The Role of Working Memory in Planning and Generating Written Sentences

Ronald T. Kellogg, Casey E. Turner, Alison P. Whiteford & Andrew Mertens

Saint Louis University, MO | USA

Abstract: Planning a sentence with concrete concepts whose referents can be mentally imaged has been shown in past work to require the limited resources of visual working memory. By contrast, grammatically encoding such concepts as lexical items in a syntactic structure requires verbal working memory. We report an experiment designed to demonstrate a double dissociation of these two stores of working memory by manipulating the difficulty of (1) planning by comparing related concepts to unrelated concepts and (2) grammatical encoding of an English sentence in active voice versus the more complex structure of the passive voice. College students (N = 46) composed sentences that were to include two noun prompts (related versus unrelated) while concurrently performing either a visual or a verbal distracting task. Instructions to produce either active or passive sentences were manipulated between groups. The results surprisingly indicated that the supposedly easier planning with related concepts made a large demand on verbal working memory, rather than unrelated concepts demanding more visual working memory. The temporal dynamics of the sentence production process appear to best account for the unexpected findings.

Keywords: sentence generation, sentence planning, working memory

journal of
WRITING RESEARCH

Kellogg, R. T., Turner, C. E., Whiteford, A. P., & Mertens, A. (2016). The role of working memory in planning and generating written sentences. Journal of Writing Research, *7*(3), 397-416. doi: 10.17239/jowr-2016.07.03.04

Contact: Ronald T. Kellogg, Department of Psychology, Saint Louis University, Morrissey Hall, 3700 Lindell Blvd., St. Louis, MO, 63108 - USA | Email: kelloggr@slu.edu

1. Background and Rationale

In a well-known model of spoken language production, planning the conceptual content of the message precedes a grammatical encoding stage of processing in which lexical representations are selected and positioned in a syntactic structure (Bock & Levelt, 1994). Phonological and orthographic encoding of the words to be typed are subsequent stages of processing that occur downstream from planning and grammatical encoding. However, the psycholinguistics literature often uses the term planning with reference to either the semantic content or the grammatical structure of the sentence (Clark & Clark, 1977). Similarly, with respect to written sentence production, some investigators employ the term grammatical planning (e.g., Nottbusch, 2010) while others restrict planning to the conceptual representations of the message or content to be translated into syntactic structures (e.g., Hayes & Flower, 1980). Here, the term grammatical encoding will be used to distinguish between planning the semantic content versus selecting the lexical and syntactic form of the sentence.

A model of how working memory (WM) supports the planning of ideas, the translation of ideas into written sentences through grammatical, phonological, and orthographic encoding, and the reviewing the ideas and text already produced was proposed by Kellogg (1996). In writing as well as other complex cognitive tasks, working memory provides a means for transiently holding knowledge in an accessible form so it can be effectively used. For example, knowledge about the writing topic and the specific language in which the text will be written must not only be available in long-term memory, but also must be retrieved and accessible for use in solving the content and rhetorical problems at hand. The model specified the demands of planning ideas, translating ideas into sentences, and reviewing ideas or sentences on the central executive, phonological loop, and visuo-spatial sketchpad based on the evidence then available. It thus integrated Baddeley's (1986) model of working memory with the seminal Hayes and Flower (1980) model of written composition.

The phonological loop of Baddeley's (1986) model has the function of storing and maintaining verbal representations such as words, phrases, and whole sentences in phonological form. Levy and Marek (1999) have shown that irrelevant speech causes errors in both number and tense during sentence generation. Importantly, they were able to show that same effects were observed with scrambled unattended speech as with words in a meaningful order. Thus, it was the phonological rather than the semantic properties of the speech that made a difference, consistent with Baddeley's model. Similarly, Chenoweth and Hayes (2003) used articulatory suppression to preclude the possibility of using covert or inner speech for linguistic encoding processes during translation. This technique blocks the phonological loop by repeating an irrelevant word over and over again aloud, precluding silent speech articulation. The number of words produced per second was reliably impaired by this concurrent task relative to a control condition.

Other investigators have used the term verbal working memory to contrast the phonological storage of words with the storage of visual and spatial representations (e.g., Smith & Jonides, 1997). Brain images taken while participants perform a task requiring the short-term maintenance of verbal information reveals the activation of left hemisphere regions including Broca's area and a parietal region. These regions support speech production and reflected the covert speech of the phonological loop as people covertly use language in the form of inner speech. The term verbal WM is also used as a contrast to semantic WM (Martin, Sheldon, & Yaffee, 1994). The latter short-term store of WM supports the planning of the conceptual content of the message instead of grammatical encoding (Kellogg, Whiteford, Turner, Cahill, & Mertens, 2013). Smith and Jonides (1997) further showed that Baddeley's visuo-spatial sketchpad must be fractionated into two separate components based on the neuroimaging results. Maintaining visual objects activated regions in the left hemisphere that were distinct from those involved with verbal information, whereas the right hemisphere was activated when maintaining spatial locations. Because the visuo-spatial sketchpad is best conceived in terms of a visual WM and spatial WM, the phonological loop will be referred to here, for parallelism, as verbal WM.

Kellogg (1996) proposed that verbal working memory (WM) is a necessary resource for all grammatical encoding operations in written sentence production, whereas planning draws on visual WM. In support of this hypothesis, Kellogg, Olive and Piolat (2007) found that generating a written sentence from two noun prompts interfered with a concurrent verbal task. Such interference was observed regardless of whether the nouns named abstract or concrete concepts, as would be expected if both types of concepts required grammatical encoding following the planning stage of processing. Visual WM was needed in planning the conceptual content of a sentence when the nouns named concrete concepts that invoked visual imagery of their referents. Abstract noun prompts, on the other hand, did not make demands on visual WM, as evidenced by a lack of interference with a concurrent visual task. Thus, the demands on verbal WM appear to be obligatory in composing written sentences given that it supports grammatical encoding required in selecting either concrete or abstract words for positioning in a syntactic structure. By contrast, planning makes selective use of visual WM, depending on whether the concepts involved evoke visual imagery. In an updated review of the literature, Kellogg et al. (2013) cited further support for this pattern of demands on verbal WM and visual WM during written sentence production.

The present research attempted to further test the assumption that planning can in certain circumstances rely on visual WM whereas grammatical encoding necessarily makes demands on verbal WM. We sought to manipulate the amount of planning that would be required as a way of creating either a relatively low demand on visual WM versus a high demand. We employed concrete nouns for all sentence prompts, but the degree of semantic relatedness of the noun pairs was manipulated. When two nouns are given as prompts to compose a written sentence, it is known that unrelated nouns require more planning in comparison with related nouns (Kellogg, 2004; Rosenberg, 1977). Strong semantic associations between the nouns (e.g., chair-table) minimize the amount of planning in the conceptual domain needed to create a proposition to be expressed in a sentence. It takes about a half second longer to initiate typing a sentence when the nouns are weak semantic associates (e.g., bride-eagle), because more conceptual planning is needed to form a proposition linking the two ideas (Kellogg, 2004). Once a proposition is created, however, the grammatical encoding demands ought to be the same for either related or unrelated items. This follows from the assumption that grammatical encoding is stage of composition that follows the planning stage and is independent of it. According to the model, the grammatical encoding stage as well as other stages of written sentence generation (i.e., phonological and orthographic encoding) depend on verbal WM, not visual. Thus, unrelated concrete nouns were expected to demand more visual WM during planning compared with related concrete nouns, but have no effect on verbal WM.

On the other hand, translating ideas into passive sentences ought to demand more verbal WM relative to active sentences, but leave visual WM unaffected. Passive structures presumably are more complex syntactically compared with active sentences. The justification for this assertion is in part theoretical and in part empirical. In terms of linguistic theory, Chomsky's (1965) transformational grammar and in his successive revisions argue that the passive surface structure is derived from an active form of the sentence. In the original model, for example, an active deep structure had to be transformed to produce a passive surface structure. As an empirical fact, the evidence shows that passive sentences typically require more time to comprehend compared with active sentences (Gough, 1965). This is consistent with the linguistic analysis that the passive voice is the more complex of the two. Because grammatical encoding presumably requires verbal WM alone, composing a passive sentence ought to make greater demands on verbal WM compared with active sentences. This manipulation of grammatical structure ought to have no impact on visual WM, according to the model. In short, it should be possible to demonstrate a double dissociation between planning and grammatical encoding with respect to the demands that they place on visual versus verbal WM.

2. Method

2.1 Tasks and Design

Participants wrote sentences in response to prompts, presented on a computer screen, with instructions to either create a sentence in the active voice or in the passive voice. For each sentence, two nouns were presented as prompts that were to be included in the sentence. For example, for the prompts chair-table, the participant might write "The student sat in the chair behind the table" as an active voice response. As a passive construction, the participant might write "The chair behind the table was sat in by the student." The noun prompts were either closely related in meaning or were semantically unrelated.

In the control condition, the participants wrote the sentences without any distraction from a second task. In the visual WM condition, participants had to retain six visual symbols in memory while composing each sentence. The symbols were presented at the beginning of the trial for study and then were removed to be replaced by the two noun prompts. After writing a sentence that included the nouns, the participant pressed the escape key (ESC), which caused a test string of symbols to be presented. The participant then responded as to whether the test symbols were identical or different from the study symbols. In the verbal WM condition, the same sequence of events was used except that six digits were studied. The digits could be coded verbally whereas the symbols were designed to be not readily named. A block of control trials of the WM task was also included, when the symbol and digit tasks were performed without the need to compose a written sentence. By subtracting accuracy on the WM task when it was combined with sentence production from accuracy on the WM control trials, a measure of task interference was derived.

Without WM control trials, the present experiment employed a 2 x 3 x 2 mixed design, crossing the variables of sentence task complexity (active voice vs. passive voice), WM task type (no WM task vs. verbal WM task vs. visual WM task), and word pair relatedness (related vs. unrelated).

2.2 Stimuli

The experimental stimuli consisted of the following: Nine distinct symbols, taken from the SPSS Marker Set font in Microsoft Word, for the symbol condition; nine digits (1-9), for the digit condition; and 60 noun pairs, drawn from the norms of Nelson, McEvoy, & Schreiber (1998). Thirty of the noun pairs were related (e.g., door-knob) and 30 were unrelated (e.g., ice-jail). The norms provide a cue (door) that elicits a target (knob) from a large percentage of individuals; in other words there was a high degree of associative strength between the semantically related word pairs. By contrast, unrelated pairs were selected so that there was little, if any, association between them based on the normative data. Words selected for all noun pairs were 3-7 letters in length, familiar words in common print usage based on the raw print frequencies provided in the Nelson et al. (2007) norms and were generally rated as highly concrete or easy to image; ratings of 5 or 6 on the 1-7 point concreteness scale are considered concrete and esasy to image. Note that 7 selected words were missing concreteness ratings in the norms. The complete list of pairs is shown in Appendix 1. The related pairs did not differ from the unrelated pairs in either raw frequencies or rated concreteness.

2.3 Procedures

All experimental procedures were controlled using E-Prime stimulus presentation software (Psychology Software Tools, Pittsburgh, PA). The experiment consisted of three 20-item blocks of sentence writing trials and one 20-item block of visual WM and verbal WM trials only (no sentence block). The no sentence block was displayed first. The no sentence block was followed by each of the three sentence writing blocks combined with either a verbal WM (digit strings) task, a visual WM (symbol strings) task, or a control task that consisted of a single repeated digit or symbol. The order of the three sentence writing-WM task blocks was counterbalanced across participants. Each sentence writing block displayed a total of 20 related and unrelated word pairs. Participants were randomly assigned to either the active voice condition or the passive voice condition as a between subjects manipulation. A total of 80 trials were completed by each participant. The instructions for the active and passive sentences are given below. The specific examples were adopted from Warriner and Griffith's (1977; p. 167) *English Grammar and Composition*. In each case, participants were provided with four examples of each sentence type.

> *Active Instructions*
> You will be constructing active sentences. An active sentence is one with a verb that expresses action performed by the subject. The subject performs the action of the verb, and if there is a receiver of the action, it is expressed by the object of the verb. For example: The raging flood waters destroyed the bridge. The manager closes the theater every Wednesday. It will reopen on Thursday. No one had reported the fire.
>
> *Passive Instructions*
> You will be constructing passive sentences. The verb in a passive sentence is always a verb phrase consisting of some form of 'be' ('is', 'was', etc.) plus the past participle. In this way, the object of the sentence becomes the subject. For example: The bridge was destroyed by the raging flood waters. The theater is closed every Wednesday. It will be reopened on Thursday. The fire had not been reported.

Sentence trials began with participants viewing a study string of 6 digits or 6 non-verbal symbols that were displayed for 15 seconds. Participants were instructed to remember the string for a later recognition test. Participants then viewed a pair of related or unrelated concrete nouns and generated a sentence in either active voice (e.g., "*He turned the knob on the door*") or passive voice (e.g., "*The knob on the door was turned by him*"), depending on the experimental condition. Participants were given unlimited time to generate and type the sentence. Upon completing their sentence, participants pressed the ESC key instead of typing a period. The sentence production time was measured by E-Prime from the appearance of the noun prompts to the pressing of the ESC key. The ESC key may have been less practiced than hitting a period at the end of each sentence slightly inflating the total production time, but this factor presumably was constant across conditions. A new string of digits or symbols was then displayed as a test stimulus, and participants were instructed to determine whether the new string matched the string shown at the beginning of the trial by pressing Y for yes or N for no. The test string matched the study string exactly on 50% of the trials (randomly chosen) and failed to match by a change of a single pair of symbols or digits on the remaining 50%. A random pair of elements were exchanged to create the unmatched string. For example, if a study set of digits was 748625, then two examples calling for a no response would be 784625 or 248675.

2.4 Participants

Our goal was to test enough participants to yield an equal sample size (n = 23) in both the active and the passive conditions. To that end, we tested a total sample of 74 Saint Louis University undergraduate students, with 30 participants in the active sentence condition and 44 participants in the passive sentence condition. All participants were recruited through a Saint Louis University-maintained online participant pool. A common problem in studying language production is that the experimenter can design the prompts and instructions used to elicit language production, but cannot control what the participant actually says or writes (Bock, 1996). Consequently, we anticipated the need to screen participants for their compliance with the instructions; they had to produce a sentence for all or nearly all trials and the sentence had to conform to the instruction to produce either an active or a passive sentence. A total of 15 participants (8 passive and 7 active) skipped more than 8 of the trials and so were dropped from the analysis. Further, an additional 13 participants instructed to produce only passive sentences failed to comply on 8 or more of the sentences and so were also dropped from consideration. It appeared that syntactic priming played a role in these failures to follow the passive instructions on every trial: if a passive participant wrote a single active sentence he or she might then repeat with active constructions for several trials in a row. As planned, then, our final sample (N = 46) was equally divided between active and passive compliant participants.

3. Results

3.1 Production Time

Median times in milliseconds required to produce a sentence in each condition and for each participant were examined in a 2 x 3 x 2 mixed analysis of variance (ANOVA), with task complexity as a between subjects variable and WM task type and word pair relatedness as within subjects variables. Reported below are the ANOVA results with participants as a random factor averaged over the materials. An item analysis was also conducted with the materials as the random factor averaged over participants.

The mean values of the time taken to produce sentences, computed across participants in each condition, are shown in Table 1. The ANOVA revealed a main effect of word pair relatedness, F (1, 44) = 46.39, p <.001, $\eta\rho^2$ = .51. Participants generated sentences significantly faster with related words pairs (M = 12885 ms) than with unrelated word pairs (M = 15607 ms). A main effect of sentence task complexity was also observed, with sentences in active voice (M = 12640 ms) generated significantly faster than sentences in passive voice (M = 15853 ms), F (1, 44) = 6.39, $p < .05$, $\eta p^2 = .13$. In addition, a significant interaction between sentence task complexity and WM task type was revealed, F (2, 88) = 4.28, $p < .05$, $\eta p^2 = .09$. The means for this interaction are displayed in Table 1. No other main effects or interactions were statistically reliable. The item analyses showed the same effects of word pair relatedness, F' = (1, 348) = 35.15, $p < .001$, $\eta p^2 = .09$, sentence task complexity, F' = (1, 348) = 62.56, $p < .001$, $\eta p^2 = .15$, and sentence task complexity and WM task type, F' = (2, 348) = 7.24, $p < .001$, $\eta p^2 = .04$. In addition, the item analyses with their greater statistical power also revealed a main effect of WM task type, F' = (2, 348) = 3.43, $p < .05$, $\eta p^2 = .02$, with production times shortest in the visual WM condition.

Table 1: *Means (with Standard Errors) for Sentence Production Times (ms)*

	WM Task		
Sentence Task Complexity	Control (no WM Task)	Verbal	Visual
Active Sentences	14532 *(1407)*	12088 *(1026)*	11303 *(964)*
Passive Sentences	15250 *(1034)*	16631 (970)	15679 *(1031)*

As can be seen in Table 1, active sentences were produced about 4 s faster than passive sentences in the verbal and visual working memory conditions, whereas in the control condition (sentence production only), active and passive sentences did not differ significantly. Thus, in the absence of a load on working memory, it takes about the same amount of time to construct both types of sentences, despite the increased complexity of sentences written in passive voice. Of interest, participants took less time to compose an active sentence than a passive sentence only when under a concurrent load on working memory.

Because the length of sentences in terms of the total numbers of words produced varied across conditions, the effects found for sentence production time could be driven entirely by sentence length. There were in fact reliable correlations between the number of words written and the time required to produce them. However, the strengths of this relationship were relatively modest in the control condition (r = .31, $p < .001$), the verbal WM condition (r = .39, $p < .001$), and in the visual WM condition (r = .25, $p < .001$). It was important, therefore, to examine the words produced per second. The ANOVA revealed a reliable main effect of relatedness, $F(1, 44) = 44.26$, $p < .001$, $\eta\rho^2 = .50$, with reliably more words produced per second for related noun pairs ($M = .60$) compared with unrelated pairs ($M = .54$). The only other statistically significant effect was an interaction between WM task type and sentence complexity, $F(2,88) = 4.81$, $p < .05$, $\eta\rho^2 = .10$. The main effect of WM task type was also marginally significant, $F(2,88) = 2.90$, $p < .06$, $\eta\rho^2 = .06$. The same effects were also found in the item analyses: word pair relatedness, $F' = (1, 348) = 21.04$, $p < .001$, $\eta\rho^2 = .06$, WM task type and sentence complexity, $F' = (2, 348) = 8.30$, $p < .001$, $\eta\rho^2 = .05$, and the marginal effect of WM task type, $F' = (1, 348) = 21.04$, $p < .001$, $\eta\rho^2 = .06$. The item analyses, with more observations per condition, uncovered an additional main effect of sentence complexity, $F' = (1, 348) = 5.46$, $p < .05$, $\eta\rho^2 = .02$, with active sentences produced at a faster rate overall compared with passive sentences.

Table 2: Means (with Standard Errors) for Words Produced per Second

	WM Task		
Sentence Task Complexity	Control (no WM Task)	Verbal	Visual
Active Sentences	0.54 *(.036)*	0.59 *(.034)*	0.62 *(.037)*
Passive Sentences	0.57 *(0.36)*	0.51 *(.034)*	0.57 *(.037)*

The relevant means for the reliable interaction are shown in Table 2. As also seen in the sentence production times, the words of the active sentences were written at a higher rate per second in the verbal and especially the visual WM conditions relative to the control condition. For passive sentences, by contrast the slowest rate of production was observed in the verbal WM condition. This finding is consistent with the hypothesis that passive sentences place the greatest demand on verbal WM, but the key question is whether the passive sentences significantly disrupted the accuracy of performing the digit task.

3.2 Sentence Length

A straightforward index of sentence complexity is sentence length. Generally speaking, the more words in a sentence, the more complex it is in terms of grammatical structure. Certainly, passive sentences were expected to be longer than active sentences given the need to use auxiliary verbs and prepositional phrases in the passive voice. Unrelated word pairs were also predicted to result in longer sentences than related words, based on the assumption that participants must use more words to connect two unrelated words in a sentence. These hypotheses were supported by an analysis of mean sentence lengths, as shown in Figure 1. As expected, sentences constructed using unrelated word pairs were significantly longer than sentences constructed using related word pairs, F (1, 44) = 38.05, $p < .001$, $\eta p^2 = .46$. Passive sentences were also significantly longer than active sentences, F (1, 44) = 19.31, $p < .001$, $\eta p^2 = .31$. No other findings were statistically reliable. The item analyses similarly revealed a main effect of word pair relatedness, $F' = (1, 348) = 24.28$, $p < .001$, $\eta p^2 = .07$, and sentence complexity, $F' = (1, 348) = 272.97$, $p < .001$, $\eta p^2 = .44$, in the same directions found in the participant analyses. The WM task type was also reliable in the item analyses, $F' = (2, 348) = 4.76$, $p < .01$, $\eta p^2 = .03$, but this effect must be qualified by an interaction of sentence complexity and WM task type, $F' = (1, 348) = 3.27$, $p < .05$, $\eta p^2 = .02$. Overall, sentences were longest in the control condition and shortest in the visual WM condition. However, this order of sentence length held only for the active sentences. For passive sentences, the shortest sentence length was found in the verbal WM condition, similar to what was observed in the production rate data.

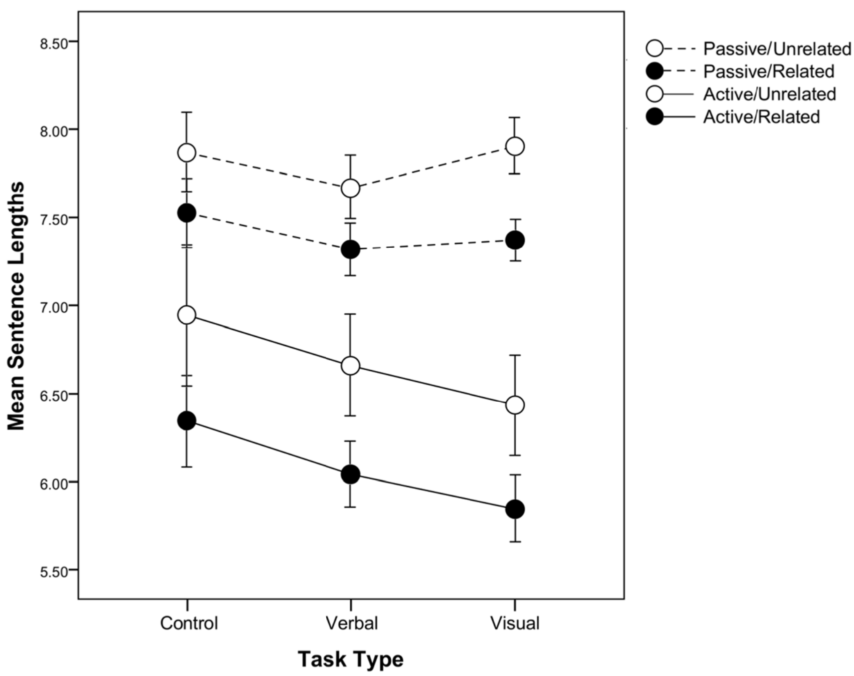

Figure 1. The mean sentence length for each condition.

3.3 Working Memory Interference

With regard to the working memory task, the control task (single task, no sentence production) accuracy scores were first analyzed to determine whether the verbal and visual tasks were comparable in difficulty. The mean accuracy score for verbal WM (*M* = .97; *SD* = .04; *n* = 46), did not differ from that of the visual task (*M* = .97; *SD* = .08; *n* = 46). An interference score was then calculated to take into account individual differences in working memory task performance. For example, if Participant A scored 80% on the verbal task while writing, while Participant B scored 90%, it might appear that written sentence production required more working memory for Participant A. But if the control task accuracy was also only 80% for Participant A when carrying it out as a single task, then in fact it should be concluded that the writing task did not interfere at all with the verbal WM task. Conversely, if Participant B was 100% accurate when the verbal task was performed in isolation, then the lower score of 90% on dual task performance would suggest a significant degree of verbal WM was required for concurrent sentence production. Thus, for each participant and task, a WM interference score was calculated by subtracting the control task accuracy from dual task accuracy. Note that it was not possible to compute a WM interference score for each word pair because the control task accuracy existed only for each participant.

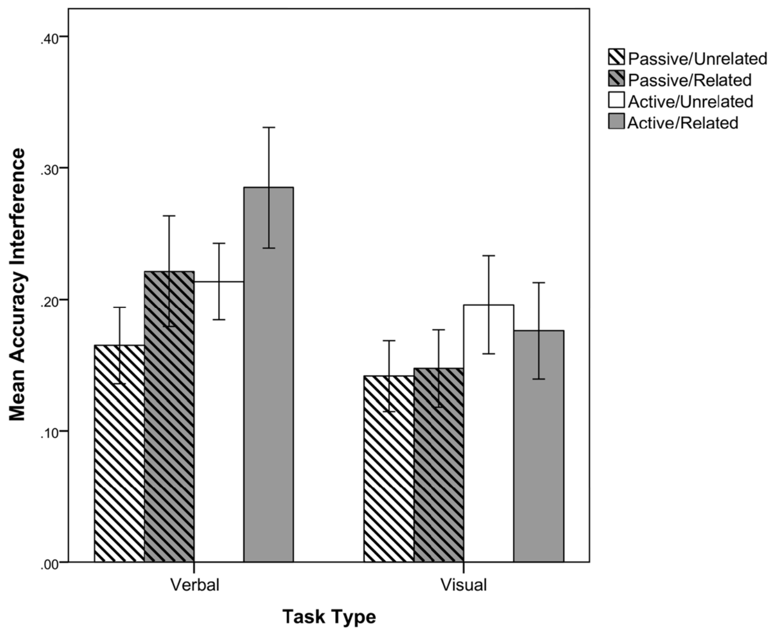

Figure 2. The mean interference score for each condition.

The mean WM task interference scores are shown in Figure 2. A main effect of WM task type was observed, F (1, 44) = 5.97, $p < .05$, $\eta\rho^2 = .12$. Writing interfered more with the verbal WM task (M = .22) compared with the visual WM task (M = .17). The other important source of variance in the 2 X 2 X 2 ANOVA was an interaction of the WM task with the relatedness of the word pairs, F (1, 44) = 3.88, $p < .06$, $\eta\rho^2 = .08$. This effect fell just short of statistical significance, but inspection of the means in Figure 2 suggests that relatedness affected the verbal WM task rather than the visual WM task, contrary to our predictions. A simple effect of relatedness was reliable in the verbal WM task, F (1, 45) = 5.51, $p < .05$, $\eta\rho^2 = .11$, but not in the visual WM task.

Overall, there was neither a reliable effect of sentence complexity nor any interactions with this factor. Contrary to our expectations, if anything it was the active sentences that appeared to demand more, rather than less, verbal WM resources compared with the more grammatically complex passive sentences (see Figure 2). The same pattern held for visual WM. Thus, the active sentences (overall M = .22) tended to produce more interference than the passive sentences (overall M = .17) for both kinds of concurrent tasks.

To explore the data further, a separate 2 X 2 ANOVA was conducted on the interference scores for the verbal WM task and another one for the visual WM task. For the verbal WM task, there was a main effect of noun relatedness, F (1, 44) = 5.40, $p < .05$, $\eta\rho^2 = .11$ and no other effects. Sentences composed with related nouns required more verbal WM resources relative to those using unrelated nouns. In the analysis of visual WM, there were no significant sources of variance.

As noted above, item analyses could not be performed for the interference analysis because an accuracy score for a single word pair was unavailable in the no sentence control condition. For each person who participated, a score was available reflecting performance when no sentence was concurrently produced. However, for each item or word pair no such score existed.

3.4 Word Frequency

We used the Nelson et al. (1998) norms to equate related and unrelated word pairs in terms of raw frequencies and concreteness. After collecting the data, it came to our attention that Balota et al. (2007) recommended the use of log frequencies based on the updated HAL print norms that include electronic media rather than older norms employed by Nelson et al. (1998). We discovered that the log HAL frequency of the related word pairs (M = 9.3; SE = .16) were in fact somewhat higher than those of the unrelated word pairs (M = 8.7; SE = .16), t (118) = 3.10. Even though all the words were relatively familiar (on the log scale, a value less than 4.0 would indicate a low frequency word), it is possible that log HAL frequency also contributed to the relatedness effects. We examined this issue with hierarchical multiple regression, with relatedness entered first in the model. With both factors entered in the equation the model was statistically reliable for sentence production time, F (2,357) = 17.11, $p <$ .001. The standardized coefficient beta ($b*$) was statistically reliable for relatedness ($b*$ = -.229, $p < .001$) and log HAL frequency ($b*$ = -.121, $p < .05$). R^2 increased from .072 to .087 with the addition of log HAL frequency as a predictor, indicating that only 1.5% of the variance was explained by adding log HAL frequency to the model. The words produced per second showed a similar pattern with R^2 increasing from .053 to .066 with log HAL frequency ($b*$ = .124, $p < .05$) added after relatedness ($b*$ = .184, $p <$ .001); F (2,357) = 12.61, $p < .001$. Sentence length, by contrast, revealed only a reliable effect of relatedness ($b*$ = -.168, $p < .01$) with a non-significant contribution from log HAL frequency ($b*$ = -.063, p = .26); F (2,357) = 7.41, $p < .001$. Such an item analysis could not be computed for interference scores because there was no control value for a single word pair. In sum, although relatedness consistently accounted for the most variance, log HAL frequency did influence to a degree sentence production time and the words produced per second. As will be discussed below, it is not obvious how an influence of log HAL frequency working in tandem with noun relatedness would help to explain the pattern of results obtained in the experiment.

4. Discussion

4.1 Findings and Interpretations

It was hypothesized that unrelated noun prompts would slow sentence writing times relative to related noun prompts, because of the extra time needed to plan a conceptual link for distantly associated nouns. Prior research has shown that initiation times to type the first keystroke of a sentence are in fact longer for unrelated compared with related nouns by about 700 ms (Kellogg, 1994; Rosenberg, 1977). Thus, the majority of the extra time (approximately 2 s) required to produce sentences in the unrelated condition compared with the related condition probably came from the linguistic encoding processes of translating ideas into a grammatical string of words and transcribing it through typing. Consistent with this interpretation, unrelated sentences were reliably longer compared with related sentences. Similarly, passive sentences took more time and included more words than did active sentences. As expected, the production of passive sentences appeared to entail more grammatical encoding because these have a longer, more complex sentence structure with more words to encode.

Using the dual task logic, we anticipated that poorer accuracy on the working memory tasks while participants were writing sentences would reflect the degree to which planning or grammatical encoding required visual versus verbal WM resources. Specifically, we predicted that more difficulty in planning would disrupt visual WM but not verbal WM. By contrast, greater difficulty in grammatical encoding ought to have impaired accuracy for verbal WM but not visual WM. Recall that we had anticipated the unrelated nouns would most disrupt the visual WM task. Instead, we found that the related—not the unrelated—noun prompts most disrupted the verbal WM task. It is difficult to attribute this effect to the planning stage of sentence generation, because the related nouns are known to require minimal planning compared with unrelated nouns (Kellogg, 2004; Rosenberg, 1977). Thus, it seems likely that the related nouns were disrupting verbal WM because of their demands on a stage of sentence generation downstream from planning. Prompt relatedness unexpectedly appeared to affect grammatical encoding or possibly, still further downstream, the phonological and orthographic encoding of the words to be typed. Of interest, the negative impact of related prompts on verbal WM was virtually identical for both active and passive sentences.

In sum, the manipulations of relatedness and grammatical voice affected the composing process as expected with respect to the product measure of sentence length and the process measure of typing time. The longer sentences and the additional time it took to type the sentence for unrelated nouns was consistent with the view that more planning was involved compared with related nouns. Yet, this manipulation impacted verbal WM rather than visual WM, contrary to our expectations from past findings and theorizing (Kellogg, 1996; Kellogg et al., 2007).

A limitation of our study is that the related and unrelated word pairs were not fully equated in terms of their familiarity, at least using the print frequency metric advocated by Balota et al. (2007). Even so, this familiarity difference would not seem to provide a ready explanation for our unexpected finding that related pairs demanded more verbal WM resources than did unrelated pairs. If anything one would expect the more familiar related pairs to be easier to maintain in verbal WM compared with less familiar unrelated pairs. Yet, it was the related pairs that caused the most disruption of the concurrent digit task. In short, we are not convinced that controlling for log HAL frequency would eliminate the theoretically troubling outcome that relatedness had on verbal WM.

The passive voice required more time and more words to complete a sentence compared with active voice, consistent with the view that their grammatical encoding demands varied as expected. At least under dual task conditions, especially when combined with a verbal WM task, the passive sentences were also generated at a slower rate in terms of words per second. Even so, the interference scores for the working memory tasks revealed no reliable differences between sentence types, but showed that, if anything, actives produced more interference than passives. Clearly, the prediction that passives would selectively disrupt a concurrent verbal WM task was not confirmed.

4.2 Theoretical Implications

It was anticipated that manipulating relatedness would impact the planning of conceptual content rather its grammatical encoding. Consistent with this view, unrelated nouns are known to slow the initiation of sentence typing compared with related nouns, suggesting an effect limited to planning. Why, then, did related nouns also apparently impair the grammatical encoding process more than the unrelated nouns? We scrutinized the sentences produced by participants to see if they yielded any clues that could explain why the related noun prompts so impaired verbal, but not visual, WM. We wondered, for example, if the sentence structures tended to be more complex when the noun prompts were related. Certainly, this did not seem plausible given that related sentences were reliably shorter in length than unrelated sentences, contrary to what one would expect with greater complexity. Neither could we intuit any consistent pattern in the grammatical structure or the word choices in the sentences generated with related as opposed to unrelated prompts.

One flaw in the Kellogg (1996) model that may be relevant is its assumption that the demands of planning on visual WM and the demands of grammatical encoding on verbal WM are static over time. Fayol (1999) argued that a complete account of writing must include an understanding of the temporal dynamics of sentence production in relation to working memory demands. For example, when it is difficult to meet the immediate demands on working memory made by several processes that must be integrated, writers can slow their rate of motor transcription or increase the duration of pauses. Such strategies would enable the writer to spread out the demands of these processes on working memory over time. How might strategies for allocating working memory resources over time help to account for the unexpected pattern of results observed here?

A plausible explanation is that the related nouns were integrated rapidly during planning into a single chunk that was then immediately cascaded forward to grammatical encoding. For example, with *door* and *knob* as related prompts, the semantic planning might well have been completed before the sentence was initiated, resulting in a large package of two integrated ideas being sent forward for grammatical encoding (*The knob on the door was hard to turn)*. The reason grammatical encoding demanded the most verbal WM for related nouns is that they were already integrated into a single conceptual chunk that needed to be expressed as a whole in words. By contrast, the temporal dynamics of translating unrelated nouns into sentences was likely much slower and more piecemeal. One of the two noun prompts could have been grammatically encoded and then typed before a plan was developed for how to finish the sentence. For example, with the unrelated prompts *ice* and *jail*, one might start the sentence (*The ice…)* before determining the semantic relationships that would drive the grammatical encoding process to completion (*…melted quickly in the hot jail*.) In short, the strategy of starting to type the opening of the sentence before the two unrelated nouns were fully integrated into a single chunk would result in a lower demand on verbal WM compared with related nouns. This does not imply that initiation time ought to be shorter for unrelated nouns. It is known from past work that unrelated nouns delay sentence initiation, but it is unknown whether semantic integration into a single chunk is completed even with the extra processing time. Instead, the long time required for planning with unrelated nouns may trigger the strategy of starting the sentence prior to integrating the two nouns successfully.

Effects of the sort outlined above have been observed in spoken language production. Of particular relevance here, Griffin (2001) found evidence of incremental grammatical encoding in speech production. When speakers produce a sentence frame of the form *The A and the B are above the C* to describe objects in a picture, the ease of naming *B* and *C* affected how long people gaze at the objects in the picture. When there were multiple alternative names for these objects and when the frequency of the dominant names was low, then the naming of *B* and *C* required additional processing resources. Of interest is the fact that the initiation of the preliminary phrase (*"The A…"*) was not impacted by the difficulties in encoding the names for *B* and *C*. Once they had encoded the name for A, they initiated production of the sentence.

The incremental selection and phonological encoding of the nouns in Griffin's (2001) experiment provide a likely temporal model for understanding the present findings. Because the unrelated nouns were difficult to link semantically during planning, participants may have initiated grammatical encoding to output the first unrelated noun and then returned to planning the semantic relationship before incrementally encoding the second noun. Indeed, it is known that unrelated nouns take more time to plan prior to the first keystroke (Kellogg, 2004), because of the difficulty of integrating them semantically. This difficulty may well not be resolved by the time the writer types the first word or phrase of the sentence. Further conceptual planning can continue even after the first words of the sentence are grammatically encoded and executed in motor output. Thus, unrelated nouns can take longer not just for planning, but also for grammatical encoding and transcription, because they are fed forward in piecemeal fashion. Put differently, the longer production times required for unrelated sentences may result not just from the fact that the sentences were longer with more words to encode and type, but may also result from the dynamics of grammatical encoding in relation to planning.

The phenomenon of incremental grammatical encoding has also been reported with typed written production. Nottbusch (2010) found that a complex phrase structure (a noun phrase with a subordinated phrase embedded within it) was produced as a unit suggesting grammatical encoding had been completed prior to the first keystroke. Here is an example, translated from the German sentences actually tested in the experiment: "The red stars with the blue circles are beside the yellow arrows" (p. 784). By contrast, coordinated phrases seemed to be produced in a piecemeal fashion. For example, "The black triangles and the yellow rectangles are beside the green stars" (Nottbusch, 2010; p. 784). It appeared that the subordinated phrase called for syntactic planning or grammatical encoding at the level of the full phrase prior to initiating sentence production. By contrast, the coordinated phrase was encoded more incrementally by comparison. Consistent with this interpretation, the writers paused longer before initiating the coordinating conjunction ("and") than they did before initiating the preposition "with." This difference clearly indicated that additional grammatical encoding was underway after typing the first words ("The black arrows"). We suspect a similar phenomenon occurred in the present experiment with the first unrelated word typed before planning the semantic content and grammatically encoding the remainder of the sentence took place.

However, other investigators have found effects supporting the grammatical encoding of larger units in spoken language rather than the simple phrase encoding reported by Griffin (2001). Smith and Wheeldon (1999) reported that speakers delayed initiating speech articulation when describing a picture called for a complex phrase at the start of a single clause sentence compared with a simple phrase; this was so even when the lexical and syntactic complexity of the sentences were matched in terms of a verb phrase and two noun phrases. To illustrate, a complex phrase might be "The dog and kite moved above the house" whereas a simple phrase might be "The dog moves above the kite and the house" (Smith and Wheeldon, 1999, p. 205). The unit of grammatical encoding thus appears to be at more than the initial article and noun but less than the full clause. On the other hand, in their replication of the task used by Griffin (2001), Torrance and Nottbusch (2012; Study 3) found the same results that she did—highly incremental grammatical encoding—for both speaking and writing.

We interpret our results as suggesting that planning the semantic content and grammatically encoding a noun phrase can proceed incrementally in an alternating fashion when it is difficult to plan the conceptual content with unrelated noun prompts. However, for related noun prompts, the semantic link between the nouns might have been first completed—imposing a greater load on verbal WM—before grammatical encoding and motor output ensued. Such an interpretation is consistent with view that the writing process is dynamically managed during written composition depending on a large variety of specific task demands (Fayol, 1999; Fayol, Foulin, Maggio, & Lété, 2012). Pause lengths and the size of units sent forward for grammatical encoding can vary with the specific demands of the task. Moreover, in some situations, low level aspects of grammatical encoding and motor output can be combined in parallel with high level semantic planning processes, while at other times they must be sequenced. In a recent review of the literature, Olive (2014) similarly argued that low level transcription processes can occur in parallel with higher level planning or sentence formulation processes, as long as both do not exceed the available executive attention of working memory. A phrase within a clause can be conceptually planned and then cascaded to the processes of grammatical encoding before the entire clause is conceptualized. Our results fit well with such a model whereby the unit of analysis cascaded forward depends on ease of initial semantic processing during planning. Unrelated noun prompts appear to result in a more piecemeal cascading of information compared with the integrated chunk resulting from two related nouns.

The use of keystroke logging might be helpful in further exploring such an interpretation of our findings. If it is correct that both related nouns (*door, knob)* are chunked and sent forward as a unit, then the first key stroke should occur relatively quickly. Moreover, the subject and predicate might already be fully planned at least for active sentences, with linguistic encoding and motor output following in single burst (e.g., *He turned the knob on the door)*. By contrast, unrelated nouns (*ice, jail)* should reveal a long pause not just before initiating production but also after the opening phrase stating the subject of the sentence is typed (e.g., *The jail...)*.

In conclusion, the present results suggest that the role of WM in written sentence production is markedly more complex than previously postulated. Manipulating the relatedness of the noun prompts does not simply affect the visual WM demands of planning, but also appears to influence the verbal WM demands of grammatical encoding. One mechanism by which this might plausibly occur is the temporal dynamics of the planning and grammatical encoding stages of processing.

Acknowledgements

The authors thank Mark Torrance and an anonymous reviewer for their helpful comments on an earlier draft of this manuscript.

References

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

Balota, D. A., Yap, M. J., Cortese, H. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445-459. http://dx.doi.org/10.3758/BF03193014

Bock, J. K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review, 3*, 395-421. http://dx.doi.org/10.3758/BF03214545

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Chenoweth, N. A., & Hayes, J. R. (2003). The inner voice in writing: *Written Communication, 20*, 99-118. http://dx.doi.org/10.1177/0741088303253572

Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics.* New York: Harcourt Brace Jovanovich.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Fayol, M. (1999). From on-line management problems to strategies in written composition. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing* (pp. 13-23). Amsterdam: Amsterdam University Press.

Fayol, M., Foulin, J-N., Maggio, S., & Lété, B. (2012). Towards a dynamic approach of how children and adults manage text production. In Grigorenko, E. L., Mambrino, E., & Preiss, D. D. (Eds.), *Writing: A mosaic of perspectives* (pp. 141-158). New York: Psychology Press.

Gough, P. B. (1965). Grammatical transformations and the speed of understanding. *Journal of Verbal Learning and Verbal Behavior, 4*, 107-111. http://dx.doi.org/10.1016/S0022-5371(65)80093-7

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*, B1-B14. http://dx.doi.org/10.1016/S0010-0277(01)00138-X

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg, & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kellogg, R. T. (2004). Components of working memory in written sentence production. *American Journal of Psychology, 117*, 341-361. http://dx.doi.org/10.2307/4149005

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy and S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382-397. http://dx.doi.org/10.1016/j.actpsy.2006.02.005

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research, 5*, 159-190. http://dx.doi.org/10.17239/jowr-2013.05.02.1

Levy, C. M., & Marek, P. (1999). Testing components of Kellogg's multicomponent model of working memory in writing: The role of the phonological loop. In M. Torrance & G. C. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory in text production* (pp. 25-41). Amsterdam: Amsterdam University Press.

Martin, R. C., Shelton, J., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33*, 83-111. http://dx.doi.org/10.1006/jmla.1994.1005

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. http://w3.usf.edu/FreeAssociation/.

Nottbusch, G. (2010). Grammatical planning, execution, and control in written sentence production. *Reading and Writing, 23*, 777-801. http://dx.doi.org/10.1007/s11145-009-9188-4

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research, 6*, 173-194. http://dx.doi.org/10.17239/jowr-2014.06.02.4

Rosenberg, S. (1977). Semantic constraints on sentence production: An experimental approach. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory*. Hillsdale, NJ: Erlbaum.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5-42. http://dx.doi.org/10.1006/cogp.1997.0658

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken language production. *Cognition, 73*, 205-246. http://dx.doi.org/10.1016/S0010-0277(99)00053-0

Torrance, M., & Nottbusch, G. (2012). Written production of single words and simple sentences. In V. W. Berninger (Ed.) *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 403-422). New York: Psychology Press.

Warriner, J. E., & Griffith, F. (1977). *English grammar and composition* (5th ed.). New York: Harcourt Brace Jovanovich

Appendix: Word Pairs

Related	Unrelated
door knob	ice jail
cow calf	lion dollar
arm leg	lip missile
iron steel	lobster radio
ball racket	corn emerald
meat butcher	spinach tennis
orange juice	mouse kite
brother sister	pimple river
key lock	flask lamb
angel devil	factory autumn
leaf tree	ski trumpet
man lady	horse lemon
baby infant	mother sky
hand glove	cider boat
wick candle	bread hurdle
fish trout	wallet swamp
army soldier	skull rocket
jewel crown	thorn menu
doctor nurse	blood square
armor knight	shin penny
animal zoo	bath cave
king queen	mink cigar
ape gorilla	mist flea
laugh clown	cork dentist
flag banner	piano pyramid
band rubber	rainbow nose
hammer nail	ankle beef
gold silver	onion pencil
apple pie	mother foam
island harbor	lettuce money