Research Article

The Role of Working Memory in Attentional Allocation and Grammatical Development under Textually-enhanced, Unenhanced and No Captioning Conditions

Minjin Lee, *Yonsei University*

https://orcid.org/0000-0002-8248-4672

mlee.minjin@gmail.com

Andrea Révész, *University College London*

https://orcid.org/0000-0003-1093-4336

a.revesz@ucl.ac.uk

ABSTRACT

This study investigated the extent to which individual differences in working memory (WM) mediate the effects of captions with or without textual enhancement on attentional allocation and L2 grammatical development, and whether L2 development is influenced by WM memory in the absence of captions. We employed a pretest-posttest-delayed posttest design, with 72 Korean learners of English randomly assigned to three groups. The groups differed as to whether they were exposed to news clips without captions, with textually-enhanced captions, or with unenhanced captions during the treatment. We measured attentional allocation with eye-tracking methodology, and assessed development with an oral production, a written production and a fill-in-the-blank test. To assess various aspects of WM, we employed measures of phonological and visual short-term memory (PSTM, VSTM) and the executive functions of updating, task-switching, and inhibitory control. We found that, in both captions groups, higher PSTM was associated with higher oral production gains. For the enhanced captions group, PSTM was also positively related to gains on the written production test. Participants in the no-captions group, however, showed a positive link between VSTM and oral production gains. Attentional location only correlated positively with updating ability and PSTM under the enhanced captions condition. These results, overall, indicate that WM can moderate the effects of captions on attention and L2 development, and various WM components may play a differential role under various captioning conditions.

Keywords: working memory, textual enhancement, captioning, eye-tracking, grammatical knowledge

INTRODUCTION

The role of working memory in second language acquisition (SLA) has been the subject of much SLA research over the past two decades. Working memory refers to an individual's cognitive ability to temporarily store and manipulate information (Baddeley, 1992; Juffs & Harrington, 2011). It is presumed to predict the ability to carry out and learn a wide range of complex cognitive activities, including the development of second language (L2) knowledge and skills. Indeed, a growing amount of empirical research suggests that there is a positive relationship between working memory capacity and L2 outcomes. Results of several narrative reviews (Jackson, 2020; Jeon & Yamashita, 2014; Juffs & Harrington, 2011; Wen, 2016; Williams, 2012) and meta-analyses (Linck et al., 2014; Shin, 2020) indicate that, overall, working memory has robust, positive links with L2 processing and learning outcomes.

Within the larger area of WM research, an expanding strand has been concerned with how working memory may relate to the effectiveness of L2 instructional interventions. Based on the assumptions that WM is linked to attentional allocation (Robinson, 1995, 2003) and that attention is essential for the processing of new linguistic information (Schmidt, 2001), researchers have hypothesized that WM plays a role in the extent to which learners benefit from instruction that aims to draw learners' attention to L2 constructions (e.g., Mackey et al., 2002). By now, substantial empirical evidence has accumulated indicating that WM may be associated with the amount of learning that results from exposure to L2 instruction. Previous findings, however, also suggest that its influence may vary according to type of instruction. Li's (2017) synthetic review, for example, found that working memory significantly correlates with the impact of explicit feedback but not with that of implicit feedback. Interestingly, however, the narrative component of Li's review yielded inconsistent patterns for the link between WM and type of feedback. Similar, Granena and Yilmaz's (2018) research synthesis of aptitude-treatment interaction studies concluded that working memory tends to facilitate L2 learning through explicit pedagogical interventions. Their results, however, were less conclusive about how WM relates to L2 development when learners are exposed to implicit instruction.

The aim of the present study was to investigate how WM may influence the effects of captions, textually enhanced or unenhanced, on the attention to and acquisition of an L2 grammatical construction. In doing so, we hoped to help to clarify the nature of the relationship between WM and implicit instruction. There are several novel aspects of this study. First, although much research is available about both the roles of captions (Montero Perez et al., 2013) and textual enhancement (Lee & Huang, 2008) in L2 development, it remains unexplored how WM may interact with their effectiveness in promoting L2 learning of grammar. Second, while researchers have already begun to examine the role of WM in attentional allocation to captions (Gass et al., 2019; Kam et al., 2020) and textually enhanced grammatical features (Indrarathne & Kormos, 2018), little is known about how working memory may influence the amount of attention paid to L2 grammatical features included in captions, enhanced or unenhanced. Finally, unlike most existing SLA studies that have tended to investigate a single function of executive control (see, however, Indrarathne & Kormos, 2018; Michel et al., 2019; Révész et al., 2017), we focused on several executive functions when operationalizing working memory capacity.

Before describing the methodology and results of our research, we turn to a review of the theoretical and empirical work serving as the background for the study.

LITERATURE REVIEW

Attention and Working Memory

In the field of SLA, attention is generally regarded as a prerequisite for the acquisition of new linguistic knowledge. Recently, cognitive psychologists have conceptualized attention as a multiple system, which includes an external and an internal component (Chun et al., 2011). External attention is associated with perceptions, and can be triggered by external stimuli (e.g., through manipulating visual or audio input). Internal attention, on the other hand, involves the selection and modulation of internally-generated information, comprising what is held in working and long-term memory. Internal attention is also associated with cognitive control and executive functioning. Although the two systems can operate separately, they are also assumed to interact. Working memory is held to provide the interface for the internal and external systems. For example,

the internal attentional system determines what part of the perceptual information attended to gets selected for encoding in working memory. Conversely, perceptual attention can be influenced by the current content of working memory. That is, there is presumed to be a strong relationship between attentional allocation and working memory.

Of the different working memory models, Baddeley and Hitch's (1974) multicomponent model of working memory has probably been the most widely applied in L2 research. This model sees working memory as consisting of three components: the central executive and two subsidiary systems, the visuo-spatial sketchpad or visual-spatial short-term memory (VSTM) and the phonological loop or phonological short-term memory (PSTM). The central executive is responsible for controlling and regulating complex cognitive operations. It is associated with executive functions such as the ability to switch between tasks; to deliberately inhibit responses when required; and to monitor, revise and update incoming information (e.g., Baddeley, 1996; Miyake et al., 2000). While these executive functions (task-switching, inhibition, updating) seem to share a common underlying construct, there is empirical evidence suggesting that they are associated with different subprocesses and thus are separable (Miyake et al., 2000). The phonological loop temporarily stores and processes verbal and acoustic information, whereas the visuo-spatial sketchpad retains and processes visual images and spatial relations. A fourth component, the episodic buffer, was later added to the model (Baddeley, 2000). Its role is to combine verbal, visual and spatial codes through integrating information from the various subsystems and long-term memory into *episodes* or multimodal units.

Working memory is generally assumed to be limited in capacity, but researchers have proposed that, depending on the type of information processing required, time-sharing across various activities will be differentially demanding. According to Wickens' (2007) multiple-resource model, for example, there are various cognitive resource pools, which can be distinguished across three dichotomous dimensions: processing stage (perception vs. response), modality (auditory/vocal vs. visual/manual), and processing code (verbal vs. spatial). Each of these dimensions relates to a distinct area of performance, and task difficulty will be determined by the amount of interference within a certain dimension. That is, competition for similar or same type of resources may result in processing difficulty. For instance, carrying out two perception tasks (e.g., listening and reading) is more likely to lead to interference than performing a perception and response task (e.g., listening and writing), as the activities in the first combination involve the same processing stage (i.e., perception). In light of this framework, we would expect that the processing of captions might lead to interference at the perception stage, as learners need to perceive several sources of input (textual, visual, and audio) simultaneously. These increased demands on perception, however, might be better handled by learners who have better phonological and/or visual short-term memory and/or more developed executive functions.

In the area of instructed SLA, most previous WM studies have focused on the roles of PSTM and/or executive control in L2 development. To measure executive control, researchers have usually utilized a single WM index (see, however, Indrarathne & Kormos, 2018), taking the form of a complex working memory task such as a reading, listening, or operation span task. These complex WM tasks require both storing and processing information, thus, besides short-term memory capacity, are thought to provide information about the central executive component of working memory (Gathercole, 1999). Operationalizing executive control in terms of a single measure, however, ignores research findings suggesting that various subconstructs associated with executive control are separable (Miyake et al., 2000). Thus, to avoid the possible danger of construct underrepresentation, SLA studies would profit from employing several, carefully chosen indices of executive functions. For example, the functions of task-switching, inhibition, and updating all seem pertinent to learning from exposure to textually enhanced captions, the pedagogical intervention investigated here. To benefit from enhanced captions, learners would likely need to be able to switch between a focus on meaning and form, inhibit interfering information to pay attention to the enhanced forms, and keep updating information through replacing irrelevant with relevant pieces of incoming input. Hence, to get a fuller picture of the role of WM in learning through enhanced captions, in the present study we included measures of these various executive functions.

To date, PSTM capacity has typically been measured by span tasks in SLA research, which involve recalling rising number of digits or nonwords of increasing length. Visual-spatial sketchpad capacity has rarely been assessed by SLA researchers, as this WM component is considered less relevant to the processing of verbal information (see however, Michel et al., 2019; Révész et al., 2017; Sachs, 2011) than PSTM. However, it would appear important to include, besides a measure of PSTM, an index of VSTM in studies exploring the processing of a combination of textual and imagery information. Given that working memory is limited in capacity and interference is likely to arise when learners need to perceive various input sources concurrently (Wickens, 2007), VSTM capacity might influence the extent to which learners can pay attention to verbal input that is presented simultaneously with images. As watching captioned videos involves dealing with multimodal input, we utilized an index of VSTM along with PSTM.

Working Memory and Attention to and Learning of L2 Grammar

The role of WM in learning grammatical constructions under instructed conditions has been the subject of a considerable number of studies. In a research synthesis of aptitude-treatment interaction research, Granena and Yilmaz (2018) reached the conclusion that working memory promotes learning under explicit pedagogical conditions, but the patterns are less clear for the link between WM and L2 development when learners receive implicit instruction. Indeed, when we consider some specific areas of instructed SLA research, empirical evidence for the role of WM in L2 development also appears inconclusive. For example, in some studies a positive relationship has been detected between performance on complex working memory capacity tasks and the extent to which L2 learners benefit from implicit corrective feedback (e.g., Goo, 2012; Mackey et al., 2002; Mackey et al., 2010; Mackey & Sachs, 2012; Révész, 2012). Other studies of implicit corrective feedback, however, yielded no evidence for an association between complex working memory capacity and development in the knowledge of L2 grammar (Trofimovich et al., 2007). In an attempt to synthesize these findings, Li (2017) found significant but weak correlations between WM and the effects of explicit feedback, but no significant WM effects

emerged for implicit feedback. In the narrative component of his synthesis, however, Li also concluded that the results are inconsistent. Similarly, some researchers found that, depending on learners' WM, task manipulations differentially affect learners' development in targeted grammatical constructions (e.g., Kim et al., 2015). Others, however, identified little difference in the extent to which learners developed their grammatical knowledge under various task conditions as a function of their WM (e.g., Baralt, 2010; Jung, 2017).

It is also worth noting that the few studies which have included separate measures of PSTM and executive functions often found distinct results for the different indices (e.g., Li & Roshan, 2019; Mackey & Sachs, 2012; Révész, 2012; Trofimovich et al., 2017). For example, Révész (2012), in a study of recasts and grammatical development, observed that non-word and digit span scores were positively related to learners' gains on an oral production test, whereas a reading span test had positive links with participants' gains on a written production and grammaticality judgement test. Based on these results, Révész speculated that, high PSTM capacity might have enabled learners to keep the grammatical information entailed in recasts longer in short-term memory, leading to increased likelihood that long-term memory traces were encoded. Then, these long-term memory traces could be subjected to data-driven processes, promoting development in procedural knowledge (N. Ellis, 2005). On the other hand, high complex working memory capacity might have made learners more capable of paying conscious attention to the feedback they received, leading to more improvement in declarative knowledge. In turn, Révész argued, increased procedural and declarative knowledge likely helped learners to perform better on the oral and written tests respectively. Similarly, Li and Roshan (2019) found that phonological short-term memory and complex working memory may relate differently to the effectiveness of written feedback. More research is needed to reach firmer conclusions about how various components of WM may be linked to L2 grammar learning.

One way to shed light on the relationship between WM and L2 grammatical development is to investigate how attentional allocation is influenced by working memory in instructed settings, given the assumption that attention is a prerequisite for the encoding of new L2 knowledge

(Schmidt, 2001). To date, very few instructed SLA studies have considered how WM capacity may influence the amount of attention paid to L2 grammatical constructions. A notable exception is Indrarathne and Kormos's (2018) recent research, which investigated how attentional allocation and L2 grammar learning is affected by WM memory capacity under various instructional conditions. The authors, unlike most previous SLA research, employed a broad range of WM measures, including an index of PSTM (forward digit-span test) and measures of the updating (Keep Track task), task-switching (Plus Minus task) and inhibitory control (Stroop task) functions of the central executive. Informed by the results of a factor analysis, the researchers used a composite score of WM (based on participants' Forward-digit, Keep Track and Stroop scores) and a measure of task-switching ability in their analyses. Amount of attention paid to the target feature was assessed by the means of eye-tracking methodology. Particularly relevant to the present study are Indrarathne and Kormos' findings that the composite WM scores had medium to strong correlations with participants' pretest-posttest gains under both a textually unenhanced and enhanced instructed condition, with the correlations being stronger for the group who received textually enhanced input. Interestingly, however, WM only had a strong correlation with fixation durations when learners were exposed to textual enhancement, no such link was found under the unenhanced condition. The researchers speculated that this might indicate that WM is not only involved in explicit but also implicit learning of grammar.

Although the results of Indrarathne and Kormos' (2018) study are informative, more research is necessary to confirm their findings. Also, it is important to explore how WM may relate to attentional allocation to grammatical features under different types of instructional conditions. Therefore, the aim of this study was to examine how WM may influence learning L2 grammar through exposure to captioned videos, a means of L2 instruction growing in popularity with multimedia materials becoming more and more available. We also compared whether, depending on WM, learners benefit differentially from textually unenhanced or unenhanced captions. We formed the following research questions:

1. To what extent do individual differences in working memory capacity moderate L2 grammatical development when learners are exposed to news clips with textually unenhanced captions, textually enhanced captions, or no captions?

2. To what extent does learners' attention allocated to the target linguistic construction relate to their working memory capacity? Is this relationship influenced by whether learners are exposed to textually unenhanced or enhanced captions?

METHOD

Design

We conducted this study as part of a larger project, in which we investigated the effects of textually enhanced and unenhanced captioning on attentional allocation and second language development. We were also interested in exploring how these relationships might be influenced by individual differences in working memory capacity. In the present article, we report our results for working memory capacity (see Lee & Révész, 2020, for a description of the effects of captioning on allocation of attention and L2 learning).

We employed a pretest–posttest–delayed posttest design, with 72 L2 learners of English being randomly assigned to a no-captions group ($n = 24$), an unenhanced captions group ($n = 24$), and an enhanced captions group ($n = 24$). At each testing session, participants were administered an oral production test, a written production test, and a fill-in-the-blank test. We employed eye-tracking technology to record participants eye-gaze behaviors during the treatment under the captioning conditions. Besides the tests and treatment, all participants completed a background questionnaire, the Oxford Placement Test (OPT), five working memory measures, and an exit questionnaire.

Participants

Our initial pool of participants included 93 Korean learners of L2 English, but we had to exclude 21 participants either because they did not complete the delayed posttest ($n = 4$)

or because of data loss or technical issues during eye-tracking ($n = 17$). The remaining 72 participants were all university students, with a mean age of 21.86 ($SD = 1.42$). Forty-five participants were female and 27 were male. According to a one-way ANOVA run on the OPT scores, the three groups had comparable levels of proficiency (listening section: $F (2, 69) = 1.23$, $p = .23$, $\eta^2 = .03$; grammar section: $F (2, 69) = 1.12$, $p = .33$, $\eta^2 = .03$; total: $F (2, 69) = 1.14$, $p = .33$, $\eta^2 = .03$). Participants' average OPT scores ranged from 166 to 187, indicating that they were at level C1 and above in terms of the Common European Framework for Reference (CEFR).

Target Linguistic Construction

The target construction was the use of the present perfect when reporting news. It has been suggested that learners of English find it difficult to master properties of the English tense and aspect system, especially if morphosemantic differences exist between the first and second language (e.g., Bardovi-Harlig, 2001; Gabriele, 2009; Montrul & Slabakova, 2002). In light of this, we expected participants to experience difficulty with the target construction, as the present perfect only occurs with a few stative and intransitive verbs in Korean. In most cases, the past simple form is used to denote meanings associated with the English present perfect construction in Korean (Chang, 1996; Sohn, 1995). Probably as a result, Korean learners of English often overuse the English past simple form in present perfect contexts (Han & Hong, 2015).

Treatment Task

In the treatment tasks, we asked participants to imagine that they work as an editor in a newsroom. Their task was to preview news items and assess the appropriacy of titles and categories for them. If both the news title and category were appropriate, we instructed them to press 'z' on the keyboard. If either of them was inappropriate, they had to press 'm'. The instructions were provided on the computer screen. While performing the treatment tasks, the participants were exposed to the target linguistic construction through aural and/or textual mode according to the conditions they were assigned to. For the unenhanced captions group, we provided non-manipulated captions. For the enhanced captions group, both the present perfect and past simple constructions were enhanced through using yellow fonts. Our rationale for enhancing the past simple forms was to decrease the chance that the treatment leads to overuse of the present perfect by learners. The no captions group watched the news clips without captions (control group).

We developed 24 treatment tasks altogether. They were based on news clips we collected from various news channels. Each clip lasted between 20-50 seconds. In all clips, the present perfect construction was used to introduce the topic, then details were provided using the past simple. The clips contained the same number of passive and active uses of the present perfect. The captions were added with the software Camtasia 8.0.

Outcome Measures

Three types of outcome measures were employed to examine learners' development in the use of the target constructions: an oral production test, a written production test, and a fill-in-the-blank test. Our rationale for including these types of tests was to assess learners' development in different modalities (oral vs. written) and in less and more controlled tasks (production vs. fill-in-the-blank). We developed three sets of each test, which were counterbalanced across participants, modality, and testing sessions using a Latin Square design.

The oral and written production tests, whose format only differed in modality, required the participants to watch news items in their native language (Korean) and report the content in English. All test items, written and oral, differed in content. We included five non-captioned news clips in both the oral and written production tests. The lengths of the clips were similar to those in the treatment tasks. To assess the performances, we applied a partial scoring system. When participants supplied a correct, partially correct, or incorrect present perfect form in obligatory contexts, they received 2, 1, and 0 points respectively. Given the very small number of partial scores awarded, we converted the scores into a dichotomous scale (1 point for correct, 0 points for incorrect). The maximum score on both production tests was 5 points.

In the fill-in-the-blank test, we included 10 target items along with 30 distractors. In the target items, there were two

blanks in each sentence to complete, with one blank targeting the use of the present perfect and the other the past simple. We used the same partial scoring system to assess participants' use of the present perfect as for the production tasks, but the data were again recoded into a dichotomous scale (1 point for correct, 0 points for incorrect) due to the low incidence of partial scores. Thus, the maximum score for the present perfect items was 10 points respectively.

Measures of Attention

The amount of attention allocated to the target linguistic constructions was measured with eye-tracking methodology. While performing the treatment tasks, the participants' eye-movements in the two captions groups were recorded. We used a remote eye-tracker with a temporal resolution of 60 Hz (Tobii X2-60), which was mounted on a 15-inch screen laptop. The visual angle was about 22 degrees. We employed a nine-point calibration procedure to calibrate the eye-tracking system, which was repeated after every eight treatment tasks.

For the purposes of this study, our area of interest was the present perfect form in the captions (see Figure 1). Four eye-tracking measurements were used to assess the amount of attention allocated to the target linguistic construction: (a) number of visits, where a visit refers to all the fixations made within an area of interest from the time a participant's eyes first enter that area of interest until they leave; (b) first pass reading time, which is the sum of all fixation durations during the first visit to the area of interest; (c) second pass reading time, which is defined as the sum of fixation durations when the eyes return to an area of interest for the first time after an initial visit; and (d) skipping rate, which is the proportion words skipped during first pass reading (Conklin & Pellicer-Sánchez, 2016). Before further analyses, we cleaned the eye-tracking data following recommendations outlined by Conklin and Pellicer-Sánchez (2016).

Working Memory Measures

We employed five working memory measures to test different constructs associated with working memory capacity: a non-word span task, a forward Corsi block test, a color shape task, a stop signal task, and an automated

operation span (AOSPAN) task. Except for the non-word span task, which was presented using PowerPoint, all working memory measures were administered on a computer using Inquisit 4 Lab (Millisecond, 2015).

Figure 1. *Area of Interest*

Nonword Repetition Span

We administered a Korean version of the non-word span task (Jung, 2017) to assess participants' PSTM. The participants were presented with 32 nonwords, which were developed based on the phonotactic rules of Korean. The words varied in length from 4 to 11 syllables. There were four sets for each syllable length. The nonwords were presented aurally in a random order, and the participants had to listen and recall them orally. Approximately 10 seconds were allowed for the participants to give their responses, which were audio recorded. Each recall was scored as either correct or incorrect. Span length was determined as the maximum number of syllables that participants correctly recalled at least twice for each syllable-length, resulting in a range of scores from 4 to 11.

Forward Corsi Block Task

We used the forward Corsi block task to measure participants' visual-spatial short-term memory capacity. First, participants were presented with 2 to 9 identical blocks highlighted in different orders on a computer screen, each highlighted block appearing as a blue square. Then, the participants had to click the blocks in the same order as they had been highlighted. There were two trials for each block

length. Total score was calculated by counting the number of correctly repeated sequences until the test ended (Kessels et al., 2000).

Color Shape Task

We employed the color shape task to measure task-switching ability, an executive function (Miyake et al., 2004). The task involved two different blocks, non-switching and switching blocks. In the non-switching blocks, there were two separate sub-blocks, a color or a shape block. In the color blocks, the participants were asked to provide a response depending on the color of the shape they had seen. They were instructed to press "A" for green and "L" for red. In the shape blocks, they had to press "A" if they had seen a triangle and "L" if they had seen a circle. In switching blocks, however, participants had to switch between making a decision based on either a color or a shape, according to a cue letter that appeared on the screen (C for color and S for shape). Performance on the task was assessed in terms of switching cost, which was calculated based on the difference in mean reaction times between non-switching and switching blocks (e.g., Altgassen et al., 2014; Friedman et al., 2006; Gold et al., 2013; Miyake et al., 2004). Reaction times of individual participants were trimmed to values within two standard deviations above and below the mean.

Stop Signal Task

We used a version of the stop signal task as a measure of inhibitory control, which is another executive function. The participants were asked to respond to an arrow stimulus as quickly and correctly as they could. The arrow stimulus was displayed on the screen pointing either left or right. The participants had to respond by pressing 'D' on the keyboard if the arrow pointed to the left and 'K' if the arrow pointed to the right. Some arrow stimuli were accompanied by an auditory signal (a beep). In this case, the participants were asked to withhold response. Performance on the task was measured in terms of the stop signal reaction time (SSRT), which refers to the time required for an individual to inhibit the response when the auditory signal occurred. Participants with slower SSRT were more likely to exhibit difficulties in inhibiting their response. SSRTs were trimmed to values
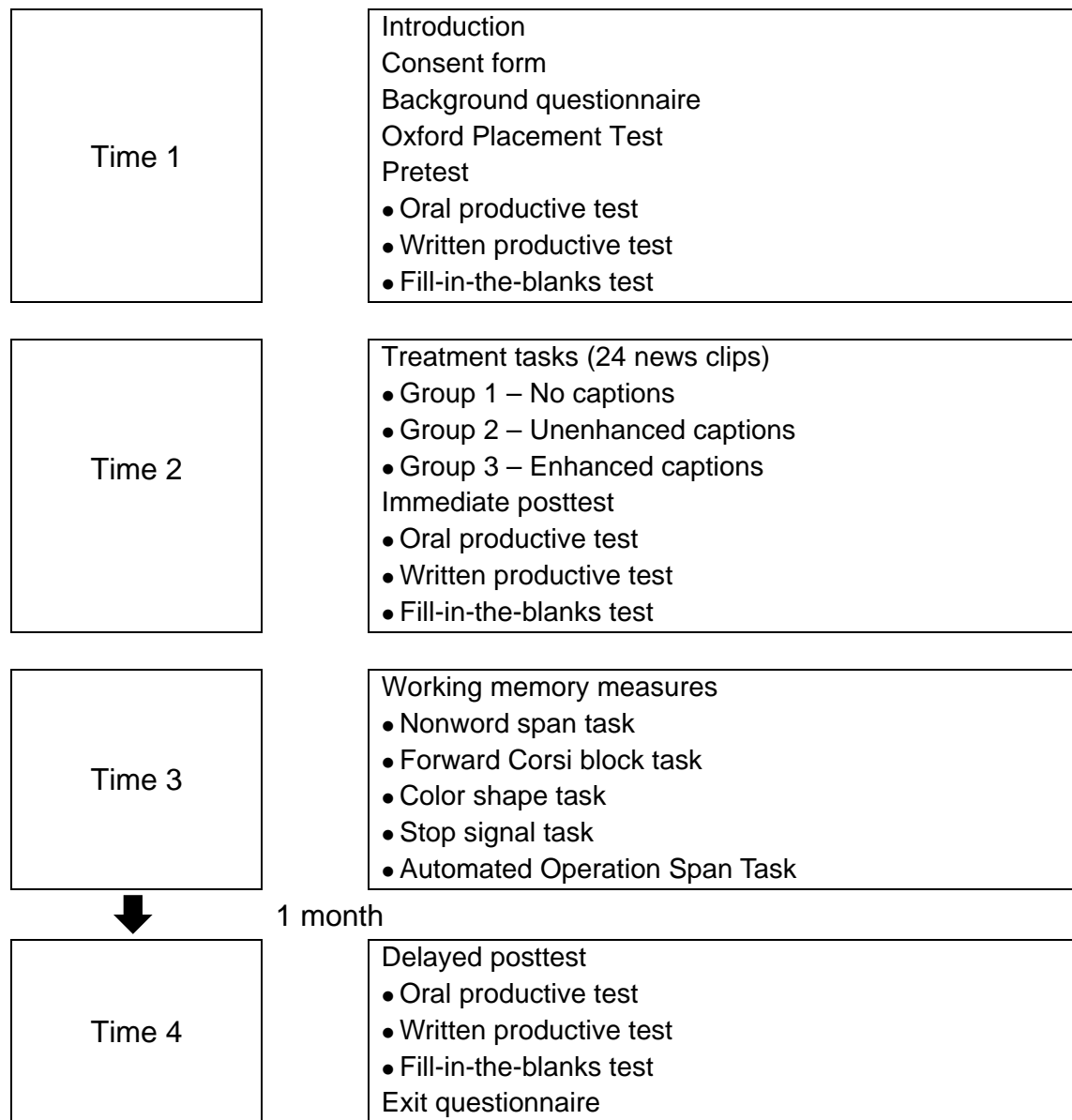
within two standard deviations above or below the mean for each individual (Congdon et al., 2012; Enticott et al., 2006).

Automated Operation Span Task (AOSPAN)

We employed an automated version of the operation span task (Turner & Engle, 1989) to assess the executive function of updating. First, a simple arithmetic equation was presented on the screen, followed by a letter. Then, the participants were asked to remember the letter and make a judgement about whether a given answer to the arithmetic equation was correct or incorrect. After completing each equation, feedback on the accuracy of participants' responses was provided automatically. At recall, 12 possible letters in a 4 x 3 matrix were displayed, and the participants' task was to select the letters in the order they had appeared. The number of correctly recalled letters, ranging from 3 to 7, was defined as the set size. The task included three sets of each set size, which were presented in a random order. As an index of AOSPAN, we used the total score, which was calculated based on the total number of letters recalled in their correct positions within a particular string (e.g., Friedman & Miyake, 2005, Miyake, 2001; Miyake et al., 1999). The maximum score was 75.

Data Collection Procedure

The experimental schedule is summarized in Figure 2. First, participants completed a consent form (15 min), a background questionnaire (10 min), the OPT (40 min), and the pretest. The pretest included a version of the oral production test (15-18 min), the written production test (15-18 min), and the fill-in-the-gap test (40 min). Second (2 days later), they carried out 24 treatment tasks, followed by the immediate posttest. Third (2 days later), the working memory tests were administered to the participants. The duration of the WM tests varied, with the non-word span task taking 9 to 10 minutes, the Corsi block test 4 to 5 minutes, the color-shape task 40 to 45 minutes, the stop signal task 9 to 10 minutes, and the AOSPAN 30 to 40 minutes. The order of the working memory tests was counterbalanced across participants. Finally (4 weeks after the posttest), participants were asked to complete a delayed posttest and an exit questionnaire. Each session lasted approximately 2 to 3 hours.

Figure 2. *Experimental Schedule*

Time 1	Introduction Consent form Background questionnaire Oxford Placement Test Pretest • Oral productive test • Written productive test • Fill-in-the-blanks test
Time 2	Treatment tasks (24 news clips) • Group 1 – No captions • Group 2 – Unenhanced captions • Group 3 – Enhanced captions Immediate posttest • Oral productive test • Written productive test • Fill-in-the-blanks test
Time 3	Working memory measures • Nonword span task • Forward Corsi block task • Color shape task • Stop signal task • Automated Operation Span Task

↓ 1 month

Time 4	Delayed posttest • Oral productive test • Written productive test • Fill-in-the-blanks test Exit questionnaire

Statistical Analyses

To address research question 1, the statistical analyses involved conducting a series of mixed-effects models using the *lme4* package in the R statistical environment (R Core Team, 2016). We used the *glmer* function to construct logistic mixed effects models, as our outcomes measures were binary due to dichotomous scoring. As follow-up analyses, Pearson correlations were computed. To address research question 2, we conducted additional correlational analyses between the eye-tracking and working memory indices. Following Plonsky and Oswald (2014), we classified *r* values of .25, .40, and .60 as small, medium, and large-size correlations. An alpha level of $p < .05$ was set for all tests.

RESULTS

Preliminary Analyses

The descriptive statistics for the developmental outcome measures and eye-tracking measurements are summarized in Tables 1 and 2 (also presented in Lee & Révész, 2020). To assess whether there were differences among the three groups on the oral production, written production, and fill-in-the-blank pretests, we carried out a series of logistic mixed-effects regression analyses. In all models, the fixed effect was group, participant and item served as random effects, and participants' scores on one of the pre-tests was the dependent variable. As shown in Table 3 (also presented in Lee & Révész, 2020), the analyses found no significant difference in the three group's performance, that is, they had comparable pretest scores.

Table 1. *Descriptive Statistics for Developmental Outcome Measures*

| | | Oral Production Test | | | | Written Production Test | | | | Fill-in-the-blank Test | | | |
| | | | | 95% CI | | | | 95% CI | | | | 95% CI | |
Captioning	N	M	SD	Low	Up	M	SD	Low	Up	M	SD	Low	Up
No captions													
Pretest	24	.92	1.14	.44	1.40	.79	1.05	.35	1.24	1.21	.36	.46	1.95
Posttest	24	1.54	1.59	.87	2.21	.81	.86	1.50	3.00	1.37	.28	.79	1.96
Del. posttest	24	1.12	1.19	.62	1.63	.79	.88	.42	1.16	1.40	.30	.78	2.01
Unenhanced													
Pretest	24	.83	1.01	.41	1.26	.96	.27	.39	1.52	1.42	1.45	.80	2.03
Posttest	24	1.87	1.70	1.16	2.59	2.25	.36	1.50	3.00	3.04	.48	2.05	4.03
Del. posttest	24	1.50	1.61	.82	2.18	2.12	1.89	1.32	2.92	2.60	2.63	1.49	3.72
Enhanced													
Pretest	24	.87	.88	.48	1.22	.83	.22	.37	1.30	1.40	1.61	.72	2.08
Posttest	24	3.43	1.30	2.89	3.98	4.00	1.35	3.43	4.57	6.60	2.27	5.64	7.56
Del. posttest	24	4.00	1.32	3.44	4.56	3.25	1.77	2.50	4.00	6.44	2.25	5.49	7.39

Table 2. *Descriptive Statistics for Eye-tracking Measurements*

| | | | | 95% CI | |
	N	M	SD	Low	Upper
Number of visits					
Unenhanced captions	24	1.62	.68	1.33	1.91
Enhanced captions	24	2.20	.47	2.01	2.40
First-pass reading time					
Unenhanced captions	24	131	62	105	158
Enhanced captions	24	175	52	153	197
Second-pass reading time					
Unenhanced captions	24	90	76	57	122
Enhanced captions	24	270	82	235	304
Skipping rate					
Unenhanced captions	24	.24	.24	.14	.34
Enhanced captions	24	.07	.17	< .01	.14

Table 3. *Results for Logistic Mixed-effects Models Examining Pretest Performances*

| | Fixed effects | | | | Random effects | | | |
| | | | | | by participant | | by item | |
	Estimate	SE	z	p	variance	SD	variance	SD
Oral productive								
Intercept	−1.77	.35	−4.96	< .01**	.88	.94	.02	.14
Unenhanced captions	−.27	.47	−.58	.56				
Enhanced captions	−.10	.46	−.21	.83				
Written productive								
Intercept	−2.36	.50	−4.68	< .01**	2.06	1.44	.07	.26
Unenhanced captions	0.31	.60	.51	.61				
Enhanced captions	−.08	.61	−.13	.90				
Fill-in-the-blank								
Intercept	−2.59	.44	−5.85	< .01**	.30	.55	.17	.41
Unenhanced captions	.62	.46	1.35	.18				
Enhanced captions	.47	.47	1.01	.31				

** p < .01, * p < .05

Table 4. *Descriptive Statistics for Working Memory Measures*

| | | N | M | SD | 95% CI | |
					Lower	Upper
Non-word span	No captions	24	8.42	1.86	7.63	9.20
	Unenhanced captions	24	8.33	1.86	7.55	9.12
	Enhanced captions	24	8.96	1.65	8.20	9.66
Corsi block	No captions	24	62.62	15.45	56.09	69.16
	Unenhanced captions	24	58.75	18.99	50.73	66.77
	Enhanced captions	24	63.46	16.74	56.39	70.53
Color shape	No captions	24	410	203	324	496
	Unenhanced captions	24	476	259	366	585
	Enhanced captions	24	371	185	293	449
Stop signal	No captions	24	300	74	268	331
	Unenhanced captions	24	280	64	253	308
	Enhanced captions	24	291	89	253	329
AOSPAN	No captions	24	53.12	10.48	51.23	61.53
	Unenhanced captions	24	56.37	12.02	49.64	58.95
	Enhanced captions	24	54.29	11.18	51.97	57.23

The descriptive statistics for the working memory measures are presented in Table 4. Note that the non-word span maximum total score was 11, the Corsi block maximum total score was 88, and the AOSPAN maximum total score was 75. A series of one-way ANOVAs found no significant difference among the three groups on any of WM measurements (non-word span: $F_{(2, 69)} = .86$, $p = .43$, $\eta^2 = .02$; Corsi block: $F_{(2, 69)} = .52$, $p = .60$, $\eta^2 = .01$; color-shape: $F_{(2, 69)} = 1.41$, $p = .25$, $\eta^2 = .04$; stop signal: $F_{(2, 69)} = .38$, $p = .65$, $\eta^2 = .01$; AOSPAN: $F_{(2, 69)} = .51$, p

= .60, η^2 = .01). We also conducted a series of Pearson correlations among the various working memory measures. As shown in Table 5, we only found a medium-size positive correlation between the nonword span task and AOSPAN results. These results suggest that non-word span and AOSPAN tapped somewhat overlapping constructs, but overall, the five WM tasks assessed different aspects of working memory.

Table 5. *Correlations among Working Memory Measures*

		Non-word	Corsi block	Color shape	Stop signal
Corsi block	r	−.05	-		
	p	.68	-		
Color shape	r	−.08	.05		
	p	.50	.67		
Stop signal	r	−.26	−.20	.16	
	p	.05	.09	.17	
AOSPAN	r	.40**	−.00	.10	−.13
	p	< .01	.99	.40	.29

** *p* < .01, * *p* < .05

RQ1: Moderating Effects of Working Memory on Development in L2 Grammatical Knowledge

To address the first research question, we ran a series of logistic mixed-effects models. In each model, the fixed effects were group, time, one of the five working memory measures, and their interactions; participant and item served as random effects, and participants' performance on one of the outcome measures was the dependent variable (see Tables 1–15 in the supplementary information online for the results of the full models). For the fixed effects, we also included by-participant and by-item random slopes (time-by-participant and group-by-item) to create a maximum model (Barr et al., 2013). When the maximal model did not converge, we removed the random effect explaining the least variance till convergence was achieved (Blom et al., 2012). Our predictors of interest were the three-way interactions between time, group, and the working memory measure. A significant three-way interaction would mean that the extent of gains achieved by the three groups on the respective assessment task was differently influenced by the component of working memory in the model. Three models

yielded significant interaction effects: the models including the non-word span scores as a fixed effect and the written or oral production test scores as the dependent variable and the model constructed with the Corsi block task and the oral production test results.

To investigate the interaction effects, we ran a series of Pearson correlational analyses. As shown in table 6, on the oral production test, participants' non-word span scores had a large-size correlation with the unenhanced captions group's pretest-posttest and pretest-delayed posttest gain scores and a medium-size correlation with the enhanced captions group's pretest-delayed posttest gains. On the oral production test, we also found medium- size correlations between the no captions' group Corsi block scores and their pretest-posttest and pretest-delayed posttest gains. On the written production test, medium-size correlations were observed between the enhanced group's non-word span scores and their pretest-posttest and pretest-delayed posttest gain scores.

RQ2: Relationship between Attention and Working Memory

To examine the relationship of attention allocated to the target linguistic construction in captions with working memory capacity, we ran a series of Pearson correlation analyses between the various eye-tracking and working memory measures for the unenhanced and enhanced captions groups separately. As shown in table 7, for the unenhanced captions group, no significant correlations were observed. For the enhanced captions group, on the other hand, three out of four eye-movement indices (number of visits, second pass reading time, skipping rate) positively correlated with AOSPAN to a medium degree. In addition, second pass reading time also had a medium-size positive correlation with participants' non-word span scores. These results indicate that learners who had better updating ability were less likely to skip the enhanced present perfect construction during their first visit. Also, they were more likely to pay more frequent visits to the enhanced form and spend more time fixating on it during their second visit. In addition, there was greater likelihood that participants with better PSTM fixate longer on the enhanced construction during their second visit.

Table 6. *Results of Follow-up Correlation Analysis*

Test/Group		Pretest-Posttest Gains	Pretest-Delayed Posttest Gains
Oral Production Test			
No Captions Group			
Non-word span	r	.32	.30
	p	.12	.15
Unenhanced Captions Group			
Non-word span	r	.75**	.71**
	p	< .01	< .01
Enhanced Captions Group			
Non-word span	r	.38	.47*
	p	.07	.02
Written Production Test			
No Captions Group			
Non-word span	r	−.17	−.14
	p	.43	.51
Unenhanced Captions Group			
Non-word span	r	.21	.27
	p	.33	.21
Enhanced Captions Group			
Non-word span	r	.58**	.52**
	p	< .01	< .01
Oral Production Test			
No Captions Group			
Corsi block	r	.45*	.50*
	p	.03	.01
Unenhanced Captions Group			
Corsi block	r	−.20	−.15
	p	.34	.48
Enhanced Captions Group			
Corsi block	r	−.17	−.19
	p	.42	.38

Table 7. *Results of Pearson Correlations between Eye-tracking and Working Memory Measures*

		Number of visits	First Pass	Second pass	Skipping rate
Unenhanced captions					
Non-word	r	.10	.17	.11	−.10
	p	.66	.44	.61	.63
Corsi block	r	−.12	.25	.21	−.22
	p	.58	.23	.32	.30
Color shape	r	−.29	−.17	−.09	.18
	p	.17	.43	.69	.40
Stop signal	r	−.30	−.28	−.34	−.25
	p	.15	.19	.10	.24
AOSPAN	r	.24	.37	.25	−.26
	p	.26	.07	.24	.22

Enhanced captions					
Non-word	r	.32	.29	.42*	−.20
	p	.13	.17	.04	.35
Corsi block	r	.02	.01	.12	−.10
	p	.94	.99	.57	.65
Color shape	r	−.11	−.17	.00	−.12
	p	.61	.43	.99	.58
Stop signal	r	−.10	−.16	−.16	.00
	p	.63	.44	.47	.99
AOSPAN	r	.41*	.33	.42*	−.44*
	p	.04	.11	.04	.03

** p < .01, * p < .05

DISCUSSION

Working Memory, Enhanced Captions, and Development in L2 Grammatical Knowledge

Our first question was concerned with the relationship with the moderating effects of working memory on L2 development when learners are exposed to videos with textually unenhanced, textually enhanced captions, or no captions. While we found no associations between the indices of executive functions measured by the color shape task, stop signal, and AOSPAN tasks and L2 development, some significant relationships emerged between the short-term memory measures and gain scores on the productive outcome measures. On the oral production tests, the non-word span scores had large-size correlations with the unenhanced captions' groups pretest-posttest and pretest-delayed posttest gain scores, and medium-size links with the enhanced captions group's pretest-delayed posttest development. On the written production test, the enhanced caption group's pretest-posttest and pretest-delayed posttest gains also showed medium-size correlations with non-word span scores. Interestingly, for the no captions group, we found the Corsi block scores to correlate, to a medium degree, with the oral production pretest-posttest and pretest-delayed posttest gain scores.

It was expected that we would find a positive role for PSTM in relation to the learning of grammar through captions, similar to previous research investigating other pedagogical interventions (e.g., Révész, 2012; Trofimovich et al., 2017). Those with higher PSTM were probably able to maintain the textual input in captions in short-term memory longer, which led to greater likelihood that long-term memory traces could be encoded. The resulting long-term memory traces could then serve as the basis for implicit, data-driven learning processes, promoting the proceduralization of L2 knowledge (N. Ellis, 2005). Increased procedural knowledge, in turn, might have enabled learners with high PSTM test scores to perform better on the oral production tests, as the oral assessments were likely to draw more extensively on procedural knowledge than the written tests. A reason why PSTM might have had a smaller influence under the enhanced condition could be that textual enhancement induced learners to focus on rehearsing the enhanced part of the captions only, which required less PSTM capacity than maintaining the full caption in short-term memory, enabling also lower PSTM learners to achieve considerable gains. The likely focus on the enhanced parts, by the enhanced captions group, might also explain why PSTM, for this group, was also related to gains on the written production test. Those with high PSTM were probably able to rehearse the enhanced part more extensively, and thereby more likely become aware of the target construction (Robinson, 1995) and develop their declarative knowledge. Then, their superior declarative knowledge might have enabled them to achieve better scores on the written description test, as this test allowed for more extensive reliance on declarative knowledge than the oral production test.

We also predicted that VSTM would be related to the extent to which learners benefit from the treatment.

Following Wickens' (2007) multiple-resource framework, interference probably arose at the perception stage, since processing the videos involved perceiving audio as well as visual input at the same time. Those with higher VSTM probably coped with this interference better given their superior ability to process visual information, leaving them with more capacity to pay attention to the textual input. It was surprising, however, that VSTM only moderated learners' gains in the no-captions group. One way to account for this finding might be that the captions were salient enough to draw learners' attention away from the images, making the captions groups prioritize the processing of captions over images.

The salience of captions might also provide an explanation why no significant relationships emerged between the indices of executive functions and L2 learning through captions. The captions, due to their perceptual salience and information value, might have attracted the attention of learners with low executive control as well, resulting in a ceiling effect. That is, there might have been lower need to rely on executive functions such as the ability to switch attention and inhibit information to facilitate attention to the captions. Notably, the absence of a role for executive functions in L2 development is in line with some instructed SLA studies that have investigated implicit instructional conditions. It would appear that, whether executive control influences the effectiveness of pedagogical techniques, depends on the exact nature of the intervention. Further research is needed to identify variables that may moderate this potential association.

Another finding worthy of discussion is that working memory had a negligible effect on the use of grammatical knowledge when measured by the fill-in-the-blank test. This contradicts the results reported in Indrarathne and Kormos' (2018) study, where under both explicit and implicit instructional conditions, working memory ability was found to be associated with gains in receptive knowledge of the target construction when measured by a fill-in-the-blanks test. Indrarathne and Kormos, however, calculated a composite working memory score of four measures (forward digit test, keep track task, plus minus task, and stroop task), thus their and our results are not directly comparable. Further studies are needed to clarify this relationship.

Working Memory, Attention, and Textual Enhancement in Captions

Our second research question investigated the extent to which learner attention allocated to the target linguistic construction entailed in captions related to their working memory capacity, and whether this relationship was influenced by textual enhancement. We only found significant correlations between WM and the eye-gaze behaviors of participants for the enhanced captions group. Learners with higher AOSPAN scores were less likely to skip part of the present perfect construction during first pass reading. In addition, during their second visit, they were more likely to visit the enhanced target form and fixate longer on it. When visiting the second time, participants with higher PSTM also had longer fixations at the enhanced construction.

These findings, taken together with the developmental results, suggest that those with higher AOSPAN and PSTM scores were more likely to allocate attention to the enhanced part of the captions, but it was only those with higher PSTM who showed greater development in the use of the target construction. There are at least two possible explanations for this. First, given that AOSPAN and PSTM measure overlapping construct to some degree (they were found to have a medium-size positive correlation in the present study), these results might be interpreted as suggesting that AOSPAN, in the absence of a strong PSTM, was not sufficient to trigger L2 development. In other words, high updating ability (measured by AOSPAN), by allowing learners to better monitor and code the relevance of incoming information (Miyake et al., 2000), might have facilitated learners' attention to the enhanced input as it was perceived to be relevant. This increased attention, however, only led to more learning when learners were also able to rehearse the enhanced textual input in short-term memory more extensively due to their superior PSTM.

Another possible way of explaining the finding that increased attention by high AOSPAN learners did not lead to greater L2 development relates to our previous discussion of the role of perceptual salience. As argued above, even lower-AOSPAN learners' attention might have been captured to a considerable extent given the salience of enhanced captions. Thus, although high AOSPAN learners might have overall paid more attention to the target construction, there could have been a ceiling effect in terms

of the amount of attention needed to trigger rehearsal in short-term memory, a prerequisite for subsequent L2 learning to take place.

LIMITATIONS & FUTURE DIRECTIONS

Before drawing our conclusions, it is important that we consider some limitations of the study and how these might be addressed in future research. A methodological limitation pertains to the absence of verbal protocol data. In future studies, gathering verbal protocol comments, such as stimulated recall data, would allow for investigating the conscious operations in which learners engage while processing captioned videos, making it possible to obtain a fuller picture of attentional processes and how they might relate to working memory. Another potential methodological issue concerns the precision of the eye-tracking system. We used a remote eye-tracker with a temporal resolution of 60 Hz to collect the eye-tracking data. In further research, using a more sophisticated eye-tracker would enable researchers to obtain more precise eye-tracking measurements, increasing the reliability of the results.

Besides addressing these limitations, it would be worthwhile to replicate this study with different target constructions and with learners from different linguistic and educational backgrounds. This would allow for testing the transferability of the results to other linguistic features and learner populations. It would also be interesting to explore how the effectiveness of captioning may be linked to other cognitive individual difference factors, with the ultimate aim of identifying optimal matches between learner aptitude profiles and instructional interventions.

CONCLUSION

The aim of this study was to examine how WM may relate to the effectiveness of captions, textually enhanced or unenhanced, in facilitating attention to and development in the knowledge of an L2 grammatical feature. We operationalized WM in terms of measures of various executive functions (updating ability, task-switching ability, and inhibitory control) and indices of phonological and visual short-term memory capacity. We assessed amount of attention with the help of eye-gaze measurements and gauged L2 development by the means of an oral and a written production test and a fill-in-the blanks test.

We found relatively few significant links between the WM measures and attention to and development in the knowledge of the target construction. Those with higher PSTM achieved larger gains on the oral production test when they were exposed to captioned videos. PSTM was also positively related to gains on the written production test when learners received textually enhanced captions. Better visual-short term memory, on the other hand, was linked to more oral production gains in the absence of captions in the videos. Interestingly, working memory (updating ability, PSTM) only had positive links to attentional allocation under the enhanced captions condition. These results, overall, indicate that WM can moderate the effectiveness of captions in facilitating L2 learning, and various components of WM may differentially relate to attention to and acquisition of L2 grammatical constructions.

REFERENCES

Altgassen, M., Vetter, N. C., Phillips, L. H., Akgün, C., & Kliegel, M. (2014). Theory of mind and switching predict prospective memory performance in adolescents. *Journal of Experimental Child Psychology*, *127*, 163–175. https://doi.org/10.1016/j.jecp.2014.03.009

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, *49A*, 5–28. https://doi.org/10.1080/027249896392784

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–90). Academic Press.

Baralt, M. (2010). *Task complexity, the cognition hypothesis and interaction in CMC and FTF environments*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Bardovi-Harlig, K. (2001). Another piece of the puzzle: The emergence of the present perfect. *Language Learning*, *5*, 215–264. https://doi.org/10.1111/j.1467-1770.2001.tb00018.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Blom, E., Paradis, J., & Sorenson Duncan, T. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular –s in child L2 English. *Language Learning*, *62*, 965–994. https://doi.org/10.1111/j.1467-9922.2012.00715.x

Chang, S-J. (1996). *Korean*. John Benjamins.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Congdon, E., Mumford, J. A., Cohen, J. R., Galvan, A., Canli, T., & Poldrack, R. A. (2012). Measurement and reliability of response inhibition. *Frontiers in Psychology*, *3*, 1–10. https://doi.org/10.3389/fpsyg.2012.00037

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, *32*, 453–467. https://doi.org/10.1177/0267658316637401

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305–352. https://doi.org/10.1017/S027226310505014X

Enticott, P. G., Ogloff, J. R., & Bradshaw, J. L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences*, *41*, 285–294. https://doi.org/10.1016/j.paid.2006.01.011

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, *37*, 581–590. https://doi.org/10.3758/bf03192728

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Gabriele, A. (2009). Transfer and transition in the SLA of aspect. *Studies in Second Language Acquisition*, *31*, 371–402. https://doi.org/10.1017/S0272263109090342

Gass, S., Winke, P., Isbell D. R., & Ahn, J. (2019). How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning and Technology*, *23*, 84–104. https://doi.org/10125/44684

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, *3*, 410–419. https://doi.org/10.1016/S13646613(99)01388-1

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in

aging. *The Journal of Neuroscience*, *33*, 387–396. https://doi.org/10.1523/JNEUROSCI.3837-12.2013

Goo, J. (2010). Working memory and reactivity. *Language Learning*, *60*, 712–752. https://doi.org/10.1111/j.1467-9922.2010.00573.x

Granena, G., & Yilmaz, Y. (2018). Aptitude-treatment interaction in L2 learning: A research synthesis. *Studies in English Education*, *4*, 803–830. https://doi.org/10.22275/SEE.23.4.02

Han, J., & Hong, S. (2015). The acquisition problem of English present perfect to Korean adult learners of English: L1 transfer matters. *English Language and Linguistics*, *213*, 141–164.

Indrarathne, B., & Kormos, J. (2018). The role of working memory in processing L2 input: Insights from eye-tracking. *Bilingualism: Language and Cognition*, *21*, 355–374. https://doi.org/10.1017/S1366728917000098

Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching*, *10*, 89–109. https://doi.org/10.14746/ssllt.2020.10.1.5

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*, 160–212. https://doi.org/10.1111/lang.12034

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*, 137–166. https://doi.org/10.1017/S0261444810000509

Jung, J. (2017). *Effects of task complexity, glossing and working memory on L2 reading and L2 learning*. Unpublished doctoral dissertation, University College London, London, UK.

Kam, E. F., Liu, Y. T., & Tseng, W. T. (2020). Effects of modality preference and working memory capacity on captioned videos in enhancing L2 listening outcomes. *ReCALL*, *32*, 213–230. https://doi.org/10.1017/S0958344020000014

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: standardization and normative data. *Applied Neuropsychology*, *7*, 252–258. https://doi.org/10.1207/S15324826AN0704_8

Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory: L2 question development through recasts in a laboratory setting. *Studies in Second Language Acquisition*, *37*, 549–581. https://doi.org/10.1017/S0272263114000618

Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, *42*, 625–651. https://doi.org/10.1017/S0272263120000108

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition*, *30*, 307–331. https://doi.org/10.1017/S0272263108080479

Li, S. (2017). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach* (pp. 41–70). John Benjamins.

Li, S., & Roshan, S. (2019). The associations between working memory and the effects of four different types of written corrective feedback. *Journal of Second Language Writing*, *45*, 1–15. https://doi.org/10.1016/j.jslw.2019.03.003

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*, 861–883. https://doi.org/10.3758/s13423-013-0565-2

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*, 501–533. https://doi.org/10.1111/j.1467-9922.2010.00565.x

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). John Benjamins.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, *62*, 704–740. https://doi.org/10.1111/j.1467-9922.2011.00649.x

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, *45*, 31–45. https://doi.org/10.1016/j.jslw.2019.03.002

Millisecond. (2015). Inquisit 4 Lab [Computer software]. Retrieved from http://www.millisecond.com/

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General*, *130*, 163–168. https://doi.org/10.1037/0096-3445.130.2.163

Miyake, A., Emerson, M. J., & Friedman, N. P. (1999). Good interactions are hard to find. *Brain & Behavioral Sciences*, *22*, 108–109. https://doi.org/10.1017/S0140525X99371788

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J-C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, *115*, 123–142. https://doi.org/10.1016/j.actpsy.2003.12.004

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Montero Perez, M., van den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*, 720–739. https://doi.org/10.1016/j.system.2013.07.013

Montrul, S., & Slabakova, R. (2002). The L2 acquisition of morphosyntactic and semantic properties of the aspectual tenses preterite and imperfect. In A.T. Perez-Lerouz & J. Liceras (Eds.), *The acquisition of Spanish morphosyntax: the L1-L2 connection* (pp. 113–149). Kluwer.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, *62*, 93–132. https://doi.org/10.1111/j.14679922.2011.00690.x

Révész, A, Michel, M., & Lee, M. (2017). Investigating IELTS Academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports Online Series*, 2017/3.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*, 283–331. https://doi.org/10.1111/j.1467-1770.1995.tb00441.x

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Blackwell.

Sachs, R. (2011). *Individual differences and the effectiveness of visual feedback on reflexive binding in L2 Japanese*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 1–32). Cambridge University Press.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics*, *41*, 873–900. https://doi.org/10.1017/S0142716420000272

Sohn, S-O. (1995). *Tense and aspect in Korea*. University of Hawaii Press.

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytic ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 171–195). Oxford University Press.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Multilingual Matters.

Wickens, C. D. (2007). Attention to the second language. *International Review of Applied Linguistics in Language Teaching*, *45*, 177–191. https://doi.org/10.1515/iral.2007.008

Williams, J. N. (2012). Working memory and SLA. In S.M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 427–441). Routledge.