Semantic Knowledge Constrains the Processing of Serial Order Information in Working Memory

Benjamin Kowialiewski[1,2,3], Simon Gorin[4], Steve Majerus[1,3]

[1] University of Liège, Liège, Belgium

[2] Université Grenoble Alpes, Grenoble, France

[3] Fund for Scientific Research – F.R.S.-FNRS, Brussels, Belgium

[4] University of Geneva, Faculty of Psychology, Geneva, Switzerland

Correspondence concerning this article should be addressed to Benjamin Kowialiewski, Laboratoire de Psychologie et NeuroCognition (LPNC), Université Grenoble-Alpes, Bâtiment Michel Dubois prev. BSHM, 1251 Avenue Centrale, 38400 Saint-Martin-d'Hères. E-mail: bkowialiewski@uliege.be

The authors' contributions are reported following the CRediT taxonomy (see https://casrai.org/credit/). Conceptualization: BK and SG; Formal analysis: BK; Investigation: BK; Methodology: BK, SG, and SM; Resources: SM; Software: BK and SG; Validation: SG; Visualization: BK; Writing – original draft: BK, SG, and SM; Writing – review & editing: BK, SG, and SM.

Open science statement: The scripts, data and experimental set up have been made available on the Open Science Framework: https://osf.io/a4d29/.

10.1037/xlm0001031

Abstract

Long-term memory knowledge is considered to impact short-term maintenance of item information in working memory, as opposed to short-term maintenance of serial order information. Evidence supporting an impact of semantic knowledge on serial order maintenance remains weak. In the present study, we demonstrate that semantic knowledge can impact the processing of serial order information in a robust manner. Experiment 1 manipulated semantic relatedness effect by using semantic categories presented in subgroups of items (*leaf* – *tree* – *branch* – cloud – sky – rain). This semantic grouping manipulation was compared to a temporal grouping manipulation whose impact on the processing of serial order information is well-established. Both the semantic and temporal grouping manipulations constrained the occurrence of serial order errors in a robust manner: when migrating to a non-target serial position, items tended to do so most of the time toward the position of a semantically related item or within the same temporal group. Critically, this impact of semantic knowledge on the pattern of migration errors was not observed anymore in Experiment 2, in which we broke-up the semantic groups, by presenting the semantically related items an interleaved fashion (*leaf* – cloud – *tree* – sky – *branch* – rain). Both semantic and temporal grouping factors may reflect a general mechanism through which information is represented hierarchically. Alternatively, both factors could result from the syntactic and/or semantic regularities that naturally structure linguistic information. These results support models considering direct interactions between serial order and linguistic components of WM.

Keywords: Working Memory, Serial order, Semantic knowledge

Verbal working memory (WM), the ability to store verbal information over a short period of time, is thought to emerge from the interaction between different sub-systems involved in the representation of item identity, on the one hand, and the representation of serial order information, on the other hand (Brown et al., 2000; Henson et al., 2003; Hitch et al., 1996; Majerus, 2009, 2013, 2019). This distinction is supported by behavioural (Gorin et al., 2016; Henson et al., 2003), neuropsychological and neuroimaging evidence (Kalm & Norris, 2014; Majerus et al., 2010, 2015; Papagno et al., 2017). At the same time, there is also emerging evidence for an interaction between components involved in the representation of item and serial order information (Poirier et al., 2015; Saint-Aubin et al., 2005; Tse, 2009; Tse et al., 2011). So far, evidence supporting interactions between sematic knowledge and the processing of serial order information remains weak. In this study, we examined the extent to which semantic knowledge, generally thought to support the representation of item information in WM, may also interact with the representation of serial order information in a reliable manner.

The mechanisms responsible for the maintenance of serial order information (the sequential and/or temporal order in which memoranda are presented) are generally considered to be independent from those involved in the maintenance of item information (the orthographic, phonological and lexico-semantic characteristics of the memoranda) (Majerus, 2013). Behavioural studies have shown that linguistic long-term memory knowledge impacts the recall of item information, while minimally affecting the recall of serial order information (Allen & Hulme, 2006; Campoy et al., 2015; Hulme et al., 1997, 2003; Romani et al., 2008; Roodenrys et al., 2002; Saint-Aubin et al., 2005; Tse & Altarriba, 2007; Walker & Hulme, 1999). In addition, the maintenance of serial order information appears to be more sensitive to rhythmic and articulatory interfering tasks than is the maintenance of item information (Gorin et al., 2016;

Henson et al., 2003). Neuropsychological studies have also revealed double dissociations between item and serial order retention capacities in brain-injured and neurodevelopmental populations (Attout et al., 2014; Brock & Jarrold, 2005; Majerus et al., 2015; Martinez Perez et al., 2015). Finally, neuroimaging and neurostimulation studies have shown the involvement of neural regions specifically sensitive to the maintenance of serial order versus item information (Attout et al., 2014; Guidali et al., 2019; Kalm & Norris, 2014; Majerus et al., 2010; Papagno et al., 2017). The majority of current models of WM consider the existence of specific codes responsible for the maintenance of serial order, independent from those involved in the maintenance of item identity (Abrahamse et al., 2014; Brown et al., 2000; Burgess & Hitch, 1999, 2006; Hartley et al., 2016; Henson, 1996; Lewandowsky & Murdock, 1989).

Maintenance of verbal item information, on the other side, is considered to depend on interactions with linguistic knowledge stored in long-term memory (Acheson & MacDonald, 2009; Hulme et al., 1991; Majerus, 2019; Martin et al., 1996; Schweickert, 1993). This is shown by the presence of a large number of psycholinguistic effects impacting the recall of item information, involving the different levels of language processing (phonological, lexical and semantic). At the sublexical, phonological level, it has been shown that verbal WM performance increases for nonwords containing high versus low phonotactic probability structures (Gathercole et al., 1999; Majerus et al., 2004, 2012). Similarly, rhyming lists of items are characterized by less omission errors compared to non-rhyming lists (Fallon et al., 2005; Gupta et al., 2005). At the lexical level, WM performance is higher for words as compared to nonwords (Brener, 1940; Jefferies et al., 2006), as well for high versus low frequency words (Hulme et al., 1991). The same effects are also observed at the semantic level, with concrete words being better recalled as compared to abstract words (Bourassa & Besner, 1994; Campoy et al., 2015; Kowialiewski &

Majerus, 2018; Walker & Hulme, 1999), and with semantically related words (e.g. leaf – tree – branch) being better recalled as compared to semantically unrelated words (e.g. hand – table – goat). Language-based models of WM (Acheson & MacDonald, 2009; Majerus, 2013, 2019; Martin et al., 1996; Schwering & Macdonald, 2020) interpret these results as reflecting the direct involvement of the linguistic system; items associated with richer or more robust representations receive stronger feedback stabilizing activation (Dell et al., 1997; Kowialiewski & Majerus, 2020; Yap et al., 2015), leading to increased recall performance.

However, recent evidence suggests that the representation of serial order information in WM may not be completely independent from linguistic knowledge involved in the representation of item information. While lexical knowledge has been shown to strongly support item recall, with a robust recall advantage for word lists as compared to nonword lists, a few studies have shown that the availability of lexical knowledge may actually lead to impaired serial order recall, with a higher rate of serial order errors for word lists than nonword lists (Fallon et al., 2005; Guérard & Saint-Aubin, 2012; Kowialiewski & Majerus, 2018; Saint-Aubin & Poirier, 1999b). In addition, although rhyming lists of items enhance recall performance at the item level compared to non-rhyming lists, rhyming lists are also characterized by more errors at the serial order level (Fallon et al., 2005; Gupta et al., 2005). Similarly, lists composed of words sharing semantic features (e.g. leaf – tree – branch) show in WM tasks superior overall recall performance at the item level but decreased performance at the serial order level as compared to semantically unrelated words (Poirier & Saint-Aubin, 1995; Saint-Aubin et al., 2005; Tse, 2009; Tse et al., 2011). Poirier et al. (2015) manipulated semantic relatedness by presenting triplets of semantically related items in the first half of to-be-remembered lists. In a control condition, the remaining items were semantically unrelated (e.g. <u>leaf</u> – <u>tree</u> – <u>branch</u> – wall – sky – dog). In the experimental condition, the fifth item was semantically related to a triplet in the first half of the list (e.g. leaf – tree – branch – wall – nature – dog). Compared to the control condition, the authors observed an increase of migration errors of the fifth item toward earlier serial positions, that is, towards the semantically related group of words.

These studies suggest that serial order representations, even if independent from linguistic knowledge considered to support maintenance of item information, nevertheless interact with linguistic knowledge. As noted, a few studies appeared to show an increase of serial order errors for recall of semantically related versus unrelated word lists but the effects were weak. Using a large sample (N = 152) of participants, Tse et al. (2011) found effect sizes around d = 0.278 (computed using formula 7 from Lakens, 2013). Similarly, the study reported by Poirier et al. (2015) is also associated with a similarly small effect (d = 0.262). This observation suggests that the impact of semantic knowledge on serial order processing may be anecdotal.

The aim of this study was to examine in an extended manner the interactions that may exist between semantic knowledge and the representation of serial order information in WM, by focusing on the semantic relatedness effect. Previous studies assessed the impact of semantic knowledge on the ability to recall serial order information, broadly speaking. These studies provide however little information on how exactly semantic knowledge supports and/or influences the maintenance of serial order information. In this study, we take one step further by analysing the influence of semantic knowledge in a fine-grained manner, via a detailed investigation of the pattern of serial order errors. In a first experiment, we manipulated the semantic relatedness effect by creating lists composed of semantically unrelated items or lists containing subgroups of semantically related items (e.g. *leaf* – *tree* – *branch* – cloud – sky – rain). The lists were presented for standard immediate serial recall. We examined detailed patterns of serial order errors that occurred during the recall of semantically grouped versus unrelated word lists. If semantic knowledge interacts with the processing of serial order information, migration errors should be constrained by the grouping pattern of semantic relatedness. We predicted that for semantically grouped word lists, serial order errors should involve more frequently displacements between two related words than between two unrelated words (i.e. recalling “branch – tree” instead of “tree – branch”).

Previous studies already manipulated grouping effects, such as temporal grouping (Hartley et al., 2016; Henson, 1996; Hitch et al., 1996). These studies showed that the insertion of a short temporal pause between items (e.g. cat – finger – sky – *pause* – wall – truck – grass) enhances overall recall performance compared to a condition without additional temporal pause. In addition, serial order errors typically involve increased displacements of items within the same temporal group (e.g. recalling “cat – sky – finger” instead of “cat – finger – sky”). If the effects observed for the semantic grouping manipulation reflect the effect of more general grouping mechanisms, then we should observe similar patterns of performance when manipulating temporal grouping. To further test the grouping account in a direct manner, we broke up the semantic groups by manipulating the presence of semantic categories in an interleaved fashion in a second experiment (e.g. *leaf* – cloud – *tree* – sky – *branch* – rain). In this context, the grouping account predicts that there should be little or no increase of displacements between semantically related items, as compared to the same positions for semantically unrelated lists.

Experiment 1

Method

Participants. Our resources allowed us to recruit 40 participants from the university community of the University of Liège. The final sample included thirty-nine undergraduate students (31 females) aged between 18 and 26 years (*M* = 21.31, *SD* = 2.31). All the participants were French-native speakers, reported no history of neurological disorder or learning difficulties. With the exception of one participant which did not meet these inclusion criteria, and was therefore excluded from the sample. All participants gave their written informed consent before starting the experiment. The experiment had been approved by the Ethic committee of the Faculty of Psychology, Speech and Language Therapy, and Education of the University of Liège.

Material. In the present experiment, we used a set of 180 auditory verbal stimuli consisting in monosyllabic words with a mean log-frequency of $M_{log} = 3.452$ and $SD_{log} = 1.629$. Triplets of words sharing similar semantic characteristics, such as “leaf – tree – branch”, had been selected based on the results of an online survey conducted on 136 participants (see Kowialiewski & Majerus, 2018). In that study, participants had been asked to rate the semantic relatedness characterizing pairs of words by using a scale ranging from 0 (not semantically related at all) to 5 (strongly semantically related). The semantic pairs were defined in an a priori manner and were drawn from a mix of taxonomic (e.g. figue [fig], fraise [strawberry], pêche [peach]) and thematic (e.g. toast [toast], pain [bread], beurre [butter]) categories. A previous study showed that taxonomic and thematic categories produce similar effects on item and serial order processing (Tse, 2009). On average, each pair was judged 75.25 times. Based on these results, we first constructed semantic groups of three words for which adjacent words consisted in pairs of words that received a score strictly superior to 3.5 of semantic relatedness in the online survey. In other words, we ensured that in each semantic group the first word was judged as highly related to the second word, and the second word as highly related to the third word. The stimuli were recorded by a French-speaking Belgian native male speaker and normalized to a duration of 375 milliseconds.

Design. The task consisted in the presentation of 60 lists of six words following a 2 × 2 experimental design with a 2-level semantic factor (related vs. unrelated) and a 2-level temporal grouping factor (grouped vs. ungrouped). According to the semantic factor, one half of the sequences was composed of semantically unrelated words (e.g. soir [evening], plante [plant], tête [head], force [strength], nombre [number], crabe [crab]) and the other half consisted in two groups of three semantically related words (e.g. nuage [cloud], ciel [sky], pluie [rain], tigre [tiger], hyène [hyena], lion [lion]). Concerning the temporal grouping factor, one half of the sequences were presented at a regular pace while the other half formed two groups of words separated by an increased inter-stimulus duration (see details below).

The creation of related and unrelated lists was performed by combining two semantic groups of three words while ensuring that the two semantic groups did not cover similar semantic categories. Using the same procedure, we also ensured that each word in a semantic group could not have a specific semantic relation with a word from the other semantic group in the same list. For the construction of the unrelated lists, we selected six words from our stimuli set by carefully ensuring that the words were not semantically related. Finally, each of the 180 stimuli occurred two times over the 60 trials, once in a related list and once in an unrelated list, but in different serial positions; the two occurrences of the same word had to be separated by at least two trials.

To rule out the presence of list effects on task performance, we generated eight different versions of the 60 lists. We first created four different versions of the 60 lists: For each version, the different semantic groups were combined in different pairings within a list. Next, we generated four additional versions by reordering the order of the words within the sequences of the four first versions; for semantically related sequences, this meant that only the order of the words within a semantic group was changed. No more than three lists of the same type (i.e., related, unrelated) could be presented successively.

Procedure. The lists were presented to participants through headphones at a comfortable sound level. The presentation of each list was announced via a numerical countdown starting at 3 at a pace of 500 milliseconds per digit and presented on the center of the screen. Participants were then required to listen carefully to the 6-word sequences and they had to recall the words in the correct serial order immediately after their presentation. They started recalling the list after a 440-Hertz sine wave tone of 100 millisecond duration presented after the last word of the target list. If participants did not remember the word for a given position, they were asked to say “blanc” (blank in French). To help participants in recalling all items and in their correct serial position, they were provided a 6-square horizontal grid and they indicated for each word they recalled its serial position on the grid by pointing to the corresponding square. They pressed the spacebar to move from one trial to the next. The participants’ responses were recorded via a digital audio recorder for later transcription.

The temporal grouping factor was implemented in the following manner. For the ungrouped lists, words were presented with an interstimulus interval (ISI) of 250 milliseconds. For grouped lists, the intra-group ISIs were of 100 milliseconds while the inter-group ISI (between the third and fourth item of the list) was of 750 milliseconds leading to a temporal separation of the first and second group. This procedure ensured that overall list presentation durations were identical for the two conditions (3500ms). Participants were presented three practice trials before starting the task. The practice trials contained semantically unrelated words presented at a regular presentation speed. Stimulus presentation was controlled via the OpenSesame software running on a desktop station computer (Mathôt et al., 2012).

The temporal grouping factor was blocked, with the 30 first trials being temporally ungrouped while the 30 last trials were temporally grouped, the two blocks being separated by a short pause. This was done to minimize the possibility that participants deploy grouping strategies already on the temporally ungrouped sequences (Farrell & Lewandowsky, 2004; Henson, 1999). The semantic grouping factor was manipulated within each temporal factor, with each temporal grouping condition containing the same number of semantically related and unrelated sequences.

Statistical analysis. We used Bayesian statistical analysis techniques as they have the advantage of not being influenced by the intention underlying data collection or by the use of optional stopping rules (Berger & Berry, 1988; Rouder, 2014). Furthermore, they test evidence for both the alternative (H_1) and the null (H_0) model — while classical frequentist statistical techniques test only evidence against the null model. Bayesian statistics also are not sensitive to the time of analysis (i.e. planned *versus* post hoc analysis) (Dienes, 2016).

In the present study we report results with a relative measure of evidence called the Bayes factor (BF) that quantifies the extent to which data are more likely to be observed under a model relative to another (e.g., H_1 versus H_0, see Jeffreys, 1961). For example, if after looking at the data H_1 is preferred over H_0 by a factor of 20, this indicates that the data are 20 times more likely to have been observed under H_1 than under H_0. When interpreting the strength of evidence associated to BF values, we used the following terminology (see Lee & Wagenmakers, 2014; Wagenmakers et al., 2017): BFs lesser than three, between three and 10, between 10 and 30, between 30 and 100, and higher than 100, were considered as representing anecdotal, moderate, strong, very strong, and decisive level of evidence, respectively. In Bayesian ANOVAs, we performed Bayesian model comparisons using a top-down testing procedure, which first computes the BF value for the most complex model possible (i.e. the model including all main effects and all possible interactions). The BF value for each term is then assessed by directly comparing the full model against the same model, but by dropping the term under investigation. To minimize error of model estimation, the number of Monte Carlo simulations generated was set to $N_{iterations} = 100{,}000$. For some critical contrasts of interest, we also report the 95% Bayesian Credible Intervals using the highest density intervals of the sampled posterior distribution of the model under investigation ($N_{iterations} = 100{,}000$). All the analyses were performed using the BayesFactor package (Morey & Rouder, 2014) implemented in R (R Development Core Team, 2008) using the default wide Cauchy prior distribution of $r = \frac{\sqrt{2}}{2}$ for both ANOVAs and Bayesian T-Tests.

We assessed the sensitivity of our statistical design using a Bayes Factor Design Analysis (BFDA). In BFDA, researchers can estimate the number of subjects required to obtain a given BF value for an priori defined effect size (Schönbrodt & Wagenmakers, 2018). This is done by using Monte Carlo simulations of Bayesian tests and by sequentially incrementing the sample size. In sequential designs, incrementing is performed until the desired BF value is reached. This procedure informs about the median number of participants required to reach the desired BF value. Using this procedure, the median number of participants required to obtain a value of $BF_{10} > 3$ with an a priori effect size of $\delta = 0.25$ and assuming a default Cauchy prior distribution for a Paired Sample T-Test is 16. With these parameters, 80% of the tests reached the upper boundary of $BF > 3$ with a sample size of 37 subjects. This analysis was performed using the web-application available on http://shinyapps.org/apps/BFDA/.

When presenting the data graphically, we reported the 95% confidence intervals for each mean. We follow the recommendations made by Baguley (2012). After correcting the data for between-subject variability (Cousineau, 2012; Morey, 2008), the confidence intervals of each mean j were computed using the following formula:

$$(1)\ \hat{\mu}_j \pm t_{n-1,1-a/2}\sqrt{\frac{2J}{4(J-1)}}\hat{\sigma}'_{\hat{\mu}_j}$$

where $\hat{\mu}_j$ is the j^{th} mean, $t_{n-1,1-a/2}$ is the two-tailed critical t value with $n-1$ degrees of freedom, J is the number of means included in the graph, and $\hat{\sigma}'_{\hat{\mu}_j}$ is the standard error of the j^{th} mean.

All the data and scripts have been made available on the Open Science Framework: https://osf.io/a4d29/.

Scoring procedure. To determine the relative impact of the different grouping variables on WM performance, we applied two different scoring procedures. First, we performed a *strict serial recall scoring* procedure, considering an item as correct only if it was recalled at the correct serial position. For instance, given the target sequence “Item1 – Item2 – Item3 – Item4 – Item5 – Item6” and the output sequence “Item1 – Item2 – Item4 – Item3 – blank – Item6”, items 1, 2 and 6 would be scored as correct. Second, we used an item scoring procedure by considering an item correct if recalled, regardless of its serial position in the list. In the example described above, items 1, 2, 3, 4 and 6 would be scored as correct.

We characterized the pattern of serial order errors by determining within-group transpositions. Within-group transposition errors reflect serial order errors occurring within a temporal or semantic group (e.g. transposing D and F in ABC-DEF, see **Figure 1**). We determined the proportions of within-group transposition as a function of semantic versus temporal groups, by dividing the number of within-group transpositions by the total number of transposition errors observed for a given participant, and this separately for each experimental condition. This way of computing the pattern of serial ordering errors implies that data sets with no serial ordering error were considered as invalid for this score as they would lead to a 0/0 score. These data points were therefore excluded from the analysis on serial order errors.

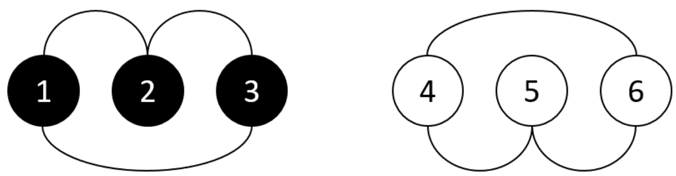

Figure 1. Graphic illustration of within-group transposition errors for Experiment 1. Items drawn the same semantic category have the same color. Items in different temporal groups are spatially separated.

Results and discussion

A first analysis assessed recall performance as a function of semantic relatedness (related vs. unrelated), temporal grouping (grouped vs. ungrouped) and serial position (position 1 through 6). Descriptive results are displayed in **Figure 2**. Using a Bayesian Repeated Measures ANOVA, we found for both item and strict serial recall criteria, decisive evidence supporting the presence of semantic relatedness, temporal grouping, and serial position effects (all BFs > 100). The interaction between semantic relatedness and temporal grouping was not clearly supported (BF_{10} = 3.433 and BF_{10} = 1.597 using an item and strict serial recall criterion, respectively). Decisive evidence supported the interaction between semantic relatedness and serial position (BFs > 100), and we observed robust evidence for the interaction between temporal grouping and serial position (BF_{10} = 82.57 and BF_{10} = 6.683e+6 using an item recall and strict serial recall criterion, respectively). The absence of triple interaction was supported by strong evidence (BF_{01} = 12.84) and anecdotal evidence (BF_{01} = 1.184), using the item and strict serial recall criteria, respectively.

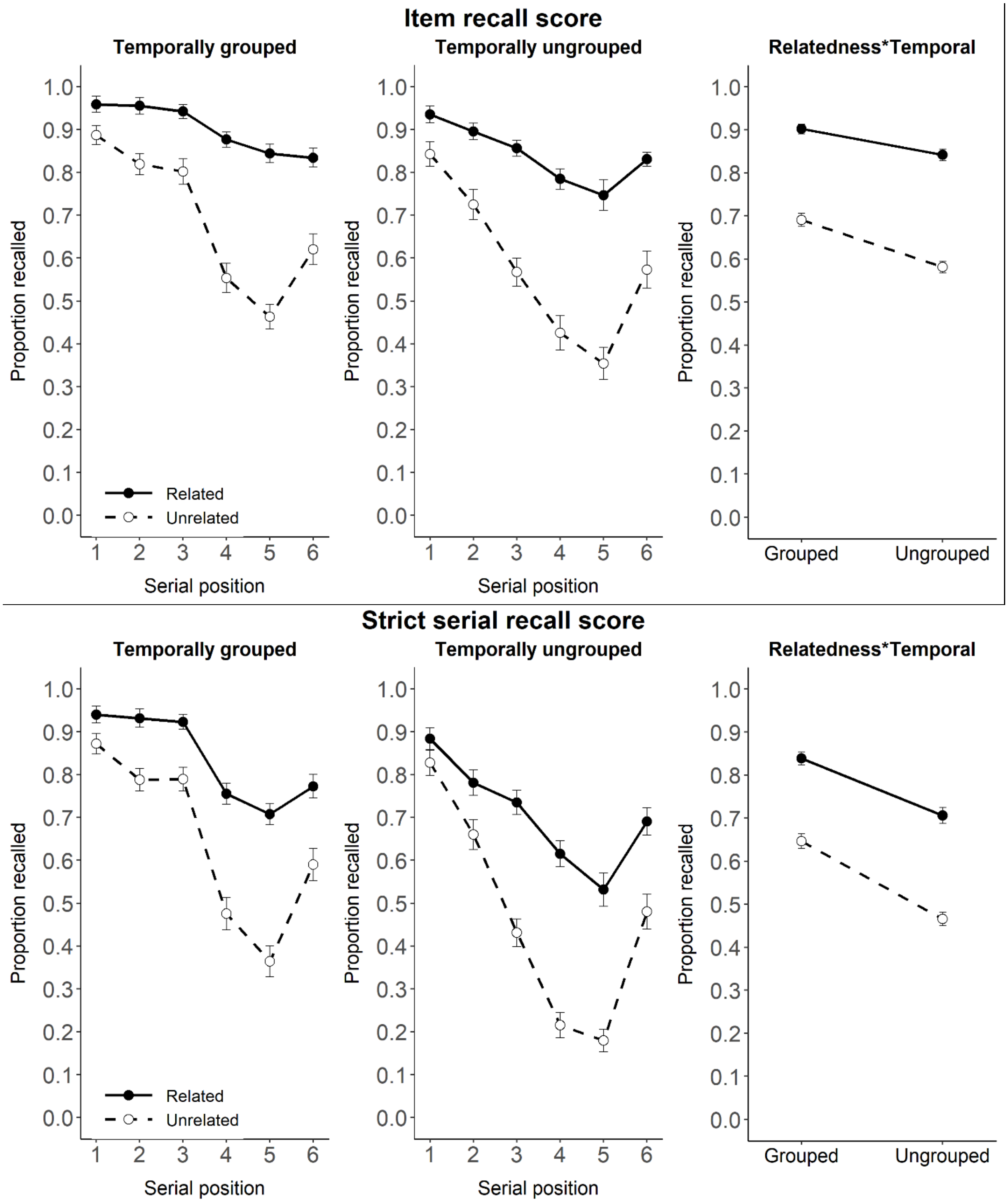

Figure 2. Proportion of items correctly recalled (y axis) as a function of serial position and

semantic relatedness (x axis), for temporally grouped (left panels) and ungrouped (middle panels) lists in Experiment 1. The right panels show the interaction between the temporal grouping and semantic conditions without the serial position factor. Upper panels: item recall criterion. Lower panels: strict serial recall criterion. Error bars represent 95% confidence intervals.

The most critical analyses were those assessing the effects of semantic relatedness and temporal grouping on the pattern of serial order errors. For these analyses, 6 data points were discarded due to some participants producing zero serial order errors, leading to a total of 150 observations. We assessed the effect of the two variables on the proportion of within-group transposition errors (see **Figure 3**). If both grouping manipulations stem from the same general mechanism, we expect them to increase within-group transpositions as compared to the ungrouped condition. We found decisive evidence supporting the presence of both semantic relatedness (BF_{10} = 7.669e+6) and temporal grouping (BF_{10} = 108.627) effects. In other words, when a serial order error occurred in grouped sequences, the vast majority of these errors involved the transposition of an item within its semantic or temporal group, as compared to equivalent positions in ungrouped sequences. The interaction term was supported by ambiguous evidence (BF_{10} = 2.492), but we nonetheless explored the impact of each condition using Bayesian Paired Sample T-Tests. The effect of semantic relatedness was moderately supported in the temporal grouping condition (BF_{10} = 3.41, $CI_{95\%}$ = [0.071; 0.772], d = 0.456, M_{diff} = 0.11) and was associated with decisive evidence in the condition without temporal grouping (BF_{10} = 5.175e+6, $CI_{95\%}$ = [0.76; 1.604], d = 1.244, M_{diff} = 0.277). Finally, the temporal grouping effect was supported in the semantically ungrouped condition (BF_{10} = 34.264, $CI_{95\%}$ = [0.205; 0.897], d

= 0.595, M_{diff} = 0.194), but not in the semantically grouped condition (BF_{01} = 4.328, $CI_{95\%}$ = [-0.21; 0.432], d = 0.122, M_{diff} = 0.028). These results show that both the semantic and temporal grouping manipulations credibly increased the presence of within-group transpositions, and this as compared to the same positions in the ungrouped condition.

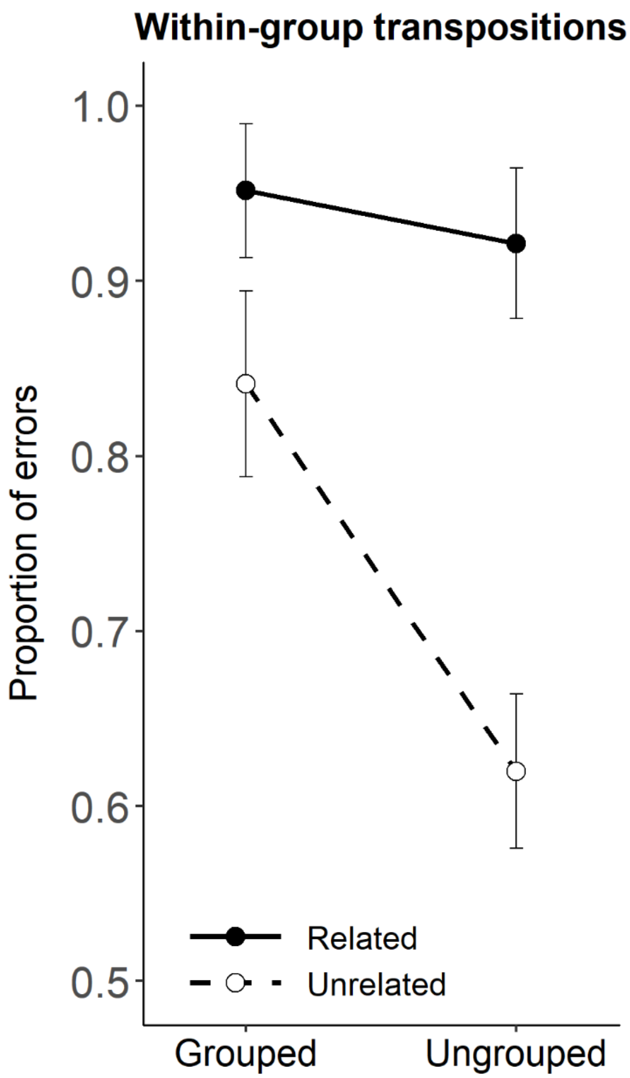

Figure 3. Proportion of within-group transpositions errors. Filled and dashed lines represent performance for related and unrelated lists, respectively. Error bars represent 95% confidence intervals.

To sum up, we observed an impact of both semantic relatedness and temporal grouping on item and serial order recall performance. The critical result of this experiment is that both the semantic and temporal grouping factors strongly constrained the pattern of transposition errors, and this is in a very similar way. When a transposition occurred, it involved serial positions associated to a same temporal or semantic group up to ~90% of the time (see **Figure 3**). Importantly, since the temporal grouping manipulation was blocked and always appeared in the second half of the experiment, the semantic grouping effect could not have been primed by the temporal manipulation itself. The fact that both grouping manipulations increased the presence of within-group transpositions suggests that they could stem from a same general grouping mechanism, and furthermore supports approaches considering that linguistic knowledge information also interacts with the processing of serial order information.

In the next experiment, we tested the grouping account in a direct manner. We used a new manipulation in which we broke up the semantic grouping structure, while keeping overall semantic relatedness unaffected. If the pattern of serial order errors observed in Experiment 1 was caused by a general grouping mechanism, then breaking up the semantic groups while leaving overall semantic relatedness unaffected should dramatically change the overall pattern of serial order errors.

Experiment 2

Experiment 1 manipulated the semantic relatedness effect by using semantic categories presented in subgroups (e.g. *leaf* – *three* – *branch* – <u>cloud</u> – <u>sky</u> – <u>rain</u>). In Experiment 2, the same manipulation was performed, but this time by presenting the categories in an interleaved fashion (e.g. *leaf* – <u>cloud</u> – *tree* – <u>sky</u> – *branch* – <u>rain</u>). Because the semantically related items were not grouped anymore, this new manipulation allowed us to test whether the constrains produced by semantic relatedness on serial order errors is compatible with a grouping account.

One potential problem with this new manipulation is that the transposition errors occurring between semantically related items now involved displacements of distance 2. It is well-known that transposition errors decrease exponentially with between-position distance (Hurlstone & Hitch, 2015). Due to this property, transposition errors between semantically related items could be more difficult to observe. To control for this potentially confounding factor, we took advantage of a phenomenon robustly observed when manipulating temporal grouping effects: the presence of *interposition errors* (Henson, 1996, 1998). These errors involve transpositions occurring between two items of different groups, but keeping the same relative within-group position (e.g. transposing “B” and “D” in “AB-CD-EF”). In a temporally grouped condition involving three groups composed of two items, these errors correspond to displacements of distance 2 and 4. If we are able to reproduce the increase of interposition errors in a temporally grouped condition, then there should be enough room to observe transposition errors between semantically related items in the semantically interleaved condition.

The changes in the manner semantic relatedness and temporal grouping were manipulated in this experiment imply that there is no distinction between within-semantic group transposition errors and between-temporal group interposition errors. Indeed, given the sequence “leaf – sky – *pause* – tree – cloud”, transposing “tree” and “leaf” is a within-semantic category transposition, but also a between-temporal group interposition. From now on, we will refer to these errors simply as *interposition* errors, reflecting the impact of the semantic variable when comparing the semantically related to the unrelated conditions, but also the impact of the temporal grouping variable when comparing the temporally grouped to the temporally ungrouped conditions.

Method

Participants. We recruited thirty-nine participants to equalize the number of participants recruited in Experiment 1. The participants were undergraduate students (26 females) aged between 18 and 30 years (*M* = 20.67, *SD* = 2.86). They were recruited from the university community of the University of Liège. All the participants were French-native speakers and reported no history of neurological disorder or learning difficulties and gave their written informed consent before starting the experiment. The experiment had been approved by the Ethics committee of the Faculty of Psychology, Speech and Language Therapy, and Education of the University of Liège.

Material. The material used was identical as in Experiment 1.

Design & procedure. The semantic relatedness manipulation involved now two different semantic categories presented in an interleaved fashion (e.g. leaf – *cloud* – tree – *sky* – branch – *rain*) in the related condition. In addition, there were three instead of two temporal groups in the temporally grouped sequences. As in Experiment 1, words in the ungrouped lists were presented with an ISI of 250 ms. For grouped lists, the ISI was 100 ms inside each temporal group. A longer ISI of 375 ms separated the second and the third words, as well as the fourth and the fifth words, leading to three two-stimulus temporal groups.

Scoring procedure. As already noted, the changes in the manner semantic relatedness and temporal grouping were manipulated in Experiment 2 imply that there is no distinction between within-semantic group transposition errors and between group interposition errors identified in Experiment 1 (see **Figure 4**). In Experiment 2, these errors will all be referred to as interposition errors reflecting either the impact of the semantic relatedness variable (when comparing the semantically related to the unrelated condition) or the temporal grouping variable (when comparing the temporally grouped to the temporally ungrouped condition). Interposition errors were considered to occur between positions 1 – 3 – 5 or positions 2 – 4 – 6. All other aspects of scoring procedures were identical to Experiment 1.

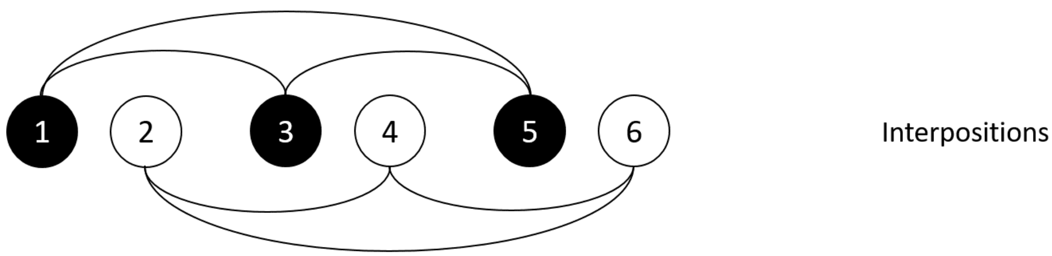

Figure 4. Graphic illustration of interposition errors for Experiment 2. Items from the same semantic category share the same color. Items in different temporal groups are spatially separated.

Results and discussion

First, we assessed the impact of semantic relatedness (related vs. unrelated) and temporal grouping (grouped vs. ungrouped) on the proportion of items recalled across serial positions (1 through 6). Using Bayesian Repeated Measures ANOVAs we observed decisive evidence supporting the three main effects of semantic relatedness, temporal grouping, and serial position (all BFs > 100). As can be seen in **Figure 5**, both semantic relatedness and temporal grouping manipulations led to increased recall performance. There was an overall absence of interaction between semantic relatedness and temporal grouping manipulations, which was supported by moderate evidence (BF_{01} = 8.382 and BF_{01} = 7.545 using an item recall and strict serial recall criterion, respectively). There was no clear evidence supporting an interaction between semantic relatedness and serial position using an item recall criterion (BF_{10} = 1.728), and strong evidence supporting the absence of this interaction using a strict serial recall criterion (BF_{01} = 36.987). The interaction between temporal grouping and serial position was supported by decisive evidence using an item recall (BF_{10} = 1.09e+5) and a strict serial recall criterion (BF_{10} = 1.597e+8). Strong evidence against the triple interaction was found using both recall criteria (BF_{01} = 44.41 and BF_{01} = 35.953, for item and strict serial recall criteria, respectively).

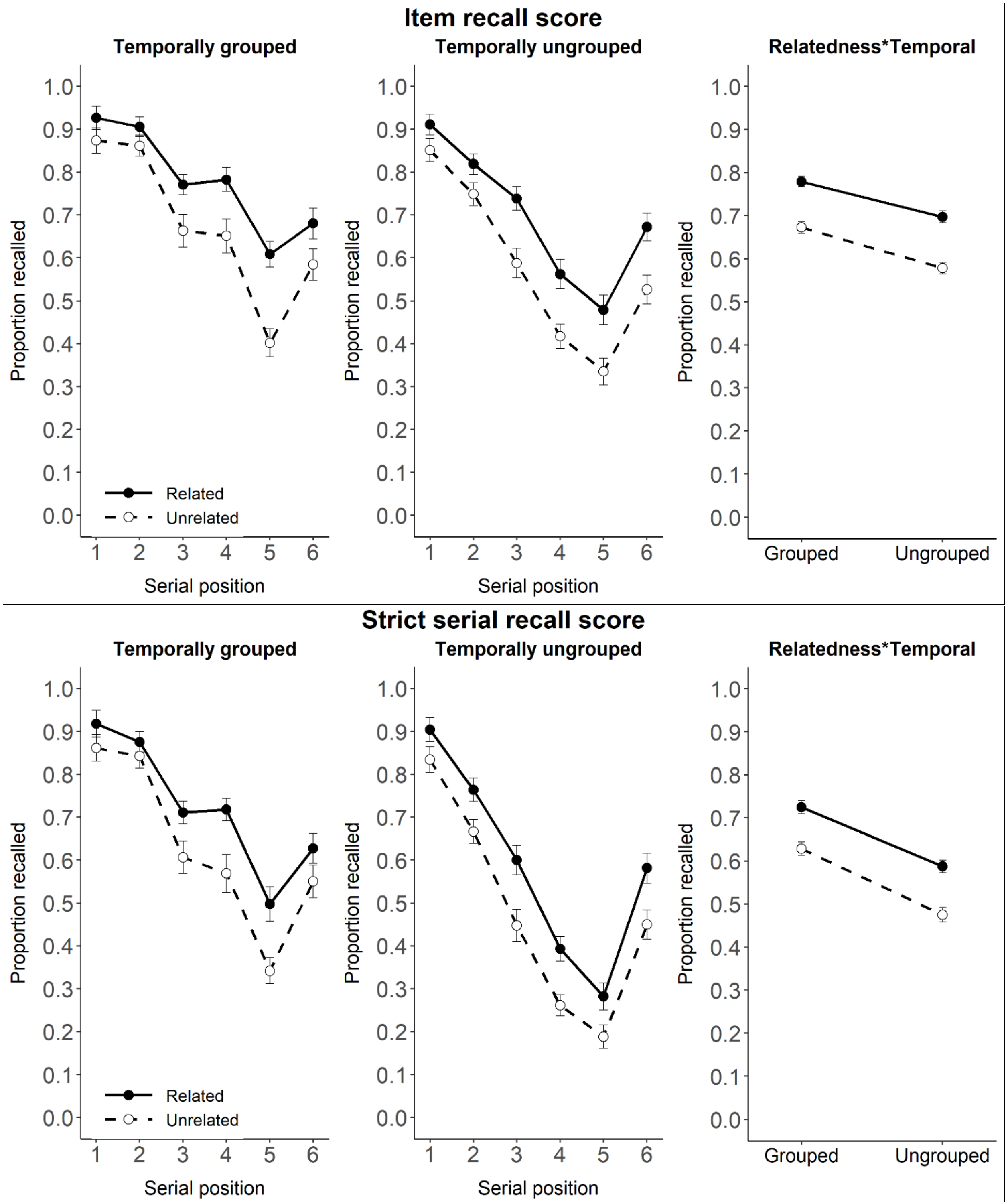

Figure 5. Proportion of items correctly recalled (y axis) as a function of serial position and semantic relatedness (x axis), for temporally grouped (left panels) and ungrouped (middle panels) lists in Experiment 2. The right panels show the interaction between the temporal

grouping and semantic conditions without the serial position factor. Upper panels: item recall criterion. Lower panels: strict serial recall criterion. Error bars represent 95% confidence intervals.

Next, we assessed the impact of semantic relatedness and temporal grouping manipulations on the proportion of interposition errors and this using a Bayesian Repeated Measures ANOVA. Note that, as in Experiment 1, some participants produced zero serial order errors in at least one of the experimental conditions. This resulted in the exclusion of 12 data points, leading to a total of 144 observations included in the analysis. If the increased migration errors between semantically related items observed in Experiment 1 stem from a general grouping mechanism, then the influence of semantic knowledge on within-group transpositions should be reduced or may even disappear when breaking up the semantic groups. Contrary to Experiment 1, the effect of interleaved semantic relatedness on interposition errors was ambiguous (BF_{10} = 1.412). Thus, there was no clear evidence supporting an increase of serial order errors between related words. However, we observed decisive evidence supporting the presence of a temporal grouping effect (BF_{10} = 4087.963). As can be seen in **Figure 6**, interposition errors were more frequent in temporally grouped sequences compared to ungrouped ones. The interaction term was associated with ambiguous evidence (BF_{01} = 2.681). Next, we explored the impact of each condition using Bayesian Paired Samples T-Tests. The impact of semantic relatedness was associated with ambiguous evidence in the temporally grouped condition (BF_{01} = 1.313, $CI_{95\%}$ = [-0.064; 0.667], d = 0.33, M_{diff} = 0.129). Moderate evidence supported the absence of semantic relatedness effect in the temporally ungrouped condition (BF_{01} = 3.464, $CI_{95\%}$ = [-0.151; 0.466], d = 0.169, M_{diff} = 0.044). In the semantically related condition, the impact of the temporal grouping manipulation was associated with very strong evidence (BF_{10} = 49.11, $CI_{95\%}$ = [0.246; 1.004], d = 0.672, M_{diff} = 0.248). In the semantically unrelated condition, the temporal grouping effect also impacted interposition errors, and was supported by strong evidence (BF_{10} = 10.83, $CI_{95\%}$ = [0.153; 0.861], d = 0.541, M_{diff} = 0.153).

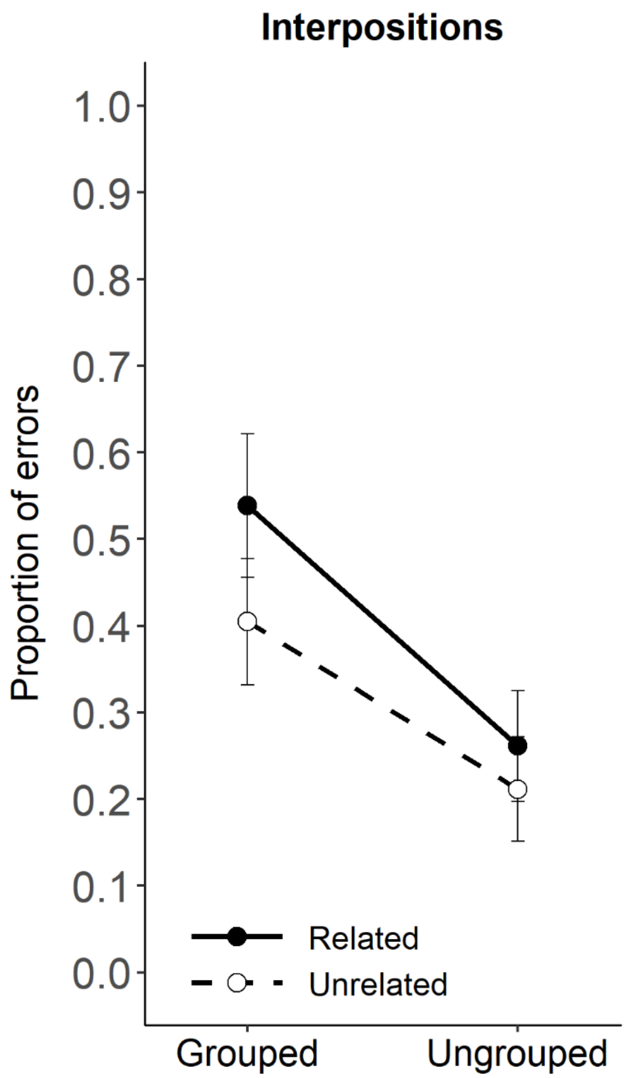

Figure 6. Proportion of interposition errors (y axis) as a function of semantic relatedness and temporal grouping. Filled and dashed lines represent performance for related and unrelated lists, respectively. Error bars represent 95% confidence intervals.

In a final analysis, we explored the differential impact of the semantically interleaved condition (Experiment 2) as compared to the semantically grouped condition (Experiment 1) on overall WM recall performance, but also on the pattern of serial ordering errors. We decided to include trials from ungrouped conditions only to discard the influence of temporal grouping. As can be seen in **Figure 7**, semantic relatedness had a stronger impact on overall recall performance under the grouped as compared to the interleaved condition. This stronger effect was supported by decisive evidence across both the item recall (BF_{10} = 2.175e+7, $CI_{95\%}$ = [1.038; 2.079], d = 1.63) and the strict serial recall (BF_{10} = 5.731e+4, $CI_{95\%}$ = [0.734; 1.729], d = 1.293) criteria, as assessed using a Bayesian Independent Samples T-Test. Similarly, the results from Experiments 1 and 2 show that semantic relatedness credibly impacted the ability to constrain serial ordering errors in the grouped but not in the interleaved condition, as also illustrated in **Figure 8**. Accordingly, decisive evidence supported a difference between both conditions, as indicated by a Bayesian Independent Samples T-Test (BF_{10} = 332.969, $CI_{95\%}$ = [0.415; 1.365], d = 0.958).

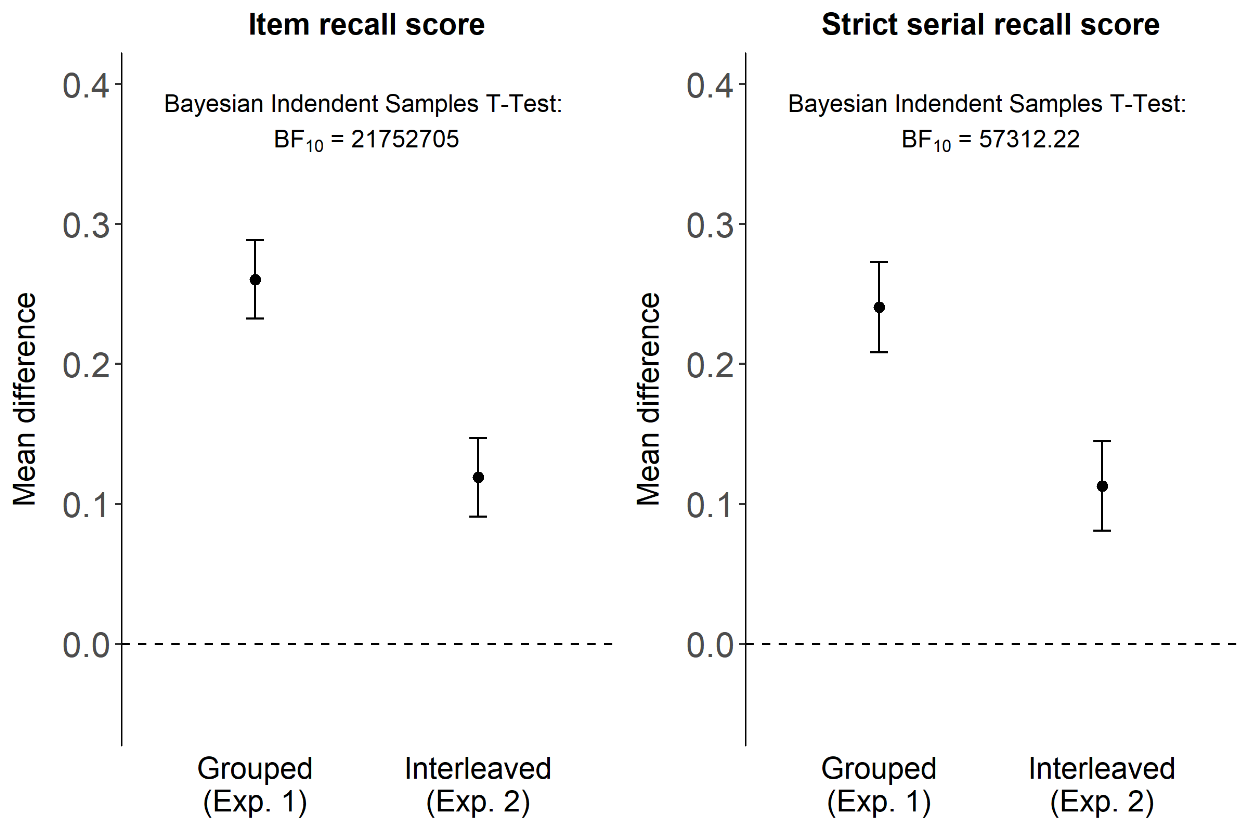

Figure 7. Direct comparison of recall performance between Experiment 1 and Experiment 2. The y axis represents the mean difference between related and unrelated lists in the temporally ungrouped conditions. Left panel: item recall criterion. Right panel: strict serial recall criterion. Error bars represent 95% confidence intervals around the mean.

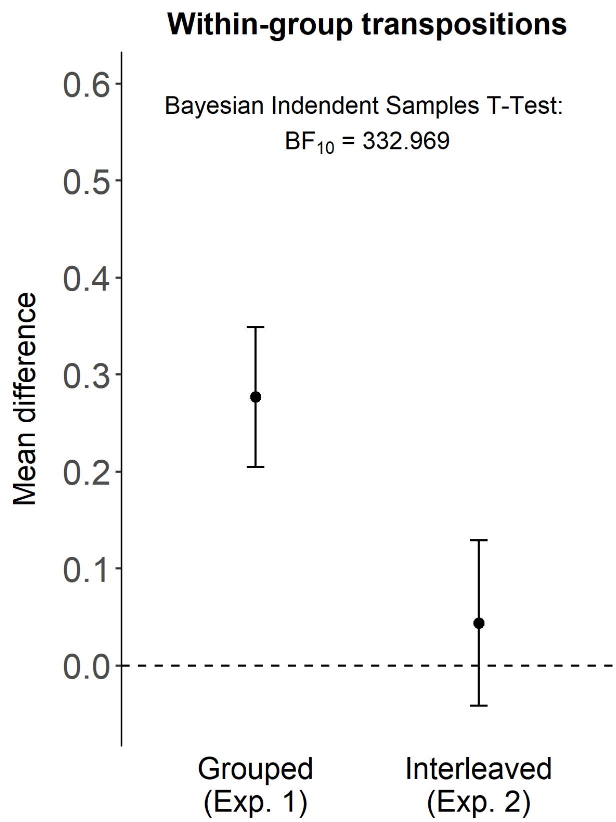

Figure 8. Direct comparison of the pattern of transposition errors between Experiment 1 and Experiment 2. The y axis represents the mean difference between related and unrelated lists in the temporally ungrouped condition. Error bars represent 95% confidence intervals around the mean.

To sum up, in Experiment 2, when presenting semantically related words in an interleaved manner, no increase of transposition errors was observed between semantically related items, contrary to Experiment 1. In addition, we found an increase of between-group interposition errors in the temporally grouped versus the temporally ungrouped condition. These results suggest that the increase of serial order errors observed between semantically related words observed in Experiment 1 was caused by the grouped presentation of semantically related items. This further strengthen the possibility that our semantic manipulation stems from a general grouping mechanism. Finally, we observed that the semantically grouped condition had a stronger impact on recall performance as compared to the semantically interleaved condition.

General discussion

This study assessed the impact of semantic knowledge on WM for serial order information by manipulating temporal grouping and the semantic relatedness of words presented for immediate serial recall. Overall, we observed a strong impact of semantic relatedness and temporal grouping on overall recall performance. Critically, semantic relatedness also impacted the pattern of serial order errors. In Experiment 1, we showed that when related lists contained two successive groups of semantically related items, the majority of transposition errors involved two items from the same semantic group. This effect disappeared when the semantically related words were presented in an interleaved manner in Experiment 2. Critically, this latter effect cannot be explained by the fact that transposition errors of distance 2 are less frequent than transposition errors of distance 1, because an increase of interposition errors of distance 2 was observed using a temporal grouping manipulation.

Does semantic knowledge impact serial order processing?

Whether semantic knowledge has a (detrimental) effect on serial order processing remains a controversial question (Saint-Aubin & Poirier, 1999a). As already noted by Saint-Aubin et al. (2005), the proportion of serial order errors usually points towards a detrimental effect although these effects tend to be small and are not always reliable at the statistical level (Tse et al., 2011). The present study demonstrates that the impact of semantic relatedness on serial order errors is a complex phenomenon. We show that when recalled at a wrong serial position, semantically related words migrate more often towards other related words, rather than towards another unrelated word. However, this was observed only when the semantic categories were presented in semantic sub-groups. Once the semantic categories were presented in an interleaved fashion, this increase of serial order errors between semantically related items was not credibly supported anymore. At the same time, the effect size of semantically-driven serial order transpositions observed in Experiment 1 was much stronger (d = 1.244) than the mean effect size observed in previous studies (d = 0.278 in the Tse et al., 2011 study and d = 0.262 in the Poirier et al., 2015 study), showing that semantic variables can also have a strong influence on serial order processing when assessed using more fine-grained methods.

How does semantic knowledge influence the retention of serial order information?

How can the effect of semantic grouping on serial order errors pattern be explained? One potential explanation boils down to a trace confusion account. According to this account, items tend to be confused because participants maintain the shared semantic content of items throughout the list and then reconstruct the degraded items based on these general semantic features (Saint-Aubin & Poirier, 1999a). This hypothesis could easily account for the increased serial order errors within semantically related lists observed in Experiment 1. However, this hypothesis, in its actual state, cannot account for the absence of a semantic relatedness effect on serial order errors observed in Experiment 2: trace reconstruction should lead to equally erroneous reconstruction of semantically related items also when those are presented in an interleaved manner.

Another possibility is that our data could be explained by a general grouping mechanism. Positional models of WM (Burgess & Hitch, 2006; Henson, 1998) explain the presence of temporal grouping effects by postulating that positional coding occurs at two different levels. The first positional coding occurs at the item level and is supposed to bind each item to a positional marker. The second one is at the group level and supposedly marks each group that compose the to-be-remembered sequence at a super-ordinate level. Due to the fact that they are represented by the same super-ordinate representation, items enclosed within the same temporal group are more similar to each other, and hence are more likely to be confused as compared to items from different temporal groups.

A similar mechanism may account for the semantic grouping effect we observed on serial ordering errors. The semantically related items could be compressed by activating and maintaining a broader and super-ordinate conceptual representation (Martin et al., 2018), leading to a reduction of WM load and better recall performance when items are presented in different subgroups. This mechanism will also lead to increased serial order errors within each subgroup, compared to other semantic groups. This is because information about the arbitrary serial order of items will get lost during the compression-decompression process of the different items from the semantic category. Recent studies suggest indeed that once a chunked representation is encoded as such, the different features that compose the chunk cannot be accessed anymore without additional attentional resources (Huang & Awh, 2018). This is however different when related items are presented in an interleaved fashion. In this case, the compression mechanism will not occur at all, or to a lesser extent; the broader semantic category will be less obvious at least for earlier presented items or might require additional controlled mechanisms to be detected. This leads to a reduced relatedness effect on overall recall accuracy, but also a reduction of serial order errors between semantically related items as compared to semantically grouped presentation. This is what we observed: the impact of the semantic relatedness dimension on recall performance was overall smaller as compared to the semantically grouped condition. Reduced inter-item relatedness effects on item recall performance as a function of within-list positional distance have previously been shown in immediate serial recall tasks (Saint-Aubin et al., 2014), which furthermore strengthen our own observations.

Finally, a more extreme account postulates the linguistic system itself as being responsible for serial order maintenance (Schwering & Macdonald, 2020), without the need for memory-specific serial order coding mechanisms. According to this account, serial order maintenance is an emerging product of the linguistic regularities occurring at each level of language processing, including sub-lexical, lexical, semantic, morphological but also syntactic. This predicts that the manipulation of linguistic properties at one of those levels should affect serial order errors. In fact, both the temporal and semantic grouping factors we manipulated could be the result of a general semantic and/or syntactic property of the linguistic system. The language system automatically regroups information as a function of semantic coherence, allowing for example the language system to predict the end of a sentence even if not yet completed: "The baby boy was very tired, the father put him to … [*bed*] " (Freunberger & Roehm, 2016). Similarly, the linguistic system automatically parses the speech input into hierarchical structures, even in the absence of explicit acoustic cues (Ding et al., 2016). In the case of the semantic grouping manipulation in this study, the detection of these semantic regularities may have prevented participants from confounding the items between two different semantic groups. Similarly, the pause in the temporal grouping manipulation may have induced suprasegmental encoding (i.e. at the sentence level), which may in turn prevent participants from confounding items in different temporal groups, because items in a same temporal group are automatically detected as being more similar between each other. This interpretation is supported by studies showing that WM recall performance and transposition errors is influenced by prosodic manipulations, such as pitch-accent type or stress (Tanida et al., 2015; Taylor et al., 2015). It is important to note that it is difficult to know whether this purely linguistic account could potentially predict the absence of semantic relatedness impact on the pattern of serial ordering errors as observed in the interleaved condition of Experiment 2. This is because purely linguistic models of WM are rare. The only model known so far that could potentially account for linguistic influences on serial order processing, without postulating the existence of specific serial order mechanisms, is the recurrent neural network model developed by Botvinick & Plaut (2006). However, without actually running this model, it is difficult to predict its behaviour, neural networks being more generally opaque and unpredictable. Note that the presence of such linguistic influence does not rule out the existence of more general serial order mechanisms (Hurlstone & Hitch, 2015). Instead, this shows the complexity of WM-linguistic interactions.

Conclusion

By using a semantic relatedness manipulation, we observed that semantic knowledge had a strong impact on the pattern of serial order errors, but only when semantically related items were presented in subgroups. These results could reflect the intervention of general grouping mechanisms, but could also stem from direct interactions between the processing of serial order mechanisms and linguistic knowledge. Overall, this study supports the existence of strong and robust WM-linguistic interactions.

References

Abrahamse, E., Dijck, J. Van, 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*. https://doi.org/10.3389/fnhum.2014.00932

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

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*(1), 64-88. https://doi.org/10.1016/j.jml.2006.02.002

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. https://doi.org/10.1371/journal.pone.0092049

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, *44*(1), 158-175. https://doi.org/10.3758/s13428-011-0123-7

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order : A recurrent neural network model. *Psychological Review*, *113*(2), 201-233. https://doi.org/10.1037/0033-295X.113.2.201

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.

Brock, J., & Jarrold, C. (2005). Serial order reconstruction in Down syndrome : Evidence for a selective deficit in verbal short-term memory. *Journal of Child Psychology and Psychiatry*, *46*(3), 304-316. https://doi.org/10.1111/j.1469-7610.2004.00352.x

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-Based Memory for Serial Order. *Psychological Review*, *107*(1), 127-181.

Burgess, N., & Hitch, G. J. (1999). Memory for Serial Order : A Network Model of the Phonological Loop and Its Timing. *Psychological review*, *106*(3), 551-581.

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

Campoy, G., Castellà, J., Provencio, V., Hitch, G. J., & Baddeley, A. D. (2015). Automatic semantic encoding in verbal short-term memory : Evidence from the concreteness effect. *The Quarterly Journal of Experimental Psychology*, *68*(4), 759-778. https://doi.org/10.1080/17470218.2014.966248

Cousineau, D. (2012). Erratum to Cousineau (2005) : Confidence intervals in within-subject designs : A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *8*(3), 182-182. https://doi.org/10.20982/tqmp.08.3.p182

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801-838.

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 78-89. https://doi.org/10.1016/j.jmp.2015.10.003

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158-164. https://doi.org/10.1038/nn.4186

Fallon, A. B., Mak, E., Tehan, G., & Daly, C. (2005). Lexicality and phonological similarity : A challenge for the retrieval-based account of serial recall? *Memory*, *13*(3/4), 349-356. https://doi.org/10.1080/09658210344000215

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies : Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*, 115-135. https://doi.org/10.1016/j.jml.2004.03.007

Freunberger, D., & Roehm, D. (2016). Semantic prediction in language comprehension : Evidence from brain potentials. *Language, Cognition and Neuroscience*, *31*(9), 1193-1205. https://doi.org/10.1080/23273798.2016.1205202

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 84-95. https://doi.org/10.1037/0278-7393.25.3.562

Gorin, S., Kowialiewski, B., & Majerus, S. (2016). Domain-generality of timing-based serial order processes in short-term memory : New insights from musical and verbal domains. *PLoS ONE*, *11*(12), 1-25. https://doi.org/10.1371/journal.pone.0168699

Guérard, K., & Saint-Aubin, J. (2012). Assessing the Effect of Lexical Variables in Backward Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(2), 312-324. https://doi.org/10.1037/a0025481

Guidali, G., Pisoni, A., Bolognini, N., & Papagno, C. (2019). Keeping order in the brain : The supramarginal gyrus and serial order in short-term memory. *Cortex*, *119*, 89-99. https://doi.org/10.1016/j.cortex.2019.04.009

Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall : The roles of type of similarity, category cuing, and item recall. *Memory & Cognition*, *33*(6), 1001-1016.

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

Henson, R. N. A. (1996). *Short-term memory for serial order*.

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

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

Henson, R. N. A., 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*, *56A*(8), 1307-1334. https://doi.org/10.1080/02724980244000747

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal Grouping Effects in Immediate Recall : A Working Memory Analysis. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *49*(1), 116-139. https://doi.org/10.1080/713755609

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports, 8*(1), 23. https://doi.org/10.1038/s41598-017-18157-5

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.

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.

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), 295-324. https://doi.org/10.1037/a0038223

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*(1), 81-98. https://doi.org/10.1016/j.jml.2005.08.001

Jeffreys, H. (1961). *Theory of Probability* (3rd éd.). Oxford University Press.

Kalm, K., & Norris, D. (2014). The Representation of Order Information in Auditory-Verbal Short-Term Memory. *Journal of Neuroscience, 34*(20), 6879-6886. https://doi.org/10.1523/JNEUROSCI.4104-13.2014

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. (2020). The varying nature of semantic effects in working memory. *Cognition*, *202*, 104278. https://doi.org/10.1016/j.cognition.2020.104278

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science : A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling : A Pratical Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lewandowsky, S., & Murdock, B. B. (1989). Memory for Serial Order. *Psychological Review*, *96*(1), 25-57.

Majerus, S. (2009). 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* (p. 244).

Majerus, S. (2013). Language repetition and short-term memory : An integrative framework. *Frontiers in Human Neuroscience*, *7*, 357. https://doi.org/10.3389/fnhum.2013.00357

Majerus, S. (2019). Verbal working memory and the phonological buffer : The question of serial order. *Cortex*, 1-12. 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*, NSYD1500380. https://doi.org/10.1016/j.neuropsychologia.2015.08.010

Majerus, S., D'Argembeau, A., Martinez Perez, T., Belayachi, S., Linden, M. Van Der, & Collette, F. (2010). The Commonality of Neural Networks for Verbal and Visual Short-term Memory. *Journal of Cognitive Neuroscience*, *22*(11), 2570-2593.

Majerus, S., Martinez, 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., 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. https://doi.org/10.1016/j.jml.2004.05.002

Martin, N., Minkina, I., Kohen, F. P., & Kalinyak-Fliszar, M. (2018). Assessment of linguistic and verbal short-term memory components of language abilities in aphasia. *Journal of Neurolinguistics*, *48*, 199-225. https://doi.org/10.1016/j.jneuroling.2018.02.006

Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia : Evidence for a relation between auditory–verbal STM capacity and lexical errors in repetition. *Brain and Language*, *52*(1), 83-113.

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. https://doi.org/10.1080/87565641.2016.1153098

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame : An open-source , graphical experiment builder for the social sciences. *Behavior research methods*, *44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

Morey, R. D. (2008). Confidence Intervals from Normalized Data : A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61-64.

Morey, R. D., & Rouder, J. N. (2014). *BayesFactor : Computation of Bayes factors for common designs (Version 0.9.2+)* (p. https://richarddmorey.github.io/BayesFactor/).

Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., Fava, E., & Bello, L. (2017). Mapping the Brain Network of the Phonological Loop. *Human Brain Mapping*, *00*, 00-00. https://doi.org/10.1002/hbm.23569

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words : Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology*, *48*(2), 384-404. https://doi.org/10.1080/14640749508401396

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. https://doi.org/10.3758/s13421-014-0470-6

R Development Core Team. (2008). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks : Implications for models of short-term memory. *The Quarterly Journal of Experimental Psychology*, *61*(2), 292-323. https://doi.org/10.1080/17470210601147747

Roodenrys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, L. M. (2002). Word-Frequency and Phonological-Neighborhood Effects on Verbal Short-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1019-1034. https://doi.org/10.1037//0278-7393.28.6.1019

Saint-Aubin, J., Guérard, K., Chamberland, C., & Malenfant, A. (2014). Delineating the contribution of long-term associations to immediate recall. *Memory*, *22*(4), 360-373. https://doi.org/10.1080/09658211.2013.794242

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. https://doi.org/10.3758/BF03196364

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

Saint-Aubin, J., & Poirier, M. (1999b). The influence of Long-term Memory Factors on Immediate Serial Recall : An Item and Order Analysis. *International Journal of Psychology*, *34*(5-6), 347-352.

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

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, *21*(2), 168-175.

Schwering, S. C., & Macdonald, M. C. (2020). Verbal Working Memory as Emergent from Language Comprehension and Production. *Frontiers in human neuroscience*, *14*, 1-19. https://doi.org/10.3389/fnhum.2020.00068

Tanida, Y., Ueno, T., Lambon Ralph, M. A., & Saito, S. (2015). The roles of long-term phonotactic and lexical prosodic knowledge in phonological short-term memory. *Memory & Cognition*, *43*(3), 500-519. https://doi.org/10.3758/s13421-014-0482-2

Taylor, J. C., Macken, B., & Jones, D. M. (2015). A matter of emphasis : Linguistic stress habits modulate serial recall. *Memory & Cognition*, *43*(3), 520-537. https://doi.org/10.3758/s13421-014-0466-2

Tse, C.-S. (2009). The role of associative strength in the semantic relatedness effect on immediate serial recall. *Memory*, *17*(8), 874-891. https://doi.org/10.1080/09658210903376250

Tse, C.-S., & Altarriba, J. (2007). Testing the associative-link hypothesis in immediate serial recall : Evidence from word frequency and word imageability effects. *Memory*, *15*(6), 675-690. https://doi.org/10.1080/09658210701467186

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. https://doi.org/10.1080/17470218.2011.604787

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1323-7

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.

Yap, M. J., Lim, G. Y., & Pexman, P. M. (2015). Semantic richness effects in lexical decision : The role of feedback. *Memory & Cognition*, *43*(8), 1148-1167. https://doi.org/10.3758/s13421-015-0536-0