model included only legality* ($\beta = 0.46$, 95% *CI* [0.28, 0.64], $BF_{10} = 1.32 \times 10^4$; see Appendix Table A1). This model was 3.76 times more likely than the model incorporating both legality* and adjective position (legality*: $\eta_p^2 = 0.232$; adjective position: $\eta_p^2 = 0.057$; legality* × adjective position: $\eta_p^2 = 0.013$). This means that serial recall scores were higher for syntactically legal* pairs (*mean* = 0.667, *SD* = 0.159) than illegal* pairs (*mean* = 0.603, *SD* = 0.174), regardless of whether adjectives were antepositional or postpositional (see Figure 1.1 and Appendix Table A.2).

Item Recall Score. For item recall scores, the same type of analysis identified the best-fitting and most parsimonious model as including legality* as a fixed factor ($\beta = 0.18$, *95% CI* [0.09, 0.28], $BF_{10} = 87.46$; see Appendix Table A1). Although the model including legality* and adjective position is 1.40 times more likely than the model including only legality*, the latter should be retained as the more parsimonious model (legality*: $\eta_p^2 = 0.109$; adjective position: $\eta_p^2 = 0.079$; legality* × adjective position: $\eta_p^2 = 0.035$). The result means that item recall performance was higher for syntactically legal* pairs (*mean* = 0.780, *SD* = 0.110) than illegal* pairs (*mean* = 0.755, *SD* = 0.107), regardless of whether adjectives were antepositional or postpositional (see Figure 1.1 and Appendix Table A.2).

Order Recall Score. For the order recall score**,** given its derivational nature, a Bayesian analysis of variance was conducted on subject-averaged scores per experimental cell. The best model was again including the effect of legality* ($\beta = 0.38$, 95% *CI* [0.23, 0.53], $BF_{10} = 1.18 \times 10^4$; see Appendix Table A1). This model was 5.76 times more likely than the model including both legality* and adjective position (legality*: $\eta_p^2 = 0.242$; adjective position: $\eta_p^2 = 0.004$; legality* × adjective position: $\eta_p^2 = 0.006$). That means order recall scores were higher for syntactically legal* pairs (*mean* = 0.848, *SD* = 0.119) than for illegal* pairs (*mean* = 0.789, *SD* = 0.154), independently of adjective positions (see Figure 1.1 and Appendix Table A.2).

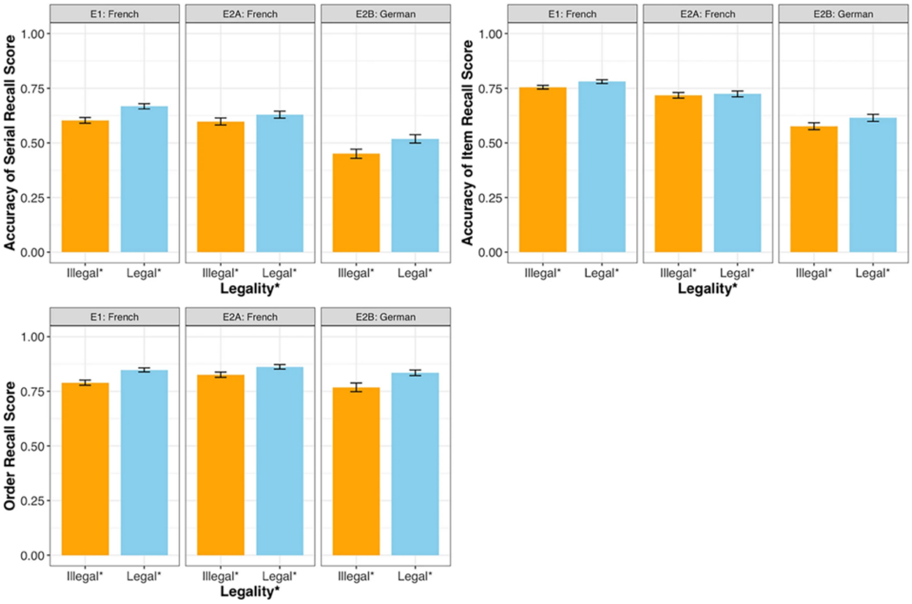

Figure 1.1. Accuracy of serial recall score, item recall score, and order recall score, as a function of legality*, across Experiment 1 (French), Experiment 2A (French), and 2B (German).

Within-Pair Order Migration Errors. Next, we analyzed error scores. Generalized linear mixed analyses of within-pair order migration errors also supported the best-fitting model as including only legality* ($\beta = -0.57$, 95% *CI* [-0.77, -0.37], $BF_{10} = 8.75 \times 10^5$; see Appendix Table A1). This model was 3.00 more likely than the model including both legality* and interaction of legality* and adjective position (legality*: $\eta_p^2 = 0.182$; adjective position: $\eta_p^2 = 5.731 \times 10^{-4}$; legality* × adjective position: $\eta_p^2 = 1.672 \times 10^{-4}$). This result means that there were more order migration errors for syntactically illegal* pairs (*mean* = 0.245, *SD* = 0.093) compared to legal* pairs (*mean* = 0.161, *SD* = 0.096), independently of adjective position, indicating that for illegal* pairs, adjectives and nouns more frequently exchanged position to occupy their legal* position (see Figure 1.2 and Appendix Table A.2).

Omission Errors. Finally, for omission errors, the best model included only legality* ($\beta = -0.23$, 95% *CI* [-0.37, -0.10], $BF_{10} = 43.78$; see Appendix Table A1). This model was 7.08 times more likely than the model including both legality* and adjective position (legality*: $\eta_p^2 = 0.090$; adjective position: $\eta_p^2 = 0.004$; legality* × adjective position: $\eta_p^2 = 0.002$). This result means that more omission errors occurred for syntactically illegal* pairs (*mean* = 0.190, *SD* = 0.086) than for legal* pairs (*mean* = 0.158, *SD* = 0.077), regardless of whether adjectives were antepositional or postpositional (see Figure 1.2 and Appendix Table A.2).

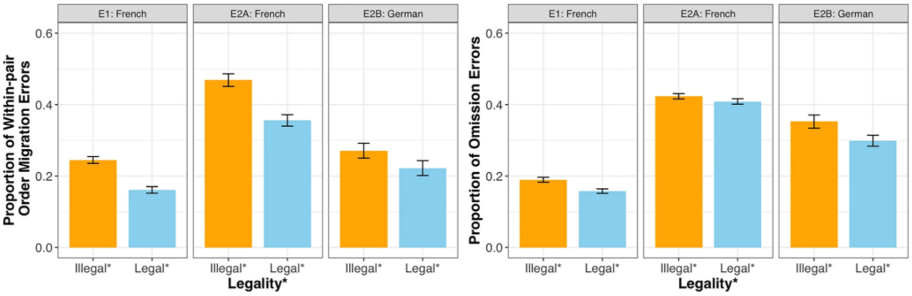

Figure 1.2. Proportion of within-pair order migration errors and omission errors, as a function of legality*, across Experiment 1 (French), Experiment 2A (French), and 2B (German).

Discussion

On the one hand, Experiment 1 replicated the results of Querella and Majerus (2024) as well as those by Schweppe and her colleagues (2022) by showing an impact of syntactic legality* on the overall serial recall and item recall scores. Critically, however, we also observed an effect of syntactic legality* on order recall, with reduced order recall accuracy for syntactic illegal* pairs accompanied by an increase in order migration errors. For illegal* pairs, these errors correspond to order regularization errors, where an adjective or a noun in an illegal* position moves to a legal* position. The present results thus show that, when the position of the adjectives (relative to the nouns) cannot be easily predicted based on temporary syntactic parsing of the list at encoding, then long-term syntactic knowledge structures can be shown to exert an impact on order recall. To determine the robustness of these novel findings, we aimed to replicate these results as well as to test their generality by extending them to the German language, in which adjective-noun order is deterministic rather than probabilistic.

Experiment 2

Experiment 2 aimed to replicate the impact of syntactic legality on order recall in WM and to extend the findings to the German language. Given that in German, only adjective anteposition is legal, mixing adjective anteposition and postposition in the same list but still comparing entirely legal and illegal lists is not possible. Therefore, reducing order predictability based on adjective position is only possible by mixing legal and illegal pairs in the same list (as adjective postposition will always be illegal in German). As this change in itself may affect the findings, we also used the same type of list setup for the replication study in the French language. Experiment 2 was thus divided into two sub-experiments. Experiment 2A aimed at replicating the findings of Experiment 1 in the French language by using an adapted list setup that could also be compared to a German version of the task. Experiment 2B then aimed at extending the findings to the German language. In order to ensure maximal comparability between the two language versions, we used for the French version only adjective postposition stimuli as legal* stimulus pairs, given that adjective postposition is more frequent than adjective anteposition and thus has a status more similar to the adjective anteposition pairs, which are the only legal type of pair in German. Both experiments were based on one-factorial designs with the within-subjects factor syntactic legality* (syntactically legal* vs. syntactically illegal* pairs). We expected effects of syntactic legality* on order recall in both languages.

Experiment 2A

Method

Transparency and Openness. We reported how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The design, hypotheses, and analysis plan for Experiment 2 were preregistered at https://osf.io/8z7t2.

Participants. Eighty-four participants (43 females and 41 males) aged from 18 to 30 (*average age* = 21.6, *SD* = 1.87) were recruited (see *Scoring and Analysis Procedure* section for justification of sample size) from the Université de Liège by advertisements or social media. Seventy-nine of them were right-handed and 5 were left-handed. The study had been approved by the ethics committee of the Faculty of Medicine of the University of Liège. Participants were informed that no financial compensation was provided. All participants have given their informed consent to participate in the study before testing.

Materials. The stimuli were selected from the stimulus pool of Experiment 1, with the selection restricted to preferred postpositional adjective-noun pairs. Consistent with Experiment 1, 16 lists, each containing three adjective-noun pairs, were constructed while controlling for phonological similarity (e.g., avoiding identical rhythmic patterns) and minimizing potential compound formations within each list. Additionally, the total number of syllables per list was controlled at 11. In total, 48 syntactically preferred postposition adjective-noun pairs (*mean preference* = 1.61, *SD* = 0.31, *range* = 1.21–2.12) were used to generate the experimental lists, 21 of which were the same pairs as in Experiment 1. To introduce syntactic unpredictability, the order of one or two adjective-noun pairs was reversed within each list (e.g., 2 adj-n + 1 n-adj or 1 adj-n + 2 n-adj per list). We first generated 16 lists (List 1) and then systematically reversed the adjective-noun order within the same pairs to create the corresponding set of lists (List 2). As in Experiment 1, 32 lists in total were presented.

Different from Experiment 1, all stimuli were presented in written format on a computer screen to eliminate potential prosodic-related differences when comparing the French and German sub-experiments. The same timing and duration parameters were used as for Experiment 1. Full stimulus lists are provided in Appendix Tables A.11 and A.12, the judgment results for each pair are in Table A.13, and the characteristics descriptions are in Table A.14.

Procedure. The procedure matched that in Experiment 1. Subjects were pseudo-randomly assigned to one of two presentation conditions (condition A: block of List 1 followed by block of List 2; condition B: block of List 2 followed by block of List 1).

Scoring and Analysis Procedure. Data scoring was identical to Experiment 1, as preregistered. The statistical design was the following: fixed factors only included legality* (legal* vs. illegal*), and the random-effect structure included random intercept of subjects, with or without random slope of legality* by subject. The same two-step Bayesian model comparison approach was used as in Experiment 1. Likewise, sample size had been determined via a Bayesian sequential sampling approach based on effect size stabilization (R. B. Anderson et al., 2022). The effect sizes of the item recall score and serial recall score were both stable after the 35th participant. As in Experiment 1, a higher number of participants was recruited to match the number of participants in the studies by Querella and Majerus (2024), and Schweppe and her colleagues (2022).

Results

Serial Recall Score. The first Bayesian generalized linear mixed model analysis on the serial recall scores indicated that the best-fitting model included legality* ($\beta = 0.30$, *95% CI* [0.17, 0.43]; see Appendix Table A.3). This model was 1.86×10^3 times more likely than the model excluding legality* (legality*: $\eta_p^2 = 0.224$). That means serial recall performance was significantly better for syntactically legal* pairs (*mean* = 0.629, *SD* = 0.016) compared to illegal* pairs (*mean* = 0.598, *SD* = 0.146), the same as in Experiment 1 (see Figure 1.1 and Appendix Table A.4).

Item Recall Score. For item recall scores, the factor legality* was associated with evidence for its absence ($BF_{01} = 5.79$; $\eta_p^2 = 0.014$; see Appendix Table A.3). The result means that item recall performance did not differ between syntactically legal* pairs (*mean* = 0.724, *SD* = 0.120) and illegal* pairs (*mean* = 0.718, *SD* = 0.117), in contrast to Experiment 1 (see Figure 1.1 and Appendix Table A.4).

Order Recall Score. For order recall scores, a Bayesian analysis of variance showed that the best-fitting model included legality* ($\beta = 0.28$, *95% CI* [0.14, 0.41]; see Appendix Table A.3). This model was 5.60×10^2 times more likely than the model excluding legality* (legality*: η_p^2 = 0.198). That means order recall performance was superior for syntactically legal* pairs (*mean* = 0.862, *SD* = 0.094) compared to illegal* pairs (*mean* = 0.826, *SD* = 0.109), consistent with Experiment 1 (see Figure 1.1 and Appendix Table A.4).

Within-Pair Order Migration Errors. Generalized linear mixed model analyses on within-pair order migration errors demonstrated that the best-fitting model included legality* (β = -0.33, *95% CI* [-0.49, -0.18]; see Appendix Table A.3). This model was 7.51×10^2 times more likely than the model excluding legality* (legality*: η_p^2 = 0.135). That means fewer order migration errors were observed for syntactically legal* pairs (*mean* = 0.356, *SD* = 0.146) than for illegal* pairs (*mean* = 0.469, *SD* = 0.162). Adjectives and nouns in syntactically illegal* positions were inclined to be recalled in their syntactically legal* positions, in line with prior findings in Experiment 1 (see Figure 1.2 and Appendix Table A.4).

Omission Errors. Finally, for omission errors, the factor legality* was associated with evidence for its absence (BF_{01} = 4.13; η_p^2 = 0.024; see Appendix Table A.3). The result means that no impact of syntactic legality* was observed for omission errors (*mean* = 0.409, *SD* = 0.070 vs. *mean* = 0.432, *SD* = 0.068), consistent with the results for item recall scores but different from Experiment 1 (see Figure 1.2 and Appendix Table A.4).

Experiment 2B

Method

Participants. Fifty participants (30 female and 20 male) aged from 20 to 31 (*average age* = 22.8, *SD* = 2.12) were recruited (see *Scoring and Analysis Procedure* section for justification of sample size) from the University of Passau. Forty-seven participants were right-handed and 3 were left-handed. Participants received course credit for their participation. Participants were treated in accordance with APA ethical standards as well as the guidelines of the German Research Foundation (DFG) and the German Psychological Society (DGPs). As the study was non-medical, low-risk research, no explicit approval was required from the responsible ethics committee at the University of Passau. All participants gave their informed consent to participate in the study before testing.

Materials and Procedure. The stimuli comprised 48 adjective-noun pairs. To align with the materials in Experiment 2A, only masculine nouns were used. To maintain maximal comparability with Experiment 2A, the stimuli were directly translated into German when the German noun was also of masculine gender (e.g., "nervöser – Busch" translated from "bus – nerveux" in French). In some cases, adjective-noun pairs were recombined to minimize phonological similarity within pairs and to equate list length. In addition, pairs with masculine nouns from Schweppe et al. (2022) were included. Unlike in French, the adjectives also needed to be inflected with masculine nouns. This resulted in a slightly longer list length than in Experiment 2A, with 13 or 14 syllables per list (*mean length* = 13.56 syllables). As in Experiment 2A, the order of one or two adjective-noun pairs was reversed within each of the 16 lists (List 1). A corresponding set of 16 lists (List 2) was created by systematically reversing the adjective-noun order within each of the pairs, resulting in 32 lists in total. The complete set of stimulus lists is provided in the Appendix (Tables A.15 and A.16), and the characteristics descriptions of stimuli are in Table A.17.

The procedure was the same as in Experiment 2A except that the experiment was conducted via jsPsych 6.3 (de Leeuw et al., 2023) with the psychophysics plugin (Kuroki, 2021).

Scoring and Analysis Procedure. The data scoring and analysis procedures were the same as in Experiment 2A.

Sample size had been determined via a Bayesian sequential sampling approach based on effect size stabilization (R. B. Anderson et al., 2022). The effect sizes of the item recall score and serial recall score were both stable after the 25th participant. As in the other experiments, a higher number of participants was recruited. Given the clear-cut results in Experiments 1 and 2A, recruitment, however, stopped at around 50 in order to avoid unnecessary participant recruitment and time investment.

Results

Serial Recall Score. We first conducted a generalized linear mixed model analysis on the serial recall scores. Results showed that the best-fitting model included legality ($\beta = 0.36$, *95% CI* [0.14, 0.59]; see Appendix Table A.5). This model was 27.26 times more likely than the model excluding legality (legality: $\eta_p^2 = 0.474$). That means serial recall performance was better for syntactically legal adjective-noun pairs (*mean* = 0.519, *SD* = 0.135) compared to illegal pairs (*mean* = 0.450, *SD* = 0.146), consistent with findings from Experiment 1 and Experiment 2A (see Figure 1.1 and Appendix Table A.6).

Item Recall Score. Next, the same analysis was performed on item recall scores. The best generalized mixed linear model included legality (β = 0.17, *95% CI* [0.09, 0.25]; see Appendix Table A.5). This model was 2.79×10^2 times more likely than the following model, excluding legality (legality: η_p^2 = 0.303). The result means that there was superior item recall performance for syntactically legal pairs (*mean* = 0.615, *SD* = 0.114) compared to illegal pairs (*mean* = 0.576, *SD* = 0.113). Interestingly, while differing from the results of Experiment 2A, this finding aligned with Experiment 1 (see Figure 1.1 and Appendix Table A.6).

Order Recall Score. For order recall scores, a Bayesian analysis of variance indicated the best-fitting model included legality (β = 0.39, *95% CI* [0.16, 0.61]; see Appendix Table A.5). This model was 53.71 times more likely than the model excluding legality (legality: η_p^2 = 0.220). That means that order recall performance of adjectives and nouns was better in syntactic legal order (*mean* = 0.835, *SD* = 0.090) than in syntactic illegal order (*mean* = 0.768, *SD* = 0.140), in line with Experiments 1 and 2A once again (see Figure 1.1 and Appendix Table A.6).

Within-Pair Order Migration Errors. Moreover, generalized linear mixed model analyses were conducted to examine errors. Analyses of within-pair order migration errors indicated that the best-fitting model included legality (β = -0.80, *95% CI* [-1.36, -0.28]; see Appendix Table A.5). This model was 51.37 times more likely than the model excluding legality (legality: η_p^2= 0.038). That means more within-pair order migration errors were observed with adjectives and nouns in syntactically illegal order (*mean* = 0.271, *SD* = 0.147) than legal order (*mean* = 0.222, *SD* = 0.147), indicating a higher frequency of adjectives and nouns in syntactically illegal order recalled at legal positions, consistent with Experiment 1 and 2A (see Figure 1.2 and Appendix Table A.6).

Omission Errors. Finally, the generalized linear mixed model analysis for omission errors identified the best-fitting model as including legality (β = -0.23, *95% CI* [-0.34, -0.12]; see Appendix Table A.5). This model was 2.99×10^2 times more likely than the model excluding legality (legality:

$\eta_p^2 = 0.290$). This result means that more items were forgotten when recalling syntactically illegal adjective-noun pairs (*mean* = 0.353, *SD* = 0.130) compared to legal pairs (*mean* = 0.299, *SD* = 0.109), aligned with Experiment 1, while contrasting to Experiment 2A (see Figure 1.2 and Appendix Table A.6).

Discussion

In sum, in both Experiment 2A and 2B, we observed a clear advantage for syntactically legal adjective-noun pairs over illegal pairs on order recall performance. Furthermore, we found the higher incidence of within-pair order migration errors for syntactically illegal pairs, where an adjective or a noun in an illegal* position has a greater possibility to move to a legal* position. One difference between Experiments 2A and 2B needs, however, to be noted. While in Experiment 2B (as well as in Experiment 1) we also observed recall performance differences on item recall and omission errors between legal* and illegal* pairs, this was not the case in Experiment 2A. This discrepancy may be explained by language-specific syntactic constraints and the specific setup of the lists in Experiment 2A. In French, as already mentioned, adjective-noun order is more flexible, allowing for greater variability in parsing and recall strategies, which may reduce the impact of syntactically illegal pairs on recall performance, particularly when only a subset of syntactic rules is being manipulated as opposed to Experiment 1. This idea is supported by an overall smaller difference between legal and illegal conditions in Experiment 2A for most of the scores, relative to Experiment 1, as well as to Experiment 2B.

General discussion

The present study replicated but also critically extends the experiments of Querella and Majerus (2024) and Schweppe et al. (2022), while introducing a crucial methodological modification: we intermixed adjective-noun and noun-adjective pairs within the same list to prevent the build-up of list-level temporary syntactic predictability rules. This methodological modification was associated with the observation of not only an impact of syntactic legality on item recall, but also, crucially, on order recall and order migration errors, in both French and German languages.

Traditional models of verbal WM, such as that of Baddeley and Hitch (1974), have suggested that serial order information is maintained via phonological mechanisms within a phonological store (Baddeley, 2000).

However, this perspective has been challenged as it does not fully account for the complexity of serial order phenomena (e.g., Hurlstone et al., 2014). Recent models of verbal working memory increasingly consider the interaction between lexico-semantic knowledge and working memory processes to explain serial recall. The Activated Network (ANet) model (Poirier et al., 2015) proposes that order information is supported by activation dynamics within long-term semantic memory: when later items are semantically related to earlier ones, their activation is boosted, increasing the likelihood of early recall. Other approaches have attempted to formalize the interface between semantic and episodic memory through computational modeling. For example, Mewhort and colleagues (2018) introduced a holographic memory model (Franklin & Mewhort, 2015) that uses BEAGLE vectors (Bound Encoding of the Aggregate Language Environment; Jones et al., 2006; Jones & Mewhort, 2007) to capture semantic similarity and explain effects such as release from proactive interference in immediate recall tasks. More recently, Guitard et al. (2025) developed the Embedded Computational Framework of Memory (eCFM), which combines Latent Semantic Analysis (Landauer & Dumais, 1997) with an episodic retrieval process based on MINERVA 2 (Hintzman, 1986) to account for the influence of semantic structure in working memory performance. However, these models have primarily focused on lexico-semantic features while largely overlooking the contribution of syntactic information to working memory. Moreover, most of them tend to treat memory performance as a unified outcome, without separating the cognitive mechanisms underlying item recall from those supporting order recall. This limits their ability to fully account for the complexity of language-based memory tasks in their current form. From another aspect, behavioral and neuroimaging evidence has suggested that serial order relies on domain-general mechanisms, such as spatial or temporal coding (e.g., Brown et al., 2000; Burgess & Hitch, 2006; Hartley et al., 2016; Majerus, 2008, 2013; van Dijck & Fias, 2011), whereas item recall is more directly linked to linguistic knowledge, particularly at the lexical-semantic level (Gathercole et al., 2001; Hulme et al., 1991, 1997; Saint-Aubin & Poirier, 1999; Walker & Hulme, 1999). However, our findings challenge the notion that the representation of serial order information is fully isolated from language processing. Instead, the observation of an impact of syntactic legality on order recall performance and order migration errors supports emergent models of verbal WM, which propose that WM and language systems are deeply interconnected and interact dynamically (Acheson & MacDonald, 2009; Buchsbaum & D'Esposito, 2019; Cowan, 1993; Hasson et al., 2015; MacDonald, 2016; Postle, 2006; Schwering & MacDonald, 2020). These models emphasize that verbal WM is the activated portion of linguistic long-term memory, and all aspects of verbal WM, including serial order, are grounded in linguistic structures (Schwering & MacDonald, 2020). The present study provides support for this perspective, with syntactic information being crucial for serial positioning even in word lists. Our findings are also in line with other recent studies that suggest an impact of linguistic variables on serial order coding. These findings contribute to a more refined understanding of the interaction between language and the representation of serial order information in verbal WM. While previous models of serial order WM have primarily emphasized the role of temporal context signals (Burgess & Hitch, 1999, 2006a; Lewandowsky & Farrell, 2008a; Oberauer et al., 2012), our results suggest that serial order encoding and maintenance are also influenced by fundamental linguistic structures, aligning with "limited emergent" approaches. Future models of serial order WM may consider the usefulness of integrating syntactic sequential knowledge structures in order to fully account for the multiple mechanisms that support coding of serial order information in WM, in particular when the memoranda do not simply represent one word class.

The present study further highlights the importance of controlling for the multiplicity of variables that can intervene in a verbal WM task when aiming at isolating the impact of a specific variable. Given the robust impact of syntactic legality on order memory in the present study, and its robust absence in previous studies (Querella & Majerus, 2024; Schweppe et al., 2022), it is indeed very likely that the inclusion of adjectives in the same position for a given list was hiding the contribution of long-term syntactic knowledge in the latter studies. Querella and Majerus (2024) had argued that this situation may have led to the build-up of temporary syntactic predictability rules, based on the creation of a particularly robust syntactic frame during the parsing of the stimuli during the encoding of the list. This temporary syntactic frame may have overcome the impact of syntactic knowledge on order recall and order migration errors, as it will constrain the order in which adjectives and nouns are expected to be output for a given list. Paradoxically, this situation may also have increased the impact of syntactic legality on item recall, as the temporary syntactic frame will predict the occurrence of an adjective for a position in which an item reflecting an adjective is not a suitable candidate based on syntactic legality, leading to an item omission error instead of an order error. In the present study, we prevented the impact of temporary syntactic parsing regularities by alternating adjective ante-/post-position in the same lists, and we were able to observe an impact of syntactic legality on order recall, and this impact appears to be even slightly more robust than the impact on item recall.

In conclusion, the present study provides direct evidence for an influence of syntactic knowledge on order recall performance in verbal WM, as opposed to item recall performance. Our findings indicate the need for a deeper integration of language and WM architectures, including for the modelling of serial order mechanisms.

Appendix

Table A.1. Bayesian comparison of each score in Experiment 1. Each value represents the BF$_{10}$ of the present model compared to the null model.

Models	Serial recall score	Item recall score	Order recall score	Within-pair order migration errors	Omission errors
Null (incl. random factors)	1.00	1.00	1.00	1.00	1.00
Legality*	1.32 e^{+4}	87.46	1.18 e^{+4}	8.75 e^{+5}	43.78
Adjective position	0.16	1.38	0.18	0.57	0.14
Legality* + Adjective position	3.51 e^{+3}	122.46	2.05 e^{+3}	7.80 e^{+4}	6.18
Legality* + Legality* × Adjective position	1.13 e^{+3}	117.05	6.94 e^{+2}	2.92 e^{+5}	2.29
Legality* + Adjective position + Legality* × Adjective position	2.81 e^{+3}	122.10	6.37 e^{+2}	2.49 e^{+4}	2.21

Note. All the models of serial recall score and within-pair order migration errors include a random intercept of subject and random slopes of legality* and the interaction of legality* and adjective position (interaction is not correlated to subject). All the models of item recall score, order recall score, and omission errors include a random intercept of subject and a random slope of legality*.

Table A.2. Descriptive statistics of the proportion of recall scores and error scores in Experiment 1

	Legality*	**Adjective position**	**Mean**	**SD**	**N**
Serial recall score	Legal*	Ante	0.678	0.171	93
		Post	0.657	0.158	93
	Illegal*	Ante	0.607	0.174	93
		Post	0.599	0.182	93
Item recall score	Legal*	Ante	0.793	0.125	93
		Post	0.768	0.108	93
	Illegal*	Ante	0.757	0.110	93
		Post	0.753	0.117	93
Order recall score	Legal*	Ante	0.848	0.128	93
		Post	0.848	0.127	93
	Illegal*	Ante	0.794	0.157	93
		Post	0.785	0.161	93
Within-pair order migration errors	Legal*	Ante	0.162	0.129	93
		Post	0.160	0.121	93
	Illegal*	Ante	0.247	0.140	93
		Post	0.243	0.121	93
Omission errors	Legal*	Ante	0.160	0.089	93
		Post	0.155	0.082	93
	Illegal*	Ante	0.190	0.101	93
		Post	0.190	0.089	93

Table A.3. Bayesian comparison of each score in Experiment 2A. Each value represents the BF_{10} of the present model compared to the null model.

Model	**Serial recall score**	**Item recall score**	**Order recall score**	**Within-pair order migration errors**	**Omission errors**
Null (incl. random factors)	1.00	1.00	1.00	1.00	1.00
Legality*	$1.86\ e^{+3}$	0.17	559.66	750.57	0.24

Note. All the models of serial recall score and within-pair order migration errors include a random intercept of subject and a random slope of legality*. All the models of item recall score, order recall score, and omission errors include only a random intercept of subject.

Table A.4. Descriptive statistics of the proportion of recall scores and error scores in Experiment 2A

	Legality*	**Mean**	**SD**	**N**
Serial recall score	Legal*	0.629	0.146	84
	Illegal*	0.598	0.147	84
Item recall score	Legal*	0.724	0.120	84
	Illegal*	0.718	0.117	84
Order recall score	Legal*	0.862	0.094	84
	Illegal*	0.826	0.109	84
Within-pair order migration errors	Legal*	0.356	0.146	84
	Illegal*	0.469	0.162	84
Omission errors	Legal*	0.409	0.070	84
	Illegal*	0.423	0.068	84

Table A.5. Bayesian comparison of each score in Experiment 2B. Each value represents the BF_{10} of the present model compared to the null model.

Model	**Serial recall score**	**Item recall score**	**Order recall score**	**Within-pair order migration errors**	**Omission errors**
Null (incl. random factors)	1.00	1.00	1.00	1.00	1.00
Legality	27.26	279.07	53.71	51.37	298.90

Note. All the models of serial recall score, within-pair order migration errors, and omission errors include a random intercept of subject and a random slope of legality. All the models of item recall score and order recall score include only a random intercept of subject.

Table A.6. Descriptive statistics of the proportion of recall scores and error scores in Experiment 2B

	Legality	**Mean**	**SD**	**N**
Serial recall score	Legal	0.519	0.135	50
	Illegal	0.450	0.146	50
Item recall score	Legal	0.615	0.114	50
	Illegal	0.576	0.113	50
Order recall score	Legal	0.835	0.090	50
	Illegal	0.768	0.140	50
Within-pair order migration errors	Legal	0.222	0.147	50
	Illegal	0.271	0.147	50
Omission errors	Legal	0.299	0.109	50
	Illegal	0.353	0.130	50

Table A.7. Lists with legal* order pairs in Experiment 1, including English translations (in parentheses) and word class [in brackets]. Materials from Querella and Majerus (2023) are in grey.

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	grand$_{\text{[adj.]}}$ (big)	titane$_{\text{[n.]}}$ (titanium)	sale$_{\text{[adj.]}}$ (dirty)	silence$_{\text{[n.]}}$ (silence)	orteil$_{\text{[n.]}}$ (toe)	agressif$_{\text{[adj.]}}$ (aggressive)
2	lourd$_{\text{[adj.]}}$ (heavy)	repos$_{\text{[n.]}}$ (rest)	tissu$_{\text{[n.]}}$ (tissue)	creux$_{\text{[adj.]}}$ (hollow)	second$_{\text{[adj.]}}$ (second)	silicone$_{\text{[n.]}}$ (silicone)
3	coussin$_{\text{[n.]}}$ (cushion)	sec$_{\text{[adj.]}}$ (dry)	jasmin$_{\text{[n.]}}$ (jasmine)	méfiant$_{\text{[adj.]}}$ (mistrustful)	petit$_{\text{[adj.]}}$ (littile)	laiton$_{\text{[n.]}}$ (brass)
4	témoin $_{\text{[n.]}}$ (witness)	naturel$_{\text{[adj.]}}$ (natural)	beau$_{\text{[adj.]}}$ (beautiful)	vecteur$_{\text{[n.]}}$ (vector)	pull$_{\text{[n.]}}$ (sweater)	jaloux$_{\text{[adj.]}}$ (jealous)
5	dernier$_{\text{[adj.]}}$ (last)	thorax$_{\text{[n.]}}$ (thrax)	brave$_{\text{[adj.]}}$ (brave)	cheveu$_{\text{[n.]}}$ (hair)	poignet$_{\text{[n.]}}$ (wrist)	gluant$_{\text{[adj.]}}$ (sticky)
6	sombre$_{\text{[adj.]}}$ (dark)	effort$_{\text{[n.]}}$ (effort)	pied$_{\text{[n.]}}$ (foot)	facile$_{\text{[adj.]}}$ (easy)	mauvais$_{\text{[adj.]}}$ (wrong)	vestibule$_{\text{[n.]}}$ (vestibule)
7	soda$_{\text{[n.]}}$ (soda)	naïf$_{\text{[adj.]}}$ (naive)	manteau$_{\text{[n.]}}$ (coat)	bavard$_{\text{[adj.]}}$ (talkative)	noble$_{\text{[adj.]}}$ (noble)	hasard$_{\text{[n.]}}$ (chance)
8	menton$_{\text{[n.]}}$ (chin)	brumeux$_{\text{[adj.]}}$ (foggy)	gros$_{\text{[adj.]}}$ (big)	pseudonyme$_{\text{[n.]}}$ (pseudonym)	lieu$_{\text{[n.]}}$ (venue)	mignon$_{\text{[adj.]}}$ (cute)
9	nouveau$_{\text{[adj.]}}$ (new)	trapèze$_{\text{[n.]}}$ (trapeze)	doux$_{\text{[adj.]}}$ (sweet)	dommage$_{\text{[n.]}}$ (pity)	fusil$_{\text{[n.]}}$ (musket)	malade$_{\text{[adj.]}}$ (sick)
10	vilain$_{\text{[adj.]}}$ (naughty)	sommeil$_{\text{[n.]}}$ (sleep)	poivron$_{\text{[n.]}}$ (pepper)	discret$_{\text{[adj.]}}$ (discreet)	faible$_{\text{[adj.]}}$ (weak)	rideau$_{\text{[n.]}}$ (curtain)
11	bus$_{\text{[n.]}}$ (bus)	nerveux$_{\text{[adj.]}}$ (nervous)	poumon$_{\text{[n.]}}$ (lung)	dingue$_{\text{[adj.]}}$ (crazy)	gentil$_{\text{[adj.]}}$ (kindly)	cerveau$_{\text{[n.]}}$ (brain)
12	vaisseau$_{\text{[n.]}}$ (vessel)	fou$_{\text{[adj.]}}$ (mad)	premier$_{\text{[adj.]}}$ (first)	terroir$_{\text{[n.]}}$ (land)	coeur$_{\text{[n.]}}$ (heart)	illégal$_{\text{[adj.]}}$ (illegal)
13	joyeux$_{\text{[adj.]}}$ (joyful)	béret$_{\text{[n.]}}$ (beret)	curieux$_{\text{[adj.]}}$ (curious)	salon$_{\text{[n.]}}$ (living room)	crâne$_{\text{[n.]}}$ (skull)	gratuit$_{\text{[adj.]}}$ (free)
14	pauvre$_{\text{[adj.]}}$ (poor)	numéro$_{\text{[n.]}}$ (number)	fruit$_{\text{[n.]}}$ (fruit)	courageux$_{\text{[adj.]}}$ (courageous)	stupide$_{\text{[adj.]}}$ (dumb)	ventre$_{\text{[n.]}}$ (belly)

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
15	vent$_{\text{[n.]}}$ (wind)	intact$_{\text{[adj.]}}$ (unbroken)	voleur$_{\text{[n.]}}$ (thief)	amer$_{\text{[adj.]}}$ (bitter)	puissant$_{\text{[adj.]}}$ (mighty)	nombril$_{\text{[n.]}}$ (navel)
16	dos$_{\text{[n.]}}$ (back)	coupable$_{\text{[adj.]}}$ (guilty)	vieux$_{\text{[adj.]}}$ (old)	pentagone$_{\text{[n.]}}$ (pentagon)	saumon$_{\text{[n.]}}$ (salmon)	ringard$_{\text{[adj.]}}$ (corny)

Table A.8. Lists with illegal* order pairs in Experiment 1, including English translations (in parentheses) and word class [in brackets]. Materials from Querella and Majerus (2023) are in grey.

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	titane$_{[n.]}$ (titanium)	grand$_{[adj.]}$ (big)	silence$_{[n.]}$ (silence)	sale$_{[adj.]}$ (dirty)	agressif$_{[adj.]}$ (aggressive)	orteil$_{[n.]}$ (toe)
2	repos$_{[n.]}$ (rest)	lourd$_{[adj.]}$ (heavy)	creux$_{[adj.]}$ (hollow)	tissu$_{[n.]}$ (tissue)	silicone$_{[n.]}$ (silicone)	second$_{[adj.]}$ (second)
3	sec$_{[adj.]}$ (dry)	coussin$_{[n.]}$ (cushion)	méfiant$_{[adj.]}$ (mistrustful)	jasmin$_{[n.]}$ (jasmine)	laiton$_{[n.]}$ (brass)	petit$_{[adj.]}$ (littile)
4	naturel$_{[adj.]}$ (natural)	témoin $_{[n.]}$ (witness)	vecteur$_{[n.]}$ (vector)	beau$_{[adj.]}$ (beautiful)	jaloux$_{[adj.]}$ (jealous)	pull$_{[n.]}$ (sweater)
5	thorax$_{[n.]}$ (thrax)	dernier$_{[adj.]}$ (last)	cheveu$_{[n.]}$ (hair)	brave$_{[adj.]}$ (brave)	gluant$_{[adj.]}$ (sticky)	poignet$_{[n.]}$ (wrist)
6	effort$_{[n.]}$ (effort)	sombre$_{[adj.]}$ (dark)	facile$_{[adj.]}$ (easy)	pied$_{[n.]}$ (foot)	vestibule$_{[n.]}$ (vestibule)	mauvais$_{[adj.]}$ (wrong)
7	naïf$_{[adj.]}$ (naive)	soda$_{[n.]}$ (soda)	bavard$_{[adj.]}$ (talkative)	manteau$_{[n.]}$ (coat)	hasard$_{[n.]}$ (chance)	noble$_{[adj.]}$ (noble)
8	brumeux$_{[adj.]}$ (foggy)	menton$_{[n.]}$ (chin)	pseudonyme$_{[n.]}$ (pseudonym)	gros$_{[adj.]}$ (big)	mignon$_{[adj.]}$ (cute)	lieu$_{[n.]}$ (venue)
9	trapèze$_{[n.]}$ (trapeze)	nouveau$_{[adj.]}$ (new)	dommage$_{[n.]}$ (pity)	doux$_{[adj.]}$ (sweet)	malade$_{[adj.]}$ (sick)	fusil$_{[n.]}$ (musket)
10	sommeil$_{[n.]}$ (sleep)	vilain$_{[adj.]}$ (naughty)	discret$_{[adj.]}$ (discreet)	poivron$_{[n.]}$ (pepper)	rideau$_{[n.]}$ (curtain)	faible$_{[adj.]}$ (weak)
11	nerveux$_{[adj.]}$ (nervous)	bus$_{[n.]}$ (bus)	dingue$_{[adj.]}$ (crazy)	poumon$_{[n.]}$ (lung)	cerveau$_{[n.]}$ (brain)	gentil$_{[adj.]}$ (kindly)
12	fou$_{[adj.]}$ (mad)	vaisseau$_{[n.]}$ (vessel)	terroir$_{[n.]}$ (land)	premier$_{[adj.]}$ (first)	illégal$_{[adj.]}$ (illegal)	coeur$_{[n.]}$ (heart)
13	béret$_{[n.]}$ (beret)	joyeux$_{[adj.]}$ (joyful)	salon$_{[n.]}$ (living room)	curieux$_{[adj.]}$ (curious)	gratuit$_{[adj.]}$ (free)	crâne$_{[n.]}$ (skull)
14	numéro$_{[n.]}$ (number)	pauvre$_{[adj.]}$ (poor)	courageux$_{[adj.]}$ (courageous)	fruit$_{[n.]}$ (fruit)	ventre$_{[n.]}$ (belly)	stupide$_{[adj.]}$ (dumb)

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
15	intact$_{[adj.]}$ (unbroken)	vent$_{[n.]}$ (wind)	amer$_{[adj.]}$ (bitter)	voleur$_{[n.]}$ (thief)	nombril$_{[n.]}$ (navel)	puissant$_{[adj.]}$ (mighty)
16	coupable$_{[adj.]}$ (guilty)	dos$_{[n.]}$ (back)	pentagone$_{[n.]}$ (pentagon)	vieux$_{[adj.]}$ (old)	ringard$_{[adj.]}$ (corny)	saumon$_{[n.]}$ (salmon)

Table A.9. Descriptive statistics of pairs in Experiment 1 according to subjects' ratings (N = 57). A higher mean score means a stronger preference for adjective-noun order, while a lower mean score means a stronger preference for noun-adjective order.

Adjective-noun pair	Mean	SD	Minimum	Maximum	Noun-adjective pair	Mean	SD	Minimum	Maximum
nouveau trapèze	4.83	0.43	3	5	orteil agressif	1.211	0.526	1	4
gros pseudonyme	4.79	0.45	3	5	tissu creux	1.211	0.526	1	4
vieux pentagone	4.79	0.59	2	5	coussin sec	1.211	0.559	1	4
beau vecteur	4.77	0.66	1	5	pied facile	1.211	0.491	1	3
petit laiton	4.75	0.61	2	5	manteau bavard	1.246	0.635	1	4
mauvais vestibule	4.74	0.67	1	5	lieu mignon	1.246	0.635	1	5
premier terroir	4.74	0.52	3	5	crâne gratuit	1.246	0.576	1	4
dernier thorax	4.72	0.53	3	5	vent intact	1.263	0.669	1	5
second silicone	4.63	0.79	2	5	vaisseau fou	1.298	0.778	1	5
grand titane	4.60	0.75	2	5	saumon ringard	1.316	0.711	1	5
noble hasard	4.54	0.91	1	5	pull jaloux	1.351	0.834	1	5
gentil cerveau	4.19	1.06	1	5	coeur illégal	1.368	0.771	1	5
vilain sommeil	4.16	1.21	1	5	fusil malade	1.368	0.957	1	5
pauvre numéro	4.07	1.27	1	5	poignet gluant	1.404	0.863	1	5
sombre effort	4.00	1.39	1	5	menton brumeux	1.421	0.823	1	4
faible rideau	3.93	1.28	1	5	dos coupable	1.421	1.017	1	5
curieux salon	3.86	1.43	1	5	soda naïf	1.474	0.947	1	5
sale silence	3.83	1.43	1	5	poivron discret	1.614	1.114	1	5
brave cheveu	3.74	1.42	1	5	bus nerveux	1.632	1.096	1	5
doux dommage	3.70	1.46	1	5	poumon dingue	1.684	1.183	1	5
lourd repos	3.67	1.46	1	5	témoin naturel	1.719	1.25	1	5
joyeux béret	3.63	1.53	1	5	voleur amer	1.86	1.274	1	5
puissant nombril	3.54	1.50	1	5	jasmin méfiant	1.895	1.305	1	5
stupide ventre	3.46	1.56	1	5	fruit courageux	1.93	1.307	1	5

Table A.10. Characteristic description of word stimuli in Experiment 1 according to a French database: http://www.lexique.org. Freq means the frequency; NSorth means orthographic neighborhood size; NSphon means phonological neighborhood size; Nphon means number of phonemes; Nsyll means number of syllables.

Adjective	Freq	NS orth	NS phon	N phon	N syll	Noun	Freq	NS orth	NS phon	N phon	N syll
agressif	5.76	0	2	7	3	béret	1.19	1	8	4	2
amer	5.82	3	15	4	2	bus	0.5	19	25	2	1
bavard	2.39	3	9	5	2	cerveau	57.67	2	8	5	2
beau	281.23	4	24	2	1	cheveu	5.11	1	4	4	2
brave	24.55	11	10	4	1	coeur	224.98	1	13	3	1
brumeux	1.04	0	0	5	2	coussin	2.44	2	6	4	2
coupable	46.45	0	1	6	2	crâne	26.88	3	14	4	1
courageux	20.22	0	0	6	3	dommage	59.43	2	6	5	2
creux	2.83	3	10	3	1	dos	100.34	15	24	2	1
curieux	26.11	1	1	5	2	effort	23.26	0	3	4	2
dernier	138.57	1	2	6	2	fruit	15.99	2	2	4	1
dingue	53.01	2	7	3	1	fusil	36.52	0	7	4	2
discret	9.52	0	1	6	2	hasard	46.98	2	10	4	2
doux	37.66	6	28	2	1	jasmin	1.57	1	1	5	2
facile	153.57	0	3	5	2	laiton	0.21	0	9	4	2
faible	31.03	2	2	4	1	lieu	153.12	8	13	3	1
fou	181.51	13	27	2	1	manteau	36.16	2	17	4	2
gentil	134.11	0	6	4	2	menton	6.45	4	20	4	2
gluant	1.23	1	1	4	2	nombril	4.26	0	0	6	2
grand	338.27	3	15	3	1	numéro	162.08	0	0	6	3
gratuit	12.15	0	0	6	2	orteil	4.04	0	0	5	2
gros	180.91	6	15	3	1	pentagone	0.25	0	0	7	3
illégal	11.93	0	1	6	3	pied	105.51	5	17	3	1
intact	6.21	0	1	5	2	poignet	6.38	2	7	5	2
jaloux	29.87	0	2	4	2	poivron	0.51	1	4	6	2

Adjective	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**	**Noun**	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**
joyeux	31.81	3	3	5	2	poumon	4.55	1	5	4	2
lourd	10.15	2	23	3	1	pseudonyme	1.46	0	0	8	3
malade	147.5	5	6	5	2	pull	11.41	2	20	3	1
mauvais	138.04	2	8	4	2	repos	42.29	2	6	4	2
méfiant	0.09	3	5	5	2	rideau	10.81	1	10	4	2
mignon	46.14	3	10	4	2	salon	37.06	10	21	4	2
naïf	4.22	0	1	4	2	saumon	5.28	1	11	4	2
naturel	19.01	1	0	7	3	silence	105.53	0	1	5	2
nerveux	32.52	1	4	5	2	silicone	2.42	2	0	7	3
noble	23.73	1	1	4	1	soda	9.12	7	14	4	2
nouveau	106.48	0	1	4	2	sommeil	44.51	0	2	5	2
pauvre	148.93	0	0	4	1	témoin	49.35	0	0	5	2
petit	573.72	1	5	4	2	terroir	0.17	0	1	6	2
premier	146.12	0	0	6	2	thorax	3.62	0	1	6	2
puissant	22.65	3	3	5	2	tissu	9.21	3	5	4	2
ringard	1.99	0	1	5	2	titane	1.69	4	3	5	2
sale	120.13	16	25	3	1	trapèze	1.1	0	0	6	2
sec	27.4	13	28	3	1	vaisseau	67.11	0	3	4	2
second	15.32	0	3	4	2	vecteur	0.62	3	3	6	2
sombre	24.66	4	3	4	1	vent	71.5	13	27	2	1
stupide	60.06	0	0	6	2	ventre	46.07	5	16	4	1
vieux	180.08	7	9	3	1	vestibule	0.96	0	0	8	3
vilain	11.03	0	2	4	2	voleur	41.39	2	3	5	2

Table A.11. Lists in Experiment 2A (List 1), including English translations (in parentheses) and word class [in brackets]. Materials from Querella and Majerus (2023) are in grey.

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	gaz$_{\text{[n.]}}$ (gas)	moche$_{\text{[adj.]}}$ (ugly)	océan$_{\text{[n.]}}$ (ocean)	inutile$_{\text{[adj.]}}$ (useless)	sensible$_{\text{[adj.]}}$ (sensitive)	piège$_{\text{[n.]}}$ (trap)
2	soda$_{\text{[n.]}}$ (soda)	naïf$_{\text{[adj.]}}$ (naive)	entier$_{\text{[adj.]}}$ (entire)	progrès$_{\text{[n.]}}$ (progress)	thé$_{\text{[n.]}}$ (tea)	content$_{\text{[adj.]}}$ (thrilled)
3	compétent$_{\text{[adj.]}}$ (competent)	tiroir$_{\text{[n.]}}$ (drawer)	fidèle$_{\text{[adj.]}}$ (faithful)	souffle$_{\text{[n.]}}$ (breath)	pied$_{\text{[n.]}}$ (foot)	facile$_{\text{[adj.]}}$ (easy)
4	étroit$_{\text{[adj.]}}$ (narrow)	travail$_{\text{[n.]}}$ (work)	poumon$_{\text{[n.]}}$ (lung)	dingue$_{\text{[adj.]}}$ (crazy)	silencieux$_{\text{[adj.]}}$ (silent)	nez$_{\text{[n.]}}$ (nose)
5	arrêt$_{\text{[n.]}}$ (stop)	malin$_{\text{[adj.]}}$ (clever)	pot$_{\text{[n.]}}$ (jar)	furieux$_{\text{[adj.]}}$ (furious)	radin$_{\text{[adj.]}}$ (stingy)	passeport$_{\text{[n.]}}$ (passport)
6	signal$_{\text{[n.]}}$ (signal)	inquiet$_{\text{[adj.]}}$ (worried)	creux$_{\text{[adj.]}}$ (hollow)	tissu$_{\text{[n.]}}$ (fabric)	âge$_{\text{[n.]}}$ (age)	délicat$_{\text{[adj.]}}$ (delicate)
7	intact$_{\text{[adj.]}}$ (intact)	vent$_{\text{[n.]}}$ (wind)	malade$_{\text{[adj.]}}$ (sick)	fusil$_{\text{[n.]}}$ (gun)	terrain$_{\text{[n.]}}$ (ground)	urgent$_{\text{[adj.]}}$ (urgent)
8	capable$_{\text{[adj.]}}$ (able)	portrait$_{\text{[n.]}}$ (portrait)	vaisseau$_{\text{[n.]}}$ (vessel)	fou$_{\text{[adj.]}}$ (mad)	gluant$_{\text{[adj.]}}$ (sticky)	poignet$_{\text{[n.]}}$ (wrist)
9	manteau$_{\text{[n.]}}$ (coat)	bavard$_{\text{[adj.]}}$ (talkative)	goût$_{\text{[n.]}}$ (taste)	calme$_{\text{[adj.]}}$ (calm)	agressif$_{\text{[adj.]}}$ (aggressive)	orteil$_{\text{[n.]}}$ (toe)
10	fromage$_{\text{[n.]}}$ (cheese)	émotif$_{\text{[adj.]}}$ (emotional)	disponible$_{\text{[adj.]}}$ (available)	crime$_{\text{[n.]}}$ (crime)	bruit$_{\text{[n.]}}$ (noise)	chaud$_{\text{[adj.]}}$ (hot)
11	absent$_{\text{[adj.]}}$ (absent)	papier$_{\text{[n.]}}$ (paper)	courageux$_{\text{[adj.]}}$ (brave)	fruit$_{\text{[n.]}}$ (fruit)	dos$_{\text{[n.]}}$ (back)	coupable$_{\text{[adj.]}}$ (guilty)
12	gratuit$_{\text{[adj.]}}$ (free)	crâne$_{\text{[n.]}}$ (skull)	coeur$_{\text{[n.]}}$ (heart)	illégal$_{\text{[adj.]}}$ (illegal)	discret$_{\text{[adj.]}}$ (discreet)	poivron$_{\text{[n.]}}$ (pepper)
13	téléphone$_{\text{[n.]}}$ (phone)	aveugle$_{\text{[adj.]}}$ (blind)	foyer$_{\text{[n.]}}$ (hearth)	ivre$_{\text{[adj.]}}$ (drunk)	ringard$_{\text{[adj.]}}$ (cheesy)	saumon$_{\text{[n.]}}$ (salmon)
14	coussin$_{\text{[n.]}}$ (cushion)	sec$_{\text{[adj.]}}$ (dry)	sérieux$_{\text{[adj.]}}$ (serious)	cercle$_{\text{[n.]}}$ (circle)	muguet$_{\text{[n.]}}$ (thrush)	familier$_{\text{[adj.]}}$ (familiar)
15	gourmand$_{\text{[adj.]}}$ (greedy)	balcon$_{\text{[n.]}}$ (balcony)	jaloux$_{\text{[adj.]}}$ (jealous)	pull$_{\text{[n.]}}$ (sweater)	menton$_{\text{[n.]}}$ (chin)	brumeux$_{\text{[adj.]}}$ (foggy)

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
16	amer$_{[adj.]}$ (bitter)	voleur$_{[n.]}$ (thief)	bus$_{[n.]}$ (bus)	nerveux$_{[adj.]}$ (nervous)	glissant$_{[adj.]}$ (sliding)	champagne$_{[n.]}$ (champagne)

Table A.12. Lists in Experiment 2A (List 2), including English translations (in parentheses) and word class [in brackets]. Materials from Querella and Majerus (2023) are in grey.

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	moche$_{[adj.]}$ (ugly)	gaz$_{[n.]}$ (gas)	inutile$_{[adj.]}$ (useless)	océan$_{[n.]}$ (ocean)	piège$_{[n.]}$ (trap)	sensible$_{[adj.]}$ (sensitive)
2	naïf$_{[adj.]}$ (naive)	soda$_{[n.]}$ (soda)	progrès$_{[n.]}$ (progress)	entier$_{[adj.]}$ (entire)	content$_{[adj.]}$ (thrilled)	thé$_{[n.]}$ (tea)
3	tiroir$_{[n.]}$ (drawer)	compétent$_{[adj.]}$ (competent)	souffle$_{[n.]}$ (breath)	fidèle$_{[adj.]}$ (faithful)	facile$_{[adj.]}$ (easy)	pied$_{[n.]}$ (foot)
4	travail$_{[n.]}$ (work)	étroit$_{[adj.]}$ (narrow)	dingue$_{[adj.]}$ (crazy)	poumon$_{[n.]}$ (lung)	nez$_{[n.]}$ (nose)	silencieux$_{[adj.]}$ (silent)
5	malin$_{[adj.]}$ (clever)	arrêt$_{[n.]}$ (stop)	furieux$_{[adj.]}$ (furious)	pot$_{[n.]}$ (jar)	passeport$_{[n.]}$ (passport)	radin$_{[adj.]}$ (stingy)
6	inquiet$_{[adj.]}$ (worried)	signal$_{[n.]}$ (signal)	tissu$_{[n.]}$ (fabric)	creux$_{[adj.]}$ (hollow)	délicat$_{[adj.]}$ (delicate)	âge$_{[n.]}$ (age)
7	vent$_{[n.]}$ (wind)	intact$_{[adj.]}$ (intact)	fusil$_{[n.]}$ (gun)	malade$_{[adj.]}$ (sick)	urgent$_{[adj.]}$ (urgent)	terrain$_{[n.]}$ (ground)
8	portrait$_{[n.]}$ (portrait)	capable$_{[adj.]}$ (able)	fou$_{[adj.]}$ (mad)	vaisseau$_{[n.]}$ (vessel)	poignet$_{[n.]}$ (wrist)	gluant$_{[adj.]}$ (sticky)
9	bavard$_{[adj.]}$ (talkative)	manteau$_{[n.]}$ (coat)	calme$_{[adj.]}$ (calm)	goût$_{[n.]}$ (taste)	orteil$_{[n.]}$ (toe)	agressif$_{[adj.]}$ (aggressive)
10	émotif$_{[adj.]}$ (emotional)	fromage$_{[n.]}$ (cheese)	crime$_{[n.]}$ (crime)	disponible$_{[adj.]}$ (available)	chaud$_{[adj.]}$ (hot)	bruit$_{[n.]}$ (noise)
11	papier$_{[n.]}$ (paper)	absent$_{[adj.]}$ (absent)	fruit$_{[n.]}$ (fruit)	courageux$_{[adj.]}$ (brave)	coupable$_{[adj.]}$ (guilty)	dos$_{[n.]}$ (back)
12	crâne$_{[n.]}$ (skull)	gratuit$_{[adj.]}$ (free)	illégal$_{[adj.]}$ (illegal)	coeur$_{[n.]}$ (heart)	poivron$_{[n.]}$ (pepper)	discret$_{[adj.]}$ (discreet)
13	aveugle$_{[adj.]}$ (blind)	téléphone$_{[n.]}$ (phone)	ivre$_{[adj.]}$ (drunk)	foyer$_{[n.]}$ (hearth)	saumon$_{[n.]}$ (salmon)	ringard$_{[adj.]}$ (cheesy)
14	sec$_{[adj.]}$ (dry)	coussin$_{[n.]}$ (cushion)	cercle$_{[n.]}$ (circle)	sérieux$_{[adj.]}$ (serious)	familier$_{[adj.]}$ (familiar)	muguet$_{[n.]}$ (thrush)
15	balcon$_{[n.]}$ (balcony)	gourmand$_{[adj.]}$ (greedy)	pull$_{[n.]}$ (sweater)	jaloux$_{[adj.]}$ (jealous)	brumeux$_{[adj.]}$ (foggy)	menton$_{[n.]}$ (chin)

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
16	voleur$_{[n.]}$ (thief)	amer$_{[adj.]}$ (bitter)	nerveux$_{[adj.]}$ (nervous)	bus$_{[n.]}$ (bus)	champagne$_{[n.]}$ (champagne)	glissant$_{[adj.]}$ (sliding)

Table A.13. Descriptive statistics of pairs in Experiment 2A according to subjects' ratings (N = 57).

Noun-adjective pair	Mean	SD	Minimum	Maximum	Noun-adjective pair	Mean	SD	Minimum	Maximum
coussin sec	1.21	0.56	1	4	champagne glissant	1.54	1.02	1	5
orteil agressif	1.21	0.53	1	4	âge délicat	1.58	0.98	1	5
pied facile	1.21	0.49	1	3	poivron discret	1.61	1.11	1	5
tissu creux	1.21	0.53	1	4	bus nerveux	1.63	1.10	1	5
crâne gratuit	1.25	0.58	1	4	thé content	1.63	1.23	1	5
manteau bavard	1.25	0.63	1	4	terrain urgent	1.65	1.14	1	5
vent intact	1.26	0.67	1	5	gaz moche	1.67	1.19	1	5
crime disponible	1.28	0.53	1	3	portrait capable	1.68	1.02	1	5
vaisseau fou	1.30	0.78	1	5	poumon dingue	1.68	1.18	1	5
saumon ringard	1.32	0.71	1	5	goût calme	1.82	1.27	1	5
bruit chaud	1.35	0.74	1	5	voleur amer	1.86	1.27	1	5
pull jaloux	1.35	0.83	1	5	fruit courageux	1.93	1.31	1	5
coeur illégal	1.37	0.77	1	5	pot furieux	2.00	1.41	1	5
fromage émotif	1.37	0.72	1	4	progrès entier	2.00	1.38	1	5
fusil malade	1.37	0.96	1	5	balcon gourmand	2.04	1.44	1	5
téléphone aveugle	1.39	0.62	1	4	piège sensible	2.04	1.51	1	5
poignet gluant	1.40	0.86	1	5	arrêt malin	2.05	1.42	1	5
dos coupable	1.42	1.02	1	5	océan inutile	2.05	1.29	1	5
menton brumeux	1.42	0.82	1	4	signal inquiet	2.05	1.47	1	5
muguet familier	1.44	1.04	1	5	travail étroit	2.05	1.33	1	5
tiroir compétent	1.46	1.02	1	5	passeport radin	2.07	1.36	1	5
soda naïf	1.47	0.95	1	5	foyer ivre	2.07	1.46	1	5
papier absent	1.49	0.95	1	5	cercle sérieux	2.11	1.36	1	5
nez silencieux	1.51	1.04	1	5	souffle fidèle	2.12	1.48	1	5

Table A.14. Characteristic description of word stimuli in Experiment 2A according to a French database: http://www.lexique.org. Freq means the frequency; NSorth means orthographic neighborhood size; NSphon means phonological neighborhood size; Nphon means number of phonemes; Nsyll means number of syllables.

Adjective	Freq	NS orth	NS phon	N phon	N syll	Noun	Freq	NS orth	NS phon	N phon	N syll
absent	9.46	0	5	4	2	âge	150.45	4	18	2	1
agressif	5.76	0	2	7	3	arrêt	46.8	0	28	3	2
amer	5.82	3	15	4	2	balcon	9.9	1	0	5	2
aveugle	33.85	2	0	5	2	bruit	78.94	2	3	4	1
bavard	2.39	3	9	5	2	bus	0.5	19	25	2	1
brumeux	1.04	0	0	5	2	cercle	17.77	1	1	5	1
calme	58.78	9	16	4	1	champagne	32.38	0	1	5	2
capable	65.15	1	3	6	2	coeur	224.98	1	13	3	1
chaud	50.2	3	20	2	1	coussin	2.44	2	6	4	2
compétent	4.45	1	0	6	3	crâne	26.88	3	14	4	1
content	114.75	8	13	4	2	crime	81.77	8	15	4	1
coupable	46.45	0	1	6	2	dos	100.34	15	24	2	1
courageux	20.22	0	0	6	3	foyer	25.57	5	11	5	2
creux	2.83	3	10	3	1	fromage	25.68	0	2	6	2
délicat	9.6	0	1	6	3	fruit	15.99	2	2	4	1
dingue	53.01	2	7	3	1	fusil	36.52	0	7	4	2
discret	9.52	0	1	6	2	gaz	36.33	7	25	3	1
disponible	8.74	0	0	9	3	goût	50.51	3	26	2	1
émotif	1.16	0	1	6	3	manteau	36.16	2	17	4	2
entier	42.15	2	4	4	2	menton	6.45	4	20	4	2
étroit	5.94	0	1	5	2	muguet	0.38	0	5	4	2
facile	153.57	0	3	5	2	nez	75.18	9	28	2	1
familier	7.77	0	1	7	3	océan	22.86	0	0	4	3
fidèle	20.04	0	2	5	2	orteil	4.04	0	0	5	2
fou	181.51	13	27	2	1	papier	56.32	4	8	5	2

Adjective	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**	**Noun**	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**
furieux	18.06	1	2	5	2	passeport	19.81	0	0	6	2
glissant	0.7	3	7	5	2	pied	105.51	5	17	3	1
gluant	1.23	1	1	4	2	piège	27.53	4	5	4	1
gourmand	1.81	0	3	5	2	poignet	6.38	2	7	5	2
gratuit	12.15	0	0	6	2	poivron	0.51	1	4	6	2
illégal	11.93	0	1	6	3	portrait	22.64	0	0	6	2
inquiet	17.9	0	1	4	2	pot	25.72	17	24	2	1
intact	6.21	0	1	5	2	poumon	4.55	1	5	4	2
inutile	64.05	0	0	6	3	progrès	18.52	0	1	6	2
ivre	16.59	0	4	3	1	pull	11.41	2	20	3	1
jaloux	29.87	0	2	4	2	saumon	5.28	1	11	4	2
malade	147.5	5	6	5	2	signal	33.98	1	5	5	2
malin	33.11	3	10	4	2	soda	9.12	7	14	4	2
moche	25.64	11	20	3	1	souffle	26.55	3	5	4	1
naïf	4.22	0	1	4	2	téléphone	155.68	2	0	7	3
nerveux	32.52	1	4	5	2	terrain	49.12	2	14	4	2
radin	2.69	8	12	4	2	thé	67.84	2	27	2	1
ringard	1.99	0	1	5	2	tiroir	12.18	1	2	6	2
sec	27.4	13	28	3	1	tissu	9.21	3	5	4	2
sensible	21.11	0	0	6	2	travail	367.43	1	2	6	2
sérieux	80.99	0	2	5	2	vaisseau	67.11	0	3	4	2
silencieux	8.3	0	0	7	3	vent	71.5	13	27	2	1
urgent	26.75	1	4	4	2	voleur	41.39	2	3	5	2

Table A.15. Lists in Experiment 2B (List 1), including English translations (in parentheses) and word class [in brackets].

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	wässriger$_{\text{[adj.]}}$ (watery)	Löwe$_{\text{[n.]}}$ (lion)	eckiger$_{\text{adj.]}}$ (angular)	Knoblauch$_{\text{[n.]}}$ (garlic)	Busch$_{\text{[n.]}}$ (bush)	nervöser$_{\text{[adj.]}}$ (nervous)
2	fröhlicher$_{\text{[adj.]}}$ (cheerful)	Hering$_{\text{[n.]}}$ (herring)	Wodka$_{\text{[n.]}}$ (vodka)	verrückter $_{\text{[adj.]}}$ (crazy)	lockerer$_{\text{[adj.]}}$ (loose)	Tisch$_{\text{[n.]}}$ (table)
3	Delfin$_{\text{n.]}}$ (dolphin)	witziger$_{\text{[adj.]}}$ (funny)	Zucker$_{\text{[n.]}}$ (sugar)	schwacher$_{\text{[adj.]}}$ (weak)	roter$_{\text{[adj.]}}$ (red)	Tequila$_{\text{[n.]}}$ (tequila)
4	Samt$_{\text{[n.]}}$ (velvet)	loyaler$_{\text{[adj.]}}$ (loyal)	tiefer$_{\text{adj.]}}$ (deep)	Elefant$_{\text{[n.]}}$ (elephant)	Rock$_{\text{[n.]}}$ (skirt)	aktiver$_{\text{[adj.]}}$ (active)
5	froher$_{\text{[adj.]}}$ (glad)	Hals$_{\text{[n.]}}$ (neck)	sonniger$_{\text{[adj.]}}$ (sunny)	Pinguin$_{\text{[n.]}}$ (penguin)	Motor$_{\text{[n.]}}$ (motor)	bewölkter$_{\text{[adj.]}}$ (cloudy)
6	glücklicher$_{\text{adj.]}}$ (happy)	Schuh$_{\text{[n.]}}$ (shoe)	Gepard$_{\text{[n.]}}$ (cheetah)	beschämter$_{\text{[adj.]}}$ (ashamed)	mieser$_{\text{[adj.]}}$ (dire)	Rücken$_{\text{[n.]}}$ (back)
7	Korridor$_{\text{n.]}}$ (corridor)	unschuldiger$_{\text{[adj.]}}$ (innocent)	Lachs$_{\text{[n.]}}$ (salmon)	frecher$_{\text{[adj.]}}$ (cheeky)	verkrampfter$_{\text{[adj.]}}$ (tense)	Schnaps$_{\text{[n.]}}$ (schnapps)
8	Karpfen$_{\text{[n.]}}$ (carp)	müder$_{\text{[adj.]}}$ (tired)	leichter$_{\text{[adj.]}}$ (light)	Gürtel$_{\text{[n.]}}$ (belt)	Essig$_{\text{[n.]}}$ (vinegar)	hohler$_{\text{[adj.]}}$ (hollow)
9	alberner$_{\text{[adj.]}}$ (silly)	Spiegel$_{\text{[n.]}}$ (mirror)	eifriger$_{\text{[adj.]}}$ (eager)	Falke$_{\text{[n.]}}$ (falcon)	Wein$_{\text{[n.]}}$ (wine)	brutaler$_{\text{[adj.]}}$ (brutal)
10	staubiger$_{\text{[adj.]}}$ (dusty)	Hai$_{\text{[n.]}}$ (shark)	Mantel$_{\text{[n.]}}$ (coat)	besiegter$_{\text{[adj.]}}$ (defeated)	heiterer$_{\text{[adj.]}}$ (cheerful)	Zeh$_{\text{[n.]}}$ (toe)
11	Adler$_{\text{[n.]}}$ (eagle)	schicker$_{\text{[adj.]}}$ (fancy)	Balkon$_{\text{[n.]}}$ (balcony)	gezielter$_{\text{[adj.]}}$ (targeted)	tapsiger$_{\text{[adj.]}}$ (lumbering)	Kopf$_{\text{[n.]}}$ (head)
12	Ozean$_{\text{[n.]}}$ (ocean)	nutzloser$_{\text{[adj.]}}$ (useless)	erregter$_{\text{[adj.]}}$ (aroused)	Tee$_{\text{[n.]}}$ (tea)	Fuß$_{\text{[n.]}}$ (foot)	einfacher$_{\text{[adj.]}}$ (simple)
13	gläubiger$_{\text{[adj.]}}$ (religious)	Atem$_{\text{[n.]}}$ (breath)	wütender$_{\text{[adj.]}}$ (angry)	Krug$_{\text{[n.]}}$ (jug)	Boden$_{\text{[n.]}}$ (floor)	dringender$_{\text{[adj.]}}$ (urgent)
14	intakter$_{\text{[adj.]}}$ (intact)	Wind$_{\text{[n.]}}$ (wind)	Pfeffer$_{\text{[n.]}}$ (pepper)	schuldiger$_{\text{[adj.]}}$ (guilty)	geiziger$_{\text{[adj.]}}$ (stingy)	Ausweis$_{\text{[n.]}}$ (passport)
15	Bus$_{\text{[n.]}}$ (bus)	unruhiger$_{\text{[adj.]}}$ (agitated)	Schädel$_{\text{[n.]}}$ (skull)	freier$_{\text{[adj.]}}$ (free)	kluger$_{\text{[adj.]}}$ (clever)	Pullover$_{\text{[n.]}}$ (pullover)
16	Geschmack$_{\text{[n.]}}$ (taste)	stummer$_{\text{[adj.]}}$ (mute)	gieriger$_{\text{[adj.]}}$ (greedy)	Belag$_{\text{[n.]}}$ (topping)	Käse$_{\text{[n.]}}$ (cheese)	diskreter$_{\text{[adj.]}}$ (discreet)

Table A.16. Lists in Experiment 2B (List 2), including English translations (in parentheses) and word class [in brackets].

List	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
1	Löwe[n.] (lion)	wässriger[adj.] (watery)	Knoblauch[n.] (garlic)	eckigeradj.] (angular)	nervöser[adj.] (nervous)	Busch[n.] (bush)
2	Hering[n.] (herring)	fröhlicher[adj.] (cheerful)	verrückter [adj.] (crazy)	Wodka[n.] (vodka)	Tisch[n.] (table)	lockerer[adj.] (loose)
3	witziger[adj.] (funny)	Delfinn.] (dolphin)	schwacher[adj.] (weak)	Zucker[n.] (sugar)	Tequila[n.] (tequila)	roter[adj.] (red)
4	loyaler[adj.] (loyal)	Samt[n.] (velvet)	Elefant[n.] (elephant)	tieferadj.] (deep)	aktiver[adj.] (active)	Rock[n.] (skirt)
5	Hals[n.] (neck)	froher[adj.] (glad)	Pinguin[n.] (penguin)	sonniger[adj.] (sunny)	bewölkter[adj.] (cloudy)	Motor[n.] (motor)
6	Schuh[n.] (shoe)	glücklicheradj.] (happy)	beschämter[adj.] (ashamed)	Gepard[n.] (cheetah)	Rücken[n.] (back)	mieser[adj.] (dire)
7	unschuldiger[adj.] (innocent)	Korridorn.] (corridor)	frecher[adj.] (cheeky)	Lachs[n.] (salmon)	Schnaps[n.] (schnapps)	verkrampfter[adj.] (tense)
8	müder[adj.] (tired)	Karpfen[n.] (carp)	Gürtel[n.] (belt)	leichter[adj.] (light)	hohler[adj.] (hollow)	Essig[n.] (vinegar)
9	Spiegel[n.] (mirror)	alberner[adj.] (silly)	Falke[n.] (falcon)	eifriger[adj.] (eager)	brutaler[adj.] (brutal)	Wein[n.] (wine)
10	Hai[n.] (shark)	staubiger[adj.] (dusty)	besiegter[adj.] (defeated)	Mantel[n.] (coat)	Zeh[n.] (toe)	heiterer[adj.] (cheerful)
11	schicker[adj.] (fancy)	Adler[n.] (eagle)	gezielter[adj.] (targeted)	Balkon[n.] (balcony)	Kopf[n.] (head)	tapsiger[adj.] (lumbering)
12	nutzloser[adj.] (useless)	Ozean[n.] (ocean)	Tee[n.] (tea)	erregter[adj.] (aroused)	einfacher[adj.] (simple)	Fuß[n.] (foot)
13	Atem[n.] (breath)	gläubiger[adj.] (religious)	Krug[n.] (jug)	wütender[adj.] (angry)	dringender[adj.] (urgent)	Boden[n.] (floor)
14	Wind[n.] (wind)	intakter[adj.] (intact)	schuldiger[adj.] (guilty)	Pfeffer[n.] (pepper)	Ausweis[n.] (passport)	geiziger[adj.] (stingy)
15	unruhiger[adj.] (agitated)	Bus[n.] (bus)	freier[adj.] (free)	Schädel[n.] (skull)	Pullover[n.] (pullover)	kluger[adj.] (clever)
16	stummer[adj.] (mute)	Geschmack[n.] (taste)	Belag[n.] (topping)	gieriger[adj.] (greedy)	diskreter[adj.] (discreet)	Käse[n.] (cheese)

Table A.17. Characteristic description of word stimuli in Experiment 2B according to a database: https://clearpond.northwestern.edu. Freq means the frequency per million; NSorth means orthographic neighborhood size; NSphon means phonological neighborhood size; Nphon means number of phonemes; Nsyll means number of syllables.

Adjective	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**	**Noun**	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**
aktiver	2	2	1	6	3	Adler	6.42	3	3	4	2
alberner	1.10	3	2	6	3	Atem	15.55	3	5	4	2
beschämter	NA	NA	NA	NA	3	Ausweis	19.45	1	0	5	2
besiegter	NA	NA	NA	NA	3	Balkon	6.50	3	3	6	2
bewölkter	NA	NA	NA	NA	3	Belag	NA	NA	NA	NA	2
brutaler	2.01	3	2	7	3	Boden	98.78	7	10	5	2
diskreter	NA	NA	NA	NA	3	Bus	47.80	9	18	3	1
dringender	8.00	3	1	8	3	Busch	6.54	4	7	3	1
eckiger	NA	NA	NA	NA	3	Delfin	1.42	1	2	6	2
eifriger	NA	NA	NA	NA	3	Elefant	6.77	1	1	7	3
einfacher	31.06	4	2	6	3	Essig	1.65	1	2	4	2
erregter	NA	NA	NA	NA	3	Falke	2.17	3	5	5	2
frecher	3.00	3	3	5	2	Fuß	62.29	1	4	3	1
freier	7.00	7	6	4	2	Gepard	NA	NA	NA	NA	2
froher	4.00	4	4	4	2	Geschmack	22.95	0	0	6	2
fröhlicher	3.62	3	2	7	3	Gürtel	8.50	0	0	6	2
geiziger	NA	NA	NA	NA	3	Hai	7.56	16	23	2	1
gezielter	NA	NA	NA	NA	3	Hals	65.16	9	12	4	1
gieriger	0.79	1	1	6	3	Hering	1.18	2	0	5	2
gläubiger	2.01	2	1	7	3	Karpfen	1.61	0	0	7	2
glücklicher	8.00	3	2	8	3	Käse	19.61	1	1	4	2
heiterer	NA	NA	NA	NA	3	Knoblauch	3.50	0	0	7	2
hohler	NA	NA	NA	NA	2	Kopf	286.70	3	3	4	1
intakter	NA	NA	NA	NA	3	Korridor	6.14	0	0	7	3
kluger	6.00	6	4	5	2	Krug	6.42	2	5	4	1

Adjective	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**	**Noun**	**Freq**	**NS orth**	**NS phon**	**N phon**	**N syll**
leichter	24.45	4	4	5	2	Lachs	3.23	6	18	4	1
lockerer	2.60	2	2	6	3	Löwe	5.12	3	4	4	2
loyaler	0.00	0	1	6	3	Mantel	22.21	3	2	6	2
mieser	6.00	6	7	4	2	Motor	35.51	3	2	5	2
müder	NA	NA	NA	NA	2	Ozean	8.70	1	1	5	3
nervöser	NA	NA	NA	NA	3	Pfeffer	2.76	1	1	4	2
nutzloser	1.97	3	1	7	3	Pinguin	2.17	1	1	6	3
roter	9.00	9	9	4	2	Pullover	5.35	0	1	6	3
schicker	1.38	3	6	4	2	Rock	28.19	7	16	3	1
schuldiger	NA	NA	NA	NA	3	Rücken	65.40	8	8	5	2
schwacher	2.28	4	4	5	2	Samt	4.96	11	15	4	1
sonniger	NA	NA	NA	NA	3	Schädel	22.60	1	2	5	2
staubiger	NA	NA	NA	NA	3	Schnaps	7.91	1	4	5	1
stummer	NA	NA	NA	NA	2	Schuh	13.27	3	18	2	1
tapsiger	NA	NA	NA	NA	3	Spiegel	24.76	3	3	6	2
tiefer	5.00	5	5	4	2	Tee	87.17	9	40	2	1
unruhiger	NA	NA	NA	NA	4	Tequila	5.95	0	0	7	3
unschuldiger	9.00	3	1	9	4	Tisch	96.62	5	12	3	1
verkrampfter	NA	NA	NA	NA	3	Wein	45.36	19	31	3	1
verrückter	8.00	3	2	8	3	Wind	47.13	10	17	4	1
wässriger	NA	NA	NA	NA	3	Wodka	8.46	0	0	5	2
witziger	4.25	3	2	6	3	Zeh	6.22	8	38	2	1
wütender	2.36	2	1	7	3	Zucker	21.58	1	1	4	2

Study 2

Effect of Spatial Context on Serial Order Recall in Verbal Working Memory

Hong Xiao, Lucie Attout, and Steve Majerus

In preparation

Abstract. Contemporary theories of verbal working memory (WM) propose that position markers are used for coding serial order information. One type of marker involves spatial coding, which suggests that serial order is encoded along a mental spatial pattern, such as a mental line. However, while empirical evidence for co-activation of mental spatial codes in serial order WM has been obtained, the direct impact of the use of these codes on serial order recall accuracy remains uncertain. The present study manipulated the setup of canonical vs. non-canonical spatial codes of the information during WM encoding/recall and their impact on serial order WM performance. In Experiments 1A, and 1B, we manipulated the spatial codes during the encoding phase by using an immediate serial recall task in which digits were presented sequentially from left to right (canonical spatial pattern) or from right to left (non-canonical spatial pattern). Experiment 2 manipulated the spatial codes during the recall phase by presenting items centrally during the encoding and contrasting the need to recall the items in a canonical vs. non-canonical spatial direction. Finally, Experiments 3A and 3B maximally raised participants' awareness about the spatial processing dimensions by manipulating both presentation and recall spatial directions. This manipulation led to an advantage of recall performance for memory lists presented in a canonical, left-to-right format. These findings indicate that the impact of spatial coding on serial verbal WM is highly context- and task-specific, suggesting that spatial coding may be an ancillary rather than a fundamental dimension of serial order coding.

Introduction

Maintaining the serial order of verbal information is a core function of working memory (WM), with critical implications for language processing, reasoning, and learning (Baddeley & Hitch, 2019; Majerus et al., 2008). Despite extensive research, the mechanisms underlying the encoding and maintenance of serial order information remain a subject of ongoing debate (Majerus, 2019). Broadly speaking, current theoretical models focusing on how serial order is coded in WM can be divided into two classes: associative chaining and position marker models.

Associative chaining (Ebbinghaus, 1964) was one of the earliest theoretical approaches, positing that serial order is maintained via learned associations between successive items, with each item serving as a retrieval cue for the next (e.g., Lashley, 1951). However, chaining models faces several empirical limitations. For instance, they predict that recall should break down following an error in the sequence, which is often not the case (Henson, 1996). Moreover, they struggle to explain distance effects in WM, where order judgments are more difficult for items that are close in the sequence than for those that are farther apart (Attout et al., 2014; Marshuetz et al., 2000b). As a result, chaining accounts have largely been abandoned in favor of position marker models (Burgess & Hitch, 2006; Farrell & Lewandowsky, 2002; Henson, 1996). However, recently, a model integrating both chaining and position coding processes has been proposed, highlighting the possible coexistence of these two processes to maintain information (Logan & Cox, 2021).

More specifically, position marker models propose that serial order is represented by the association between items and some independent and varying contextual representations of their position within a sequence of items. These models differ in whether position markers are implemented using temporal, absolute, or relative codes (Henson, 1999). The oscillator-based associative recall (OSCAR) model (Brown et al., 2000) is one of the models relying on a temporal coding scheme. It is supposed that serial order is linked with the different states of a time-varying context signal driven by sets of temporal oscillators operating at various frequencies. Absolute coding models, such as the original model proposed by Burgess and Hitch (1992), consider that serial order is coded via associations between items and a gradually evolving temporal context signal, allowing sequential recall through context reinstatement. Relative coding models, for example, the Start-End Model (SEM) by Henson (1998), assume that items are coded relative to both start and end markers of a sequence. These computational context-based models can account for many of the behavioral phenomena that characterize serial order memory, such as the temporal grouping effect or the locality constraint on transpositions (items are more likely to move to the position adjacent to their target position) (Henson, 1998; Lewandowsky & Farrell, 2008; Oberauer et al., 2018). Moreover, neuroimaging studies support these item-position association models by showing that serial order memory recruits specific fronto-parietal cortices involving the intraparietal sulcus as compared to item memory, which recruits to a larger extent fronto-temporal cortices (Henson et al., 2000; Majerus et al., 2006, 2010; Marshuetz et al., 2000, 2006). Kalm and Norris (2014) showed that multivariate brain patterns in these brain areas during encoding and recall of verbal WM sequences were best predicted by a positional model of neural similarity for items and their associations within a sequence. However, these positional models differ in fundamental ways when it comes to defining the nature of the codes of serial order information in a more concrete manner (Majerus & Attout, 2018).

One specific variant of positional models assumes these markers could be spatial by nature, proposing the existence of spatial codes, in the form of an internal spatial map such as a mental line, with memory items getting associated with different portions of the mental space/lines, from left to right as a function of the temporal succession of the items within the list (Abrahamse et al., 2014; Ginsburg et al., 2014, 2017; Ginsburg & Gevers, 2015; van Dijck & Fias, 2011). In a seminal study, van Dijck and Fias (2011) employed a dual-task paradigm in which participants judged digit parity for items maintained in WM. They found that parity judgements for items early in the memory list facilitated left-hand responses, whereas judgments for later items facilitated right-hand responses—a pattern suggesting the use of an internal left-to-right spatial template. van Dijck et al. (2013) further showed that reactivating start-of-list WM items induced a leftward attentional spatial bias in a concurrent dot detection task, while end-of-list items induced a rightward spatial bias. Moreover, WM recall performance can interact with spatial cues presented during retrieval, a spatial cue on the left of the screen leading to faster retrieval of the beginning WM items and a spatial cue on the right of the screen leading to faster retrieval of the last WM items (De Belder et al., 2015). Similarly, Guida et al. (2016) adopted a recognition task rather than the dual task with auditorily presented consonants and found that faster left-hand "yes" responses were made for items from the beginning of a memorized sequence, and right-hand "yes" responses for items from the end. These findings have been conceptualized within the *mental whiteboard hypothesis*, which proposes that sequential items in verbal WM are associated with spatial position markers along an internalized spatial axis (Abrahamse et al., 2014; 2017). These spatialization effects[1] in verbal WM are also referred to as the *ordinal position effect* (Ginsburg et al., 2014) or the *Spatial–Positional Association of Response Codes* (SPoARC; Guida et al., 2016; Guida & Lavielle-Guida, 2014). In Western participants, these spatialization effects typically involve a left-to-right dimension, but cultural reading habits modulate this spatial mapping: for example, reversed spatial mappings have been observed in right-to-left readers (Guida, Megreya, et al., 2018; Park et al., 2024; Shaki & Fischer, 2012). At the same time, top-down vertical spatialization effects have also been shown in verbal WM tasks (Hartmann et al., 2021). These behavioral effects are also supported by neuroimaging and eye-tracking studies. At the neural level, Cristoforetti et al. (2022) adapted the original paradigm of van Dijck and Fias (2011) to examine the multivariate neural codes associated with serial order coding (see also Attout et al., 2014; Attout et al., 2022; Rasoulzadeh et al., 2021; Tian & Fischer-Baum, 2025). They observed that the retrieval of items from early vs. late serial positions was associated with specific neural patterns in the intraparietal sulcus and the hippocampus, two regions that have been associated with spatial representations and navigation. By studying oculomotor movements, Sahan et al. (2022) found that retrieving verbal items at a specific position in a sequence in WM was accompanied by horizontal eye movements, which the gaze diverted more to the left side of space when searching items from initial parts of the memorized sequence and

---

[1] The use of the term "spatialization effect" in the present study will specifically refers to the research that has shown spatial-position association effects (also called SPoARC or ordinal position effect) or spatial biases in the context of verbal WM tasks based on judgment tasks (i.e., the paradigms proposed by van Dijck & Fias (2011) as well as Guida et al. (2016)), in order to distinguish this effect from the more direct impact of spatial coding on actual verbal WM performance as investigated in the present study.

more to the right side for later parts, supporting the idea that serial order information in verbal WM is spatially coded and governed by spatial attention. A similar effect was also reported in a recent recall study, where gaze shifts in visual space correlated with spatial shifts of attention along a left-to-right mapping of serial positions (Schroth et al., 2025).

A further study by Guida, Abrahamse, et al. (2020), however, indicates that left-to-right spatial codes may not be a default coding format for serial order (see also Guida, Mosinski, et al., 2020). In their study, they manipulated the spatial presentation format (left-to-right, central, right-to-left) of consonant sequences in a recognition task. Participants were asked to indicate whether the probe was in the sequence or not by pushing either the "yes" or "no" button with their left or right hand. They observed the expected faster left-hand reaction on beginning items and faster right-hand reaction on later items in central and left-to-right presentation (encoding) formats, but this effect reversed in the right-to-left presentation condition. These findings led to an assumption of the *principle of economy*, whereby spatial codes are flexibly recruited to provide a functional advantage in serial recall.

At the same time, although spatialization effects in verbal WM appear to be relatively robust (Guida et al., 2026), very few studies have directly examined the extent to which spatial coding provides real support to the encoding and maintenance of serial order information in verbal WM, or whether these effects are merely mental by-products with no direct impact on recall performance. A few studies have manipulated spatial variables during the encoding of verbal information in WM, although it should be noted that the aim of most of these studies was not to directly test the hypothesis of spatial coding of serial order information in WM.

For example, Hitch and Morton (1975) contrasted lists presented from left to right versus in identical positions on the left of the display without observing any difference in recall performance. Battacchi et al. (1990) found a stronger recency effect when sequences were presented in spatially distributed formats (e.g., left-to-right or up-to-down) compared to a central, non-distributed format. However, LeCompte (1992) attempted to replicate this advantage but failed to observe any effect. Similarly, Li and Lewandowsky (1995) found no recall benefit when comparing centrally presented items versus items presented in random spatial positions. More recently, studies directly comparing spatial coding with alternative coding mechanisms have reported little to no evidence for spatial coding. For instance, Ordonez et al. (2022) found better recall performance under simultaneous than sequential presentation, suggesting the use of distinct encoding strategies. The sequential presentation may have induced a non-spatial coding strategy, the encoding presentation not inducing a spatialization, leading to lower performance. Similarly, using a dual-task paradigm, Attout et al. (2025) did not observe the expected facilitation effect of external spatial cues. Beyond the inconsistency of the findings, a key problem of these studies is that they compare situations that provide an encoding structure versus no structure, and hence, it is difficult to draw any conclusions regarding the specific impact of spatial variables or their absence on verbal WM recall performance. Two studies contrasted two spatially organized encoding conditions by comparing recall performance for left-to-right vs right-to-left presentation of the memoranda during encoding. Fischer-Baum and Benjamin (2014) reported an advantage for the canonical, left-to-right condition, but the advantage was only restricted to the final list positions. By testing more specific questions regarding an advantage of memory span for Arabic numerals versus digit words, Chincotta et al. (1999) showed that this advantage disappeared when the items were presented in a right-to-left encoding direction, suggesting that at least the encoding of sequences of Arabic numerals relies on spatial codes.

In sum, few studies have examined the impact of spatial variables on verbal WM recall performance, and even fewer studies have used spatial manipulations to directly examine whether spatial codes support coding and recall performance of serial order information in verbal WM. The present study provides a comprehensive assessment of spatial coding manipulations on serial order recall performance in verbal WM tasks. Experiment 1 aimed at replicating the results observed by Fischer-Baum and Benjamin (2014), by having participants encode verbal memoranda either in a (for Westerners) canonical left-right or in a non-canonical right-left encoding direction. We furthermore used longer lists (seven items) to increase task sensitivity and thereby maximize the intervention of spatial codes, if these are the essential codes used for representing serial order information. In Experiment 2, items were presented centrally, which—based on the evidence from the spatialization effect literature (Guida et al., 2026)—should lead to the setup of a left-to-right mental spatial representation of the memory sequence. If this spatial structure is essential for serial order maintenance and recall, then participants should have difficulty if they are instructed to recall the items in a direction opposite to this “spontaneous” spatial structure (i.e., from right to left). Experiment 3 provided a particularly stringent test of the intervention of spatial dimensions during encoding and recall by manipulating both spatial encoding direction and the intervention of spatial cues during recall. In order to examine the impact of spatial manipulations on serial order recall performance in the most direct manner in all experiments, we examined verbal WM performance for standard serial recall performance (items recalled in correct serial position), but also item recall performance (items recalled irrespective of serial position) and, most critically, order recall performance (items recalled in correct serial conditioned by overall item recall performance). If spatial codes specifically support serial order memory, then the strongest impact of the spatial manipulations should be observed for order recall performance.

Experiment 1

Experiment 1 was a conceptual replication of the studies by Chincotta et al. (1999) and Fischer-Baum and Benjamin (2014), by using an immediate serial recall task and manipulating the encoding direction. As already noted, we used longer list lengths (seven digits) as in the study by Fischer-Baum and Benjamin (2014) (six consonants), but used the same contrast of canonical vs. non-canonical presentation of items during encoding. In Experiment 1A, a standard presentation rate of one digit per second was used. To address the possibility that the setup of spatial codes is a time-demanding process, we increased the presentation rate to 1500ms per item in Experiment 1B. This experiment was part of a larger study that also involved manipulating temporal parameters during encoding via a separate task. Participants were subjected to both spatial and temporal manipulations in separate task blocks; the results of the temporal manipulations are reported in a dedicated manuscript.

Experiment 1A

Method

Participants. Fifty-seven participants (36 females, 21 males; *mean age* = 21.21 years, *SD* = 2.29) were recruited via social media from the University of Liège community (see *Scoring and Analysis Procedure* for sample size justification). All were native French speakers with no reported history of language, learning, neuropsychological, or neurological disorders. Participants also reported no current drug use (e.g., cannabis) or alcohol abuse. The study was approved by the Ethics Committee of the Faculty of Psychology, Language, and Education Sciences at the University of Liège. Participants provided informed consent before testing and were informed that no financial compensation was offered.

Materials. Thirty-four seven-digit lists were created by pseudo-randomly sampling from the digits 1 to 7, including four practice lists. Lists could not include adjacent digit pairs (e.g., 1-2 or 7-6) or other structured digit combinations (e.g., 1-3-5 or 6-4-2). Additionally, we balanced the distribution of the seven digits across lists and conditions. Half of the lists were presented for the left-to-right condition, and the other half for the right-to-left condition. The different lists were presented in random order.

Procedures. The task was programmed using the Gorilla platform (https://gorilla.sc). Participants were presented with sequences of seven non-repeating digits (ranging from 1 to 7) and were required to recall the digits in their original order of presentation. Stimuli were displayed sequentially along a horizontal array comprising seven equally spaced positions, extending symmetrically from the center of the screen and covering approximately half of its width. Each trial began with the presentation of a central fixation cross for 500ms, followed by a 500-ms blank interval. The seven digits were then displayed individually for 500ms each, with an interstimulus interval of 500ms. The key manipulation involved the direction of spatial presentation, which varied randomly across trials: digits were presented either from left to right or from right to left (See Figure 2.1).

Following a 500-ms blank screen after the final digit, a recall prompt (“rappel”, which means “recall” in English) appeared at the center of the screen. Participants were instructed to recall the digit sequence aloud in the original order of presentation. If they were unable to retrieve a digit at a given position, they were asked to say “oublié” (“forgotten”) to maintain positional integrity in the recall protocol. Participants initiated the next trial by pressing the spacebar once they had completed their response.

Before the experimental phase, participants completed a short series of practice trials with feedback to ensure accurate comprehension of the task instructions. The experimental phase consisted of 30 trials, equally divided between the two presentation direction conditions (15 trials per condition), presented in a randomized order. There was no time constraint for recall. Participants were tested individually in silent testing rooms. All verbal responses were recorded via the Gorilla website and subsequently transcribed for scoring.

The full experiment lasted approximately 20 minutes.

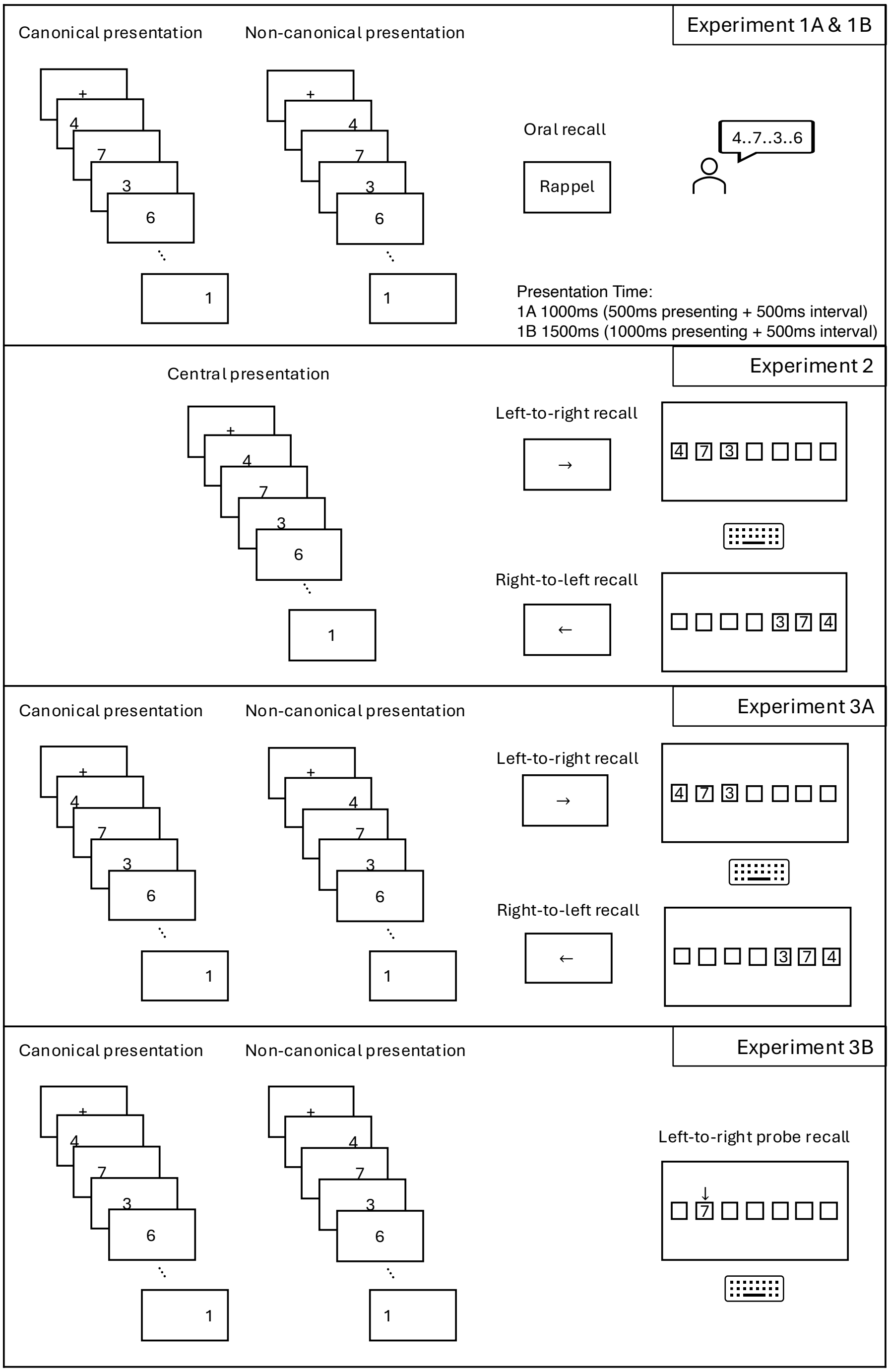

Figure 2.1. Illustration of the different experimental paradigms. In Experiment 1A and 1B, digits were presented either from left to right or from right to left, and participants were asked to orally recall the list in

presentation order. In Experiment 2, digits were presented at the center of the screen, and participants were asked to type the digits in the corresponding boxes, with the temporal recall direction always being direct forward recall, but by starting either with the most leftward or the most rightward box for typing the first digit of the list. In Experiment 3A, digits were presented either from left to right or from right to left, and participants were asked to type the digits in the corresponding boxes, with temporal recall direction being always direct forward recall but by starting either with most leftward or most rightward box for typing the first digit of the list (combination of the paradigms used in Experiment 1B and Experiment 2). In Experiment 3B, digits were presented either from left to right or from right to left, and participants were asked to recall the list in forward order by entering the digit corresponding to the serial position of the box being cued.

Scoring and Analysis Procedure. Two standard recall accuracy measures were computed: A **serial recall score**, defined as the number of items correctly recalled in their exact serial positions, and an **item recall score**, defined as the number of items correctly recalled regardless of position. Additionally, an **order recall score** was calculated as the ratio of the serial recall score to the item recall score, providing a measure of how well participants retained the order of items relative to their overall item recall performance.

To further quantify serial recall performance, **transposition error** was computed, measuring the distance between the input and output positions of a recalled item. Depending on whether an item was recalled before or after its correct position, a negative or positive value was recorded. For instance, if the sequence “4 6 1 3” were recalled as “4 3 6 7,” the transposition gradients would be “0, -2, 1, NA” indicating the respective positional deviations relative to the input position of the recalled items.

All analyses were conducted using a Bayesian statistical approach (see, e.g., Dienes, 2011; Morey & Rouder, 2011). This approach has the advantage of relying on a model comparison rationale to select and quantify the strength of evidence associated with each model, and crucially, allows for testing the strength of evidence for and against an effect of interest (i.e., positive evidence for the null hypothesis). The Bayesian framework does not involve traditional *p-values*, thereby avoiding multiple testing problems such as alpha inflation (Wagenmakers et al., 2008). All analyses are based on *Bayes factors (BF)*, which can be considered as a relative measure of statistical evidence (Morey et al., 2016). The BF represents the degree to which the observed data update the initial belief in favor of one hypothesis relative to another. The *BF* is the likelihood ratio of a given model, the best-fitting model being the one with the highest *BF*. BF_{01} indicates evidence in favor of the null hypothesis, while BF_{10} indicates evidence in favor of the alternative hypothesis. Although there are no fixed thresholds for *BF* values, we used the following categories for describing strength of evidence: a *BF* of at least 1 is considered to indicate anecdotal evidence, a *BF* of at least 3 is considered to indicate moderate evidence, a *BF* of at least 10 is considered to provide strong evidence, a *BF* of at least 30 is considered to provide very strong evidence, and a *BF* of at least 100 is considered to indicate decisive evidence (Jeffreys & Jeffreys, 1998).

Bayesian generalized linear mixed models were implemented using the R statistical platform and the brms package (Bürkner, 2017). The dependent measures included the recall accuracy scores and error types described above. Fixed effects comprised presentation direction (canonical vs. non-canonical), serial position (1 to 7), and their interaction. For the analysis of the transposition error, serial position was replaced by transposition distance (-6 to +6). The model included random intercepts for subjects to account for individual variability. To determine the inclusion of random slopes, we first compared two models: one with only the random intercept (subject) and one additionally including a random slope for the presentation direction factor. The random slope was retained in subsequent models if the *Bayes Factor (BF)* exceeded 3, indicating substantial evidence for its inclusion compared to the null model containing only the intercept. Random slopes were not considered for the serial position factor as they could not be estimated in a reliable manner due to the imbalance of variability of performance associated with the different serial positions[2]. *Estimated*

---

[2] The data were re-analyzed using efficient leave-one-out cross-validation for Bayesian models, an alternative approach based on model prediction (LOO; Vehtari, Gelman, & Gabry, 2017) closer to non-parametric methods. The outcome of results in terms of optimal models to be retained remained the same as reported in this manuscript.

marginal means (EMMs) with *credible intervals* were reported when exploring the effects included in the winning model via the emmeans() function.

Sample size was monitored via a Bayesian sequential sampling approach (Schönbrodt & Wagenmakers, 2018) combined with an effect size stabilization procedure (R. B. Anderson et al., 2022). Beginning from the 20th participant, we computed, for each additional participant, the effect sizes for the effect of interest (presentation direction - signed *Cohen's d*) and continued sampling at least until the effect size, for both item and serial order recall scores, stabilized (a priori defined minimal absolute change of effect size <.05 over 5 consecutive analyses; Anderson et al., 2022). After the 25th participant, the effect sizes of both item and serial order recall scores were stable. However, data collection was continued to maximize the statistical sensitivity of our study, given that additional participants were available. Kowialiewski (2024) recently showed that the effect size stabilization procedure does not guarantee the detection of a true effect in the population if the stabilized effect size is small.

To facilitate data visualization and interpretation, recall performance and error rates were analyzed using proportional measures. **Item recall** and **serial recall proportions** were defined as the number of correctly recalled items, regardless of order (item recall) or in correct serial position (serial recall), divided by the total number of presented items. For the **transposition gradients**, the proportions were weighted by the number of transposition errors divided by the sum of item recall scores.

Results

Serial Recall Score. A first Bayesian generalized linear mixed model (GLMM) analysis of serial recall scores revealed that the best-fitting model included only the serial position factor. Although this model was 1.35 times more likely than the model including both presentation direction and serial position factors, it was retained as the more parsimonious model (Presentation direction: $\eta_p^2 = 0.010$; Serial position: $\eta_p^2 = 0.618$; Presentation direction × Serial position: $\eta_p^2 = 0.033$). These results provide evidence that tends *to favor* the absence of an effect of presentation direction ($BF_{01} = 1.28$), and evidence strongly *favors* the absence of its interaction with the serial position ($BF_{01} = 15.09$) (see Appendix Tables B.1, 2, 7, and Figure 2.2).

Item Recall Score. For the item recall score, the best-fitting model included serial position only. This model was 4.71 times more likely than the next highest model, which included serial position and presentation direction factors (Presentation direction: η_p^2 = 0.007; Serial position: η_p^2 = 0.479; Presentation direction × Serial position: η_p^2 = 0.026). These results provide evidence *against* the effect of the presentation direction (BF_{01} = 5.06), and evidence tending towards the absence of an interaction with serial position (BF_{01} = 1.85) (see Appendix Tables B.1, 2, 8, and Figure 2.3).

Order Recall Score. For the order recall score, given its derivational nature based on average item and serial recall scores, a repeated measures Bayesian analysis of variance (ANOVA) was conducted on subject-averaged scores instead of a GLMM. The best-fitting model included only the serial position factor. This model was 7.17 times more likely than the following model, which included both presentation direction and serial position factors (Presentation direction: $\eta_p^2 = 6.755 \times 10^{-4}$; Serial position: η_p^2 = 0.454; Presentation direction × Serial position: η_p^2 = 0.010). These results provide evidence *against* an effect of presentation direction (BF_{01} = 6.04) or its interaction with the serial position (BF_{01} = 27.35) (see Appendix Tables B.1, 2, 9, and Figure 2.4).

Transposition Gradient. A repeated measures Bayesian ANOVA analysis on the transposition gradients revealed that the best-fitting model included only the transposition distance factor. This model was 206.86 times more likely than the following model, including both transposition distance and presentation direction factors (Presentation direction: $\eta_p^2 = -2.396 \times 10^{-4}$; Transposition distance: η_p^2 = 0.980; Presentation direction × Transposition distance: η_p^2 = 0.007). These results indicated evidence *against* an effect of the presentation direction (BF_{01} = 33.55) or its interaction with transposition distance ($BF_{01} = 2.75 \times 10^{20}$) (see Appendix Tables B.1, 2, 11, and Figure 2.5).

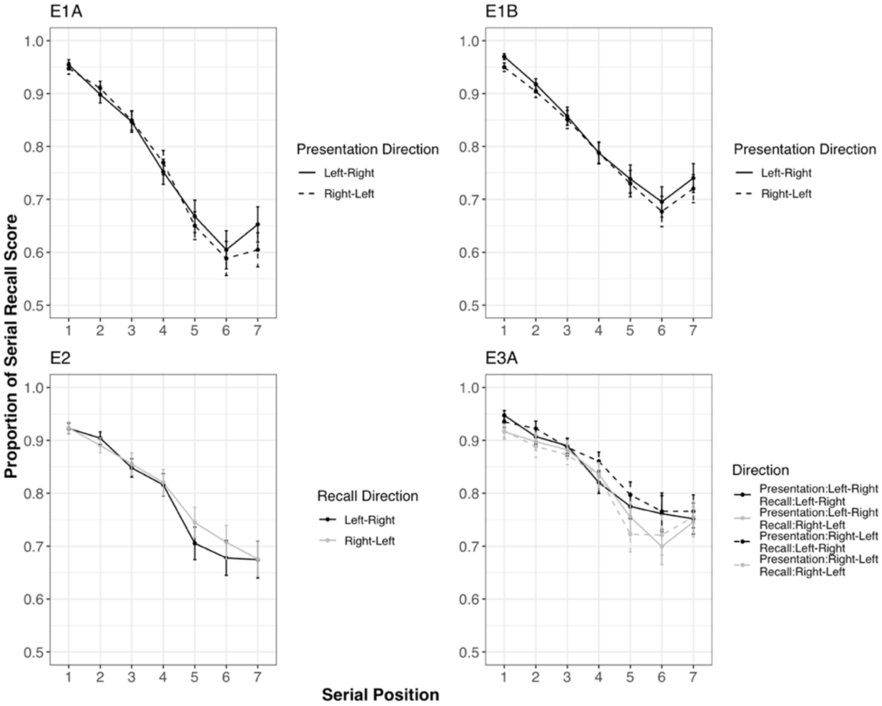

Figure 2.2. Serial order recall performance, as a function of serial position and presentation direction (Experiment 1A and 1B), recall direction (Experiment 2), or both presentation direction and recall direction (Experiment 3A). Means with standard errors are depicted.

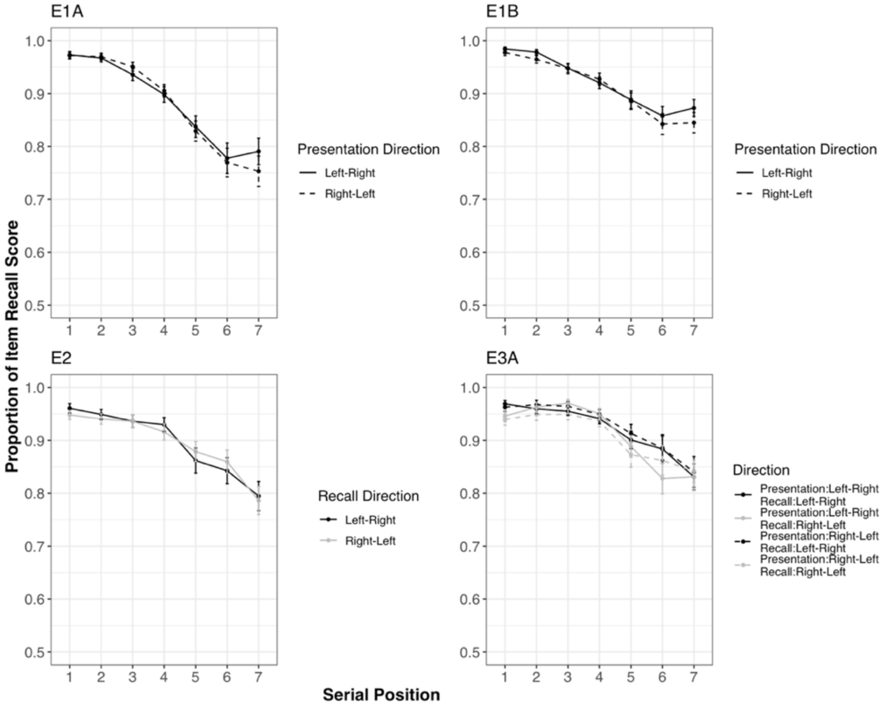

Figure 2.3. Item recall performance, as a function of serial position and presentation direction (Experiment 1A and 1B), recall direction (Experiment 2), or both presentation direction and recall direction (Experiment 3A). Means with standard errors are depicted.

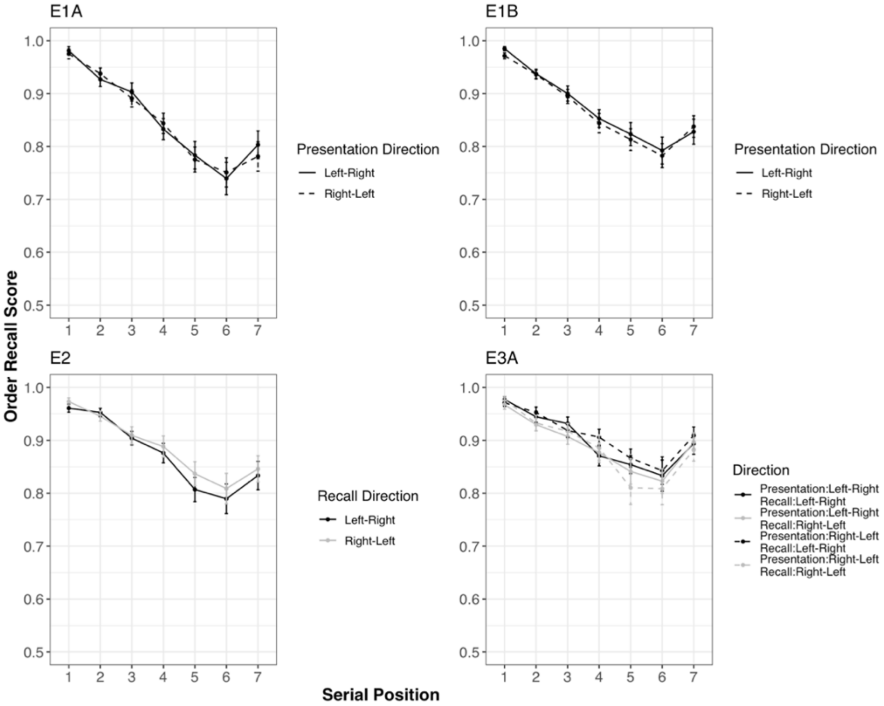

Figure 2.4. Order recall performance, as a function of serial position and presentation direction (Experiment 1A and 1B), recall direction (Experiment 2), or both presentation direction and recall direction (Experiment 3A). Means with standard errors are depicted.

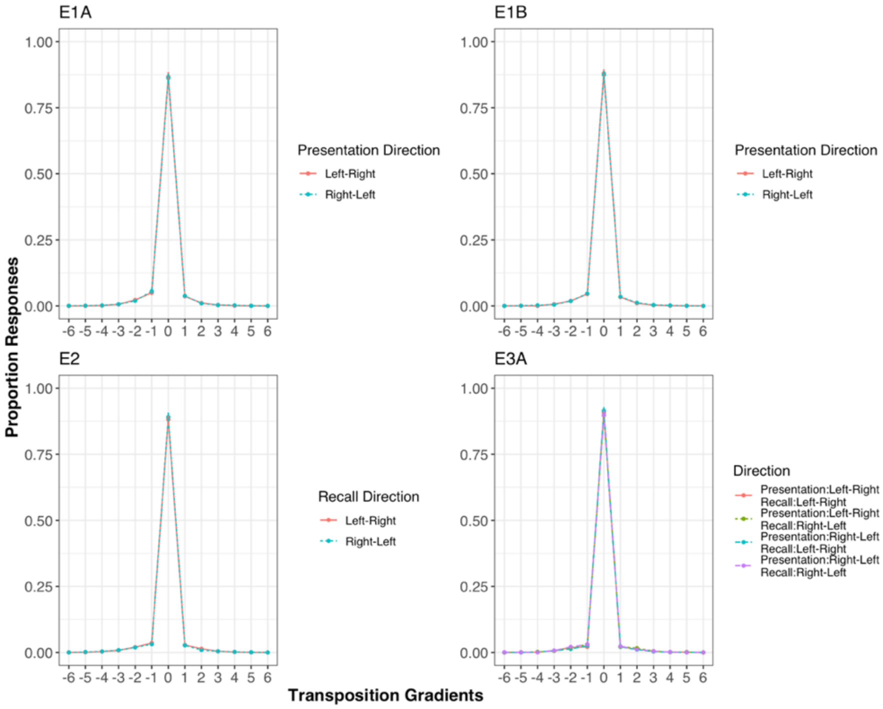

Figure 2.5. Transposition gradients, reporting the proportion of order errors as a function of transposition distance and presentation direction (Experiment 1A and 1B), recall direction (Experiment 2), or both presentation direction and recall direction (Experiment 3A). Means with standard errors are depicted.

Discussion

In the first experiment, we manipulated the spatial order of encoding of items in an immediate serial recall task by presenting lists of digits in either a canonical or non-canonical direction. Overall, no main effect of presentation direction or an interaction with serial position was observed. The rapid presentation rate may have prevented the buildup of a robust spatial mental representation of the items, leading to the highly limited spatial manipulation effect observed in Experiment 1A. We therefore conducted Experiment 1B using a more extended presentation duration of 1000ms per item during encoding. This duration is also closer to the encoding duration used in some of the studies examining spatialization effects in WM (Shi et al., 2020; Wang et al., 2018; Guida, Abrahamse et al., 2020).

Experiment 1B

Method

Participants. Seventy-seven participants were recruited (see Scoring and Analysis Procedure for justification of sample size) via social networks. Participants were between 18 and 28 years old (*average age* = 21.87, *SD* = 2.04); 39 participants were female.

Materials. The materials were as same as in Experiment 1A.

Procedures. The design and procedures were as same as in Experiment 1A, except for the presentation time extended to 1000ms per item, while the inter-stimulus interval remained constant at 500ms (See Figure 2.1).

Scoring and Analysis Procedure. Data scores and analyses were the same as in Experiment 1A. After the 25th participant, the effect sizes of both item and serial order recall scores were stable. As in Experiment 1A, data collection was continued to maximize the statistical sensitivity of our study as additional participants were available.

Results

Serial Recall Score. A Bayesian GLMM analysis of the serial recall score revealed that the best-fitting model included only the serial position factor. Although this model was 1.16 times more likely than the model including both presentation direction and serial position factors, it was retained as the more parsimonious model (Presentation direction: $\eta_p^2 = 0.029$; Serial position: $\eta_p^2 = 0.482$; Presentation direction × Serial position: $\eta_p^2 = 0.005$). Evidence tends *to favor the absence* of an effect of presentation direction ($BF_{01} = 1.17$) and evidence favor the absence of an interaction with serial position ($BF_{01} = 18.02$) (see Appendix Tables B.1, 3, 7, and Figure 2.2).

Item Recall Score. For the item recall score, the best-fitting model included serial position. Although the model including serial position and presentation direction factors was 1.54 times more likely than the model excluding the presentation direction factor, the latter was retained as the more parsimonious model (Presentation direction: $\eta_p^2 = 0.041$; Serial position: $\eta_p^2 = 0.340$; Presentation direction × Serial position: $\eta_p^2 = 0.016$). Evidence *does not support* an effect of presentation direction ($BF_{01} = 0.77$) or its interaction with serial position (BF_{01} = 1.18) (see Appendix Tables B.1, 3, 8, and Figure 2.3).

Order Recall Score. For the order recall score, a repeated measures Bayesian ANOVA analysis on subject-averaged scores showed that the best-fitting model included only the serial position factor. This model was 8.19 times more likely than the following model, including both presentation direction and serial position factors (Presentation direction: η_p^2 = 0.007; Serial position: η_p^2 = 0.369; Presentation direction × Serial position: η_p^2 = 0.005). These results provide evidence *against* an effect of presentation direction (BF_{01} = 5.59) or its interaction with serial position (BF_{01} = 176.13) (see Appendix Tables B.1, 3, 9, and Figure 2.4).

Transposition Gradient. For transposition gradients, the best-fitting model included only the transposition distance factor. This model was 261.38 times more likely than the following model, including both presentation direction and transposition distance factors (Presentation direction: η_p^2 = 0.003; Transposition distance: η_p^2 = 0.980; Presentation direction × Transposition distance: η_p^2 = 0.006). These results indicate once more evidence *against* an effect of presentation direction (BF_{01} = 38.35) or its interaction with transposition distance (BF_{01} = 1.83×10^{21}) (see Appendix Tables B.1, 3, 11, and Figure 2.5).

Discussion

Experiment 1B revealed the same pattern of results as Experiment 1A, with a consistent absence of the effect of presentation direction on serial recall and item recall performance.

Experiment 1A and Experiment 1B show that a canonical vs. non-canonical spatial presentation direction does not appear to exert a major impact on verbal WM recall performance. One possibility is that, as suggested by Guida and his colleagues (2020), mental spatial representations are flexible and can be rapidly updated as a function of the contextual cues that are provided, and therefore, encoding items from left to right or from right to left provides equally supportive spatial encoding structures, especially when no specific spatial recall direction requirements are defined. If this is the case, then constraining recall direction may be a more optimal condition to unveil the impact of spatial codes on verbal WM. Experiment 2 examined this hypothesis by presenting items centrally and by assuming that central presentation leads to a left-to-right spatial encoding structure by default in our participants of Western origin, as robustly shown by the literature on spatialization effects in verbal WM (Guida et al. 2026). At the same time, participants were requested to recall in a left-to-right or a right-to-left direction while, importantly, maintaining a direct forward temporal recall direction (i.e., the first presented items either needed to be recalled in the most leftward or in the most rightward recall position of a horizontal response display). At the same time, we maintained an encoding duration of 1000ms per item to maximize the chances for the buildup of spatial representations, as this duration is closer to the encoding duration used in some of the studies examining spatialization effects in WM (Shi et al., 2020; Wang et al., 2018; Guida, Abrahamse et al., 2020). We reasoned that if serial order codes are associated with spatial codes organized by default from left to right and if these codes are used for encoding and retrieving items, then allowing participants to use this spatial direction also during recall may improve their recall performance; on the other hand, participants may have difficulties retrieving the items when required to recall in a right-to-left direction.

Experiment 2

Experiment 2 manipulated spatial processes during the recall stage by presenting the stimuli at the center of the screen and explicitly requesting participants to output the digits in canonical or non-canonical direction while keeping a forward recall direction (participants were not asked to do backward recall relative to the temporal order of encoding of the items).

Method

Participants. Forty-six participants were recruited (see Scoring and Analysis Procedure for justification of sample size) among the first-year undergraduate students of the Faculty of Psychology, Speech and Language Pathology, and Educational Sciences of the University of Liège, based on a course credit scheme. Participants were between 18 and 28 years old (*average age* = 19.02, *SD* = 2.20); 39 participants were female.

Materials. Forty-four seven-digit lists were created by pseudo-randomly sampling from the digits 1 to 7, including four practice lists. Lists could not include adjacent digit pairs (e.g., 1-2 or 7-6) or other structured digit combinations (e.g., 1-3-5 or 6-4-2). In addition, we balanced the position of occurrence of the seven digits across lists and conditions. Half of the lists were presented for the left-to-right recall condition, and the other half for the right-to-left recall condition. For each participant, there was a different random order of the list presentation.

Procedures. The experiment was implemented using the OpenSesame platform (https://osdoc.cogsci.nl).

Each trial commenced with a fixation cross displayed at the center of the screen for 500ms to ensure attentional focus. Following a 500-ms blank interval, digits were presented sequentially, each appearing for 1000ms, with a 500-ms interstimulus blank interval. After the final digit, a directional arrow (pointing either to the left or to the right) was displayed for 1000ms, followed by the appearance of seven empty response boxes evenly spaced along a centered horizontal axis. The arrow direction was varied randomly across trials. Participants were instructed to enter the digits into the response boxes according to the direction indicated by the arrow, one by one, while maintaining the correct temporal order. Responses were provided via mouse clicks to select a response box, followed by manual digit entry using the keyboard. They were asked to fill in the current box and enter the following one, and they were not allowed to skip the following box or click on other boxes. If participants were unable to recall a digit at the given position, they were required to enter "0" as a placeholder (See Figure 2.1).

Before administration of the 40 experimental trials (20 left-to-right and 20 right-to-left), participants completed four practice trials with corrective feedback from experimenters to ensure full comprehension of the task instructions. There was no time limit for responses. The total duration of the experiment was approximately 30 minutes per participant.

Scoring and Analysis Procedure. Data scoring and analyses were the same as in Experiment 1. The only difference was that the fixed effect was recall direction instead of presentation direction.

After the 25th participant, the effect sizes of both item and serial order recall scores were stable. As for Experiment 1, data collection was continued as additional participants were available.

Results

Serial Recall Score. A Bayesian GLMM analysis of the serial recall score revealed that the best-fitting model included only the serial position factor. Although this model was 2.05 times more likely than the model including recall direction and serial position factors, it was retained as the more parsimonious model (Recall direction: η_p^2 = 0.030; Serial position: η_p^2 = 0.489; Recall direction × Serial position: η_p^2 = 0.037). These results tend to support the *absence* of an effect recall direction (BF_{01} = 2.39) and strongly support the *absence* of its interaction with serial position (BF_{01} = 32.53) (see Appendix Tables B.1, 4, 7, and Figure 2.2).

Item Recall Score. For the item recall score, the best-fitting model included the serial position factor. This model was 7.12 times more likely than the following model, including both recall direction and serial position factors (Recall direction: η_p^2 = 0.002; Serial position: η_p^2 = 0.383; Recall direction × Serial position: η_p^2 = 0.024). These results provide again evidence against an effect of recall direction (BF_{01} = 6.48) or its interaction with serial position (BF_{01} = 4.31) (see Appendix Tables B.1, 4, 8, and Figure 2.3).

Order Recall Score. For the order recall score, a repeated measures Bayesian ANOVA analysis on subject-averaged scores showed that the best-fitting model included only the serial position factor. This model was 5.70 times more likely than the following model, including both recall direction and serial position factors (Recall direction: η_p^2 = 0.045; Serial position: η_p^2 = 0.355; Recall direction × Serial position: η_p^2 = 0.018). These results indicate evidence against an effect of the recall direction (BF_{01} = 5.73) or its interaction with serial position (BF_{01} = 15.80) (see Appendix Tables B.1, 4, 9, and Figure 2.4).

Transposition Gradients. For transposition gradients, a repeated measures Bayesian ANOVA analysis revealed that the best-fitting model included the transposition distance factor. This model was 227.37 times more likely than the following model, including both recall direction and transposition distance factors (Recall direction: η_p^2 = 0.120; Transposition distance: η_p^2 = 0.986; Recall direction × Transposition distance: η_p^2 = 0.024). These results indicate once more evidence against an effect of the recall direction factor (BF_{01} = 29.80), or its interaction with the transposition distance factor ($BF_{01} = 2.63 \times 10^{20}$) (see Appendix Tables B.1, 4, 11, and Figure 2.5).

Discussion

By constraining recall rather than encoding direction, we still observed no effect of spatial manipulation on verbal WM performance, for any of the measures that were used. One possible explanation is that participants relied primarily on phonological coding rather than spatial coding to encode the memorized sequence, given that the spatial direction at recall was not predictable. If the setup of spatial mental representations is flexible, as a function of contextual and task constraints, then participants may have chosen not to use spatial codes at all in this specific experimental setup, given that spatial codes will not be helpful for retrieving items in half of the trials.

If this is the case, then an effect of spatial manipulations on WM recall performance may be observable only if spatial coding is manipulated both at encoding and recall stages. This logic was implemented in subsequent experiments by manipulating both presentation direction and recall direction. Specifically, stimuli were presented in different spatial orientations, and participants were required to recall the sequence either in the same direction as during presentation (congruent condition) or in the opposite direction (incongruent condition). If spatial coding plays an essential role in the coding of serial order information in verbal WM, we expect lower performance when presentation and recall directions are misaligned. Moreover, these manipulations allowed us to examine whether this effect is stronger when the congruent conditions align with participants' habitual reading direction, i.e., for the left-to-right encoding and recall directions. According to Guida et al.'s (2020) study, this should not necessarily be the case.

Experiment 3

In Experiment 3A and Experiment 3B, participants encoded lists of digits in either a left-to-right or a right-to-left direction, as in Experiments 1A and 1B, and their recall direction was in addition constrained (while keeping each time a forward recall direction relative to the temporal order of list presentation, as in previous experiments), as in Experiment 2. Experiment 3A was a direct combination of the paradigms of Experiment 1B and Experiment 2, by asking participants to recall items either by starting from the most leftward or the most rightward response box, as in Experiment 2. If spatial codes determine serial order in WM, then recall performance should be higher when encoding and recall directions match (for a left-to-right encoding direction, putting the first item in the most leftward response box will also match the most leftward spatial position of this item during encoding). However, if spatial codes are used flexibly and optionally rather than in an absolute manner, then this experimental setup may not yet be the most optimal setup for unveiling spatial effects on serial order WM, as participants may decide not to use spatial codes at all, given the unpredictability of the match between the spatial directions at encoding and recall. Therefore, we ran an additional, final experiment, Experiment 3B, in which participants needed to use spatial information for retrieving serial position and associated item information by using a spatial cued recall procedure instead of a full recall procedure. Specifically, at recall, response boxes were again aligned from left to right, but the serial position of the item to be retrieved was indicated each time by pointing an arrow to a specific box. The spatial position of the cued box needed to be used to retrieve the corresponding serial position of the item. If spatial codes intervene in the coding of serial order, then this specific setup should lead to an advantage for left-to-right vs. right-to-left presentation directions. Indeed, in that case, the spatial cue (e.g., the second box from the left being cued) will correspond to the spatial position of the item during encoding (the second most leftward position). For the right-to-left presentation condition, there will be an incongruence: a cue on the second most leftward recall box corresponding to the second most rightward encoding position (see Figure 2.1).

Experiment 3A

Method

Participants. Fifty-seven participants were recruited (see Scoring and Analysis Procedure for justification of sample size) among the first-year undergraduate students in the Faculty of Psychology, Speech and Language Pathology, and Educational Sciences of the University of Liège, based on a course credit scheme. Fourteen participants were excluded from the analysis: seven of them had been exposed to a second language using a right-to-left reading direction (e.g., Arabic); three of them were left-handed; two of them had a language-related disorder (e.g., dyslexia); two of them did not fully comply with task instructions. Data from 43 participants were included in this experiment (*average age* = 18.62, *SD* = 1.15; 33 females).

Materials. Sixty-four seven-digit lists were created by pseudo-randomly sampling from the digits 1 to 7, including four practice lists. Lists could not include adjacent digit pairs (e.g., 1-2 or 7-6) or other structured digit combinations (e.g., 1-3-5 or 6-4-2). Additionally, we balanced the distribution of the seven digits across lists and conditions. Half of the lists were presented for the canonical condition, and the other half for the non-canonical condition. For each presentation condition list, half of them were recalled in the left-to-right condition, and the other half in the right-to-left condition. For each participant, the lists were presented in a different random order. There were 15 lists per experimental condition.

Procedures. The experiment was conducted using the OpenSesame platform (https://osdoc.cogsci.nl).

Each trial commenced with a fixation cross displayed at the center of the screen for 500ms, designed to focus participants' attention. This was followed by a 500-ms blank interval. The digits were presented one at a time, each for 1000ms, with a 500-ms blank screen interval between items. The spatial position of each successive digit progressed either from left to right or right to left, with the presentation direction varying randomly across trials. After the final digit was displayed, an arrow appeared, followed by seven response boxes positioned at the locations corresponding to the previous digit presentations. Participants were instructed to recall the digits in the direction indicated by the arrow, entering the digits in temporal order either from left to right or right to left, by clicking the appropriate box and typing the digit for each position. Upon completion of the recall, participants clicked "Suivant" ("Next") to proceed to the next trial. If participants were unable to remember the digit at that position, they were instructed to enter "0" in place of the forgotten digit (See Figure 2.1).

Before starting the 60 experimental trials (with 15 trials per condition), participants completed four practice trials with corrective feedback to ensure they understood the task instructions. There was no time limit for responses. The total duration of the experiment was approximately 30 minutes per participant.

Scoring and Analysis Procedure. Data scoring and analysis were the same as in Experiments 1 and 2, except for the addition of the recall direction factor or the presentation direction factor.

Sample size was monitored using the same analysis method as in previous experiments. After the 25th participant, the effect sizes of both item and serial order recall scores were stable. Data collection was continued as additional participants were available to maximize the statistical sensitivity of our study.

Results

Serial Recall Score. A Bayesian GLMM analysis of the serial recall score revealed that the best-fitting model included both presentation direction and serial position factors. Although this model was 2.22 times more likely than the following model, including presentation direction, recall direction and serial position factors, the former one was retained as the more parsimonious model (Presentation direction: η_p^2 = 0.155; Recall direction: η_p^2 = 0.006; Serial position: η_p^2 = 0.455; Presentation direction × Recall direction: η_p^2 = 0.026; Presentation direction × Serial position: η_p^2 = 0.049; Recall direction × Serial position: η_p^2 = 0.019; Presentation direction × Recall direction × Serial position: η_p^2 = 0.026). These results provide evidence that tends *against* an effect of recall direction (BF_{01} = 0.82), and evidence strongly *against* the interactions of presentation direction with recall direction, presentation direction with serial position, and recall direction with serial position (BF_{01} = 4.90; BF_{01} = 12.41; BF_{01} = 190.75). However, contrary to previous experiments, better serial recall performance was observed for a canonical presentation direction, regardless of the direction of recall. (see Appendix Tables B.1, 5, 7, and Figures 2.2, 6).

Item Recall Score. The same analysis on the item recall score showed that the best-fitting model included the presentation direction and serial position factors, as well as their interaction. This model was 3.63 times more likely than the model not including the interaction (Presentation direction: η_p^2 = 0.265; Recall direction: η_p^2 = 0.001; Serial position: η_p^2 = 0.322; Presentation direction × Recall direction: η_p^2 = 0.037; Presentation direction × Serial position: η_p^2 = 0.063; Recall direction × Serial position: η_p^2 = 0.022; Presentation direction × Recall direction × Serial position: η_p^2 = 0.029). These results provide evidence strongly *against* an effect of recall direction (BF_{01} = 6.88) or the interaction between recall direction and serial position (BF_{01} = 11.65). However, evidence supports a main effect of presentation direction (BF_{10} = 3.63) as well as its interaction with serial position (BF_{10} = 3.63). An exploration of the interaction showed that the effect of presentation direction was most pronounced for serial positions 5 and 6, although note that the effect remained small, the credible intervals showing a non-negligeable overlap for the two presentation directions (Position 1: *estimated marginal mean* = 0.976, *95%CI* [0.966, 0.985] vs. 0.959, *95%CI* [0.942, 0.973]; Position 5: 0.930, *95%CI* [0.905, 0.951] vs. 0.907, *95%CI* [0.878, 0.935]; Position 6: 0.911, *95%CI* [0.881, 0.936] vs. 0.875, *95%CI* [0.836, 0.909]) (see Appendix Tables B.1, 5, 8, and Figures 2.3, 6).

Order Recall Score. A repeated measures Bayesian ANOVA analysis on group-averaged scores on the order recall score showed that the strongest model included only the serial position factor. Although the model with serial position and presentation direction was 1.17 times more likely than this model, the latter one was retained as the more parsimonious model (Presentation direction: η_p^2 = 0.093; Recall direction: η_p^2 = 0.002; Serial position: η_p^2 = 0.416; Presentation direction × Recall direction: η_p^2 = 0.026; Presentation direction × Serial position: η_p^2 = 0.018; Recall direction × Serial position: η_p^2 = 0.017; Presentation direction × Recall direction × Serial position: η_p^2 = 0.026). These results tend to support the absence of a presentation direction effect (BF_{01} = 1.56) or its interaction with recall direction (BF_{01} = 0.50), and strongly supports the absence of an effect of recall direction (BF_{01} = 7.54), as well as of the interactions of serial position with presentation direction or recall direction (BF_{01} = 83.93; BF_{01} =141.04) (see Appendix Tables B.1, 5, 9, and Figure 2.4).

Transposition Gradients. For transposition gradients, the repeated measures Bayesian ANOVA analysis showed that the best-fitting model included only the transposition distance factor (Presentation direction: η_p^2 = -0.033; Recall direction: η_p^2 = -1.184×10^{-4}; Transposition distance: η_p^2 = 0.992; Presentation direction × Recall direction: η_p^2 = 0.008; Presentation direction × Transposition distance: η_p^2 = 0.073; Recall direction × Transposition distance: η_p^2 = 0.007; Presentation direction × Recall direction × Transposition distance: η_p^2 = 0.019). These results indicate again evidence *against* an effect of presentation direction (BF_{01} = 39.55), recall direction (BF_{01} = 39.93), the interaction of presentation direction and recall direction (BF_{01} = 19.54), or their interaction with serial position (BF_{01} = 5.13×10^{19}; BF_{01} = 1.44×10^{23}) (see Appendix Tables B.1, 5, 11, and Figure 2.5).

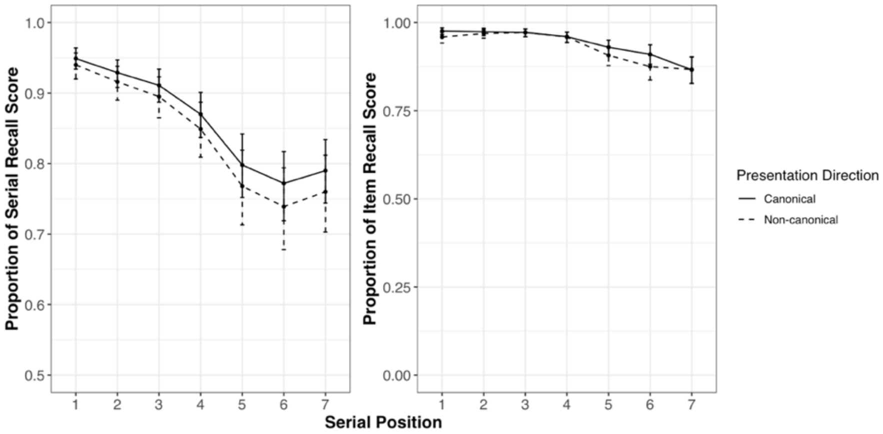

Figure 2.6. Serial recall performance and item recall performance, as a function of presentation direction and serial position in Experiment 3A. Estimated marginal means (EMMs) with credible intervals (95%CI) are depicted.

Discussion

In Experiment 3A, participants performed an immediate serial recall task in which both presentation and recall directions were manipulated, maximally raising their awareness about the spatial dimensions associated with the serial recall task. Contrary to previous experiments, a main effect of presentation direction was observed for both serial recall and item recall scores. At the same time, the fact that presentation and recall direction did not interact suggests that the spatial dimensions associated with recall direction did not influence the left-right presentation direction advantage. One possible interpretation is that participants did not take into account the spatial requirements of the recall structure, in line with a flexible and adaptive use of spatial codes. Furthermore, the main effect of presentation direction was only observed for the serial and item recall measures, while we would have expected the strongest effect on the order recall measure. Experiment 3B contrasted left-to-right vs. right-to-left presentation directions as in Experiment 3A but controlled the use of spatial codes during the recall stage in a more stringent manner. We used a probe recall task in which the serial positions to be retrieved were spatially cued, and hence, the only way to access the target WM representation was to take into account these spatial cues.

Experiment 3B

Method

Participants. Sixty-nine participants were recruited (see Scoring and Analysis Procedure for justification of sample size) among the first-year undergraduate students of the Faculty of Psychology, Speech and Language Pathology, and Educational Sciences of the University of Liège, based on a course credit scheme. Sixteen participants were excluded from the data analyses: five of them had a mother tongue different from French; two of them could speak and read Arabic fluently, a language using a right-to-left reading direction; two of them had a language-related disorder (e.g., dyslexia); and seven of them did not comply with task instructions. Valid data for fifty-three participants were included in this experiment (*average age* = 19.87, *SD* = 1.65; 41 females).

Materials. Materials were the same as in Experiment 3A, but with seventy-four seven-digit lists. Half of the lists were presented for the left-to-right condition, and the other half for the right-to-left condition. For each participant, the lists were presented in a different random order.

Procedures. The experiment was implemented using the OpenSesame platform (https://osdoc.cogsci.nl).

Each trial began with a fixation cross displayed at the center of the screen for 500ms to ensure attentional focus, followed by a 500-ms blank interval. Digits were then presented sequentially, from left to right or from right to left, with presentation direction varying randomly across trials. Each item appeared for 1000ms, with a 500-ms blank interval between items. Following a 500-ms blank interval after the final digit, a response screen appeared, displaying seven empty response boxes evenly spaced along the same horizontal axis. One of these boxes was marked with a target symbol, indicating the position of the digit to be recalled. Participants were required to determine the corresponding digit based on its temporal order while simultaneously counting positions from left to right. Responses were entered by inputting the recalled digit via the keyboard. If participants were unable to recall the digit, they were instructed to enter "0" as a placeholder (See Figure 2.1).

Before administration of the 70 experimental trials (35 canonical and 35 non-canonical), participants completed four practice trials with corrective feedback to ensure full comprehension of the task instructions. There was no time limit for responses. The total duration of the experiment was approximately 30 minutes per participant.

Scoring and Analysis Procedure. We determined recall accuracy (correct item in target position) and omission error (when participants entered "0").

The same Bayesian generalized linear mixed model method was used as in the preceding experiments. Sample size was monitored using the same analysis method as in previous experiments. After the 25th participant, the effect sizes of accuracy were stable. As in previous experiments, data collection continued as additional participants became available to maximize the statistical sensitivity of our study.

Results

Accuracy. A Bayesian GLMM analysis on the accuracy score revealed that the best-fitting model included the serial position factor and the interaction of serial position and presentation direction factors. Although this model was 1.04 times more likely than the model also including the presentation direction factor, it was retained as the more parsimonious model (Presentation direction: η_p^2 = 0.049; Serial position: η_p^2 = 0.193; Presentation direction × Serial position: η_p^2 = 0.062). Although evidence tended to support the *absence* of a main effect of presentation direction (BF_{01} = 1.47), the presence of its interaction with serial position ($BF_{10} = 1.98 \times 10^4$) was strongly supported. An exploration of the interaction showed that there was an advantage for a canonical vs. non-canonical presentation direction specifically at Position 6 (β = 0.57, *95%CI* [0.15, 0.99]; *estimated marginal mean* = 0.865, *95%CI* [0.819, 0.910] vs. 0.783, *95%CI* [0.726, 0.844]) and Position 7 (β = 0.95, *95%CI* [0.43, 1.48]; *estimated marginal mean* = 0.930, *95%CI* [0.897, 0.957] vs. 0.838, *95%CI* [0.786, 0.886]) (see Appendix Tables B.1, 6, 10 and Figures 2.7).

Omission Error. For omission errors, the best-fitting model also included serial position and the interaction of serial position and presentation direction factors. Although this model was 1.03 times more likely than the model also including the presentation direction factor, it was retained as the more parsimonious model (Presentation direction: $\eta_p^2 = 0.016$; Serial position: $\eta_p^2 = 0.086$; Presentation direction × Serial position: $\eta_p^2 = 0.042$). Again, evidence supported the *absence* of a main effect of presentation direction ($BF_{01} = 2.06$) but strongly supported its interaction with serial position ($BF_{10} = 2.05 \times 10^3$). The interaction was characterized by a selective increase of omission errors for non-canonical presentation direction at Position 5 (Position 5: 0.035, *95%CI* [0.016, 0.061] vs. 0.070, *95%CI* [0.036, 0.110]) (see Appendix Tables B.1, 6, 10, and Figures 2.7).

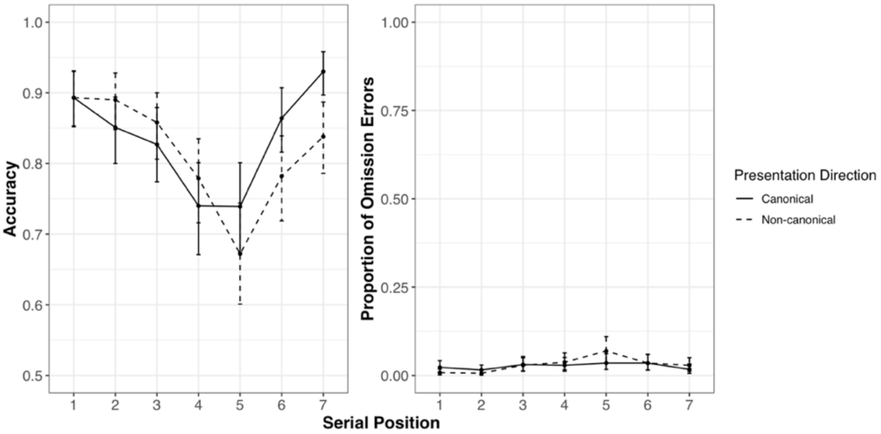

Figure 2.7. Accuracy and proportion of omission error, as a function of presentation direction. Estimated marginal means (*EMMs*) with credible intervals (*95%CI*) are depicted.

Discussion

The final Experiment 3B provides further support for a task and context-specific impact of spatial variables on verbal WM by showing that, when spatial cues need to be used for retrieving memoranda at the moment of recall, an advantage for a canonical versus non-canonical encoding direction can be observed, at least over final list positions.

General Discussion

This study investigated whether spatial codes support serial order memory by manipulating the setup of canonical vs. non-canonical spatial codes during WM encoding/recall and their impact on serial order WM performance. Across three sets of experiments, we varied presentation and recall direction in immediate serial recall and position probe tasks. Overall, we found that spatial manipulations produced very limited effects on verbal WM performance, irrespective of the type of measure used. An advantage for a canonical encoding direction was only observed in Experiments 3A and 3B, in which spatial variables were manipulated both at encoding and recall stages.

Overall, the present data provide a highly nuanced support for the role of spatial variables during encoding and recall of information in verbal WM. The clearest evidence for spatial contributions to order memory comes from Experiment 3B, where canonical left-to-right presentation of the lists during encoding improved recall performance at recency positions. This finding aligns with earlier research showing that spatial manipulations can influence recall performance, and this also mostly occurs for end-of-list positions in WM (Fischer-Baum & Benjamin, 2014; Battacchi et al., 1990). However, in the present study, spatial manipulations had an effect only when participants were maximally incentivized to process spatial dimensions by manipulating spatial variables at both the encoding and recall stages. Only manipulating spatial dimensions during encoding or recall did not lead to any robust effects on recall performance, indicating that spatial codes do not support verbal WM performance in an obligatory and universal manner. In line with this conclusion, we should also note here that in Experiment 3A, we did not observe the expected interaction between encoding direction and recall direction, despite the fact that a main effect of presentation direction was observed.

What could be the factors explaining this highly nuanced impact of spatial variables on verbal WM performance? One possible explanation is that contrary to our prediction of enhanced spatial effects when the number of serial positions to be encoded is high, higher list lengths may actually lead to a reduced reliance on spatial codes. For example, Guida, Carnet, et al. (2018) found no spatialization effects with sequences of 15 items, but it could be argued that encoding 15 items largely exceeds standard WM capacity. Huber et al. (2016) observed weaker spatialization effects as list length increased from four to six items. Recently, Vivion et al. (2025) reported spatialization effects only for short sequences (four items) but not for longer ones (five items). We, however, need to note here that spatial codes *can* be used for longer list lengths, when this can be useful, as indicated by the spatial effects observed in Experiment 3A and 3B of this study, which used lists of seven items.

Task-related factors are indeed likely to be the most plausible factors for explaining the occasional observation of spatial effects on verbal WM recall performance. As already noted, in our study, these effects were only observed in Experiments 3A and 3B in which participants were explicitly made aware of the spatial manipulations by not only manipulating encoding direction, but also by manipulating recall direction or by cueing the retrieval of memoranda via spatial cues. In Experiments 1A, 1B, and 2, participants may instead have chosen to rely on phonological/temporal codes for maintaining and recalling memoranda, despite the spatial manipulations that occurred exclusively at the encoding or at the recall stages. This interpretation could also explain the discrepancy of results between our Experiments 1A and 1B and the study by Fischer-Baum and Benjamin (2014), of which our Experiments 1A and 1B were supposed to be a conceptual replication. While our studies manipulated left-to-right vs right-to-left encoding directions, as did the study by Fischer-Baum and Benjamin (2014), the latter study additionally provided explicit directional cues before presenting the lists. This was needed to inform the participants about the forward vs. backward type of recall that was additionally manipulated in that study. More generally, phonological coding (and rehearsal) may mask potential spatial effects unless tasks or instructions reduce the dominance of phonological processes.

Another, related explanation for the limited impact of spatial manipulations on verbal WM performance in the present study is that the extent of spatial coding may vary on an inter-individual basis, as also indicated by our analyses of random slopes, which reveal participant-specific variability in the impact of spatial manipulations. This aligns with recent findings showing that only a minority of participants exhibit robust spatialization effects (Bottini & Doeller, 2020; van Dijck et al., 2022). For instance, van Dijck et al. (2022) showed that, while spatialization effects can be observed at the group level in verbal WM, only about 20% of participants will demonstrate an individually robust left-to-right spatialization effect.

At a theoretical level, *the mental whiteboard hypothesis* (Abrahamse et al., 2014) and *the principle of economy* (Guida & Campitelli, 2019; Guida, Abrahamse, et al., 2020) propose that temporal sequences are automatically mapped onto a spatial axis, with the specific spatial layout shaped by contextual information during encoding. The present findings are consistent with this view in showing that spatial coding can be recruited when task context provides informative or relevant spatial cues. At the same time, the results indicate that spatial coding is neither universal nor obligatory. Furthermore, our data suggest that spatialization depends not only on encoding, but also on retrieval demands. Taken together, these findings suggest that current formulations of the mental whiteboard hypothesis may overestimate the automaticity of spatial coding. Rather, our results support a more flexible, context-dependent account of spatial coding of order information in verbal WM.

Several limitations should be noted. First, we did not directly control for rehearsal or other phonological strategies used or not used by our participants. Although Ginsburg et al. (2017) observed spatialization effects even when introducing a secondary articulatory suppression task, either at encoding or retrieval of memoranda, the meta-analysis by Guida et al. (2024) suggests that spatialization effects can be moderated by phonological encoding and maintenance processes. Second, we did not run our own studies by using shorter list lengths, for which stronger spatialization effects may have been observed. Future work should systematically examine how phonological processes and sequence length interact with spatial coding.

In conclusion, the present findings indicate that spatial coding is not a fundamental or obligatory mechanism for serial order representation in verbal WM. Rather, it appears to be a flexible, context-dependent mechanism that can scaffold serial order memory under very specific task conditions, which still need to be elucidated in a more specific manner.

Appendix

Table B.1. The results of Bayesian analysis on random slopes. The factor was included in the model as a random slope if the Bayes factor comparing the null model was above 3.

	Model	Serial recall score	Item recall score	Order recall score	Transposition gradient
E1A	None	1	1	1	1
	Presentation	2.32×10^{4}	0.24	0.09	0.03
E1B	None	1	1	1	1
	Presentation	5.38×10^{2}	0.30	0.07	0.03
E2	None	1	1	1	1
	Recall	0.28	0.05	0.08	0.03
E3A	None	1	1		
	Presentation	0.11	0.08		
E3B	None	1	1	1	1
	Presentation	3.04×10^{8}	0.21	37.73	0.00
	Recall	3.97×10^{2}	0.07	0.14	0.00
	Presentation × Recall	2.77×10^{10}	0.00	0.58	0.00

Table B.2. Results of serial recall score, item recall score, order recall score, and transposition gradient in Experiment 1A

Model	BF$_{10}$ Serial recall	BF$_{10}$ Item recall	BF$_{10}$ Order recall	BF$_{10}$ Transposition gradient*
None (incl. random effects)	1	1	1	0
Presentation direction	0.78	0.20	0.17	0
Serial position	3.40×10^{297}	4.02×10^{182}	1.04×10^{30}	1
Presentation direction + Serial position	2.62×10^{297}	9.34×10^{181}	1.49×10^{29}	4.83×10^{-3}
Serial position + Presentation direction × Serial position	1.82×10^{296}	4.82×10^{181}	5.47×10^{27}	1.78×10^{-23}
Presentation direction + Serial position + Presentation direction × Serial position	1.44×10^{296}	4.98×10^{181}	5.36×10^{27}	1.78×10^{-23}

* For "Transposition gradient", "Serial position" means "Transposition distance".

Table B.3. Results of serial recall score, item recall score, order recall score, and transposition gradient in Experiment 1B

Model	BF_{10}			
	Serial recall	Item recall	Order recall	Transposition gradient*
None (incl. random effects)	1	1	1	0
Presentation direction	0.85	1.30	0.18	0
Serial position	5.18×10^{257}	1.20×10^{126}	2.32×10^{45}	1
Presentation direction + Serial position	4.85×10^{257}	1.85×10^{126}	3.06×10^{44}	3.82×10^{-3}
Serial position + Presentation direction × Serial position	2.40×10^{256}	1.52×10^{126}	1.63×10^{42}	2.06×10^{-24}
Presentation direction + Serial position + Presentation direction × Serial position	2.76×10^{256}	1.55×10^{126}	1.63×10^{42}	2.05×10^{-24}

* For “Transposition error”, “Serial position” means “Transposition distance”.

Table B.4. Results for the serial recall score, the item recall score, the order recall score, and transposition gradient in Experiment 2

Model				BF_{10}
	Serial recall	Item recall	Order recall	Transposition gradient*
None (incl. random effects)	1	1	1	0
Recall direction	0.42	0.15	0.17	0
Serial position	5.77×10^{168}	5.80×10^{94}	2.92×10^{35}	1
Presentation direction + Serial position	2.72×10^{168}	8.52×10^{93}	5.36×10^{34}	4.40×10^{-3}
Serial position + Presentation direction × Serial position	9.06×10^{166}	2.05×10^{93}	3.46×10^{33}	1.62×10^{-23}
Presentation direction + Serial position + Presentation direction × Serial position	8.83×10^{166}	1.92×10^{93}	3.23×10^{33}	1.69×10^{-23}

* For “Transposition error”, “Serial position” means “Transposition distance”.

Table B.5. Results of serial recall score, item recall score, order recall score, and transposition gradient in Experiment 3A

Model	BF_{10}			
	Serial recall	Item recall	Order recall	Transposition gradient[*]
None (incl. random effects)	1	1	1	0
Presentation direction	2.08	63.29	0.64	0
Recall direction	1.22	0.15	0.13	0
Serial position	3.99×10^{163}	8.76×10^{113}	1.45×10^{28}	1
Presentation direction+ Recall direction	1.27	5.37	0.09	0
Presentation direction+ Recall direction + Presentation direction × Recall direction	0.43	8.76	0.18	0
Serial position + Presentation direction	2.49×10^{164}	7.19×10^{115}	1.80×10^{28}	2.69×10^{-3}
Serial position + Presentation direction + Serial position × Presentation direction	3.21×10^{163}	2.67×10^{116}	2.08×10^{26}	5.48×10^{-23}
Serial position + Recall direction	1.77×10^{164}	1.26×10^{113}	1.73×10^{27}	3.03×10^{-3}
Serial position + Recall direction + Serial position × Recall direction	8.66×10^{161}	1.11×10^{112}	1.30×10^{25}	2.10×10^{-26}

[*] For “Transposition error”, “Serial position” means “Transposition distance”.

Table B.6. Results of accuracy and omission error in Experiment 3B

Model	BF_{10}	
	Accuracy	Omission errors
None (incl. random effects)	1	1
Presentation direction	0.68	0.49
Serial position	1.75×10^{21}	2.20×10^{7}
Presentation direction + Serial position	1.25×10^{21}	1.10×10^{7}
Serial position + Presentation direction × Serial position	2.55×10^{25}	2.33×10^{10}
Presentation direction + Serial position + Presentation direction × Serial position	2.47×10^{25}	2.26×10^{10}

Table B.7. Statistical description of the serial recall score

	Presentation direction	Recall direction	Serial position							N
			1	2	3	4	5	6	7	
E1A	Canonical		0.954 (0.074)	0.898 (0.120)	0.847 (0.149)	0.752 (0.179)	0.668 (0.231)	0.605 (0.271)	0.653 (0.251)	57
	Non-canonical		0.947 (0.081)	0.910 (0.100)	0.849 (0.139)	0.770 (0.173)	0.650 (0.198)	0.588 (0.243)	0.605 (0.244)	57
E1B	Canonical		0.970 (0.049)	0.918 (0.091)	0.857 (0.150)	0.788 (0.173)	0.739 (0.233)	0.695 (0.251)	0.740 (0.240)	77
	Non-canonical		0.950 (0.070)	0.904 (0.100)	0.851 (0.151)	0.788 (0.182)	0.730 (0.219)	0.677 (0.252)	0.720 (0.233)	77
E2		Left-to-right	0.922 (0.070)	0.904 (0.081)	0.848 (0.119)	0.816 (0.145)	0.705 (0.212)	0.678 (0.226)	0.674 (0.238)	47
		Right-to-left	0.923 (0.079)	0.890 (0.096)	0.855 (0.144)	0.820 (0.169)	0.745 (0.196)	0.707 (0.218)	0.676 (0.225)	47
E3A	Canonical	Left-to-right	0.947 (0.059)	0.907 (0.091)	0.890 (0.093)	0.820 (0.135)	0.775 (0.155)	0.761 (0.224)	0.752 (0.193)	43
		Right-to-left	0.935 (0.067)	0.922 (0.091)	0.887 (0.108)	0.860 (0.114)	0.797 (0.161)	0.766 (0.226)	0.766 (0.203)	43
	Non-canonical	Left-to-right	0.916 (0.094)	0.898 (0.108)	0.882 (0.114)	0.836 (0.143)	0.755 (0.192)	0.699 (0.223)	0.746 (0.184)	43
		Right-to-left	0.918 (0.081)	0.888 (0.133)	0.873 (0.124)	0.833 (0.167)	0.722 (0.218)	0.721 (0.248)	0.755 (0.196)	43

Table B.8. Statistical description of the item recall score

	Presentation direction	Recall direction	Serial position							N
			1	2	3	4	5	6	7	
E1A	Canonical		0.973	0.967	0.936	0.898	0.837	0.778	0.791	57
			(0.049)	(0.057)	(0.084)	(0.111)	(0.156)	(0.218)	(0.189)	
	Non-canonical		0.972	0.970	0.951	0.905	0.829	0.770	0.753	57
			(0.049)	(0.051)	(0.062)	(0.088)	(0.144)	(0.204)	(0.217)	
E1B	Canonical		0.984	0.978	0.948	0.920	0.888	0.858	0.873	77
			(0.032)	(0.042)	(0.070)	(0.097)	(0.147)	(0.155)	(0.140)	
	Non-canonical		0.977	0.965	0.947	0.927	0.886	0.842	0.845	77
			(0.050)	(0.060)	(0.085)	(0.102)	(0.137)	(0.173)	(0.169)	
E2		Left-to-right	0.961	0.949	0.936	0.930	0.862	0.843	0.795	47
			(0.061)	(0.065)	(0.083)	(0.089)	(0.163)	(0.169)	(0.189)	
		Right-to-left	0.948	0.940	0.936	0.916	0.879	0.860	0.786	47
			(0.057)	(0.066)	(0.085)	(0.105)	(0.131)	(0.150)	(0.182)	
E3A	Canonical	Left-to-right	0.969	0.960	0.955	0.941	0.901	0.884	0.831	43
			(0.042)	(0.058)	(0.052)	(0.064)	(0.110)	(0.177)	(0.161)	
		Right-to-left	0.963	0.967	0.964	0.949	0.913	0.885	0.840	43
			(0.053)	(0.055)	(0.061)	(0.061)	(0.112)	(0.151)	(0.190)	
	Non-canonical	Left-to-right	0.946	0.963	0.971	0.950	0.888	0.828	0.831	43
			(0.072)	(0.055)	(0.047)	(0.065)	(0.134)	(0.191)	(0.170)	
		Right-to-left	0.940	0.949	0.949	0.936	0.873	0.862	0.842	43
			(0.069)	(0.073)	(0.066)	(0.068)	(0.155)	(0.183)	(0.136)	

Table B.9. Statistical description of the order recall score

	Presentation direction	Recall direction	Serial position							N
			1	2	3	4	5	6	7	
E1A	Canonical		0.981 (0.059)	0.927 (0.101)	0.903 (0.128)	0.833 (0.150)	0.783 (0.198)	0.739 (0.231)	0.803 (0.199)	57
	Non-canonical		0.975 (0.070)	0.938 (0.083)	0.891 (0.123)	0.844 (0.145)	0.775 (0.178)	0.751 (0.208)	0.781 (0.210)	57
E1B	Canonical		0.985 (0.035)	0.937 (0.077)	0.900 (0.125)	0.853 (0.147)	0.823 (0.192)	0.792 (0.223)	0.828 (0.208)	77
	Non-canonical		0.971 (0.049)	0.936 (0.072)	0.895 (0.116)	0.844 (0.160)	0.813 (0.180)	0.783 (0.198)	0.838 (0.180)	77
E2		Left-to-right	0.961 (0.052)	0.953 (0.053)	0.904 (0.088)	0.876 (0.126)	0.807 (0.157)	0.790 (0.192)	0.833 (0.184)	47
		Right-to-left	0.973 (0.046)	0.946 (0.065)	0.909 (0.112)	0.888 (0.137)	0.837 (0.155)	0.809 (0.197)	0.846 (0.170)	47
E3A	Canonical	Left-to-right	0.977 (0.037)	0.945 (0.069)	0.932 (0.081)	0.871 (0.123)	0.854 (0.113)	0.833 (0.193)	0.894 (0.132)	43
		Right-to-left	0.971 (0.045)	0.953 (0.068)	0.919 (0.087)	0.906 (0.098)	0.866 (0.113)	0.843 (0.171)	0.909 (0.106)	43
	Non-canonical	Left-to-right	0.967 (0.055)	0.930 (0.076)	0.907 (0.096)	0.877 (0.122)	0.841 (0.157)	0.823 (0.154)	0.897 (0.112)	43
		Right-to-left	0.977 (0.044)	0.933 (0.095)	0.918 (0.094)	0.886 (0.153)	0.810 (0.207)	0.809 (0.198)	0.882 (0.142)	43

Table B.10. Statistical description of accuracy and proportion of omission error in E3B

	Presentation direction	Serial position							N
		1	2	3	4	5	6	7	
Accuracy	Canonical	0.875 (0.181)	0.823 (0.183)	0.811 (0.216)	0.706 (0.256)	0.721 (0.233)	0.838 (0.208)	0.909 (0.139)	53
	Non-canonical	0.864 (0.170)	0.868 (0.200)	0.830 (0.202)	0.743 (0.246)	0.649 (0.271)	0.762 (0.242)	0.808 (0.184)	53
Omission error	Canonical	0.042 (0.091)	0.030 (0.072)	0.060 (0.101)	0.060 (0.115)	0.068 (0.130)	0.072 (0.162)	0.038 (0.088)	53
	Non-canonical	0.019 (0.059)	0.015 (0.053)	0.057 (0.132)	0.072 (0.131)	0.128 (0.180)	0.068 (0.146)	0.053 (0.105)	53

Table B.11. Statistical description of the proportion of transposition gradient

	Presentation direction	Recall direction	Transposition gradient -6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	N
E1A	Canonical		0.000 (0.000)	0.000 (0.002)	0.001 (0.005)	0.006 (0.014)	0.022 (0.026)	0.050 (0.049)	0.867 (0.113)	0.037 (0.037)	0.011 (0.013)	0.003 (0.007)	0.001 (0.003)	0.001 (0.002)	0.000 (0.000)	57
	Non-canonical		0.001 (0.002)	0.001 (0.003)	0.002 (0.006)	0.006 (0.010)	0.019 (0.023)	0.055 (0.052)	0.862 (0.104)	0.038 (0.031)	0.010 (0.010)	0.003 (0.007)	0.002 (0.006)	0.001 (0.002)	0.000 (0.001)	57
E1B	Canonical		0.000 (0.001)	0.001 (0.003)	0.000 (0.002)	0.007 (0.013)	0.019 (0.023)	0.045 (0.045)	0.880 (0.112)	0.034 (0.036)	0.010 (0.012)	0.003 (0.005)	0.001 (0.004)	0.001 (0.003)	0.000 (0.002)	77
	Non-canonical		0.000 (0.001)	0.001 (0.003)	0.002 (0.006)	0.005 (0.009)	0.018 (0.024)	0.047 (0.045)	0.875 (0.109)	0.034 (0.030)	0.012 (0.014)	0.003 (0.007)	0.002 (0.005)	0.000 (0.002)	0.000 (0.000)	77
E2		Left-to-right	0.000 (0.002)	0.001 (0.003)	0.004 (0.008)	0.008 (0.011)	0.020 (0.016)	0.036 (0.038)	0.882 (0.090)	0.028 (0.030)	0.014 (0.017)	0.005 (0.008)	0.002 (0.005)	0.001 (0.002)	0.000 (0.002)	47
		Right-to-left	0.000 (0.001)	0.002 (0.006)	0.004 (0.007)	0.009 (0.012)	0.019 (0.022	0.031 (0.031)	0.891 (0.096)	0.027 (0.029)	0.009 (0.012)	0.005 (0.007)	0.002 (0.005)	0.001 (0.003)	0.000 (0.000)	47
E3A	Canonical	Left-to-right	0.001 (0.003)	0.001 (0.003)	0.001 (0.004)	0.007 (0.009)	0.019 (0.018)	0.021 (0.023)	0.908 (0.070)	0.022 (0.021)	0.012 (0.012)	0.005 (0.008)	0.002 (0.004)	0.002 (0.004)	0.000 (0.001)	43
		Right-to-left	0.000 (0.000)	0.000 (0.002)	0.002 (0.005)	0.007 (0.010)	0.013 (0.017)	0.025 (0.024)	0.915 (0.073)	0.021 (0.023)	0.011 (0.016)	0.003 (0.007)	0.001 (0.003)	0.001 (0.003)	0.000 (0.002)	43
	Non-canonical	Left-to-right	0.000 (0.000)	0.001 (0.005)	0.002 (0.006)	0.006 (0.009)	0.018 (0.019)	0.031 (0.029)	0.897 (0.077)	0.022 (0.023)	0.016 (0.015)	0.005 (0.009)	0.001 (0.004)	0.001 (0.004)	0.000 (0.000)	43
		Right-to-left	0.001 (0.003)	0.001 (0.003)	0.001 (0.002)	0.007 (0.014)	0.021 (0.026)	0.030 (0.033)	0.897 (0.099)	0.023 (0.027)	0.012 (0.016)	0.005 (0.011)	0.002 (0.005)	0.000 (0.002)	0.000 (0.001)	43

Study 3

Effects of Temporal Rhythm on Serial Order in Verbal Working Memory

Hong Xiao, Lucie Attout, and Steve Majerus

In preparation

Abstract. Verbal working memory (WM) is sensitive to the temporal structure of item presentation, yet the impact of within-sequence rhythmic regularity on serial order encoding remains unclear. Across immediate serial recall experiments, participants were presented with seven-digit sequences in either regular or irregular temporal patterns. Two types of irregular rhythms were created, involving either small discrepancies (<100 ms) or large discrepancies (at least ~200 ms) between successive inter-stimulus intervals. Accuracy in terms of item (digit recalled) and serial order (correct digit in the right position) precision was assessed. Results revealed an effect of temporal manipulation on WM performance: recall was more accurate in the regular condition, but only when temporal discrepancies were sufficiently large. However, this effect was small and limited to item recall in contrast to serial order recall measures. These findings provide limited support for temporal models of serial order, considering that temporal encoding is driven by purely bottom-up temporal signals.

Introduction

Serial order memory is a core cognitive mechanism that plays an important role in language acquisition, mathematical calculations, and speaking (A. Baddeley et al., 1998). The nature of the code supporting serial order in verbal working memory (WM) is still a central question, almost 75 years after Lashley's inaugural paper about the problem of serial order (Lashley, 1951b). Broadly speaking, current theoretical models of how serial order is coded in WM can be divided into two classes: associative chaining and position marker models.

Associative chaining (Ebbinghaus, 1964) was one of the earliest theoretical approaches, positing that serial order is maintained via learned associations between successive items, with each item serving as a retrieval cue for the next (e.g., Lashley, 1951). However, chaining models face several empirical limitations. For instance, they predict that recall should break down following an error in the sequence, which is often not the case (Henson, 1996). Moreover, they struggle to explain distance effects in WM, where order judgments are more difficult for items that are close in the sequence than for those that are farther apart (Attout et al., 2014; Marshuetz et al., 2000a). As a result, chaining accounts have largely been abandoned in favor of position marker models (Burgess & Hitch, 2006a; Farrell & Lewandowsky, 2002; R. N. Henson, 1996). Although note that more recently, a model integrating chaining and position coding processes has been proposed (Logan & Cox, 2021).

Position marker models propose that serial order is represented by associations between items and independent, varying contextual representations of their position within a sequence. These models differ in whether position markers are implemented using relative or absolute temporal codes, whose nature is often assumed to be temporal explicitly or implicitly (Henson, 1999b). Relative coding models, such as the Start-End Model (SEM) by Henson (1998), assume that items are coded relative to both start and end markers of a sequence. Absolute coding models, such as the original model proposed by Burgess and Hitch (1992), posit that serial order is coded via associations between items and a gradually evolving contextual signal, allowing sequential recall through context reinstatement. Other models explicitly assume that contextual signals are represented via temporal oscillators whose states evolve as a function of time, with serial order being encoded through associations between items and the oscillatory state present at encoding (Brown et al., 2000; Hartley et al., 2016). The present study provides an empirical test of the temporal nature of codes used for representing serial order information in verbal WM.

The oscillator-based associative recall (OSCAR) model (Brown et al., 2000) is one of the earliest models proposing a temporal coding scheme. Drawing on an analogy with a clock face, it proposes that endogenous hierarchical oscillatory signals provide a temporal context for item encoding and retrieval. At recall, these oscillators are reset, and items are retrieved according to their similarity to the current oscillatory state. More recently, Hartley et al. (2016) proposed the bottom-up multi-scale population oscillator (BUMP) model, which derives temporal context directly from the physical properties of stimuli as they unfold in time. This multi-oscillator filter mechanism incorporates bottom-up sensitivity to stimulus timing, rather than relying on top-down expectations or prior knowledge of temporal structure.

There is a vast literature highlighting the sensitivity of verbal WM performance to temporal factors. Temporal grouping effects have been most mainly examined in serial order memory, which interpret that memory performance improves when stimuli are grouped by short inter-stimulus intervals (ISIs) and separated by longer between-group intervals (Frankish, 1989; Hartley et al., 2016; Hitch, 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b), and this advantage is accompanied by systematic changes in order errors (Ryan, 1969a, 1969b). Specifically, within-group primacy and recency effects emerge in addition to the main primacy and recency effects that characterize the entire lists (Frankish, 1989; Hartley et al., 2016; Hitch, 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b). Temporal grouping also influences transposition errors, giving rise to interposition errors (Henson, 1996): serial order errors tending to involve within-group items rather than items stemming from different groups (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Hartley et al., 2016; Henson, 1999a; Ng & Maybery, 2002; Ryan, 1969a, 1969b). In these cases, temporal grouping effects are supposed to align with serial order memory models that posit a hierarchical temporal structure, with order encoded simultaneously at multiple levels, one tracking item position within groups and another encoding group positions at the whole-list level (Brown et al., 2000; Hartley et al., 2016). Instead of temporal grouping effects, additional evidence of temporal coding comes from temporal distinctiveness effects, in which items isolated by larger temporal gaps at encoding are recalled more accurately (Geiger & Lewandowsky, 2008; Lewandowsky et al., 2008; Morin et al., 2010), as well as the research on the relationships of verbal WM performance and the ability to reproduce rhythmical sequences (Saito, 2001), the temporal precision of rehearsal (Gilbert et al., 2017), and the length of maintenance intervals (De Belder et al., 2017).

Although the findings mentioned above highlight a general role of temporal factors on verbal WM performance, they do not yet directly demonstrate that serial order information is represented via temporal codes. Recently, a growing body of research has specifically examined the role of temporal variables in encoding and retrieving serial order information in WM. One set of studies has shown that presenting isochronous rhythms during the retention phase of a serial order WM task could improve recall performance. For example, Plancher et al. (2018) observed an improved serial recall performance when presenting a regular rhythm beat during maintenance, compared to silent or irregular conditions. Additionally, Fanuel et al. (2018) replicated this rhythm-related benefit during maintenance under both low and high cognitive load conditions, suggesting an automatic benefit of this regular temporal rhythm. While, in contrast, when temporal regularity was introduced during the encoding phase, a detrimental effect on serial order memory was observed. For instance, Henson et al. (2003) instructed participants to tap during list presentation and found that paced tapping impaired performance on a list probe task, regardless of whether tapping followed a regular or complex rhythm. Hall and Gathercole (2011) similarly reported reduced WM performance when a regular-paced sound was presented during encoding and maintenance phases, compared to the silent control condition. Moreover, Parmentier and Beaman (2015) found that the presence of rhythmic distractors during encoding and maintenance phases impaired recall performance compared to silence, and that regular rhythms were even more disruptive than irregular ones. These studies reveal contradictory results regarding the nature of the impact of a regular rhythm. One possibility here could be the nature of external rhythm itself, which, albeit regular, is non-isochronous with the item presentation rhythm. In this case, an isochronous rhythm presented during encoding is supposed to lead to facilitated encoding of memoranda and their serial order, or, at the least, not impair verbal WM recall performance. Studies have shown to be in line with this hypothesis. For example, Pannell et al. (2023) found comparable recall performance when presenting silence or an extra synchronized regular single tone during encoding. Similarly, Attout et al. (2025) observed improved recall performance when the encoding of memoranda was associated with an isochronous vs. random rhythm, and the isochronous rhythm condition led to the same level of performance as the silent baseline condition.

The recent findings are considered to support the temporal coding account; however, their results are complex and partly contradictory. One difficulty associated with most of the previous studies is that they involve a concurrent task paradigm, in which participants have to encode and maintain memoranda while also conducting a rhythm reproduction task. These dual task demands may lead to interactions due to the division of attentional resources during encoding, which may partly explain the detrimental effects of both regular and irregular rhythms. It should also be noted that Plancher et al. (2018) interpreted their findings as evidence for an attentional refreshing mechanism, considering that the rhythmic cue during maintenance favors the sequential refreshing of items. In order to circumvent these difficulties, a more direct approach would be to manipulate the rhythm of encoding of the memoranda, instead of adding a stream of rhythmic information.

To date, only one study has implemented this approach by comparing recall performance under regular versus irregular rhythms Gorin (2020). In his study, ISIs were manipulated by comparing regular, irregular (e.g., the time between even-numbered items of the memory list remained constant while it varied for odd-numbered items), or completely random rhythms. Despite these manipulations, no reliable differences in serial order performance emerged between regular and irregular pacing. One explanation for this null effect is the potential influence of temporal grouping effects. In irregular conditions, the fact that some of the items were nevertheless presented at fixed time points may have introduced a pattern in the lists, leading to spontaneous, facilitatory grouping effects, masking negative effects of temporal irregularity (see also Hartley et al., 2016; as well as Ryan, 1969a, for related findings). Furthermore, although Gorin used a fully irregular rhythm in one of the experiments, the large discrepancy between the different ISIs may also have created potential for item marking and grouping processes. Finally, the study used nine-item lists, which exceed the typical WM span, potentially reducing the sensitivity of the design for measuring the effect of temporal irregularity.

In the present study, we examined the impact of temporal (ir)regularity during encoding on serial order recall performance by using a paradigm similar to the one used by Gorin (2020), while ensuring the use of fully irregular ISIs in the light of the above-mentioned concerns. Participants were asked to recall seven-digit lists in either regular or irregular rhythms. To manipulate rhythm regularity while minimizing the potential for temporal grouping effects, two specific sets of ISIs were designed. In Experiment 1, irregular ISIs deviated only slightly (450-550ms) from the baseline pace(500ms), thereby avoiding relatively long intervals that could induce marking and grouping effects. Importantly, ISIs between all items were unpredictable. In Experiment 2, the same procedure was used, but the irregular ISIs varied on a slightly larger time scale (e.g., 10ms, 206ms, 402ms, 598ms, 794ms, 990ms).

We predicted that the temporal regularity of the presentation of the memoranda during encoding should lead to detectable effects on serial order recall performance, with recall performance deteriorating under the irregular encoding rhythm condition. To assess the impact of temporal manipulations on serial order memory in the most direct manner, we analyzed verbal WM performance across three measures, which, standard serial recall (items recalled in correct serial position), item recall (items recalled irrespective of serial position) and, most critically, order recall (items recalled in correct serial conditioned by overall item recall performance). If temporal codes specifically support serial order memory, the strongest effects of the temporal rhythm manipulation should be observed in order recall performance.

Experiment 1

Experiment 1 manipulated the regularity of ISIs during the presentation of the memoranda by contrasting a regular vs. fully irregular presentation rhythm, and by avoiding excessive contrasts between irregular ISIs that could lead to temporal grouping effects. We used seven-digit lists and introduced irregular rhythms with only small deviations (±50ms) from a regular 500-ms baseline ISI. Two presentation durations were compared: in Experiment 1A, each item was displayed for 500ms, while in Experiment 1B, the display duration was increased to 1000ms, to rule out the possibility of temporal encoding requiring a minimal amount of processing time. This experiment was part of a larger study that also involved manipulating spatial parameters during encoding via a separate task. Participants were subjected to both spatial and temporal manipulations in separate task blocks; the results of the spatial manipulations are reported in a dedicated manuscript.

Experiment 1A

Method

Participants. Fifty-seven participants (36 females, 21 males; mean age = 21.21 years, SD = 2.29) were recruited via social media from the University of Liège community (see *Scoring and Analysis Procedure* for sample size justification). All were native French speakers with no reported history of language, learning, neuropsychological, or neurological disorders. Participants also reported no current drug use (e.g., cannabis) or alcohol abuse. The study was approved by the Ethics Committee of the Faculty of Psychology, Language, and Education Sciences at the University of Liège. Participants provided informed consent before testing and were informed that no financial compensation was offered.

Materials. Thirty-four seven-digit lists were created by pseudo-randomly sampling from the digits 1 to 7, including four practice lists. Lists could not include adjacent digit pairs (e.g., 1-2 or 7-6) or other structured digit combinations (e.g., 1-3-5 or 6-4-2). Additionally, we balanced the distribution of the seven digits across lists and conditions. Half of the lists were presented for the regular condition, and the other half for the irregular condition. The inter-stimulus interval of the regular rhythmic condition was 500ms, while for the irregular rhythmic condition it varied from 450 to 550ms. The ISIs in the irregular condition were generated randomly. Therefore, the pace of each list in the irregular condition was totally unpredictable. The sum of the ISIs was equal across the regular and irregular rhythmic conditions (3000ms). Besides, thirty lists were presented in a random order.

Procedures. The task was programmed using the Gorilla platform (https://gorilla.sc). Stimuli were displayed sequentially at the center of the screen. Each trial began with the presentation of a central fixation cross for 500ms, followed by a 500-ms blank interval. The seven digits were then displayed individually for 500ms each, with an ISI according to the condition.

Following a 500-ms blank screen after the final digit, a recall prompt (“rappel”, which means “recall” in English) appeared at the center of the screen. Participants were instructed to recall the digit sequence aloud in the original presentation order. If they were unable to retrieve a digit at a given position, they were asked to say “oublié” (“forgotten”) to maintain positional integrity in the recall protocol. Participants initiated the next trial by pressing the spacebar once they had completed their response.

Before the experimental phase, participants completed four practice trials with feedback to ensure accurate comprehension of the task instructions. The experimental phase consisted of 30 trials, equally divided between the two presentation direction conditions (15 trials per condition), presented in a randomized order. There was no time constraint for recall. All verbal responses were recorded via the Gorilla website and subsequently transcribed for scoring.

The full experiment lasted approximately 20 minutes.

Scoring and Analysis. Two standard recall accuracy measures were computed: Serial recall score, defined as the number of items correctly recalled in their exact serial positions, and item recall score, defined as the number of items correctly recalled regardless of position. Additionally, an order recall score was calculated as the ratio of the serial recall score to the item recall score, providing a measure of how well participants retained the order of items relative to their overall item recall performance.

To further quantify serial recall performance, a transposition gradient was computed, measuring the distance between the input and output positions of a recalled item. Depending on whether an item was recalled before or after its correct position, a negative or positive value was recorded. For instance, if the sequence “4 6 1 3” were recalled as “4 3 6 7,” the transposition gradients would be “0, -2, 1, NA” indicating the respective positional deviations relative to the input position of the recalled items.

All analyses were conducted using a Bayesian statistical approach (see, e.g., Dienes, 2011; Morey & Rouder, 2011). This approach has the advantage of relying on a model comparison rationale to select and quantify the strength of evidence associated with each model, and crucially, allows for testing the strength of evidence for and against an effect of interest (i.e., positive evidence for the null hypothesis). The Bayesian framework does not involve traditional *p-values*, thereby avoiding multiple testing problems such as alpha inflation (Wagenmakers et al., 2008). All analyses are based on *Bayes factors (BF)*, which can be considered as a relative measure of statistical evidence (Morey et al., 2016). The BF represents the degree to which the observed data update the initial belief in favor of one hypothesis relative to another. The *BF* is the likelihood ratio of a given model, the best-fitting model being the one with the highest *BF*. BF_{01} indicates evidence in favor of the null hypothesis, while BF_{10} indicates evidence in favor of the alternative hypothesis. Although there are no fixed thresholds for *BF* values, we used the following categories for describing strength of evidence: a *BF* of at least 1 is considered to indicate anecdotal evidence, a *BF* of at least 3 is considered to indicate moderate evidence, a *BF* of at least 10 is considered to provide strong evidence, a *BF* of at least 30 is considered to provide very strong evidence, and a *BF* of at least 100 is considered to indicate decisive evidence (Jeffreys & Jeffreys, 1998).

Bayesian generalized linear mixed models were implemented using the R statistical platform and the brms package (Bürkner, 2017). The dependent measures included the recall accuracy scores and error types described above. Fixed effects comprised temporal rhythm (regular vs. irregular), serial position (1 to 7), and their interaction. For the analysis of the transposition gradient, serial position was replaced by transposition distance (-6 to +6). The model included random intercepts for subjects to account for individual variability. To determine the inclusion of random slopes, we first compared two models: one with only the random intercept (subject) and one additionally including a random slope for temporal rhythm. A specific random slope was retained in subsequent models if the *Bayes Factor (BF)* exceeded 3, indicating substantial evidence for its inclusion compared to the null model containing only the intercept. The serial position factor was not included in the analysis of the random effect to reduce the complexity of the models. To better interpret the best model, except for the basic reports on the means and standard errors of the raw data, *estimated* *marginal means (EMMs)* with *credible intervals* on the best models were further reported by the emmeans() function.

Sample size was monitored via a Bayesian sequential sampling approach (Schönbrodt & Wagenmakers, 2018) combined with an effect size stabilization procedure (R. B. Anderson et al., 2022). Beginning from the 20th participant, we computed, for each additional participant, the effect sizes for the effect of interest (temporal rhythm - signed *Cohen's d*) and continued sampling at least until the effect size, for both item and serial order recall scores, stabilized (a priori defined minimal absolute change of effect size <.05 over 5 consecutive analyses; Anderson et al., 2022). After the 25th participant, the effect sizes of both item and serial order recall scores were stable. However, data collection was continued to maximize the statistical sensitivity of our study, given that additional participants were available. Kowialiewski (2024) recently showed that the effect size stabilization procedure does not guarantee the detection of a true effect in the population if the stabilized effect size is small.

To facilitate data visualization and interpretation, recall performance and error rates were displayed using proportional measures. Item recall and serial recall proportions were defined as the number of correctly recalled items, regardless of order (item recall) or in correct serial position (serial recall), divided by the total number of presented items. For the transposition gradients, the proportions were weighted by the number of transposition gradients divided by the sum of transposition errors.

Results

Serial Recall Score. A first Bayesian generalized linear mixed model (GLMM) analysis of serial recall scores revealed that the best-fitting model included only the serial position factor. This model was 5.10 times more likely than the model with the next highest *BF* value, which included both temporal rhythm and serial position factors (Temporal rhythm: $\eta_p^2 = 4.486 \times 10^{-6}$; Serial position: $\eta_p^2 = 0.522$; Temporal rhythm × Serial position: $\eta_p^2 = 0.010$). These results indicate evidence *against* an effect of temporal rhythm ($BF_{01} = 5.31$) or its interaction with the serial position ($BF_{01} = 32.53$) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Item Recall Score. For the item recall score, the best-fitting model included the serial position factor. Although the model was only 1.85 times more likely than the model also including the temporal rhythm factor, the former was retained as the more parsimonious model (Temporal rhythm: $\eta_p^2 = 0.029$; Serial position: $\eta_p^2 = 0.412$; Temporal rhythm × Serial position: $\eta_p^2 = 0.013$). These results provide evidence that tends *against* the alternative hypothesis over the null hypothesis of an effect of the temporal rhythm ($BF_{01} = 2.07$) and evidence strongly *against* the effect of its interaction with serial position ($BF_{01} = 11.38$) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Order Recall Score. For the order recall score, given its derivational nature based on average item and serial recall scores, a repeated measures Bayesian analysis of variance (ANOVA) was conducted on subject-averaged scores instead of a GLMM. Results presented that the best-fitting model included only the serial position factor. Although the model also including the temporal rhythm factor was 1.03 times more likely than the model excluding it, the latter was retained as the more parsimonious model (Temporal rhythm: $\eta_p^2 = 0.032$; Serial position: $\eta_p^2 = 0.372$; Temporal rhythm × Serial position: $\eta_p^2 = 0.020$). These results provide evidence that tends *against* the alternative hypothesis over the null hypothesis of an effect of the temporal rhythm ($BF_{01} = 1.99$) and evidence strongly *against* the effect of its interaction with the serial position ($BF_{01} = 64.16$) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Transposition Gradient. A repeated measures Bayesian ANOVA analysis on the subject-averaged proportions for the different transposition distances revealed that the best-fitting model included only the transposition distance factor. This model was 207.28 times more likely than the following model, including both transposition distance and temporal rhythm factors (Temporal rhythm: $\eta_p^2 = 4.369 \times 10^{-5}$; Transposition distance: $\eta_p^2 = 0.977$; Temporal rhythm × Transposition distance: $\eta_p^2 = 0.020$). These results indicated evidence *against* an effect of the temporal rhythm ($BF_{01} = 33.33$) or its interaction with transposition distance ($BF_{01} = 9.92 \times 10^{19}$) (see Appendix Tables C.1, 2, 4, and Figure 3.1).

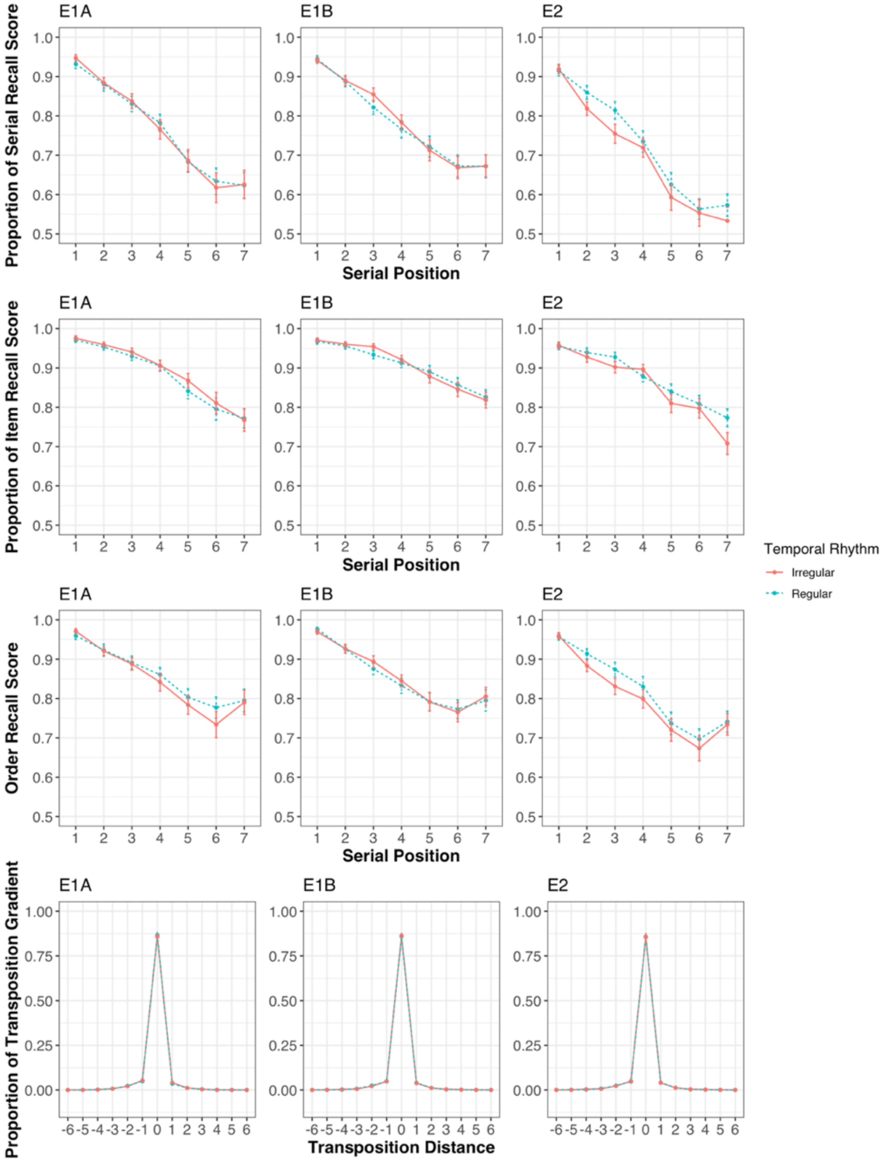

Figure 3.1. Serial recall scores, item recall scores, order recall scores, and transposition gradient, as a function of temporal rhythm and serial position or transposition distance, across experiments. Means with standard errors are depicted.

Discussion

In Experiment 1A, we manipulated the regularity of the ISI by using small discrepancies (<100ms) between ISIs aimed at minimizing potential for the intervention of temporal grouping effects that may hide the negative impact of temporal irregularity. However, results showed no significant impact of the manipulation of presentation rhythm on either order or item recall performance. Our findings are thus congruent with those obtained by Gorin (2020). Experiment 1B used the same setup as Experiment 1A but with a slower average pace (one item every 1000ms instead of 500ms) to rule out the possibility that the overall fast encoding pace prevented the subtle temporal deviations from exerting an effect in the irregular pace condition.

Experiment 1B

Method

Participant. Seventy-seven participants were recruited (see Scoring and Analysis Procedure for justification of sample size) via social networks. Participants were between 18 and 28 years old (average age = 21.87, SD = 2.04); 39 participants were female.

Materials. The materials were as same as in Experiment 1A.

Procedures. The experiment design and procedures were the same as in Experiment 1A, except that the presentation time for each digit was extended to 1000ms.

Scoring and Analysis Procedure. Data scoring and analyses were the same as in Experiment 1A. After the 25th participant, the effect sizes of both item and serial order recall scores were stable. As in Experiment 1A, data collection was continued to maximize the statistical sensitivity of our study as additional participants were available.

Results

Serial Recall Score. A first Bayesian GLMM analysis of serial recall scores revealed that the best-fitting model included only the serial position factor. This model was 5.29 times more likely than the model with the next highest *BF* value, which included both temporal rhythm and serial position factors (Temporal rhythm: η_p^2 = 0.005; Serial position: η_p^2 = 0.507; Temporal rhythm × Serial position: η_p^2 = 0.018). These results indicated evidence *against* an effect of temporal rhythm (BF_{01} = 5.83) or its interaction with the serial position (BF_{01} = 24.61) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Item Recall Score. For the item recall score, the best-fitting model included the serial position factor. Although this model was only 2.54 times more likely than the model also including the temporal rhythm factor, the former was retained as the more parsimonious model (Temporal rhythm: η_p^2 = 3.781×10^{-4}; Serial position: η_p^2 = 0.370; Temporal rhythm × Serial position: η_p^2 = 0.016). These results provide evidence that tends *against* the alternative hypothesis over the null hypothesis of an effect of temporal rhythm (BF_{01} = 2.70) or its interaction with the serial position (BF_{01} = 1.13) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Order Recall Score. For the order recall score, a repeated measures Bayesian ANOVA analysis presented that the best-fitting model included only the serial position factor. This model was 5.93 times more likely than the following model, which included both temporal rhythm and serial position factors (Temporal rhythm: η_p^2 = 0.004; Serial position: η_p^2 = 0.385; Temporal rhythm × Serial position: η_p^2 = 0.006). These results indicate evidence *against* an effect of temporal rhythm (BF_{01} = 5.43) or its interaction with the serial position (BF_{01} = 31.50) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Transposition Gradient. A repeated measures Bayesian ANOVA analysis revealed that the best-fitting model included only the transposition distance factor. This model was 238.50 times more likely than the following model, including both transposition distance and temporal rhythm factors (Temporal rhythm: η_p^2 = 2.833×10^{-4}; Transposition distance: η_p^2 = 0.977; Temporal rhythm × Transposition distance: η_p^2 = 0.007). These results indicated evidence *against* an effect of the temporal rhythm (BF_{01} = 38.11) or its interaction with transposition distance (BF_{01} = 1.14×10^{21}) (see Appendix Tables C.1, 2, 4, and Figure 3.1).

Discussion

Experiment 1B did not reveal any difference in either order or item recall performance between regular and irregular presentation rhythms, even when the average presentation pace was slowed down. The results of Experiments 1A and 1B are consistent with the findings of Gorin (2020), who concluded that item presentation rhythm does not influence serial recall performance. One could argue that these results provide evidence for accounts of serial order memory that involve time-independent processes to encode serial order information (e.g., Lewandowsky & Farrell, 2008), or relative rather than absolute temporal signals (Henson, 1998). Another possibility, however, is that the small differences between the ISIs in this experiment were too small to perturb the operation of a temporal encoder such as the temporal oscillators assumed in temporal models of serial order (G. D. A. Brown et al., 2000; Hartley et al., 2016).

In Experiment 2, we therefore used larger variations of the ISIs while keeping overall shorter ISIs as compared to Gorin (2020).

Experiment 2

Experiment 2 used the same task setup as Experiment 1B, but the ISIs varied on a slightly larger scale (e.g., 10ms, 206ms, 402ms, 598ms, 794ms, 990ms), leading to a stronger discrepancy between the regularity of a hypothetical temporal signal and the irregularity of the ISIs.

Method

Participant. Forty-seven participants were recruited (see Scoring and Analysis Procedure for justification of sample size) via social networks. Participants were between 18 and 28 years old (average age = 22.18, SD = 2.91); 38 participants were female.

Materials. The materials were as same as in Experiment 1B.

Procedures. The design was identical to the one used in Experiment 1B, with the notable difference in the temporal rhythm manipulation. More specifically, the ISI of the regular condition was kept constant at 500ms, while the differences between the ISIs for the irregular condition were increased. To increase the variation of the ISIs while ensuring that the sum of ISIs remains the same across the regular and irregular conditions (i.e., 3000ms), we used ISIs ranging between 10ms and 990ms, with intermediate ISIs of 206ms, 402ms, 598ms, and 794ms. Each item was presented for 1000ms as in Experiment 1B.

Scoring and Analysis Procedure. Scoring and analysis procedures were the same as in Experiments 1A and 1B.

Results

Serial Recall Score. A first Bayesian GLMM analysis of serial recall scores revealed that the best-fitting model included serial position. Although the model also including the temporal rhythm factor was 1.88 times more likely than the model excluding it, the latter was retained as the more parsimonious model (Temporal rhythm: η_p^2 = 0.085; Serial position: η_p^2 = 0.686; Temporal rhythm × Serial position: η_p^2 = 0.032). These results provide evidence that tends *against* the alternative hypothesis over the null hypothesis of an effect of the temporal rhythm (BF_{01} = 0.55) and evidence strongly *against* the effect of its interaction with serial position (BF_{01} = 6.48) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Item Recall Score. For the item recall score, the best-fitting model included the temporal rhythm and serial position factors. Although the model also including their interaction was 1.96 times more likely than the model without it, the latter was retained as the more parsimonious model (Temporal rhythm: η_p^2 = 0.089; Serial position: η_p^2 = 0.504; Temporal rhythm × Serial position: η_p^2 = 0.061). These results provide evidence that tends *against* an interaction of temporal rhythm and serial position (BF_{01} = 0.51). In contrast, evidence in favour of an advantage for the regular vs. irregular temporal rhythm condition was observed (β = 0.167, *95%CI* [0.045, 0.293]; *estimated marginal means* = 0.918, *95%CI* [0.893, 0.940] vs. 0.905, *95%CI* [0.877, 0.930]; see Appendix Tables C.1, 2, 3, and Figures 3.1 and 3.2). Note, however, that the effect remained small and that the credible interval of the estimated marginal means for both conditions showed a non-negligible overlap.

Order Recall Score. For the order recall score, a repeated measures Bayesian ANOVA analysis showed the best-fitting model included only the serial position factor. Although the model including serial position and temporal rhythm factors was 1.27 times more likely than the model including only serial position, the latter was retained as the more parsimonious model (Temporal rhythm: η_p^2 = 0.063; Serial position: η_p^2 = 0.557; Temporal rhythm × Serial position: η_p^2 = 0.013). These results provide evidence tending *against* an effect of temporal rhythm (BF_{01} = 2.60) and evidence strongly *against* the effect of its interaction with the serial position (BF_{01} = 48.40) (see Appendix Tables C.1, 2, 3, and Figure 3.1).

Transposition Gradient. A repeated measures Bayesian ANOVA analysis revealed that the best-fitting model included only the transposition distance factor. This model was 181.14 times more likely than the following model, including both transposition distance and temporal rhythm factors (Temporal rhythm: η_p^2 = -2.590 × 10^{-4}; Transposition distance: η_p^2 = 0.976; Temporal rhythm × Transposition distance: η_p^2 = 0.050). These results indicated evidence *against* an effect of the temporal rhythm (BF_{01} = 30.41) or its interaction with transposition distance (BF_{01} = 6.16 × 10^{19}) (see Appendix Tables C.1, 2, 4, and Figure 3.1).

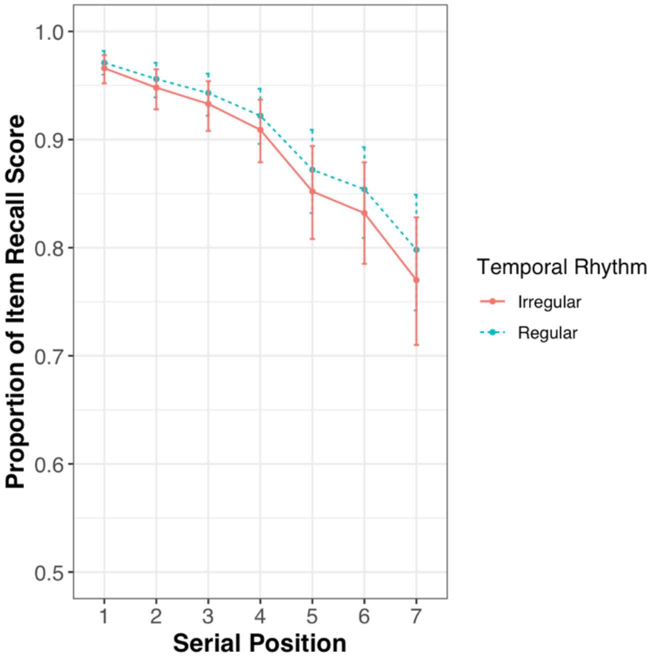

Figure 3.2. Item recall scores, as a function of temporal rhythm and serial position in Experiment 2. Estimated marginal means (EMMs) with credible intervals (95%CI) are depicted.

Discussion

By using more widely spaced ISIs as in Experiments 1A and 1B, Experiment 2 showed an effect of temporal rhythm manipulation, albeit a small one, and only for the item recall measure.

General Discussion

The present study investigated the role of temporal coding in the representation of serial order information in verbal WM. By directly manipulating the temporal rhythm of the presentation rate of stimuli during WM encoding, we observed impaired item recall performance when lists were presented with a noticeable irregular rhythm, while no such effect was observed when the rhythm varied more subtly. On the other hand, no impact of temporal rhythm manipulations was observed for the order recall measure in any of the experiments. Furthermore, the impact of the temporal rhythm manipulation on item recall that was observed in Experiment 2 remained small.

First of all, the absence of an impact of temporal regularity on serial order memory in two of our experiments is in line with Gorin (2020). Gorin (2020) consistently reported no differences between regular and irregular presentation rhythms for different serial recall measures, including recall latencies. Our study, by using more constrained irregular ISI parameters, further confirms that these null effects cannot be (only) attributed to confounds such as temporal grouping effects that may have occurred due to the particular design of the irregular ISIs used in the study by Gorin (2020). We started to observe an impact of temporal rhythm manipulations only when using slightly larger ISI variations for the irregular temporal rhythm condition, and this effect remained small and confined to the item recall measure.

Although our results are not surprising in light of the previous studies by Gorin (2020), they are more difficult to reconcile with time-based models of serial order WM and other empirical results. As already noted, time-based computational models of serial order WM hypothesize that serial order would be retained via associations between the to-be-memorized items and a context signal that changes progressively during the presentation of a sequence. At recall, successive states of the temporal context signal are replayed, and the learned associations reactivate the items (Brown et al., 2000; Hartley et al., 2016). More specifically and recently, Hartley et al. (2016) proposed the BUMP model, which suggests that serial order is encoded by a population of oscillators driven in a bottom-up manner by the structure of auditory-verbal input and sensitive to local variations in its temporal structure. In light of this theoretical account, we can predict changes in serial recall pattern for memoranda presented at a regular vs. irregular rhythm, given that in the latter case, the irregular bottom-up temporal information will have difficulties in engaging one or several oscillators in a coherent manner during encoding of the WM sequence. A similar prediction had also been proposed earlier in the context of the Dynamic Attending Theory (Jones & Boltz, 1989; Large & Jones, 1999). This theory considers that temporal stimulus regularities, either of the memoranda or of an associated external stimulus stream, can entrain internal oscillators that guide attention over time and induce temporal expectations about future events, thus facilitating sequencing and event processing for temporally regular stimulus sequences. These predictions are in line with a range of empirical studies that have shown that the regularity of an external stimulus associated with a list of memoranda (during encoding or maintenance) can have an impact on serial recall performance (Attout et al., 2025; Fanuel et al., 2018; Hall & Gathercole, 2011; R. Henson et al., 2003; Parmentier & Beaman, 2015; Plancher et al., 2018).

This thus leads us to the critical question of why an effect of temporal rhythm can be observed (although in a sometimes-inconsistent manner; see Introduction section) in WM tasks when the temporal regularity is manipulated via an external stream of stimuli, but much less when the temporal rhythm is directly manipulated via the temporal regularity of the presentation of memoranda during encoding. One possible explanation may rely on the strength of temporal interference created by the irregular rhythm. This strength may be more pronounced for external temporal streams as an additional signal is presented, leading to competition between an oscillator following the regular temporal rhythm of the memory sequence and a set of oscillators trying to capture the temporal structure of the temporally irregular external stream. In the case of the present study, only oscillators trying to capture the temporal structure of the temporally irregular memory sequence would be engaged, and they may have no measurable difficulty in encoding the memory sequence when the temporal irregularity is rather subtle. In line with this reasoning, Hartley et al. (2016) proposed a distinction between oscillators driven in a bottom-up vs. a top-down manner. When the bottom-up temporal structure provides a weakly irregular signal, temporal encoding of items may be supported by an internal top-down temporal oscillator base. However, when the external signal becomes too irregular and marked, then the internally engaged top-down oscillator cannot provide any effective support anymore to temporal item encoding. A related factor that may have interacted with the manipulation of temporal irregularity of the memory sequence in the present study is the fact that the temporal regularity conditions were presented in a fully randomized manner, i.e., temporally regular and irregular sequence conditions changed randomly from one trial to another (also like in Gorin, 2020). Although this specific experimental setup was intended to allow for the testing of a strong temporal coding account, by hypothesized that item encoding is highly sensitive to temporal parameters in a bottom-up, non-strategic, and automatic manner. However, time processing is also associated with top-down temporal prediction, as mentioned above. By presenting the different temporal regularity conditions in a fully random manner, top-down temporal prediction may be overall diminished even when a specific trial is presented in a perfectly regular manner, leading to a reduced temporal encoding advantage of the regular vs. irregular memory sequences. Studies in the time processing literature have also shown that temporally irregular sequences can influence the encoding of subsequent regular trials when presented in close succession (Beaudry et al., 2014; Keppel & Underwood, 1962). In the WM domain, most studies that observed an impact of temporal regularity manipulations used indeed designs in which the temporal regularity conditions were presented in a blocked manner (Attout et al., 2025; Fanuel et al., 2018; Plancher et al., 2018).

Several limitations should be acknowledged. First, as we mentioned above, by presenting regular and irregular rhythm list conditions in a random rather than blocked succession could have impacted the temporal expectations of the participants, limiting the entrainment of the temporal oscillator also in the regular list conditions. Future studies need to replicate the present study, but by presenting the different temporal rhythm conditions in separate blocks. Second, our reliance on visually presented digits may underestimate the potential role of temporal regularity compared with auditory materials, as auditory materials have been proven to be more sensitive to temporal manipulations than visual ones (Attout et al., 2024; Henson, 1996). Alternatively, individual differences may play a role.

In conclusion, the present study does not provide evidence for a strong temporal encoding account of serial order, in which the temporal parameters of the memory sequence drive serial order encoding in a purely external, bottom-up driven manner. They remain, however, compatible with a temporal encoding account that involves the intervention of both bottom-up and internally, top-down driven temporal encoding structures, leading to a detriment of WM performance when the internal temporal encoding structures are superseded by markedly irregular externally driven temporal signals.

Appendix

Table C.1. The results of Bayesian analysis on random slopes. The factor was included in the model as a random slope if the Bayes factor comparing the null model was above 3.

	Factor	Serial recall score	Item recall score	Order recall score	Transposition gradient
E1A	Temporal rhythm	284.86	0.24	0.08	0.03
E1B	Temporal rhythm	877.25	8.46	0.07	0.03
E2	Temporal rhythm	2.59×10^4	1.01	0.04	0.03

Table C.2. Results of serial recall score, item recall score, order recall score, and transposition gradient in three experiments

	Model	Serial recall	Item recall	Order recall	Transposition gradient[*]
E1A	Null	1	1	1	0
	Temporal rhythm	0.19	0.48	0.50	0
	Serial position	2.31×10^{233}	5.14×10^{146}	4.04×10^{33}	1
	Temporal rhythm + Serial position	4.74×10^{232}	2.82×10^{146}	4.09×10^{33}	4.82×10^{-3}
	Serial position + Temporal rhythm × Serial position	/	/	/	4.78×10^{-23}
	Serial position + Temporal rhythm + Temporal rhythm × Serial position	1.45×10^{231}	2.54×10^{145}	6.36×10^{31}	5.06×10^{-23}
E1B	Null	1	1	1	0
	Temporal rhythm	0.17	0.37	0.18	0
	Serial position	1.98×10^{255}	4.43×10^{118}	8.46×10^{53}	1
	Temporal rhythm + Serial position	3.82×10^{254}	1.77×10^{118}	1.46×10^{53}	4.19×10^{-3}
	Serial position + Temporal rhythm × Serial position	/	/	/	3.58×10^{-24}
	Serial position + Temporal rhythm + Temporal rhythm × Serial position	1.51×10^{253}	1.57×10^{118}	4.52×10^{51}	3.53×10^{-24}
E2	Null	1	1	1	0
	Temporal rhythm	1.81	4.18	0.38	0
	Serial position	1.42×10^{213}	2.60×10^{100}	1.80×10^{49}	1
	Temporal rhythm + Serial position	2.64×10^{213}	1.41×10^{101}	2.31×10^{49}	5.52×10^{-3}
	Serial position + Temporal rhythm × Serial position	/	/	/	9.12×10^{-23}
	Serial position + Temporal rhythm + Temporal rhythm × Serial position	4.09×10^{212}	2.81×10^{101}	4.68×10^{47}	8.90×10^{-23}

[*] For "Transposition gradient", "Serial position" means "Transposition distance".

Table C.3. Statistical description of recall scores and errors in all the experiments.

		Temporal rhythm	Serial position							N
			1	2	3	4	5	6	7	
Serial recall score	E1A	Regular	0.932	0.881	0.830	0.782	0.683	0.634	0.623	57
			(0.087)	(0.130)	(0.148)	(0.163)	(0.200)	(0.255)	(0.247)	
		Irregular	0.947	0.884	0.837	0.766	0.687	0.618	0.626	57
			(0.064)	(0.101)	(0.140)	(0.188)	(0.211)	(0.286)	(0.272)	
	E1B	Regular	0.945	0.887	0.822	0.766	0.722	0.673	0.672	77
			(0.072)	(0.117)	(0.157)	(0.195)	(0.233)	(0.245)	(0.258)	
		Irregular	0.941	0.890	0.855	0.784	0.713	0.668	0.673	77
			(0.067)	(0.110)	(0.147)	(0.167)	(0.235)	(0.248)	(0.248)	
	E2	Regular	0.915	0.860	0.814	0.735	0.626	0.563	0.573	47
			(0.090)	(0.119)	(0.152)	(0.187)	(0.208)	(0.177)	(0.195)	
		Irregular	0.921	0.818	0.755	0.719	0.593	0.553	0.535	47
			(0.082)	(0.119)	(0.165)	(0.165)	(0.224)	(0.232)	(0.233)	
Item recall score	E1A	Regular	0.971	0.953	0.930	0.905	0.841	0.795	0.772	57
			(0.047)	(0.062)	(0.082)	(0.105)	(0.147)	(0.213)	(0.186)	
		Irregular	0.975	0.959	0.940	0.906	0.868	0.811	0.767	57
			(0.045)	(0.052)	(0.075)	(0.101)	(0.137)	(0.207)	(0.214)	
	E1B	Regular	0.967	0.956	0.933	0.913	0.890	0.857	0.826	77
			(0.057)	(0.068)	(0.087)	(0.102)	(0.141)	(0.158)	(0.162)	
		Irregular	0.971	0.960	0.954	0.921	0.879	0.846	0.818	77
			(0.049)	(0.057)	(0.064)	(0.092)	(0.149)	(0.163)	(0.178)	
	E2	Regular	0.955	0.939	0.928	0.878	0.840	0.809	0.773	47
			(0.056)	(0.081)	(0.082)	(0.090)	(0.135)	(0.148)	(0.156)	
		Irregular	0.957	0.928	0.902	0.896	0.810	0.797	0.708	47
			(0.060)	(0.090)	(0.101)	(0.087)	(0.162)	(0.170)	(0.190)	
Order recall score	E1A	Regular	0.960	0.923	0.891	0.861	0.803	0.777	0.795	57
			(0.072)	(0.115)	(0.126)	(0.133)	(0.163)	(0.200)	(0.218)	
		Irregular	0.971	0.921	0.888	0.842	0.784	0.734	0.790	57
			(0.050)	(0.084)	(0.118)	(0.171)	(0.180)	(0.250)	(0.235)	

E1B	Regular	0.977	0.926	0.875	0.833	0.792	0.773	0.795	77
		(0.043)	(0.081)	(0.123)	(0.168)	(0.208)	(0.209)	(0.235)	
	Irregular	0.970	0.927	0.894	0.845	0.792	0.765	0.806	77
		(0.052)	(0.098)	(0.132)	(0.133)	(0.203)	(0.214)	(0.199)	
E2	Regular	0.957	0.914	0.874	0.831	0.737	0.697	0.741	47
		(0.059)	(0.086)	(0.124)	(0.176)	(0.197)	(0.177)	(0.184)	
	Irregular	0.961	0.883	0.831	0.799	0.720	0.673	0.737	47
		(0.054)	(0.102)	(0.141)	(0.159)	(0.194)	(0.216)	(0.189)	

Table C.4. Statistical description of the proportion of transposition gradient

	Temporal rhythm	Gradient													N
		-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	
E1A	Regular	0.000 (0.000)	0.000 (0.002)	0.002 (0.006)	0.007 (0.011)	0.023 (0.029)	0.049 (0.044)	0.866 (0.106)	0.035 (0.035)	0.011 (0.013)	0.004 (0.007)	0.001 (0.003)	0.000 (0.003)	0.000 (0.000)	57
	Irregular	0.000 (0.002)	0.000 (0.002)	0.002 (0.005)	0.007 (0.011)	0.021 (0.026)	0.054 (0.056)	0.858 (0.117)	0.041 (0.038)	0.012 (0.016)	0.004 (0.007)	0.001 (0.002)	0.000 (0.002)	0.000 (0.002)	57
E1B	Regular	0.001 (0.002)	0.001 (0.003)	0.002 (0.004)	0.007 (0.011)	0.025 (0.025)	0.049 (0.049)	0.859 (0.118)	0.040 (0.039)	0.012 (0.014)	0.003 (0.006)	0.001 (0.004)	0.000 (0.002)	0.000 (0.002)	77
	Irregular	0.000 (0.001)	0.001 (0.003)	0.003 (0.007)	0.005 (0.009)	0.021 (0.024)	0.048 (0.047)	0.864 (0.112)	0.038 (0.032)	0.011 (0.016)	0.004 (0.007)	0.002 (0.005)	0.001 (0.004)	0.000 (0.000)	77
E2	Regular	0.000 (0.002)	0.001 (0.004)	0.001 (0.005)	0.012 (0.017)	0.023 (0.028)	0.066 (0.052)	0.832 (0.114)	0.039 (0.035)	0.017 (0.018)	0.007 (0.010)	0.001 (0.004)	0.000 (0.002)	0.000 (0.000)	47
	Irregular	0.000 (0.000)	0.001 (0.003)	0.004 (0.009)	0.012 (0.016)	0.031 (0.028)	0.064 (0.043)	0.813 (0.111)	0.048 (0.038)	0.016 (0.014)	0.007 (0.009)	0.003 (0.009)	0.001 (0.003)	0.000 (0.002)	47

General Discussion

Discussion

This doctoral thesis examined how linguistic, spatial, and temporal contextual signals shape serial order coding in verbal WM. The General Discussion synthesizes the three empirical studies, evaluates their theoretical implications, integrates them within a hybrid, context-sensitive framework, and outlines methodological limitations and future directions.

PART I OVERVIEW OF THE EMPIRICAL FINDINGS

Across three complementary studies, the thesis investigated how distinct contextual dimensions – syntactic structure, spatial orientation, and temporal rhythm – affect the encoding and retrieval of serial order information in verbal WM.

Study 1: Syntactic Knowledge and Serial Order Recall

Study 1 examined whether syntactic knowledge influences serial order when participants encode adjective-noun word pairs that vary in syntactic legality. Three experiments were conducted, two in French and one in German, allowing examination across languages with different constraints on adjective-noun ordering. A central objective was to distinguish effects on item recall from effects on serial order recall, given their known differential sensitivity to linguistic structure (e.g., Gathercole et al., 2001; Hulme et al., 1991, 1997; Majerus, 2013; Saint-Aubin & Poirier, 1999). Across experiments, syntactic legality consistently exerted a robust effect on serial order recall. Participants recalled the correct order of words more accurately when adjective-noun pairs followed syntactically legal structures than when they were syntactically illegal. This pattern was reflected in both order-accuracy scores and migration errors, with illegal lists showing higher rates of positional displacements. These effects emerged in all three experiments despite variations in stimulus materials and language. At the level of item recall, the pattern was less uniform. In the first French experiment and in the German experiment, syntactically legal lists yielded higher item recall scores than syntactically illegal lists. In contrast, in the second French experiment, syntactic legality did not significantly affect item recall, although the effect on serial order was replicated. This suggests that item-level benefits may be more sensitive to language-specific structural constraints, whereas the impact of syntactic legality on order memory is more robust.

Taken together, the empirical findings of **Study 1** demonstrate that syntactic structure reliably affects serial order recall across languages with fixed (German) and partially flexible (French) adjective-noun ordering. They further show that syntactic legality influences both item and order memory, albeit with more consistent effects on the latter. These observations provide a strong empirical basis for examining how long-term linguistic knowledge contributes to the organization of serial order in verbal WM.

Study 2: Serial Manipulation and Conditions for Spatialization

Study 2 investigated whether spatial manipulations influence serial order recall and under which conditions spatial frames become functionally relevant during encoding and retrieval. Across three experiments, the direction of stimulus presentation (left-to-right vs. right-to-left) and the direction required at the retrieval were systematically manipulated. Spatial cues were embedded either at encoding, at retrieval, or at both stages, allowing assessment of how stable versus fragmented spatial frames shape recall performance. Across the full set of experiments, spatial influences on overall recall performance emerged only when spatial cues were made salient and consistent at both encoding and retrieval. In the final experiment, recall scores were higher for left-to-right presentation than for right-to-left presentation, replicating patterns reported by Fischer and Benjamin (2014). This directional advantage was most pronounced at later serial positions. In addition, performance improved when presentation and recall directions were congruent, suggesting that alignment of spatial frames across task phases facilitated memory performance. By contrast, when spatial cues were present at only one stage – either encoding or retrieval – no reliable spatial effects were observed. In these conditions, the presence of a unidirectional spatial cue was insufficient to elicit robust changes in recall performance. Critically, across all experiments, spatial manipulations did not significantly influence serial order-specific measures, including order-accuracy scores and transposition gradients. No evidence was found that spatial manipulation enhanced positional recall for any spatial condition. Moreover, the predicted interaction between encoding and retrieval directions, which would indicate spatialization tuned to the directional structure encountered during encoding (Guida, Abrahamse, et al., 2020), was not observed. Additional analyses considered the potential moderating influence of list length, the type of memoranda (digits), and task demands. Spatial effects were numerically small and inconsistent when list length or task load was relatively high. The use of digits may also have introduced interference from long-term spatial-numerical associations (e.g., SNARC-like effects; Ginsburg & Gevers, 2015; Huber et al., 2016). As highlighted in the recent meta-analysis by Guida et al. (2026), the nature and size of the memoranda, as well as the specific task conditions, are key moderators of SPoARC (Spatial-Ordinal Association in Response Codes) effects.

Taken together, the empirical findings from **Study 2** indicate that spatial manipulation can influence overall recall performance, but only under conditions where spatial cues are salient, stable, and present at both encoding and retrieval. The absence of effects on serial order-specific measures suggests that, under the present conditions, spatial cues did not substantially contribute to the encoding of serial order itself.

Study 3: Temporal Regularity and Its Limited Influence on Serial Order

Study 3 assessed whether the temporal rhythm of stimulus presentation contributes to serial order memory, as predicted by time-based models of serial recall (e.g., Brown et al., 2000; Hartley et al., 2016). Across three experiments, sequences of visually presented digits were encoded under conditions of regular (isochronous) or irregular ISIs. Irregularity was parametrically manipulated to distinguish small versus large deviations from regularity. Across experiments, temporal rhythm affected item recall only when temporal irregularity was pronounced. Under strongly irregular ISIs, participants recalled fewer items overall compared to in regular sequences, indicating that highly irregular timing can impair general maintenance processes. By contrast, subtle irregularity did not significantly impair item recall, suggesting that temporal manipulations must reach a salience threshold before influencing performance. Crucially, no effects of temporal regularity on serial order recall were observed. Neither order-accuracy scores nor transposition gradients differed significantly between regular and irregular ISI conditions in any experiment, despite substantial variation in irregular magnitude across studies. These findings replicate the null effects reported by Gorin (2020), who similarly found that temporal irregularity did not influence order reconstruction or recall latencies. Exploratory analyses indicated that removing the random intercept for participants from the mixed-effects models yielded a significant effect of rhythm on serial order scores, suggesting inter-individual variability in sensitivity to temporal structure. The experimental design, however, did not include measures specifically targeting such variability.

Overall, the empirical findings from **Study 3** indicate that temporal rhythm influenced item memory under strongly irregular conditions, but did not reliably affect serial order memory. These results provide an empirical basis for discussing the conditions under which temporal information can influence verbal WM and for evaluating theoretical models that treat temporal context as a primary or obligatory code for serial order.

PART II THEORETICAL IMPLICATIONS

Serial order in verbal WM has long been assumed to rely on multiple representational codes – phonological, spatial, temporal, and linguistic. The findings from the present thesis clarify the conditions under which each type of code becomes relevant. Although the empirical results reveal variability across linguistic, spatial, and temporal manipulations, they converge on a broader theoretical principle: serial order in verbal WM is supported by an adaptive interplay between long-term linguistic knowledge and short-term contextual cues.

The Role of Syntactic Knowledge in Serial Order Coding

The findings of **Study 1** speak directly to long-standing debates regarding the relationship between linguistic knowledge and verbal WM. Classical distinctions between item memory and order memory have often been interpreted as reflecting separable representational systems, with item recall grounded in lexico-semantic structure and serial order contributed to more domain-general positional mechanisms (e.g., Brown et al., 2000; Burgess & Hitch, 2006; Hartley et al., 2016; Majerus, 2008, 2013; van Dijck & Fias, 2011). The consistent syntactic legality effects observed in **Study 1** challenge such a separation by demonstrating that syntactic structure –

traditionally conceived as part of the long-term language system – exerts a reliable influence not only on item-level representation but also, and more robustly, on serial order.

This pattern is theoretically significant in several respects. First, it extends evidence suggesting that linguistic knowledge plays a broader role in WM than often assumed. Much of the modern literature on semantic effects in WM emphasizes interference: semantic relatedness increases item migrations and reduces order accuracy (e.g., Poirier et al., 2015; Saint-Aubin et al., 2005, 2023; Tse, 2009; Tse et al., 2011). By contrast, **Study 1** revealed a positive, facilitatory influence of syntactic knowledge, indicating that the language system can provide a structural scaffold for order maintenance. This pattern converges with recent findings of conditional semantic effects benefits for serial order (Kowialiewski et al., 2024), suggesting that linguistic knowledge does not uniformly interfere with short-term ordering but can also constrain and support it.

Second, **Study 1** refines long-term memory-based models of verbal WM. Accounts such as ANet (Poirier et al., 2015), holographic memory models (Franklin & Mewhort, 2015; Mewhort et al., 2018), and the recent Embedded Computational Framework of Memory (eCFM; Guitard et al., 2025) emphasize the role of lexical-semantic similarity in shaping WM dynamics. The present findings broaden these frameworks by showing that syntactic structure operates as an additional knowledge-based constraint. Table 1.1 formalizes this view by distinguishing how syntactic information contributes to long-term order templates that shape recall, particularly when temporary structural cues are minimized.

Table 1.1. Updated levels of item and serial order information encoded in verbal WM tasks.

Item level	Long-term serial order level	Short-term serial order level
Word strings	Serial position of words with syntactic legality	Serial position of words in a list
Nonword strings	Serial position of phonemes within words	Serial position of nonword strings in a list
	Serial position of phonemes for nonwords with highly familiar phoneme co-occurrences	Serial position of phonemes for nonwords with less familiar or novel phoneme co-occurrences

Third, the results support emergent accounts of verbal WM in which the WM system is conceptualized as an activation-based extension of the language system, rather than a modular buffer with dedicated storage for sequential information (Acheson & MacDonald, 2009; Buchsbaum & D'Esposito, 2019; Cowan, 1993; Hasson et al., 2015; MacDonald, 2016; Majerus, 2013, 2018; Majerus et al., 2009, 2010; Postle, 2006; Schwering & MacDonald, 2020). Under these models, syntactic structure is part of the activated long-term representational substrate that shapes short-term performance. The cross-linguistic syntactic legality effects observed in **Study 1** imply that participants recruit long-term syntactic regularities to guide sequential organization even when list content is minimal and structurally sparse.

At the same time, **Study 1** clarifies the relationship between long-term syntactic knowledge and temporary syntactic regularities. Previous studies using uniform adjective-noun structures (Querella & Majerus, 2024; Schweppe et al., 2022) may have created short-lived compositional frames that participants relied on at recall, thereby overshadowing long-term syntactic effects. By disrupting list-internal predictability, **Study 1** revealed long-term syntactic contributions that had previously been obscured. This result fits within temporal-context models (Burgess & Hitch, 1999, 2006a; Lewandowsky & Farrell, 2008a; Oberauer et al., 2012), which posit that contextual signals during encoding can reflect both enduring knowledge and emergent regularities. When temporary regularities are minimized, long-term syntactic constraints become more apparent.

Finally, the results refine the A-O-WM model (Majerus, 2008, 2013, 2018). While this model proposes interactions between activated linguistic representations, order representations, and attentional mechanisms, **Study 1** suggests that syntactic structure should be incorporated explicitly as part of the activated linguistic system contributing to order reconstruction. Syntactic knowledge appears capable of guiding attentional selection toward structurally plausible sequences and stabilizing internal order representations.

In summary, **Study 1** shows that serial order coding draws on long-term syntactic knowledge, that syntactic influences become visible when temporary structural cues do not overshadow long-term constraints, and that models of verbal WM must integrate syntactic representations as core contributors to serial ordering processes.

The Role of Spatial Manipulation in Serial Order Recall

The findings of **Study 2** clarify when and how spatial representations contribute to verbal WM. Theoretical models of the SPoARC effect propose that serial order may be mapped onto a spatial axis, typically left-to-right for Western readers (or "ordinal position effect"; Ginsburg et al., 2014). The mental whiteboard hypothesis (Abrahamse et al., 2014) suggests that sequential information is spontaneously projected onto visuospatial coordinates, whereas complementary accounts such as the expertise account and principle of economy (Guida, Abrahamse, et al., 2020; Guida & Campitelli, 2019) emphasize the interplay between long-term spatial expertise (e.g., reading direction) and task-specific contextual cues.

Study 2 challenges strong versions of these theories by showing that spatialization is neither obligatory nor reliably implicated in serial order under typical serial recall conditions. Instead, spatial coding required a sufficiently stable and behaviorally relevant spatial frame. When spatial cues were available at both encoding and retrieval, recall performance improved – particularly for canonical left-to-right sequences – consistent with the role of long-term reading-direction expertise. However, these effects did not extend to serial order-specific measures, such as transposition gradients or positional accuracy.

The results, therefore, suggest that spatial coding may operate as an auxiliary rather than a core representational dimension. When multiple codes are available (phonological, syntactic, spatial, temporal), spatialization seems to be recruited only when spatial cues align across task phases and provide consistent retrieval scaffolding. This is in line with the principle of economy, according to which WM preferentially draws on the representational resources that are most accessible and efficient given the task constraints.

In addition, competition from long-term number-space association is likely to have attenuated SPoARC effects in **Study 2**. Digit stimuli strongly activate spatial-numerical associations (e.g., SNARC; Ginsburg & Gevers, 2015), which may compete with spatialization driven by sequential structure. This interpretation is compatible with the meta-analytic results of (Guida et al., 2026), which identify the nature of the memoranda as a key moderator of spatialization.

Cognitive load further constrains the emergence of spatialization. Seven-item lists approach or exceed the focus-of-attention capacity described in embedded-processes accounts (Cowan, 1999). Under such high load, participants may prioritize phonological maintenance over spatial mapping. This is consistent with studies showing that spatialization declines as list length increases (Huber et al., 2016; Vivion et al., 2025) and disappears for very long lists (Guida, Carnet, et al., 2018).

Finally, the results suggest that spatial codes require coordinated engagement at both encoding and retrieval. Spatialization does not appear to be automatically reinstated unless retrieval demands explicitly re-engage spatial frames, which challenges the assumption that mapping sequences onto space is an automatic and ubiquitous process in serial recall tasks.

In summary, **Study 2** indicates that spatial coding in verbal WM is flexible, contingent on task demands, and constrained by long-term spatial knowledge and competing representations. Spatialization operates as a conditional, context-dependent mechanism rather than as a universal basis of serial order.

The Role of Temporal Context in Serial Order Recall

Study 3 provides a stringent test of temporal-context accounts of serial order. Internal oscillator models, such as OSCAR (Brown et al., 2000), propose that item–context associations synchronize with internally generated temporal cycles, whereas models such as BUMP (Hartley et al., 2016) assume that temporal context is derived from the timing structure of the input. Both approaches predict performance differences between regular and irregular sequences.

The empirical results from **Study 3** reveal a more restricted impact of temporal structure. Clear rhythm effects appeared at the item level only under strongly irregular ISIs, while serial order remained unaffected. Temporal regularity thus did not automatically feed into order representations, at least under visual presentation and randomized trial sequencing.

These findings suggest that temporal coding may not constitute a primary dimension of serial order in verbal WM. Unlike syntactic structure, temporal rhythm did not shape positional accuracy or transposition patterns. Temporal context, therefore, appears not to be obligatorily bound to order representation. Instead, temporal information may influence performance primarily through attentional allocation, for example, by modulating arousal, rhythmic expectation, or resource distribution.

The results also highlight the importance of the interaction between internal and external oscillators. Dynamic Attending Theory (DAT; Jones & Boltz, 1989; Large & Jones, 1999) proposes that external rhythms entrain internal oscillators and form temporal predictions. In **Study 3**, however, entrainment opportunities were limited by trial-by-trial randomization of regular and irregular rhythms, reducing the stability of the temporal context and thereby the likelihood of observing rhythm-based effects.

Modality constraints provide another explanation for the attenuated temporal effects. Visual presentation offers weaker temporal markers than auditory input (Goodfellow, 1934; Grondin, 1993). As a result, oscillatory entrainment may have been less robust, further limiting the contribution of temporal coding.

Finally, the exploration analyses that revealed rhythm effects on serial order after removing participant-level random intercepts suggest that temporal sensitivity may vary substantially across individuals (Grahn &

McAuley, 2009; Saito, 2001). This raises the possibility that temporal coding mechanisms are used heterogeneously, with some individuals relying more heavily on temporal cues than others.

In summary, **Study 3** indicates that temporal structure can influence item recall but does not reliably serve as a representational substrate for serial order under the conditions tested. Temporal coding appears to function mainly through attentional allocation rather than as a core structural code for sequence representation.

An Integrative Hybrid Mechanism of Serial Order Coding

Across the three studies, a coherent picture emerges regarding the nature of serial order coding in verbal WM. Serial order does not arise from a single, dedicated representational format; instead, it reflects the coordinated activity of multiple knowledge sources and contextual cues, whose relative contribution varies with task demands and environmental constraints.

Study 1 showed that long-term syntactic knowledge provides a powerful internal structure that facilitates the encoding and reconstruction of order information, especially when temporary syntactic frames within the sequence are minimized. This demonstrates that the language system does not merely support item-level processing but actively shapes the organization of sequential information.

Study 2 showed that spatialization operates under more constrained conditions. Spatial codes become influential only when the task environment establishes a sufficiently coherent and behaviorally meaningful spatial reference frame, and even then, their impact remains sensitive to properties of the memoranda (such as digits) and to competition from long-term spatial-numerical associations. Spatial coding thus appears to function as a supplementary and situationally invoked mechanism rather than as a ubiquitous representational foundation for serial order.

Study 3 further emphasized the context dependence of order coding. Temporal irregularity exerted only modest influences, affecting item recall under pronounced irregularity while leaving serial order recall largely unchanged. These temporal effects themselves depended on methodological parameters such as visual presentation, the absence of articulatory suppression, and trial randomization, all of which limited the formation of stable syntactic expectation. Temporal structure, therefore, appears to shape performance mainly via attentional alignment and rhythmic expectation rather than by providing an obligatory temporal code for serial position.

Together, these findings converge on a hybrid, context-sensitive mechanism of serial order coding in verbal WM. Within this framework, serial order representations arise from the dynamic coordination of long-term linguistic structures and short-term contextual signals. Long-term syntactic knowledge provides a particularly robust and readily accessible scaffold, shaping order representations even under minimal structural overlap between list elements. Spatial and temporal cues contribute more conditionally, supporting order maintenance when they are stable, salient, and behaviorally relevant, but receding when other representational resources (e.g., phonological or syntactic codes) offer more reliable guidance for recall.

This multi-source perspective is consistent with broader theoretical models of WM. For instance, the embedded-processes model (Cowan, 1999, 2019; see Figure 1.1) conceptualizes WM not as a separate storage, but as an activated subset of long-term memory governed by a limited-capacity focus of attention (3-5 items). Activated long-term memory includes both pre-existing knowledge and newly encoded representations that remain temporarily active, while the central executive directs attentional resources, either voluntary or involuntary, toward external input or internal knowledge structures. Within this architecture, long-term structures such as syntactic regularities, phonotactic patterns, and culturally acquired reading/writing spatial layouts can be treated as long-term knowledge structures that, when activated, constrain serial order coding in verbal WM, whereas temporal regularities occupy a more ambiguous status at the interface between long-term knowledge and dynamic attentional processes[1] (Large & Jones, 1999; for review, Nobre & Van Ede, 2018).

---

[1] In the case of temporal information, it is necessary to distinguish between learned temporal schemas and the momentary oscillatory mechanisms that implement them. Repeated exposure to regular temporal patterns, for example, in blocked rhythmic designs or in highly practiced musical and speech rhythms, can give rise to relatively stable temporal expectations that are stored as temporal schemas in long-term memory (Jones, 1976; Nissen & Bullemer, 1987; O'Reilly et al., 2008; Shin & Ivry, 2002). By contrast, the oscillatory dynamics that align internal attentional rhythms with a given sequence at a

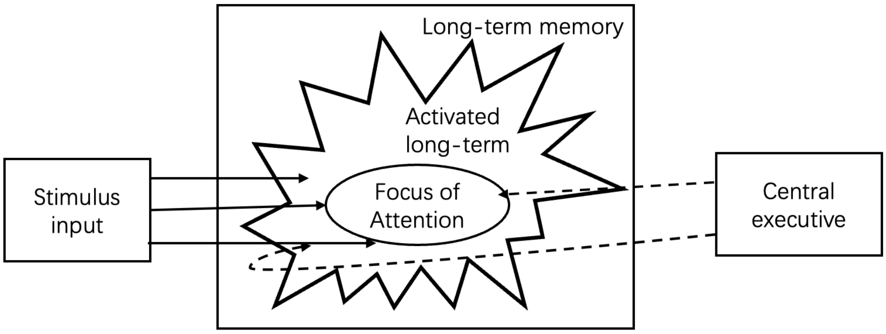

Figure 1.1. Outline of Cowan's (1988) embedded-processes model.

A related perspective is offered by the Dynamic-Processing Model of WM (Rose, 2020), which emphasizes that WM consists of dynamic representational states that flexibly recruit perceptual, attentional, and long-term memory systems. Information may be maintained through persistent neural firing or via activity-silent traces that can be reactivated when attention shifts back to them. Representational formats are continuously updated and reconfigured in response to contextual changes. This framework parallels the present findings in suggesting that verbal WM relies on the ongoing transformation and coordination of sensory, conceptual, and abstract codes rather than on fixed storage buffers.

Viewing WM as a constellation of activated long-term memory representations under attentional controls implies that the representational format of serial order is not fixed, but is constructed from the interaction between activated linguistic knowledge, attentional selection, and the immediate contextual structures available during encoding and retrieval. The present findings contribute to this theoretical landscape by illustrating how multiple sources of information operate in concert: syntactic, spatial, temporal, and phonological codes may co-exist, but they do not exert equal

---

particular moment are better understood as part of the attentional control system operating over the active knowledge base, rather than as long-term representational structures themselves. In this sense, Cowan's model can accommodate temporal coding at two intertwined levels, as learned temporal regularities in long-term memory and as dynamic attending processes that exploit these regularities to support the encoding and retrieval of serial order information.

or constant influence. Instead, the cognitive system selectively recruits those cues that are most diagnostic and reliable in the moment.

In this sense, serial order in verbal WM should be understood not as the output of a single underlying computation, but as the emergent product of a dynamic system that continually evaluates and integrates diverse information signals. This view accommodates the robust syntactic effects of **Study 1**, the conditional spatial effects of **Study 2**, and the attenuated temporal effects of **Study 3**, while offering a unified framework for understanding how verbal sequences are maintained in both experimental and more naturalistic contexts.

Limitations and Perspectives

The studies conducted in this thesis provide new insights into the dynamic, multi-source, and context-sensitive nature of serial order coding in verbal WM. As with any empirical contribution, several limitations must be acknowledged. Rather than undermining the findings, these limitations highlight conceptual and methodological issues that can guide future research toward a more comprehensive account of how serial order representations emerge from the interaction of long-term knowledge and short-term contextual structure.

Inter-Individual Variability

A first limitation concerns the restricted examination of inter-individual variability across the three empirical studies. Substantial differences in cognitive abilities – particularly attention control and working memory capacity (WMC) – and in long-term linguistic knowledge are likely to have contributed to variability in performance. Individuals with higher WMC are typically better able to maintain task goals and resist intrusion from irrelevant information, whereas those with lower capacity are more vulnerable to distraction and interference (Barrett et al., 2004; Dong et al., 2015). Differences in vocabulary size, semantic richness, and syntactic proficiency likewise influence both item memory and serial order performance in verbal tasks (e.g., Daneman & Green, 1986; Marton & Eichorn, 2015). These considerations reinforce a central idea of this thesis: verbal WM is not a standalone buffer but is deeply intertwined with the representational and structural resources of long-term memory. More specifically, the findings suggest that serial order coding is jointly shaped by long-term knowledge – such as syntactic structure, culturally acquired reading-writing direction, and expertise-based schemas - and short-term contextual information provided by the task. Individuals with more advanced syntactic processing abilities may exploit structural regularities more efficiently (e.g., Artuso & Palladino, 2019). Spatialization effects also appear to vary markedly across individuals. The work of van Dijck et al. (2022), showing that only about 20% of participants exhibit a robust left-to-right spatial bias in verbal WM tasks, underscores the considerable variability in how individuals map ordered sequences onto spatial frames. Temporal perception similarly shows inter-individual variability: WMC predicts temporal discrimination performance (Broadway & Engle, 2011), and individuals differ in their preferred timing strategies (beat-based vs. interval-based), each associated with distinct neural signatures (Grahn & McAuley, 2009). Musical expertise further enhances temporal sensitivity. Although the present studies did not directly measure linguistic proficiency, WMC, or timing strategies, mixed-effects models including random intercepts and slopes were applied to statistically account for individual differences. The main findings were generally robust, but one notable exception emerged in **Study 3**: when random intercepts were removed, a rhythm effect on serial order recall became visible. This pattern suggests that temporal sensitivity may vary substantially across participants, and that individuals may rely on different cognitive strategies or representational formats to manage serial order. Future research should therefore integrate explicit assessments of syntactic proficiency, spatial ability, and WMC in order to identify more precisely how individual skills modulate the mechanisms underlying serial order coding.

Predominance of Item-Level Effects over Order-Specific Measures

A second limitation concerns the predominance of effects on item recall relative to order-specific measures, particularly in **Studies 2 and 3**. Spatial and temporal manipulations produced reliable differences in item recall scores (and, in some analyses, in serial recall scores), whereas order-specific measures such as order recall scores and transposition gradients did not show corresponding effects. One intuitive explanation is that these order measures may be insufficiently sensitive or "pure" to detect subtle modulations in serial order coding. However, the results of **Study 1** argue against a purely methodological account, as the same analyses revealed robust effects of syntactic legality on serial order. Several characteristics of **Studies 2 and 3** may help to explain why spatial and temporal manipulations manifested more clearly at the item level. First, **Studies 2 and 3** both adopted digits as memoranda, which are highly overlearned and strongly phonological. In addition, in French, it is plausible that some participants recoded adjacent digits into multi-digit numbers that are named holistically (e.g., "6–1" encoded as "soixante-et-un"). If such chunking occurs, processing may shift away from individual items bound to positions and toward larger units, inducing the possibility that digits omitted together produce clear differences in item and serial recall but relatively little graded variation in order-specific measures. In addition, spatial congruency (**Study 2**) and pronounced temporal irregularity (**Study 3**) may have primarily influenced the robustness of item or chunk maintenance, via attentional allocation and overall resource demands, rather than directly on positional bindings among the items that remain available. Thus, the predominance of item-level effects in **Studies 2 and 3** should not be interpreted as evidence against spatial or temporal contributions to serial order; rather, in digit-based and relatively demanding tasks, these manipulations may chiefly modulate whether items or chunks are preserved in memory, whereas the internal positional code is comparatively less affected.

Task and Output Constraints: Absence of Recency Effects

Another limitation concerns the expression of serial position effects, recency in particular, across different task formats. In **Study 2**, only limited recency was observed in the oral recall tasks for both item and serial recall scores, whereas a clear recency effect emerged when participants responded by clicking boxes and typing the digits. A similar absence of recency in oral recall was observed in **Study 3**. These patterns suggest that recency effects were highly dependent on output demands rather than reflecting a fixed property of the underlying memory trace. First, recency is known to be strongly modulated by output interference and output order, that items are more likely to be recalled correctly when retrieved earlier in the response sequence (Grenfell-Essam & Ward, 2012; Madigan & McCabe, 1971; Oberauer et al., 2018; Ward et al., 2010), and the relative strength of recency varies with retrieval dynamics. Seconds, classic work on the modality effect shows that, compared to silently read visual stimuli, spoken stimuli typically produce enhanced recency (Conrad & Hull, 1968; Gardiner & Gregg, 1979; Murdock

& Walker, 1969; Oberauer et al., 2018; M. J. Watkins et al., 1974). In the present studies, oral recall of visually presented digits combines a continuous verbal response stream with visually encoded materials, a combination that may be especially vulnerable to output interference at the end of the list. Third, with digit memoranda, especially if there is possibility that adjacent digits are chunked into multi-digit numbers, failures at later serial positions are likely to involve the loss of an entire chunk rather than a single item, thereby flattening performance at the end of the list and attenuating recency in both item and serial recall scores, as observed in the oral conditions of **Studies 2 and 3**. The visual layout of the click-and-type task used in **Study 2** may have supported positional anchoring (Tremblay et al., 2006), even though participants were instructed to recall in order, allowing higher accessibility of end-of-list items to be expressed and then yielding a clearer recency effect. Overall, these considerations suggest that the lack of recency in the oral tasks should be interpreted with caution: it likely reflects constraints imposed by the modality and response dynamics rather than a complete absence of recency at the representational level.

Influence of Phonological Rehearsal

A second limitation concerns the lack of control over articulatory rehearsal across studies. Throughout this thesis, serial order coding has been conceptualized as a multi-source phenomenon drawing on spatial, temporal, syntactic, and phonological dimensions. When participants are free to rehearse aloud or subvocally, they may preferentially rely on phonological strategies, thereby reducing the likelihood that spatial or temporal codes will measurably contribute to performance. This issue is particularly salient in **Studies 2 and 3**, which used digits as memoranda. Digits are highly familiar and strongly phonological stimuli and tend to elicit shallow semantic processing and strong reliance on rehearsal-based maintenance. Ginsburg et al. (2017) showed that spatialization effects are more pronounced when stimuli undergo deeper semantic processing, and Guida et al. (2026) demonstrated that spatialization is moderated by the strength of phonological maintenance. In temporal tasks, phonological rehearsal may entrain internal oscillators, smoothing irregular temporal patterns and thereby reducing detectable differences between regular and irregular rhythms. Future studies could address these concerns by using letters or words rather than digits to encourage deeper semantic processing, implementing articulatory suppression to limit reliance on phonological rehearsal, and presenting memoranda auditorily to enhance sensitivity to temporal structure (Attout et al., 2025; Henson, 1996). Such methodological refinements would help isolate how spatial and temporal codes contribute to serial order when phonological rehearsal is constrained.

Task-Specific Constraints in Spatial Studies

A further limitation concerns the task parameters used in **Study 2**, particularly list length. Previous research suggests that spatial coding is most prominent under conditions of relatively low cognitive load, where attentional and spatial resources can be deployed effectively. Spatialization effects weaken or disappear as list length increases, diminishing between four and six items (Huber et al., 2016; Vivion et al., 2025) and vanishing entirely for very long lists (Guida, Carnet, et al., 2018). This pattern aligns with Cowan's embedded-processes model (1999, 2019), which proposes that only three to four items can reliably fall within the focus of attention. Although spatialization can re-emerge for longer sequences when task demands are modest (Unsworth & Engle, 2007), the seven-digit lists and full serial recall procedure used in **Study 2** likely imposed a relatively high cognitive load. Under these conditions, potential spatial contributions may have been overshadowed by more dominant representational formats, especially phonological coding. Moreover, the use of digits introduces spatial-numerical associations (SNARC-like tendencies), which may have further competed with the experimentally imposed spatial frames. Further studies would benefit from using shorter lists, probe-recall paradigms, or alternative procedures that reduce output interference and lower cognitive load, thereby creating more favorable conditions for spatial coding to manifest.

Methodological Limitations in Temporal Studies

Study 3 also presents methodological limitations related to the organization of the rhythmic context. Regular and irregular trials were intermixed randomly, whereas studies reporting robust temporal effects typically employ blocked designs that allow participants to form stable rhythmic expectations (Attout et al., 2025; Elbaz & Yeshurun, 2020; Fanuel et al., 2018; Plancher et al., 2018). Randomization likely undermined the build-up of entrainment and reduced the detectability of temporal coding effects.

The use of visually presented digits may also have attenuated sensitivity to temporal cues. A substantial body of research has demonstrated that humans exhibit higher temporal sensitivity in the auditory modality and auditory input more directly engages neural oscillatory dynamics (Gault & Goodfellow, 1938; Glenberg & Jona, 1991; Goodfellow, 1934; Grondin, 1993; Grondin et al., 1998; Macken et al., 2016). Further work should therefore preferentially employ auditory presentation formats, combined with articulatory suppression, to provide a more rigorous test of whether temporal structure can support serial order. Overall, these methodological considerations emphasize the importance of designing experiments that provide stable, salient, and continuous temporal contexts when evaluating theoretical predictions derived from oscillatory timing models.

Remaining Questions and Future Directions

The present findings suggest several promising directions for future research. One avenue is to investigate spatial coding under more fine-grained and theoretically constrained conditions, combining shorter lists, articulatory suppression, and explicit spatial cues to determine when and how spatialization supports serial order. A second avenue concerns the role of temporal regularity: blocked rhythmic designs, auditory materials, and rehearsal controls are needed to clarify how external rhythms interact with internal oscillatory mechanisms to support - or fail to support - temporal coding. A third line of research involves examining the interplay between long-term knowledge structures and short-term contextual cues. How do enduring syntactic and spatial schemas interact with temporary task configurations to shape serial order representations? This question remains largely unaddressed but is central to hybrid models of verbal WM. Finally, future work should explore how different representational codes – syntactic, phonological, spatial, and temporal – interact within the same task. Prior findings (Attout et al., 2025) suggest that temporal cues can overshadow spatial ones, pointing to potential competitive or hierarchical organization among codes. Understanding how these interactions unfold in more naturalistic contexts, such as reading or sentence processing, represents an important theoretical challenge.

Conclusion

This PhD thesis set out to investigate the nature of serial order coding in verbal WM through three complementary dimensions - linguistic, spatial, and temporal. Across three studies, a coherent picture emerged: serial order memory does not rely on a single representational mechanism, but instead reflects the dynamic interplay between long-term linguistic structures and short-term, task-dependent attentional processes. The findings contribute to ongoing debates concerning whether WM constitutes an autonomous system or a temporary activation of long-term knowledge, and they provide evidence in favor of hybrid models in which multiple representational sources jointly support the maintenance and reconstruction of ordered information.

The first empirical contribution concerns the role of syntactic structure. **Study 1** demonstrated that long-term syntactic knowledge exerts a clear and consistent influence on serial order recall. This influence is not restricted to familiar sentences or complex linguistic materials but extends to simple word pairs, showing that syntactic legality supports the organization of ordered sequences even under minimal surface structure. These results challenge models that locate serial order mechanisms exclusively outside the language system, and instead support accounts in which verbal WM emerges from the activation of long-term linguistic representations. Syntactic knowledge, alongside phonotactic and lexical-semantic information, thus guides the formation of sequential structure during both encoding and retrieval.

The second contribution concerns spatial coding. **Study 2** revealed that spatialization is not an obligatory feature of serial order but a contextually recruited mechanism that depends on the stability and salience of the spatial reference frame. Spatial effects emerged only when spatial cues were presented at both encoding and retrieval, indicating that spatial codes are engaged when they are coherent and behaviorally relevant, and are otherwise largely overshadowed by other representational formats. These findings align with models that treat spatialization as an auxiliary representational resource. They also caution against assuming that spatial mechanisms are uniformly accessible or consistently used across tasks and individuals.

The third empirical contribution concerns temporal structure. **Study 3** revealed that temporal regularity influenced item recall but provided limited evidence for direct contributions to serial order representation. Temporal irregularity impaired performance only when deviations from regularity were pronounced, suggesting that temporal coding depends on the salience and stability of the temporal context. The absence of reliable effects on serial order likely reflects methodological features – particularly visual presentation and trial-by-trial randomization – that attenuated rhythmic entrainment. These findings underscore that temporal signals do not automatically support order maintenance; instead, they influence performance indirectly by modulating attentional engagement, arousal, or temporal expectations.

Taken together, the three studies support a hybrid and context-sensitive conception of serial order coding in verbal WM. Serial order emerges from the coordinated activity of multiple representational sources: (1) long-term linguistic knowledge, such as syntactic structure, phonotactic patterns, or culturally learned spatial schemas, and (2) short-term contextual cues, including temporary structural regularities in the memoranda, spatial orientation during encoding and retrieval, and temporal rhythms in stimulus presentation. These sources do not operate in isolation; they interact continuously, and their relative contributions shift as a function of task demands, individual cognitive strategies, and the availability of diagnostic cues.

More broadly, the findings reinforce theoretical perspectives that view WM as a dynamic system grounded in long-term memory and shaped by attentional processes. The embedded-processes model (Cowan, 1999, 2019) and the Dynamic-Processing Model (Rose, 2020) both characterize WM as a constellation of activated representations modulated by attentional focus. The present results extend these frameworks by specifying how serial order – a defining but often under-theorized feature of WM – emerges from the joint operation of linguistic, spatial, and temporal structures. Serial order is thus neither a fixed code nor a dedicated buffer, but an adaptive construction, flexibly assembled from the most reliable cues available at the given moment.

In this way, the thesis advances the view that serial order in verbal WM is best conceived as a multi-source, flexible, and adaptively assembled representation. This perspective provides a unifying framework capable of accounting for observed variability across tasks, individuals, and modalities, and lays the groundwork for future models that more fully integrate long-term linguistic structures with the dynamic, context-sensitive operations of attentional control.

Statement on the use of generative AI

Statement: During the preparation of this thesis, the author used the generative AI tool ChatGPT (version 5.1, OpenAI) for limited language-support prupose, specifically to (a) suggest words or phrases that enhance clarity/readability of an existing sentence in Experimental Part, and (b) edit existing text for grammar, spelling, or organization in General Introduction and General Discussion. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the thesis.

References

Aaronson, D. (1968). Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology*, *76*(1p1), 129.

Abrahamse, E., Van Dijck, J.-P., Majerus, S., & Fias, W. (2014). Finding the answer in space: The mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience*, *8*, 932.

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50–68. https://doi.org/10.1037/a0014411

Agam, Y., Bullock, D., & Sekuler, R. (2005). Imitating unfamiliar sequences of connected linear motions. *Journal of Neurophysiology*, *94*(4), 2832–2843.

Agam, Y., Huang, J., & Sekuler, R. (2010). Neural correlates of sequence encoding in visuomotor learning. *Journal of Neurophysiology*, *103*(3), 1418–1424.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*(4), 341–380.

Anderson, R. B., Crawford, J. C., & Bailey, M. H. (2022). Biasing the input: A yoked-scientist demonstration of the distorting effects of optional stopping on Bayesian inference. *Behavior Research Methods*, *54*(3), 1131–1147. https://doi.org/10.3758/s13428-021-01618-1

Artuso, C., & Palladino, P. (2019). Long-term memory effects on working memory updating development. *PloS One*, *14*(5), e0217697.

Attout, L., Capizzi, M., & Charras, P. (2024). Enhancing rhythmic temporal expectations: The dominance of auditory modality under spatial uncertainty. *Attention, Perception, & Psychophysics*, *86*(5), 1681–1693. https://doi.org/10.3758/s13414-024-02898-3

Attout, L., Fias, W., Salmon, E., & Majerus, S. (2014). Common neural substrates for ordinal representation in short-term memory, numerical and alphabetical cognition. *PloS One*, *9*(3), e92049.

Attout, L., Leroy, N., & Majerus, S. (2022). The neural representation of ordinal information: Domain-specific or domain-general? *Cerebral Cortex*, *32*(6), 1170–1183.

Attout, L., Ordonez Magro, L., Szmalec, A., & Majerus, S. (2019). The developmental neural substrates of item and serial order components of verbal working memory. *Human Brain Mapping*, *40*(5), 1541–1553.

Attout, L., Ordonez Magro, L., Szmalec, A., & Majerus, S. (2020). The developmental neural substrates of Hebb repetition learning and their link with reading ability. *Human Brain Mapping*, *41*(14), 3956–3969.

Attout, L., Remouchamps, R., & Majerus, S. (2025). Impact of concurrent temporal but not spatial processing on working memory for serial order. *Quarterly Journal of Experimental Psychology*, 17470218251346745.

Avons, S. (2007). Spatial span under translation: A study of reference frames. *Memory & Cognition*, *35*(3), 402–417.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

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

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (2017). The phonological loop as a language learning device. In *Exploring Working Memory*. Routledge.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, *112*, 91–106.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158–173. https://doi.org/10.1037/0033-295x.105.1.158

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*(4), 553.

Battacchi, M. W., Pelamatti, G. M., & Umiltà, C. (1990). Is there a modality effect? Evidence for visual recency and suffix effects. *Memory & Cognition*, *18*(6), 651–658.

Beaudry, O., Neath, I., Surprenant, A. M., & Tehan, G. (2014). The focus of attention is similar to other memory systems rather than uniquely different. *Frontiers in Human Neuroscience*, *8*, 56.

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition*, *37*(5), 689–713.

Bhatarah, P., Ward, G., & Tan, L. (2006). Examining the relationship between free recall and immediate serial recall: The effect of concurrent task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 215.

Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition*, *36*(1), 20–34.

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 80–97.

Bonhage, C. E., Fiebach, C. J., Bahlmann, J., & Mueller, J. L. (2014). Brain signature of working memory for sentence structure: Enriched encoding and facilitated maintenance. *Journal of Cognitive Neuroscience*, *26*(8), 1654–1671.

Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*(1), 122–125. https://doi.org/10.3758/BF03200768

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*(5), 467–482. https://doi.org/10.1037/h0061096

Broadway, J. M., & Engle, R. W. (2011). Individual differences in working memory capacity and temporal discrimination. *PloS One*, *6*(10), e25422.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*(1), 127–181. https://doi.org/10.1037/0033-295X.107.1.127

Brown, G. D., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*(1), 127.

Buchsbaum, B. R., & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. *Cortex*, *112*, 134–148. https://doi.org/10.1016/j.cortex.2018.11.010

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

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, *31*(4), 429–460.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551–581. https://doi.org/10.1037/0033-295X.106.3.551

Burgess, N., & Hitch, G. J. (2006a). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*(4), 627–652. https://doi.org/10.1016/j.jml.2006.08.005

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

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Caplan, J. B., Hennies, N., & Sommer, T. (2022). Competition between associations in memory. *Journal of Cognitive Neuroscience*, *34*(11), 2144–2167.

Chiara Fastame, M., Flude, B., & Hitch, G. J. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, *13*(3–4), 247–258.

Chincotta, D., Underwood, G., Ghani, K. A., Papadopoulou, E., & Wresinski, M. (1999). Memory span for Arabic numerals and digit words: Evidence for a limited-capacity, visuo-spatial storage system. *The Quarterly Journal of Experimental Psychology Section A*, *52*(2), 325–351.

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

Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition*, *21*(2), 162–167. https://doi.org/10.3758/BF03202728

Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2019). *Short-term memory based on activated long-term memory: A review in response to Norris (2017).*

Cristoforetti, G., Majerus, S., Sahan, M. I., Van Dijck, J.-P., & Fias, W. (2022). Neural patterns in parietal cortex and hippocampus distinguish retrieval of start versus end positions in working memory. *Journal of Cognitive Neuroscience*, *34*(7), 1230–1245.

Crowder, R. G. (1979). Similarity and order in memory. In *Psychology of learning and motivation* (Vol. 13, pp. 319–353). Elsevier.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*(1), 1–18.

De Belder, M., Abrahamse, E., Kerckhof, M., Fias, W., & van Dijck, J.-P. (2015). Serial position markers in space: Visuospatial priming of serial order working memory retrieval. *PloS One*, *10*(1), e0116469.

De Belder, M., van Dijck, J.-P., Cappelletti, M., & Fias, W. (2017). How serially organized working memory information interacts with timing. *Psychological Research*, *81*(6), 1255–1263.

de Leeuw, J., Gilbert, R., & Luchterhandt, B. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *The Journal of Open Source Software*, *8*, 5351. https://doi.org/10.21105/joss.05351

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science*, *6*(3), 274–290. https://doi.org/10.1177/1745691611406920

Dong, S., Reder, L. M., Yao, Y., Liu, Y., & Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Research*, *1616*, 146–156.

Ebbinghaus, H. (1964). *Memory*. Рипол Классик.

Elbaz, A., & Yeshurun, Y. (2020). Can rhythm-induced attention improve the perceptual representation? *Plos One*, *15*(4), e0231200.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 671.

Fanuel, L., Portrat, S., Tillmann, B., & Plancher, G. (2018). Temporal regularities allow saving time for maintenance in working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 202–211.

Farrand, P., Parmentier, F. B., & Jones, D. M. (2001). Temporal–spatial memory: Retrieval of spatial information does not reduce recency. *Acta Psychologica*, *106*(3), 285–301.

Farrell, S. (2008). Multiple roles for time in short-term memory: Evidence from serial recall of order and timing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 128.

Farrell, S., & Lelièvre, A. (2009). End anchoring in short-term order memory. *Journal of Memory and Language*, *60*(1), 209–227.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*(1), 59–79.

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*(1), 115–135.

Farrell, S., Wise, V., & Lelièvre, A. (2011). Relations between timing, position, and grouping in short-term memory. *Memory & Cognition*, *39*(4), 573–587.

Fias, W., Lammertyn, J., Caessens, B., & Orban, G. A. (2007). Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulcus. *Journal of Neuroscience*, *27*(33), 8952–8956.

Fischer-Baum, S., & Benjamin, A. S. (2014). Time, space, and memory for order. *Psychonomic Bulletin & Review*, *21*(5), 1263–1271.

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*(5), 458–462.

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

Franklin, D. R., & Mewhort, D. (2015). Memory as a hologram: An analysis of learning and recall. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *69*(1), 115.

Gardiner, J. M., & Gregg, V. H. (1979). When auditory memory is not overwritten. *Journal of Verbal Learning and Verbal Behavior*, *18*(6), 705–719.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition*, *23*(1), 83–94.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336–360.

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *The Quarterly Journal of Experimental Psychology Section A*, *54*(1), 1–30. https://doi.org/10.1080/02724980042000002

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *13*(1), 65–77.

Gault, R. H., & Goodfellow, L. D. (1938). An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them. *The Journal of General Psychology*, *18*(1), 41–47.

Geiger, S. M., & Lewandowsky, S. (2008). Temporal isolation does not facilitate forward serial recall—Or does it? *Memory & Cognition*, *36*(5), 957–967.

Gilbert, R. A., Hitch, G. J., & Hartley, T. (2017). Temporal precision and the capacity of auditory–verbal short-term memory. *Quarterly Journal of Experimental Psychology*, *70*(12), 2403–2418.

Ginsburg, V., Archambeau, K., van Dijck, J.-P., Chetail, F., & Gevers, W. (2017). Coding of serial order in verbal, visual and spatial working memory. *Journal of Experimental Psychology: General*, *146*(5), 632.

Ginsburg, V., & Gevers, W. (2015). Spatial coding of ordinal information in short-and long-term memory. *Frontiers in Human Neuroscience*, *9*, 8.

Ginsburg, V., Van Dijck, J.-P., Previtali, P., Fias, W., & Gevers, W. (2014). The impact of verbal working memory on number–space associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 976.

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition*, *19*(5), 514–522.

Goodfellow, L. D. (1934). An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time. *The American Journal of Psychology*, *46*(2), 243–258.

Gorin, S. (2020). The influence of rhythm on short-term memory for serial order. *Quarterly Journal of Experimental Psychology*, *73*(12), 2071–2092.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, *47*(4), 1894–1903.

Gregg, V. H., Freedman, C. M., & Smith, D. K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, *80*(3), 363–374. https://doi.org/10.1111/j.2044-8295.1989.tb02326.x

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: The role of list length, strategy use, and test expectancy. *Journal of Memory and Language*, *67*(1), 106–148.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics*, *54*(3), 383–394.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research*, *61*(4), 261–268.

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 556.

Guida, A., Abrahamse, E., & van Dijck, J. (2020). About the interplay between internal and external spatial codes in the mind: Implications for serial order. *Annals of the New York Academy of Sciences*, *1477*(1), 20–33.

Guida, A., & Campitelli, G. (2019). Explaining the SPoARC and SNARC effects with knowledge structures: An expertise account. *Psychonomic Bulletin & Review*, *26*(2), 434–451.

Guida, A., Carnet, S., Normandon, M., & Lavielle-Guida, M. (2018). Can spatialisation be extended to episodic memory and open sets? *Memory*, *26*(7), 922–935.

Guida, A., & Lavielle-Guida, M. (2014). 2011 space odyssey: Spatialization as a mechanism to code order allows a close encounter between memory expertise and classic immediate memory studies. *Frontiers in Psychology*, *5*, 573.

Guida, A., Leroux, A., Lavielle-Guida, M., & Noël, Y. (2016). A SPoARC in the dark: Spatialization in verbal immediate memory. *Cognitive Science*, *40*(8), 2108–2121.

Guida, A., Mathy, F., Gobet, F., Campitelli, G., & Sala, G. (2026). Linking space and ordinal position in working memory: A multi-level meta-analysis of the SPoARC effect. *Cognition, 266,* 106276. https://doi.org/10.1016/j.cognition.2025.106276

Guida, A., Megreya, A. M., Lavielle-Guida, M., Noël, Y., Mathy, F., van Dijck, J.-P., & Abrahamse, E. (2018). Spatialization in working memory is related to literacy and reading direction: Culture "literarily" directs our thoughts. *Cognition*, *175*, 96–100.

Guida, A., Mosinski, F., Cipora, K., Mathy, F., & Noël, Y. (2020). Spatialization in working memory: Can individuals reverse the cultural direction of their thoughts? *Annals of the New York Academy of Sciences*, *1477*(1), 113–125.

Guitard, D., Saint-Aubin, J., Reid, J. N., & Jamieson, R. K. (2025). An embedded computational framework of memory: The critical role of representations in veridical and false recall predictions. *Psychonomic Bulletin & Review*, 1–55.

Gupta, P., & MacWhinney, B. (1997). Vocabulary Acquisition and Verbal Short-Term Memory: Computational and Neural Bases. *Brain and Language*, *59*(2), 267–333. https://doi.org/10.1006/brln.1997.1819

Haberlandt, K., Thomas, J. G., Lawrence, H., & Krohn, T. (2005). Transposition asymmetry in immediate serial recall. *Memory*, *13*(3–4), 274–282.

Hall, D., & Gathercole, S. E. (2011). Serial recall of rhythms and verbal sequences: Impacts of concurrent tasks and irrelevant sound. *Quarterly Journal of Experimental Psychology*, *64*(8), 1580–1592.

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, *87,* 135–178. https://doi.org/10.1016/j.cogpsych.2016.05.001

Hartmann, M., Martarelli, C. S., & Sommer, N. R. (2021). Early is left and up: Saccadic responses reveal horizontal and vertical spatial associations of serial order in working memory. *Cognition*, *217*, 104908.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, *19*(6), 304–313. https://doi.org/10.1016/j.tics.2015.04.006

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13*(6), 644–655.

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *The Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1307–1334.

Henson, R. N. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 80–115.

Henson, R. N. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, *36*(2), 73–137.

Henson, R. N. (1999a). Coding position in short-term memory. *International Journal of Psychology*, *34*(5–6), 403–409.

Henson, R. N. (1999b). Positional information in short-term memory: Relative or absolute? *Memory & Cognition*, *27*(5), 915–927.

Henson, R. N. A., Burgess, N., & Frith, C. D. (2000). Recoding, storage, rehearsal and grouping in verbal short-term memory: An fMRI study. *Neuropsychologia*, *38*(4), 426–440. https://doi.org/10.1016/S0028-3932(99)00098-6

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*(5456), 1269–1272.

Hitch, G. J. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 116–139.

Hitch, G., & Morton, J. (1975). The unimportance of explicit spatial information in serial recall of visually presented lists. *Quarterly Journal of Experimental Psychology*, *27*(2), 161–164.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 923.

Huber, S., Klein, E., Moeller, K., & Willmes, K. (2016). Spatial–numerical and ordinal positional associations coexist in parallel. *Frontiers in Psychology*, *7*, 438.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, *30*(6), 685–701. https://doi.org/10.1016/0749-596X(91)90032-F

Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology*, *86*(4), 527–536. https://doi.org/10.1111/j.2044-8295.1995.tb02570.x

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1217–1232. https://doi.org/10.1037/0278-7393.23.5.1217

Hulme, C., Stuart, G., Brown, G. D. A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, *49*(4), 500–518. https://doi.org/10.1016/S0749-596X(03)00096-2

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 295.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339–373. https://doi.org/10.1037/a0034221

Jeffreys, S. H., & Jeffreys, S. H. (1998). *The Theory of Probability* (Third Edition, Third Edition). Oxford University Press.

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 1008.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459.

Jones, T., & Farrell, S. (2018). Does syntax bias serial order reconstruction of verbal short-term memory? *Journal of Memory and Language*, *100*, 98–122. https://doi.org/10.1016/j.jml.2018.02.001

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109.

Kalm, K., & Norris, D. (2014). The representation of order information in auditory-verbal short-term memory. *Journal of Neuroscience*, *34*(20), 6879–6886.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*(3), 153–161.

Kowialiewski, B., & Majerus, S. (2018). The non-strategic nature of linguistic long-term memory effects in verbal short-term memory. *Journal of Memory and Language*, *101*, 64–83. https://doi.org/10.1016/j.jml.2018.03.005

Kowialiewski, B., Majerus, S., & Oberauer, K. (2024). Does semantic similarity affect immediate memory for order? Usually not, but sometimes it does. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(1), 68–88. https://doi.org/10.1037/xlm0001279

Kuroki, D. (2021). A new jsPsych plugin for psychophysics, providing accurate display duration and stimulus onset asynchrony. *Behavior Research Methods*, *53*(1), 301–310. https://doi.org/10.3758/s13428-020-01445-w

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119.

Lashley, K. S. (1951a). *The problem of serial order in behavior* (Vol. 21). Bobbs-Merrill Oxford.

Lashley, K. S. (1951b). The problem of serial order in behavior. In *Cerebral mechanisms in behavior; the Hixon Symposium.* (pp. 112–146). Wiley.

Leclercq, A.-L., & Majerus, S. (2010). Serial-order short-term memory predicts vocabulary development: Evidence from a longitudinal study. *Developmental Psychology*, *46*(2), 417.

LeCompte, D. C. (1992). In search of a strong visual recency effect. *Memory & Cognition*, *20*(5), 563–572.

Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, *16*(4), 395–418.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 149.

Lewandowsky, S., & Farrell, S. (2008a). Short-Term Memory: New Data and a Model. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 49, pp. 1–48). Academic Press. https://doi.org/10.1016/S0079-7421(08)00001-7

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

Lewandowsky, S., & Murdock Jr, B. B. (1989). Memory for serial order. *Psychological Review*, *96*(1), 25.

Lewandowsky, S., Nimmo, L. M., & Brown, G. D. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language*, *58*(2), 415–428.

Li, S.-C., & Lewandowsky, S. (1995). Forward and backward recall: Different retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 837.

Lian, A., Karlse, P. J., & Eriksen, T. B. (2004). Opposing effects of phonological similarity on item and order memory of words and nonwords in the serial recall task. *Memory*, *12*(3), 314–337.

Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context. *Psychological Review*, *128*(6), 1197.

Logan, G. D., Cox, G. E., Lilburn, S. D., & Ulrich, J. E. (2024). No position-specific interference from prior lists in cued recognition: A challenge for position coding (and other) theories of serial memory. *Cognitive Psychology*, *149*, 101641.

MacDonald, M. C. (2016). Speak, Act, Remember: The Language-Production Basis of Serial Order and Maintenance in Verbal Memory. *Current Directions in Psychological Science*, *25*(1), 47–53. https://doi.org/10.1177/0963721415620776

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

Madigan, S. A., & McCabe, L. (1971). Perfect recall and total forgetting: A problem for models of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*(1), 101–106.

MAJERUS, S. (2008). Verbal short-term memory and temporary activation of language representations: The importance of distinguishing item and order information. In *Interactions Between Short-Term and Long-Term Memory in the Verbal Domain*. Psychology Press.

Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00357

Majerus, S. (2018). Working memory treatment in aphasia: A theoretical and quantitative review. *Journal of Neurolinguistics*, *48*, 157–175.

Majerus, S. (2019). Verbal working memory and the phonological buffer: The question of serial order. *Cortex*, *112*, 122–133. https://doi.org/10.1016/j.cortex.2018.04.016

Majerus, S., Attout, L., Artielle, M.-A., & Van der Kaa, M.-A. (2015). The heterogeneity of verbal short-term memory impairment in aphasia. *Neuropsychologia*, *77*, 165–176. https://doi.org/10.1016/j.neuropsychologia.2015.08.010

Majerus, S., Bastin, C., Poncelet, M., Van der Linden, M., Salmon, E., Collette, F., & Maquet, P. (2007). Short-term memory and the left intraparietal sulcus: Focus of attention? Further evidence from a face short-term memory paradigm. *Neuroimage*, *35*(1), 353–367.

Majerus, S., Bruno, M.-A., Schnakers, C., Giacino, J. T., & Laureys, S. (2009). The problem of aphasia in the assessment of consciousness in brain-damaged patients. *Progress in Brain Research*, *177*, 49–61.

Majerus, S., & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language*, *64*(2), 181–197. https://doi.org/10.1016/j.jml.2010.10.003

Majerus, S., D'Argembeau, A., Martinez Perez, T., Belayachi, S., Van der Linden, M., Collette, F., Salmon, E., Seurinck, R., Fias, W., & Maquet, P. (2010). The Commonality of Neural Networks for Verbal and Visual Short-term Memory. *Journal of Cognitive Neuroscience*, *22*(11), 2570–2593. https://doi.org/10.1162/jocn.2009.21378

Majerus, S., Martinez Perez, T., & Oberauer, K. (2012). Two distinct origins of long-term learning effects in verbal short-term memory. *Journal of Memory and Language*, *66*(1), 38–51. https://doi.org/10.1016/j.jml.2011.07.006

Majerus, S., Norris, D., & Patterson, K. (2007). What does a patient with semantic dementia remember in verbal short-term memory? Order and sound but not words. *Cognitive Neuropsychology*, *24*(2), 131–151. https://doi.org/10.1080/02643290600989376

Majerus, S., Poncelet, M., Elsen, B., & Van der Linden, M. (2006). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall, and item recognition. *European Journal of Cognitive Psychology*, *18*(6), 848–873.

Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., Vandewalle, G., Collette, F., & Maquet, P. (2006). The left intraparietal sulcus and verbal short-term memory: Focus of attention or serial order? *NeuroImage*, *32*(2), 880–891. https://doi.org/10.1016/j.neuroimage.2006.03.048

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

Majerus, S., & Van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, *21*(2), 303–310. https://doi.org/10.1348/026151003765264101

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, *51*(2), 297–306.

Marshuetz, C., Reuter-Lorenz, P. A., Smith, E. E., Jonides, J., & Noll, D. C. (2006). Working memory for order and the parietal cortex: An event-related functional magnetic resonance imaging study. *Neuroscience*, *139*(1), 311–316. https://doi.org/10.1016/j.neuroscience.2005.04.071

Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000a). Order Information in Working Memory: fMRI Evidence for Parietal and Prefrontal Mechanisms. *Journal of Cognitive Neuroscience*, *12*(Supplement 2), 130–144. https://doi.org/10.1162/08989290051137459

Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000b). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience*, *12*(Supplement 2), 130–144.

Martin, N., Dell, G. S., Saffran, E. M., & Schwartz, M. F. (1994). Origins of Paraphasias in Deep Dysphasia: Testing the Consequences of a Decay Impairment to an Interactive Spreading Activation Model of Lexical Retrieval. *Brain and Language*, *47*(4), 609–660. https://doi.org/10.1006/brln.1994.1061

Martinez Perez, T., Poncelet, M., Salmon, E., & Majerus, S. (2015). Functional alterations in order short-term memory networks in adults with dyslexia. *Developmental Neuropsychology*, *40*(7–8), 407–429.

Marton, K., & Eichorn, N. (2015). Interaction between working memory and long-term memory. *Zeitschrift Für Psychologie*.

Matthews, M. L., & Henderson, L. (1970). Fast presentation rates and the recall of item and order information. *Nature*, *226*(5243), 374–376.

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

McElree, B., & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, *122*(3), 291.

Mewhort, D., Shabahang, K. D., & Franklin, D. (2018). Release from PI: An analysis and a model. *Psychonomic Bulletin & Review*, *25*(3), 932–950.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.

Morey, R. D., Romeijn, J.-W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, *72*, 6–18. https://doi.org/10.1016/j.jmp.2015.11.001

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*(4), 406–419. https://doi.org/10.1037/a0024377

Morin, C., Brown, G. D., & Lewandowsky, S. (2010). Temporal isolation effects in recognition and serial recall. *Memory & Cognition*, *38*(7), 849–859.

Murdock, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, *105*(2), 191.

Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, *100*(2), 183.

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition*, *23*(5), 631–645.

Murdock Jr, B. B., & Walker, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8*(5), 665–676.

Nairne, J. S., & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language*, *50*(2), 113–133. https://doi.org/10.1016/j.jml.2003.09.005

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 702.

Ng, H. L., & Maybery, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally? *The Quarterly Journal of Experimental Psychology Section A*, *55*(2), 391–424.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32.

Nobre, A. C., & Van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., & Hurlstone, M. J. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*(5), 779–819. https://doi.org/10.3758/s13423-012-0272-4

O'Reilly, J. X., McCarthy, K. J., Capizzi, M., & Nobre, A. C. (2008). Acquisition of the temporal and ordinal structure of movement sequences in incidental learning. *Journal of Neurophysiology*, *99*(5), 2731–2735.

Osth, A. F., & Dennis, S. (2015). Sources of interference in item and associative recognition memory. *Psychological Review*, *122*(2), 260.

Osth, A. F., & Hurlstone, M. J. (2023). *Do item-dependent context representations underlie serial order in cognition? Commentary on Logan (2021).*

Pannell, B., Guitard, D., Li, Y., & Cowan, N. (2023). Can synchronised tones facilitate immediate memory for printed lists? *Memory*, *31*(9), 1163–1175.

Park, J., Gagné, C. L., & Spalding, T. L. (2024). Writing direction and language activation affect how Arabic-English bilingual speakers map time onto space. *Frontiers in Psychology*, *14*, 1356039.

Parmentier, F. B., & Beaman, C. P. (2015). Contrasting effects of changing rhythm and content on auditory distraction in immediate memory. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *69*(1), 28.

Parmentier, F. B., & Jones, D. M. (2000). Functional characteristics of auditory temporal—Spatial short-term memory: Evidence from serial order errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 222.

Parmentier, F. B., & Maybery, M. T. (2008). Equivalent effects of grouping by time, voice, and location on response timing in verbal serial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1349.

Patterson, K., Graham, N., & Hodges, J. R. (1994). The Impact of Semantic Memory Loss on Phonological Representations. *Journal of Cognitive Neuroscience*, *6*(1), 57–69. https://doi.org/10.1162/jocn.1994.6.1.57

Perham, N., Marsh, J. E., & Jones, D. M. (2009). Short Article: Syntax and Serial Recall: How Language Supports Short-Term Memory for Order. *Quarterly Journal of Experimental Psychology*, *62*(7), 1285–1293. https://doi.org/10.1080/17470210802635599

Plancher, G., Lévêque, Y., Fanuel, L., Piquandet, G., & Tillmann, B. (2018). Boosting maintenance in working memory with temporal regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(5), 812.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *50*(4), 408–412. https://doi.org/10.1037/1196-1961.50.4.408

Poirier, M., Saint-Aubin, J., Mair, A., Tehan, G., & Tolan, A. (2015). Order recall in verbal short-term memory: The role of semantic networks. *Memory & Cognition*, *43*(3), 489–499.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Querella, P., & Majerus, S. (2024). Sequential syntactic knowledge supports item but not order recall in verbal working memory. *Memory & Cognition*, *52*(8), 1737–1761. https://doi.org/10.3758/s13421-023-01476-6

Rasoulzadeh, V., Sahan, M. I., van Dijck, J.-P., Abrahamse, E., Marzecova, A., Verguts, T., & Fias, W. (2021). Spatial attention in serial order working memory: An EEG study. *Cerebral Cortex*, *31*(5), 2482–2493.

Roberts, B. M., Clarke, A., Addante, R. J., & Ranganath, C. (2018). Entrainment enhances theta oscillations and improves episodic memory. *Cognitive Neuroscience*, *9*(3–4), 181–193.

Roodenrys, S., Hulme, C., Alban, J., Ellis, A. W., & Brown, G. D. A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition*, *22*(6), 695–701. https://doi.org/10.3758/BF03209254

Rose, N. S. (2020). The dynamic-processing model of working memory. *Current Directions in Psychological Science*, *29*(4), 378–387.

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *The Quarterly Journal of Experimental Psychology*, *21*(2), 137–147.

Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *The Quarterly Journal of Experimental Psychology*, *21*(2), 148–155.

Sahan, M. I., van Dijck, J.-P., & Fias, W. (2022). Eye-movements reveal the serial position of the attended item in verbal working memory. *Psychonomic Bulletin & Review*, *29*(2), 530–540.

Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: Is there an effect on all trials. *Psychonomic Bulletin & Review*, *12*(1), 171–177.

Saint-Aubin, J., & Poirier, M. (1999). Semantic Similarity and Immediate Serial Recall: Is There a Detrimental Effect on Order Information? *The Quarterly Journal of Experimental Psychology Section A*, *52*(2), 367–394. https://doi.org/10.1080/713755814

Saint-aubin, J., & Poirier, M. (2000). Immediate serial recall of words and nonwords: Tests of the retrieval-based hypothesis. *Psychonomic Bulletin & Review*, *7*(2), 332–340. https://doi.org/10.3758/BF03212990

Saint-Aubin, J., Poirier, M., Yearsley, J. M., Robichaud, J.-M., & Guitard, D. (2023). Modeling verbal short-term memory: A walk around the neighborhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(2), 198.

Saito, S. (2001). The phonological loop and memory for rhythms: An individual differences approach. *Memory*, *9*(4–6), 313–322.

Scheerer-Neumann, G. (1981). The utilization of intraword structure in poor readers: Experimental evidence and a training program. *Psychological Research*, *43*(2), 155–178.

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Schweppe, J., Schütte, F., Machleb, F., & Hellfritsch, M. (2022). Syntax, morphosyntax, and serial recall: How language supports short-term memory. *Memory & Cognition*, *50*(1), 174–191. https://doi.org/10.3758/s13421-021-01203-z

Schwering, S. C., & MacDonald, M. C. (2020). Verbal Working Memory as Emergent from Language Comprehension and Production. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00068

Schwering, S. C., & MacDonald, M. C. (2023). Noun Sequence Statistics Affect Serial Recall and Order Recognition Memory. *Open Mind*, *7*, 550–563. https://doi.org/10.1162/opmi_a_00092

Shaki, S., & Fischer, M. H. (2012). Multiple spatial mappings in numerical cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 804.

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 445.

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual—Spatial domain: Reconstructing sequences of unfamiliar faces. *The Quarterly Journal of Experimental Psychology Section A*, *58*(5), 909–930.

Smyth, M. M., & Scholey, K. A. (1996). Serial order in spatial immediate memory. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 159–177.

Sternberg, S. (1967). Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics*, *2*(2), 45–53.

Surprenant, A. M., Kelley, M. R., Farley, L. A., & Neath, I. (2005). Fill-in and infill errors in order memory. *Memory*, *13*(3–4), 267–273.

Thomas, J. G., Milner, H. R., & Haberlandt, K. F. (2003). Forward and backward recall: Different response time patterns, same retrieval order. *Psychological Science*, *14*(2), 169–174.

Tian, Y., & Fischer-Baum, S. (2025). The role of spatial processing in verbal serial order working memory. *Cognitive, Affective, & Behavioral Neuroscience*, *25*(1), 210–239.

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, *13*(3), 452–457.

Tse, C.-S. (2009). The role of associative strength in the semantic relatedness effect on immediate serial recall. *Memory*, *17*(8), 874–891.

Tse, C.-S., Li, Y., & Altarriba, J. (2011). The effect of semantic relatedness on immediate serial recall and serial recognition. *Quarterly Journal of Experimental Psychology*, *64*(12), 2425–2437.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*(6), 1038.

van Dijck, J., Fias, W., & Cipora, K. (2022). Spatialization in working memory and its relation to math anxiety. *Annals of the New York Academy of Sciences*, *1512*(1), 192–202.

van Dijck, J.-P., Abrahamse, E. L., Majerus, S., & Fias, W. (2013). Spatial attention interacts with serial-order retrieval from verbal working memory. *Psychological Science*, *24*(9), 1854–1859.

van Dijck, J.-P., & Fias, W. (2011a). A working memory account for spatial–numerical associations. *Cognition*, *119*(1), 114–119. https://doi.org/10.1016/j.cognition.2010.12.013

van Dijck, J.-P., & Fias, W. (2011b). A working memory account for spatial–numerical associations. *Cognition*, *119*(1), 114–119.

Vivion, M., Guida, A., Ramanoël, S., & Mathy, F. (2025). Spatial–positional association of response codes is modulated by the number of items in working memory. *Memory & Cognition*, 1–28.

Wagenmakers, E.-J., Lee, M., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian Versus Frequentist Inference. In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.), *Bayesian Evaluation of Informative Hypotheses* (pp. 181–207). Springer New York. https://doi.org/10.1007/978-0-387-09612-4_9

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256–1271. https://doi.org/10.1037/0278-7393.25.5.1256

Ward, G., Avons, S., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, *13*(3–4), 308–317.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1207.

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 430–447.

Watkins, O. C., & Watkins, M. J. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(6), 712–718. https://doi.org/10.1037/0278-7393.3.6.712