How robust is the link between working memory for serial order and lexical skills in children?

Lucie Attout [a], Coline Grégoire [a] & Steve Majerus [ab]

[a] Psychology and Neuroscience of Cognition Research Unit, University of Liège

[b] Fund for Scientific Research FNRS, Brussels, Belgium

Published in Cognitive Development

DOI: 10.1016/j.cogdev.2020.100854

Corresponding author:

Dr. Lucie Attout, Psychology and Neuroscience of Cognition Research Unit, Université de Liège; Boulevard du Rectorat, B33, 4000 Liège – Belgium; Tel: 0032 4 3662226; Email: lucie.attout@uliege.be

ACKNOWLEDGMENTS

This work was supported by a grant from Fonds de la Recherche Scientifique – FRS – FNRS (PDR, FRFC, T.1003.15, Belgium). We would like to thank all the children and their school directors, teachers, pupils, and parents for their time and effort invested in this study. There is no conflict of interest in connection with this work.

Abstract

The link between verbal working memory (WM) and vocabulary development has been explored extensively. At the same time, the vast majority of studies in this field used lexical tasks that generally involved a high WM demands, leading to an unclear understanding of this link. The present study re-explored the link between WM for serial order, WM for item information and lexical abilities by administering, to 92 children aged 4-to-6 years, both standard receptive vocabulary tasks with a high WM demands and single picture naming tasks with minimal WM demands. Analyses provided strong evidence for a specific link between serial order WM and both vocabulary measures, with a particularly important link with the rare noun subtask and the absence of link with verbs. These results suggest that the link between lexical abilities and verbal WM in young children is robust and not inflated by the WM demands of specific vocabulary tasks.

150 words

Keywords: short-term memory, serial order, language, vocabulary development

INTRODUCTION

An extensive set of studies has highlighted a consistent association between estimates of verbal working memory (WM) abilities and lexical knowledge and learning abilities, in both children and adults (e.g. Avons, Wragg, Cupples, & Lovegrove, 1998; Gathercole, Willis, Baddeley, & Emslie, 1994; Gathercole, Willis, Emslie, & Baddeley, 1992; Majerus, Poncelet, Van der Linden, & Weekes, 2008; Papagno, Valentine, & Baddeley, 1991; Service, 1992). More recent studies have suggested that it is mainly the serial order component of verbal WM tasks that is causally related to lexical abilities, based on the assumption that learning word forms involves the learning of novel serial orderings of a limited number of phonemes that characterize a given language (Gupta & MacWhinney, 1997; Majerus, Poncelet, Greffe, et al., 2006; Szmalec et al., 2012). The aim of the present study is to examine the robustness of this claim, by examining the possibility that the association between verbal WM and lexical abilities may have been inflated by the intrinsic WM load of the receptive vocabulary tasks most commonly used in previous studies.

One of the first studies assessing the association between verbal WM and lexical abilities was a study conducted in children presenting developmental language impairment. Gathercole and Baddeley (1989) observed disproportionate verbal WM impairment in this language impaired children group, even relative to a younger control group matched on language abilities. Next, a series of correlational studies in typically developing children aged 4 to 13 years showed specific longitudinal and cross-sectional associations between verbal WM, mainly assessed by a nonword repetition task, and receptive vocabulary (Baddeley et al., 1998; Gathercole & Baddeley, 1990; Gathercole et al., 1992). These results were interpreted as reflecting a causal role of verbal WM in language acquisition (Baddeley et al., 1998). However, it could be argued that verbal WM tasks are also linguistic processing tasks, and hence, that language abilities drive performance in WM tasks (e.g., Metsala, 1999). Nonword repetition is indeed not only a test of short-term retention abilities for phonological information, but also measures phonological segmentation abilities as well as access to lexical and sublexical phonological knowledge structures (Gathercole, 2006). In order to distinguish between these two possible interpretations, two levels of processing in verbal WM tasks need to be considered: the retention of item information (i.e., the phonological and lexico-semantic characteristics of the items to be maintained) and serial order information (i.e., the serial order in which the items were presented) (Lee & Estes, 1981; Majerus, Poncelet, Greffe, et al., 2006). We use the term 'working memory' in a functional manner and refer to any situation involving the temporary retention of information, in line with the most common usage of this term (see Cowan, 2017, for an extensive discussion of the difficulties associated with the definition of the term 'working memory'). By using this term, we want to stress the fact that WM tasks discussed here involve the short-term retention and restitution of verbal information. The dissociation between item and order processes in WM is supported by a series of behavioral, neuropsychological, developmental and neuroimaging studies (see for example Attout, Ordonez Magro, Szmalec, & Majerus, 2019; Majerus, Norris, & Patterson, 2007b; Saint-Aubin & Poirier, 1999). This dissociation is also implemented by the majority of recent theoretical models of WM (see for example Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Gupta, 2003; Henson, 1998). While linguistic knowledge supports the recall of verbal items, it has less influence on the retention of the typically arbitrary serial order in which the words are presented (see Majerus, 2013 for a review). Hence, if there is a link between short-term retention abilities and language development, then this link should be strongest for WM tasks maximizing the retention of serial order information. On the other hand, if the link between performance on verbal WM tasks and language development is to be explained by access to long-term linguistic structures, then this link should be strongest for

WM tasks maximizing access to verbal item information. Studies adopting the item-serial order distinction (Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006; Majerus et al., 2008; Ordonez Magro et al., 2018) confirmed that there is a specific link between serial order WM capacity and lexical language abilities, either measured by current vocabulary knowledge or by novel word learning tasks. Moreover, their differential predictive power of vocabulary development has been cross-validated with other types of WM tasks. For example, Majerus et al. (2009) showed that item and serial order WM estimates based on the calculation of the proportion of item and serial order errors based on single WM task led to the same conclusions as the tasks using the single nonword delayed repetition and the serial order reconstruction tasks used in the present study: a more robust prediction of vocabulary development by the serial order WM estimate than by the item WM estimate. It was argued that learning language essentially amounts to learning of sequential information and dependencies, and that short-term retention abilities for sequential information are critically involved in the initial steps of learning new verbal sequences (Majerus, Poncelet, Elsen, et al., 2006; Ordonez Magro et al., 2018).

However, the reverse situation, the fact that the association between verbal WM and language measures could actually be driven by the WM demands of the tasks used to measure linguistic abilities, has not yet been considered. Indeed, the task most typically used to measure vocabulary knowledge in studies assessing the association between verbal WM and lexical language abilities is the Peabody Picture Vocabulary task (PPVT; Dunn, 1981). In this task, the content of four pictures has to be identified and the corresponding lexical entries have to be activated; they then have to be compared to the auditorily presented target word that is being maintained in WM; furthermore, the comparison processes requires continuous maintenance of the target item until a response has been made. The PPVT is furthermore often considered as a verbal reasoning task (Vance et al., 1989) as many of the items require analogical or other forms of reasoning (as for example required when associating the word 'adjustable' to the picture of a belt in one of the PPVT items) and analogical reasoning has been shown to involve working memory resources (Cho et al., 2007; Morrison et al., 2011). A similar comment can be made for studies investigating the association between verbal WM abilities and novel word learning (Gathercole et al., 1997; Majerus & Boukebza, 2013): the novel word learning paradigms used are typically paired associate word-nonword or word-object learning paradigms that are structurally very similar to verbal WM tasks. In these tasks, one or several verbal labels have to be encoded and maintained for subsequent immediate recall during the initial learning stage and hence also involve WM demands. Hence, another possible interpretation of the associations observed between verbal WM and lexical development is that this association reflects, at least partially, the high storage and executive control demands of multi-picture language measures such as the PPVT.

In order to examine this hypothesis, it is necessary to assess language abilities by using language tasks that have the smallest possible WM load. Single word picture naming tasks satisfy this constraint by requiring the activation and production of a single lexical form based on the presentation of a picture of the target word. Indeed, in a single picture naming task, a single picture has to be identified and the corresponding lexical entry has to be retrieved and produced; this is a direct, single-step process involving no temporary maintenance of target information or alternative solutions. Furthermore, single word picture naming tasks provide maximal sensitivity as they require a full production of the target item and the possibility of (correct) guessing responses is minimized (while it is 25% in a four picture receptive vocabulary task). Only one previous study used a single word naming task in addition to the PPVT in eight-year-old bilingual children and observed a significant correlation with a forward digit span measure and the single word naming task (Lee Swanson et al., 2015). However, the study had not been designed to directly compare WM associations with expressive and receptive vocabulary measures. Other studies explored the relationship between verbal WM and productive vocabulary using tasks requiring the production of verbal definitions of a concept represented by a picture and showed an association of performance on this task and verbal WM abilities (Edwards et al., 2004; Henry & MacLean, 2003; Swart et al., 2017). These results are however again difficult to interpret as definition tasks require executive and attentional control abilities also involved in WM tasks, in addition to access to multiple levels of language processing. An illustration of the importance of considering the WM demands of vocabulary measures comes from a recent study in individuals with Down syndrome, a genetic syndrome caused by chromosome 21 triplication. People with Down syndrome are characterized by particularly poor verbal WM abilities, their mean digit span rarely exceeding three items at an adult age (Iacono et al., 2010). Majerus and Barisnikov (2018) administered both standard multi-picture receptive vocabulary and single word naming tasks and observed a robust correlation between verbal WM abilities and vocabulary knowledge when estimated with the PPVT (having an intrinsic WM load) but not when estimated with single picture naming tasks (no WM load). That study was conducted only in individuals with a neurodevelopmental disorder and hence it is impossible to determine whether the observed results are specific to Down syndrome (Mosse & Jarrold, 2011; Næss et al., 2015) or whether they reflect a more general situation.

The aim of the present study was to assess the association between verbal WM and both multi-picture receptive vocabulary and single picture productive vocabulary abilities in typically developing children. The link between WM abilities and vocabulary development, as assessed by a receptive vocabulary task or by novel word learning tasks, has been well established in previous studies (Gathercole et al., 1991; Leclercq & Majerus, 2010). The goal of this study was to examine whether the link commonly observed between different types of verbal serial order WM tasks (e.g., immediate serial recall like word span, serial order reconstruction) and receptive vocabulary measures is due to the fact that these tasks are multi-determined tasks sharing common general cognitive processes (e.g., attentional control, resistance to interference, short-term retention) or whether a link between these WM measures and lexical development is still observed when a much less multi-determined task is used for assessing lexical development such as a single picture naming task. This methodological research question is also of major theoretical importance as it allows us to determine whether there is a genuine link between WM capacity and lexical development, as assumed by influential models of WM (Baddeley et al., 1998; Gathercole et al., 1992) or whether this link is an artefact stemming from the specific task requirements of the receptive vocabulary tasks most commonly used to investigate the links between WM and lexical development.

We furthermore distinguished between verbal WM tasks that maximize either item or serial order retention requirements given that previous studies showed that verbal WM tasks maximizing the retention of serial order information are particular robust predictors of vocabulary development (e.g., Majerus, Poncelet, Greffe, et al., 2006; Majerus, Poncelet, Van der Linden, & Weekes, 2008b; Ordonez Magro, Attout, Majerus, & Szmalec, 2018). To estimate the retention of phonological item information, we used a single nonword delayed repetition task with nonwords being new on each trial (thereby minimizing serial order retention requirements and maximizing the activation of phonological codes) and to assess the retention of the serial order, we used a serial order reconstruction task of highly familiar words sampled from a closed pool (thereby minimizing item retention requirements).

The tasks had been validated in previous studies regarding the item and serial order components they are supposed to measure (Majerus et al., 2006; Leclercq & Majerus, 2010; Majerus & Boukezba, 2016; Majerus, Poncelet, Van der Linden, & Weekes, 2008). For example, Majerus et al. (2006) showed that performance on the serial order reconstruction task correlates with serial order errors in an immediate serial recall task and an adapted version of the single nonword delayed repetition task (single item probe recognition task) correlates with item errors in an immediate serial recall task. Hence, despite the structural differences between the two WM tasks used in the present study, previous studies have shown that the two tasks are valid estimates of item and serial order WM capacity. Moreover, we should note that the reason for using a closed pool of words in the serial order reconstruction and an open set of nonwords in the nonword repetition task was that this manipulation allowed to minimize item processing requirements in the former task (items being highly familiar and predictable and only their serial order changing) and to maximize phonological item processing requirements in the latter task (items being highly unfamiliar and novel on each trial, necessitating a detailed encoding of the phonological characteristics of the items).

We used the EVIP receptive vocabulary scale (Dunn & Theriault-Whalen, 1993) which is the French adaptation of the PPVT. For productive vocabulary, we administered a single picture naming task part of the ISADYLE standardized language assessment battery (Piérart et al., 2010). The productive vocabulary task measured vocabulary knowledge for words of similar semantic and grammatical categories as the EVIP receptive vocabulary task. The productive vocabulary task measured vocabulary knowledge for words of similar semantic and grammatical categories as the EVIP receptive vocabulary task. The productive vocabulary task we selected involved naming of lexically frequent and concrete nouns, lexically infrequent nouns (rare words) and verbs. Although we did not have specific expectations as regards the correlations between the WM and the productive vocabulary task as a function of lexical category, we may nevertheless expect a stronger association between the WM and the infrequent noun category as opposed to the frequent noun category, the words of the latter category being potentially overlearned even in children with lower WM capacities due to the frequent occasions on which these words occur.

Both vocabulary measures involved not only WM load but also inhibition mechanisms. On the one hand, the retrieval of a word in a picture naming task could be subject to interference stemming from lexical competitors of the target word or from previously named pictures. At the same time, inhibitory requirements are likely to be even stronger in the receptive vocabulary task where, in addition to interference stemming from previous trials, children have to co-activate, for a given trial, the lexico-semantic content of the four pictures of the probe display and select only the picture providing the best match with the target word while inhibiting the selection of any of the other three pictures.

The tasks were administered to young children aged 4- to 6-years in which associations between performance on verbal WM and lexical measures have been most frequently studied and demonstrated.

METHOD

Participants

Valid data were obtained from 92 children aged 4- to 6-years old (53 girls; mean age: 5;7 years; range: 4;8–6;6). Thirteen additional children had been recruited, but their data could not be retained because they did not meet the inclusion criteria (see below) or because data were incomplete. The children had been recruited in various kindergarten schools of the French-speaking community of Belgium. Parental consent had been obtained for each child. A systematic inventory of the children's development milestones and medical history was completed by the parents and ensured that all children spoke French as a first language, were monolingual, had no history of neurological disorders or neurodevelopmental delay, presented normal auditory and visual acuity and had normal language development and no learning difficulties. The children were seen in their respective schools or at home. This study has been carried out in accordance with The Code of Ethics of the World Medical Association

(Declaration of Helsinki) and consents for children's participation from them and their parents were obtained.

Materials

WM tasks

Item WM was assessed using a *single nonword delayed repetition task* (Attout et al., 2014; Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006; Martinez Perez et al., 2012). This task involved the repetition of a nonword after a 3 second delay during which rehearsal was blocked via the continuous and concurrent repetition of the syllable *bla*. Additionally, the child also had to repeat the nonword once immediately after its presentation in order to ensure that the stimulus had been correctly perceived and could be accurately reproduced. The task consisted of 30 monosyllabic nonwords (see Appendix A) and respecting French phonotactic rules. The nonwords had been recorded by a female human voice. Four practice trials were presented at the beginning of the task. By presenting single nonwords with an identical syllabic structure (consonant-vowel-consonant), this task was designed to maximize the processing of phonological item information while minimizing the contribution of serial order retention abilities. Moreover, the nonwords were new at every trial and diphone combinations were of relatively low familiarity relative to the phonological structure of French (mean diphone frequency of the consonant–vowel segments: 149; range: 3–524; mean diphones frequency of the vowel–consonant segments: 129; range: 7–728; Tubach & Boë, 1990), maximizing the processing demands of phonological item information. In practice, the task was presented as a game to the child: *"You are a knight/a princess locked up in the tower of a castle. In order to find your way out of the castle, you have to open the doors by remembering passwords. More precisely, when you see a closed door, you will hear through the headphones a word from a magic language which opens the door and which you have to remember. The door opens if you repeat "blablabla . . ." during a short time and* *then, on my order, you repeat the password you just heard."* Each phoneme correctly repeated after the delay was scored one point. This task has good internal reliability (Cronbach α: =.87; Attout et al., 2014) and test-retest reliability (R= .74 - .79; Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006).

The retention of serial order WM was assessed using a *serial order reconstruction task* (Attout et al., 2014; Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006; Martinez Perez et al., 2012). This task consisted of the auditory presentation of sequences of 2 to 7 animal names. After the presentation, the child had to reconstruct the order of presentation of the words using picture cards depicting the animals from the memory list and by arranging, on the desk, the picture cards from left to right according to the order of presentation of the animals. The picture cards were always arranged in alphabetic order when given to the child. Four lists for each sequence (24 list trials in total) were presented to each child (see Appendix B). The first trials of length 2 were practice trials. All items had been recorded by a female voice on computer disk and were presented to the child through headphones. Their mean duration was 549 ms (range: 371–696), with an interstimulus interval of 650 ms. The seven animal names (/liõ/, /ʃa/, /ʃjɛ̃/, /sɛ̃Ʒ/, /kɔk/, /lu/ /uRs/) had been selected for their high lexical frequency (mean lexical frequency: 50,631; range: 16,432–90,926; Lambert & Chesnet, 2001) and their low age of acquisition (mean age of acquisition: 20 months; range: 13–24 months; (Alario & Ferrand, 1999; Ferrand & Alario, 1998). This task maximized order information retention requirements since the words used were highly familiar, known in advance and available at recall. Moreover, the words were monosyllabic in order to minimize the phonological processing demands. This task was also presented as a game: "*Every year, animals from all over the world gather to have a huge race. This year, seven animals are participating: a cat, a dog, a cock, a lion, a wolf, a bear, and a monkey. Several races take place. Sometimes only two animals are participating. Sometimes there are* *three, four, or five animals. At other times, there are big races with six or seven animals. Through the headphones, you will hear someone announce the animal's order of arrival at the finish line, from the first to the last animal. Immediately after, you have to put the pictures of the animals on the podium in their order of arrival. The animal arriving first has to be put on the highest step and the last one on the lowest step."* We considered the number of items that were placed in the correct order of presentation by pooling over all sequence lengths. For example, if the child reconstructs the target list "cat-wolf-bear-lion" as "cat-wolf-lion-bear", a score of 2 was given as there are two items placed in correct position is two. We compute the total for all lists. This task has good internal reliability (Cronbach α: =.65; Attout et al., 2014) and test-retest reliability (R= .68 - .82; Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006).

Vocabulary tasks

For assessing receptive vocabulary knowledge, we used the EVIP vocabulary scale (Dunn & Theriault-Whalen, 1993), which is the French adaptation of the PPVT (Dunn & Dunn, 1981). In this task, children heard a word spoken by the experimenter, and they had to select, among a choice of four pictures, the picture that matches the spoken word. The words used in this task were nouns, verbs or adjectives from different semantic categories (e.g. clothes, activities, objects, food, body parts or profession), with a large range in lexical frequency. According to the standard procedure, test administration was stopped after six erroneous responses on eight consecutive trials. As a dependent variable, the raw vocabulary score was retained for analysis. This task has a good test-retest reliability (R =.80; Dunn & Theriault-Whalen, 1993).

Productive vocabulary knowledge was assessed using the picture naming subtests of the ISADYLE standardized language assessment battery (Piérart et al., 2010). Naming abilities for lexically frequent and concrete nouns were assessed via 52 pictures representing various concrete objects (animals, tools, clothes, body parts, man-made objects, food). Naming abilities for lexically infrequent nouns were assessed via 25 pictures depicting man-made objects (e.g., latch). Naming abilities for verbs were assessed via 13 pictures depicting various actions (e.g., to cut, to lick). For each task, a score of 2 was given when the correct lexical target was produced, a score of 1 was given for an approximating response (e.g., train for locomotive), and a score of 0 was given for an incorrect response or a no-response. We determined the number of pictures correctly named (by disregarding phonetic deviations due to articulatory difficulties) for the entire task (global score), as well as for the frequent word, rare word and verb subtasks separately. The frequent word, rare word and verb subtasks show a high reliability (Cronbach α: .92, .86 and .71, respectively; Piérart et al., 2010).

Nonverbal intelligence task

We also administered the Raven's Colored Progressive Matrices (Raven, Court, & Raven, 1998) in order to control for overall nonverbal intellectual efficiency when investigating the association between WM and vocabulary measures. This task has a high test-retest reliability (Cronbach α: .85-.93; Abdel-Khalek, 2005; Commissaire & Besse, 2019).

Task order and statistical analysis

The different tasks were administered in pseudo-random order in 2 sessions, lasting each about 40 minutes. We avoided presenting two tasks assessing a same cognitive domain (e.g. vocabulary or WM) in the same session. We used a Bayesian approach for all statistical analyses (see, e.g., Dienes, 2011, 2016; Wagenmakers, 2007; Wagenmakers, Lee, Lodewyckx, & Iverson, 2008). Relative to frequentist statistics, the Bayesian approach has the advantage of relying on a model comparison rationale and of adopting a model selection strategy to quantify the strength of evidence associated *with* as well as *against* each model as a predictor of vocabulary abilities (Dienes, 2011; Morey & Rouder, 2011; Wagenmakers, 2007). Although there are no fixed statistical thresholds when using a Bayesian framework, we used indicative guidelines proposed by Jeffreys (1961) for interpreting Bayes factors (BF) : $BF < 1$ = no evidence, $1 < BF < 3$ = anecdotal evidence, $3 < BF < 10$ = moderate evidence, $10 < BF < 30$ = strong evidence, $30 < BF < 100$ = very strong evidence and $BF > 100$ = extreme/decisive evidence for the presence of a given effect. When reporting BFs, BF_{10} indicates evidence in favour of a specific variable/model against the null model, and BF_{01} indicates the reverse evidence. Bayesian analyses were conducted with Version 0.9.0.1 of the JASP software package, using default settings for the Cauchy prior distribution (JASP Team, 2017, jasp-stats.org).

RESULTS

Descriptive statistics are shown in Table 1. The distribution of scores for the different tasks was assessed by determining skewness and kurtosis parameters (Tabachnick & Fidell, 1996). The majority of measures had acceptable skewness and kurtosis values (values within the recommended 2 *SE* range) with a small ceiling effect for the frequent word measure and a small floor effect for the rare word measure of the productive vocabulary task.

Table 1. Descriptive statistics for all tasks.

	Mean (Standard deviation)	Mean in percent	Skewness[1]	Kurtosis[2]
Age	68.63 (3.92)		-.04	-.66
Raven's CPM raw score (total, max.=36)	17.96 (4.36)	49.89	-.33	-.04
Receptive vocabulary (total, max.=170)	70.70 (16.03)	41.59	.18	-.80
Productive vocabulary (total, max.=180)	122.20 (17.37)	68	-.15	.50

Frequent words (total, max.=104)	90.21 (8.63)	86.74	-.79	-.94
Rare words (total, max.=50)	15.65 (6.85)	31.3	.80	.03
Verbs (total, max.=26)	16.34 (4.33)	62.85	-.33	-.62
Serial order working memory (total, max.=100)	47.52 (14.21)	47.52	.09	-.32
Item working memory (total, max.=90)	69.53 (11.31)	77.25	-.64	-.59

[1]Standard error skewness cutoff = ±.50. [2]Standard error Kurtosis cutoff = ±.99.

Link between WM and vocabulary measures

A first set of Bayesian correlation analyses examined the overall relationship between the different WM and vocabulary measures (see Table 2 and see Appendix C for the Bayesian correlation Matrix of all measures). There was strong evidence in favour of a link between serial order WM and both productive and receptive vocabulary measures. At the same time, there was moderate evidence in favour of the *absence* of an association between the item WM measure and vocabulary measures (BF_{01} = 5.56 and BF_{01} = 7.14, respectively). When we considered each productive vocabulary measure separately, we observed a moderate to strong association between the order WM measure and frequent and rare word measures but inconclusive evidence regarding the verb measure. When predicting the different productive vocabulary measures by the item WM measure, there was moderate evidence in favor of the *absence* of an association between the item WM measure and the three receptive vocabulary measures (frequent words : BF_{01} = 5.65; rare words : BF_{01} = 7.67; verbs : BF_{01} = 7.19). Obviously, the two vocabulary tasks correlated strongly together (r=.57; BF_{10} = 4.25E+6). Finally, the correlation between the serial order WM task and Raven's matrices was identical to the size of the correlation between the item WM task and Raven's matrices, indicating that both WM tasks was similarly related to general estimates of intellectual efficiency.

Table 2. Bayesian correlation (Pearson r) between the WM tasks and the different vocabulary measures and Raven's CPM.

	Item WM	Order WM
Receptive vocabulary	.09 (BF_{10} = 0.18)	.34 (BF_{10} = 28.12)
Productive vocabulary (total)	-.03 (BF_{10} = 0.14)	.33 (BF_{10} = 24.02)
Productive vocabulary (frequent words)	-.08 (BF_{10} = 0.18)	.29 (BF_{10} = 5.86)
Productive vocabulary (rare words)	-.01 (BF_{10} = 0.13)	.35 (BF_{10} = 45.45)
Productive vocabulary (verbs)	.04 (BF_{10} = 0.14)	.21 (BF_{10} = 0.93)
Raven's CPM	.19 (BF_{10} = 0.61)	.18 (BF_{10} = 0.53)

Next, Bayesian linear multiple regressions investigated the specificity of the link between serial order WM and the vocabulary measures by controlling for item WM abilities, nonverbal reasoning abilities and age. When predicting the receptive vocabulary measure, the model providing the highest BF was the model including both order WM and nonverbal reasoning variables (BF_{10} = 156.73; R^2 = .18) and was 3.20 times more likely than the model with the next-highest BF which also included age (BF_{10} = 49.1; R^2 = .18) (see Table 3 for $BF_{inclusion}$ values).

Table 3. $BF_{inclusion}$ values for each predictor variable in the linear multiple regression analysis on receptive and productive total vocabulary scores.

Predictors	Receptive vocabulary	Productive vocabulary
Age	0.34	0.42
Nonverbal reasoning	5.12	0.51

Item WM	0.30	0.54
Order WM	14.51	28.05

Similar results were obtained for Bayesian linear multiple regression analysis on the total productive vocabulary measure. The model associated with the strongest evidence included the order WM measure only (BF_{10} = 26.89, R^2 = .11) and was 2.28 times more likely than the model with the next-highest BF value which also included Raven's matrices measure (BF_{10} = 11.79, R^2 = .12) (see Table 3 for $BF_{inclusion}$ values).

Detailed analysis of the association between serial order WM and productive vocabulary

Next, we examined the link between serial order WM and productive vocabulary by considering the frequent noun, rare noun and verb scores separately. For the Bayesian multiple regression on the frequent noun score, an analysis including the serial order WM score as well as the Raven's matrices score and age as predictors provided substantial evidence for a model including only the serial order WM predictor (BF_{10} = 7.19, R^2 = .08); this model was 1.46 times more likely than the model with the next-highest BF including both item and order WM measures (BF_{10} = 4.93, R^2 = .10) (see Table 3 for $BF_{inclusion}$ values). When predicting the rare word score, the model associated with the highest evidence was a model including the serial order WM measure but also Raven's matrices (BF_{10} = 90.56, R^2 = .17). This model was 1.85 times more likely than the model with the next highest BF value which included only the serial order WM measure (BF_{10} = 48.99, R^2 = .12). Finally, for verb naming, no model was associated with robust evidence in favour of an effect, all BF_{10} values being inferior to 3. $BF_{inclusion}$ in Table 3 values show that models including the serial order WM measure led to slightly higher BF values than the other models, but evidence remained at an anecdotal level (see Table 4).

Table 4. $BF_{inclusion}$ values for each predictor variable in the linear multiple regression analysis on productive vocabulary subscores.

Predictors	Frequent words	Rare words	Verbs
Age	0.49	0.30	0.78
Nonverbal reasoning	0.44	2.24	0.35
Item WM	0.84	0.48	0.32
Order WM	9.68	30.41	1.68

DISCUSSION

This study examined the robustness of the association between verbal WM and vocabulary knowledge in 4-to-6-year old typically developing children by assessing vocabulary knowledge for both multi-item receptive vocabulary tasks and single picture productive vocabulary tasks and by distinguishing between item and serial order verbal WM abilities. Bayesian regression analyses showed that serial order WM measure was a robust predictor of both receptive and productive vocabulary abilities. Furthermore, in line with previous studies, there was no strong evidence for a role of item WM abilities in both types of vocabulary measures.

First of all, our data replicate the finding of previous studies that have shown a link between verbal WM abilities and receptive vocabulary knowledge (Avons et al., 1998; Gathercole et al., 1994, 1992; Service, 1992). Our data also replicate the finding of a more robust link between vocabulary knowledge and serial order WM abilities as compared to item WM abilities (Leclercq & Majerus, 2010; Majerus, Poncelet, Greffe, et al., 2006). Moreover, it is unlikely that the lack of matching of WM load between the two WM tasks can explain the observed pattern of results. If the link between the WM and vocabulary tasks was only due to the specific processing requirements of the tasks (e.g., higher load of the serial order WM but also the receptive vocabulary tasks) then a link should only have been observed between the serial order WM and the receptive vocabulary task but not between the serial order WM and the productive vocabulary task. This was not the case, order WM being associated with both vocabulary tasks.

It is also important to note that the two WM tasks involved inhibitory demands as information from previous trials can create proactive interference and needs to be inhibited. Furthermore, in the item WM task, the articulatory suppression task presented during the retention delay will create interference with the target nonword (Oberauer & Lewandowsky, 2008). In the serial order WM task, the children need to select the different response cards in sequential order which involves, at one point of time, selecting one card for response output and inhibiting all other cards (e.g., Burgess & Hitch, 1999). Therefore, if inhibitory processes are to explain the relationship between performance on serial order WM and vocabulary tasks, then we should have observed robust correlations between the vocabulary tasks and both WM measures, with the strongest link with the receptive vocabulary task as the multiple-choice inhibitory requirements for this vocabulary task are larger than for the single picture productive vocabulary measure. This is not what we observed, as only the serial order WM task showed a robust association with both the productive and receptive vocabulary tasks.

Moreover, it could be argued that the link between order WM and vocabulary observed in this study may stem from differences in the quality of phonological codes, given that the item WM task involved articulatory suppression that may have interfered with the memoranda as mentioned in the previous paragraph. However, interference is also present in the serial order WM task the multiple memoranda for a given trial will generate between-item interference, both within and between lists. Moreover, the serial order WM task did not involve a stronger manipulation component than in item WM task since in order WM task the card cues were presented to facilitate item maintenance so that only the serial order of the items had to be retrieved. The card cues did not have to be manipulated but they just had to be output in the same order as in the memory list. Finally, we should also note that similar differences in associations between item WM, serial order WM and vocabulary measures as reported in this study have been observed in previous studies deriving item and serial order components from within the same task rather than using two structurally distinct tasks (Majerus et al., 2009). Altogether, our findings are consistent with the theoretical assumption that language learning involves the learning of sequential information and dependencies, and that short-term retention abilities for sequential information are a critical determinant of these sequential learning processes for learning verbal sequences (Gupta & MacWhinney, 1997; Majerus, Poncelet, Elsen, et al., 2006; Ordonez Magro et al., 2018).

Critically, the present study extends these results to vocabulary knowledge as assessed by productive vocabulary performance. This is a novel and important finding as it shows that the link identified between verbal WM and vocabulary knowledge is not just the result of the specific characteristics of the PPVT receptive vocabulary measure that has been most frequently used in previous studies. This task has indeed an intrinsic WM load as it requires the short-term maintenance of an auditory target word while activating and comparing the lexico-semantic content of four different, simultaneously presented pictures. These characteristics could have biased and inflated the correlation observed between verbal WM abilities and this specific measure of vocabulary knowledge. The present study shows that this concern is not warranted as the correlations between verbal WM abilities and vocabulary knowledge were at least as strong when measured with single picture, productive vocabulary tasks characterized by no or an extremely small intrinsic WM load. Moreover, the potentially different levels of task difficulty for the vocabulary measures did not seem to play a major role in the outcome of results as the associations with the serial order WM task were as robust for the receptive vocabulary task (which may be considered to be easier as it involves a recognition response) as for the productive vocabulary task.

It is also interesting to note that this association was particularly true for the rare noun productive vocabulary subtasks, followed by the frequent noun subtask; for the verb productive vocabulary subtask, there was no reliable evidence for an association with verbal WM abilities. The stronger association between rare vs. frequent noun productive vocabulary and serial order WM abilities supports the hypothesis that serial order WM is involved in lexical learning. The role of serial order WM will be more prominent for rare words as these words are, by definition, encountered only a few times: children showing higher serial order WM abilities will be more likely to correctly repeat and encode the phoneme sequences that define the word right after its first presentation while children with low serial order WM abilities are more likely not to be able to immediately repeat the novel word. For frequent words, these interindividual differences in verbal WM abilities will have a less marked effect as these words, by definition, are encountered more frequently and hence there are more opportunities to learn them; even when serial order WM abilities are reduced, the word forms for frequent words will eventually be learnt after a high number of repetitions. This assumption is in line with novel word learning experiments in patients with impaired verbal WM abilities: although these patients show impaired novel word learning abilities, they will nevertheless show progress on novel word learning tasks even if this progress will be much slower than in non-impaired individuals (e.g., Freedman & Martin, 2001; Majerus, Norris, & Patterson, 2007). Interestingly, productive vocabulary knowledge for action words (verbs) showed no reliable association with serial order (or item) WM abilities. Although we can only speculate about this finding, it has been proposed that action words are not learnt as isolated phonological sequences referring to a specific object or concept, but are an integral part of the syntactic and conceptual context in which they occur (Naigles, 1990). Moreover, at the developmental level, even if both, concrete object names and concrete action verbs are acquired at the same time, verb grammaticalization seems to develop later than noun grammaticalization, supporting this developmental distinction between both word categories (Bassano, 2000). Furthermore, in many languages, and particularly French, the phonological sequence defining a specific verb is highly variable as its specific phonological form will result from the application of phonological rules associated with time-based and person-based inflections. Hence, learning of lexical phoneme sequences for verbs (and in some extent for action-noun, see Vinson & Vigliocco, 2002) may rely to a stronger extent on direct and fast integration with specific semantic and syntactic contexts than nouns. Objects are also characterized by a larger lexicon than verbs (in French for example, there are about 19000 nouns for 4000 verbs; source: www.lexique.org, Brulex database) and the existence of multiple labels for a same object, challenging the learning process at the lexical phonological form level as distinct phonological sequences need to be learnt for a similar or sometimes even identical semantic context.

To conclude, this study shows that the link between verbal WM and vocabulary knowledge is robust and is not limited by a higher WM load that characterizes some of the vocabulary measures frequently used in this field of study. The results of this study support theoretical accounts suggesting a link between serial order WM abilities and lexical learning and therefore should be valid for other languages than French. Our results also indicate that this link may be valid only for specific linguistic categories such as nouns which are characterized by a particularly large lexicon size and stable phonological forms (only the grammatical gender and number can slightly modify the lexical form of the word while for verbs, the forms can change dramatically as a function of person, number, time, mode). However, the noun-verb dissociation observed in this study raises new questions and needs to be further addressed at the empirical level. This latter aspect of our findings may be specific to some languages in which the form of verbs varies substantially as is the case in French: for example, in French, ‘You play – we will play’ is ‘Tu joues – nous jouerons’.

REFERENCES

Abdel-Khalek, A. M. (2005). Reliability and factorial validity of the standard progressive matrices among Kuwaiti children ages 8 to 15 years. *Perceptual and Motor Skills*, *101*(2), 409–412. https://doi.org/10.2466/pms.101.2.409-412

Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 531–552. https://doi.org/10.3758/BF03200732

Attout, L., Noël, M., & Majerus, S. (2014). The relationship between working memory for serial order and numerical development: a longitudinal study. *Developmental Psychology*, *50*, 1667–1679. https://doi.org/10.1037/a0036496

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. https://doi.org/10.1002/hbm.24466

Avons, S. E., Wragg, C. A., Cupples, W. L., & Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics*, *19*(4), 583–601. https://doi.org/10.1017/S0142716400010377

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

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, *27*(3), 521–559. https://doi.org/10.1017/S0305000900004396

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

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

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory and Cognition*, *35*(6), 1445–1455. https://doi.org/10.3758/BF03193614

Commissaire, E., & Besse, A. (2019). Investigating lexical and sub-lexical orthographic processing skills in French 3rd and 5th graders. *Journal of Research in Reading*, *42*(2), 268–287. https://doi.org/10.1111/1467-9817.12268

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

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

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

Dunn, L. M., & Dunn, L. M. (1981). *PPVT-R: Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., & Theriault-Whalen, C. (1993). *Échelle de vocabulaire en images Peabody: EVIP*. Toronto, Canada: Psycan.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2004/034)

Ferrand, L., & Alario, F.-X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. *L'Année Psychologique*, *98*(4), 659–709. https://doi.org/10.3406/psy.1998.28564

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology*, *18*(3), 193–226. https://doi.org/10.1080/02643290126002

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543. https://doi.org/10.1017/S0142716406060383

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213. https://doi.org/10.1016/0749-596X(89)90044-2

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*(4), 439–454. https://doi.org/10.1111/j.2044-8295.1990.tb02371.x

Gathercole, S. E., Hitch, G. J., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*(6), 966–979. https://doi.org/10.1037/0012-1649.33.6.966

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British*

Journal of Psychology, *82*(3), 387–406. https://doi.org/10.1111/j.2044-8295.1991.tb02407.x

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*(5), 887–898. https://doi.org/10.1037/0012-1649.28.5.887

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *56*(7), 1213–1236. https://doi.org/10.1080/02724980343000071

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

Henry, L. A., & MacLean, M. (2003). Relationships between working memory, expressive vocabulary and arithmetical reasoning in children with and without intellectual disabilities. *Educational and Child Psychology*, *20*(3), 51–64.

Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, *36*(2), 73–137. https://doi.org/10.1006/cogp.1998.0685

Iacono, T., Torr, J., & Wong, H. Y. (2010). Relationships amongst age, language and related skills in adults with Down syndrome. *Research in Developmental Disabilities*, *31*(2), 568–576. https://doi.org/10.1016/j.ridd.2009.12.009

Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Clarendon.

Lambert, E., & Chesnet, D. (2001). Novlex: une base de données lexicales pour les élèves de primaire. *L'Année Psychologique*, *101*(2), 277–288. https://doi.org/10.3406/psy.2001.29557

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

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), 149–169. https://doi.org/10.1037/0278-7393.7.3.149

Lee Swanson, H., Orosco, M. J., & Lussier, C. M. (2015). Growth in literacy, cognition, and working memory in English language learners. *Journal of Experimental Child Psychology*, *132*, 155–188. https://doi.org/10.1016/j.jecp.2015.01.001

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

Majerus, S., & Barisnikov, K. (2018). Verbal short-term memory shows a specific association with receptive but not productive vocabulary measures in Down syndrome. *Journal of Intellectual Disability Research*, *62*(1), 10–20. https://doi.org/10.1111/jir.12443

Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: New evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology*, *116*(4), 811–828.

https://doi.org/10.1016/j.jecp.2013.07.014

Majerus, S., Heiligenstein, L., Gautherot, N., Poncelet, M., Linden, M. V, Van der Linden, M., & Linden, M. V. (2009). Impact of auditory selective attention on verbal short-term memory and vocabulary development. *Journal of Experimental Child Psychology*, *103*(1), 66–86. https://doi.org/10.1016/j.jecp.2008.07.004

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. https://doi.org/10.1080/09541440500446476

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology*, *93*(2), 95–119. https://doi.org/10.1016/j.jecp.2005.07.005

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. https://doi.org/10.1016/j.cognition.2007.10.003

Martinez Perez, T., Majerus, S., Poncelet, M., Martinez, T., Majerus, S., Poncelet, M., Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study.

Journal of Experimental Child Psychology, *111*(4), 708–723. https://doi.org/10.1016/j.jecp.2011.11.007

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*(1), 3–19. https://doi.org/10.1037/0022-0663.91.1.3

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

Morrison, R. G., Doumas, L. A. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: balancing inhibitory control in working memory and relational representation. *Developmental Science*, *14*(3), 516–529. https://doi.org/10.1111/j.1467-7687.2010.00999.x

Mosse, E. K., & Jarrold, C. (2011). Evidence for Preserved Novel Word Learning in Down Syndrome Suggests Multiple Routes. *Journal of Speech, Language, and Hearing Research*, *54*(4), 1137–1153. https://doi.org/10.1044/1092-4388(2010/09-0244)a

Næss, K.-A. B., Lervåg, A., Lyster, S.-A. H., & Hulme, C. (2015). Longitudinal relationships between language and verbal short-term memory skills in children with Down syndrome. *Journal of Experimental Child Psychology*, *135*(7), 43–55. https://doi.org/10.1016/j.jecp.2015.02.004

Naigles, L. (1990). Children Use Syntax To Learn Verb Meanings. *Journal of Child Language*, *17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: decay, temporal distinctiveness, or interference? *Psychological Review*, *115*(3), 544–576. https://doi.org/10.1037/0033-295X.115.3.544

Ordonez Magro, L., Attout, L., Majerus, S., & Szmalec, A. (2018). Short-and long-term memory determinants of novel word form learning. *Cognitive Development*, *47*(July 2017), 146–157. https://doi.org/10.1016/j.cogdev.2018.06.002

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331–347. https://doi.org/10.1016/0749-596X(91)90040-Q

Piérart, B., Comblain, A., Grégoire, J., & Mousty, P. (2010). *ISADYLE: Instruments pour le screening et l'approfondissement de l'examen des dysfonctionnements du langage chez l'enfant.* Marseille, France: Solal.

Saint-Aubin, J., & Poirier, M. (1999). 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. https://doi.org/10.1080/002075999399675

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A*, *45*(1), 21–50. https://doi.org/10.1080/14640749208401314

Swart, N. M., Muijselaar, M. M. L., Steenbeek-Planting, E. G., Droop, M., de Jong, P. F., & Verhoeven, L. (2017). Cognitive precursors of the developmental relation between lexical quality and reading comprehension in the intermediate elementary grades. *Learning and Individual Differences*, *59*(10), 43–54. https://doi.org/10.1016/j.lindif.2017.08.009

Szmalec, A., Page, M. P. a., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language*, *67*(3), 342–354. https://doi.org/10.1016/j.jml.2012.07.001

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics (3rd ed.)*. New York: HarperCollins.

Tubach, J. P., & Boë, L. J. (1990). *Un Corpus de transcription phonétique (300000 phones): constitution et exploitation statistique*. Paris, France: France télécom.

Vance, B., West, R., & Kutsick, K. (1989). Prediction of Wechsler Preschool and Primary Scale of Intelligence IQ scores for preschool children using the Peabody Picture Vocabulary Test-R and the Expressive One Word Picture Vocabulary Test. *Journal of Clinical Psychology*, *45*(4), 642–644. https://doi.org/10.1002/1097-4679(198907)45:4<642::AID-JCLP2270450421>3.0.CO;2-Q

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: Semantic representations of object nouns, action nouns and action verbs. In *Journal of Neurolinguistics* (Vol. 15, Issues 3–5, pp. 317–351). Elsevier BV. https://doi.org/10.1016/S0911-6044(01)00037-9

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

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). New-York, USA: Springer.

APPENDICES

Appendix A. Nonwords used to assess the item WM abilities.

/vyz/
/bœz/
/kov/
/rug/
/ʒom/
/møg/
/tof/
/ruf/
/kyʃ/
/ʒɔz/
/bɛ̃g/
/gum/
/ʃid/
/pɑ̃b/
/ʃɔd/
/ryv/
/zœz/
/nov/
/zuk/
/non/
/løg/
/tob/
/suf/
/ʃyʃ/
/vɔz/
/pɛ̃z/
/fum/
/pig/
/fɑ̃p/
/vɔv/

Appendix B. Lists of words used to assess the order WM abilities.

Lenght 2 (trials)							
list 1	Loup	singe					
list 2	chien	lion					
Lenght 3							
list 1	singe	chien	chat				
list 2	lion	ours	loup				
list 3	loup	chien	singe				
list 4	chat	ours	coq				
Lenght 4							
list 1	coq	chat	loup	chien			
list 2	ours	lion	chien	singe			
list 3	lion	chien	singe	ours			
list 4	singe	coq	loup	chat			
Lenght 5							
list 1	ours	coq	singe	loup	chat		
list 2	chat	lion	coq	ours	singe		
list 3	coq	singe	lion	chien	ours		
list 4	chien	lion	ours	coq	loup		
Lenght 6							
list 1	chien	coq	ours	lion	loup	chat	
list 2	singe	loup	chat	coq	ours	lion	
list 3	coq	chat	lion	singe	chien	loup	
list 4	chien	ours	singe	loup	lion	coq	
Lenght 7							
list 1	coq	singe	lion	loup	chat	ours	chien
list 2	chat	coq	chien	ours	singe	loup	lion
list 3	singe	ours	loup	lion	coq	chien	chat
list 4	loup	chien	chat	lion	singe	coq	ours
Translation in english	wolf	dog	cat	lion	monkey	cock	bear

Appendix C. Bayesian correlation matrix for all measures.

	Age	Raven CPM	Serial order WM	Item WM	Receptive vocabulary	Productive Vocabulary (total)	Frequent words	Rare words	Verbs
Age	/								
Raven CPM	.19 (BF_{10}=.66)	/							
Serial order WM	.02 (BF_{10}=.13)	.18 (BF_{10}=.53)	/						
Item WM	.23 (BF_{10}=1.34)	.19 (BF_{10}=.61)	.21 (BF_{10}=.89)	/					
Receptive vocabulary	.10 (BF_{10}=.20)	.31 (BF_{10}=10.49)	.34 (BF_{10}=28.12)	.09 (BF_{10}=.18)	/				
Productive Vocabulary (total)	.09 (BF_{10}=.19)	.17 (BF_{10}=.45)	.33 (BF_{10}=24.02)	-.03 (BF_{10}=.14)	.57 (BF_{10}=4.25E+6)	/			
Frequent words	.10 (BF_{10}=.19)	.13 (BF_{10}=.28)	.29 (BF_{10}=5.86)	-.08 (BF_{10}=.18)	.51 (BF_{10}=81874.66)	.91 (BF_{10}=1.77E+33)	/		
Rare Words	.01 (BF_{10}=.13)	.27 (BF_{10}=3.27)	.35 (BF_{10}=45.45)	-.01 (BF_{10}=.13)	.50 (BF_{10}=49790.45)	.85 (BF_{10}=2.07E+23)	.60 (BF_{10}=3.98E+7)	/	
Verbs	.15 (BF_{10}=.37)	-.01 (BF_{10}=.13)	.21 (BF_{10}=.93)	.04 (BF_{10}=.14)	.47 (BF_{10}=8351.12)	.85 (BF_{10}=9.69E+33)	.72 (BF_{10}=1.23E+13)	.62 (BF_{10}=3.51E+8)	/