THE JOURNAL OF ASIA TEFL
Vol. 20, No. 1, Spring, 2023, 64-88
http://dx.doi.org/10.18823/asiatefl.2023.20.1.5.64

The Journal of Asia TEFL

http://journal.asiatefl.org/

e-ISSN 2466-1511 © 2004 AsiaTEFL.org. All rights reserved.

The Cognitive Process with Intensive and Rapid Vocabulary Acquisition Tasks Using Multimedia Glosses

Meilin Jiang
Hong Kong Metropolitan University

Di Zou
The Education University of Hong Kong

Lap-Kei Lee
Hong Kong Metropolitan University

Fu Lee Wang
Hong Kong Metropolitan University

Nga-In Wu
The Hong Kong Polytechnic University

Digital learning environments afford intensive and rapid interaction for learners, which leads to a learning duration within a short learning period. Therefore, it is essential to know how to design information presentations when the learning duration is short and learning tasks are intensive. The current study designed learning tasks with different types of vocabulary glosses (e.g., corresponding pictures or sounds) and learning duration (3, 5, or 7 seconds per word). Learning results were analyzed considering learners' word recognition and passive recall of the meaning of vocabulary with descriptive analysis. Learners' characteristics (i.e., working memory capacity, executive function, and strategy in allocating information) were collected and analyzed through three types of phonological working memory capacity tests, one visuospatial working memory capacity test, and a vocabulary encoding survey. The result indicates that text-only glosses with 3s per word (15s whole learning task) learning duration led to the highest learning result in supporting learners' word recognition. Meanwhile, text + picture glosses with 5s per word (25s whole learning task) lead to the highest learning result in learners' passive recall of vocabulary meaning. However, when the learning duration is extended, image glosses may become redundant information. Besides, the result indicates that learners with higher phonological working memory, may store more redundant phonological information which disturbs the information processes. However, if learners have a better executive function, the negative effects may be relieved.

数字学习环境提供的密集而快速的交互，让学习时长被限制在了较短的时间内。因此，了解如何在学习时间较短、学习任务较多的情况下，更好的展示信息是非常必要的。当前研究设计了不同类型的词汇注释（例如，与词汇相关的图片或声音）和不同的单词学习时长（每个单词3秒、5秒或7秒）的学习任务。描述性分析被应用于分析学生的接受性词汇的词汇识别和词汇意义记忆成果。本研究通过三种类型的语音工作记忆容量测试、一种视觉空间工作记忆容量测试和学生词汇记忆策略的调查，收集和分析了学习者的特征（即工作记忆容量、大脑的执行功

能和信息分配策略）。结果表明，每个单词3秒（整个学习任务15秒）的纯文本注释，在辅助单词识别方面效果最好。同时，每个单词5秒（整个学习任务25秒）的文本+图片的注释在词汇意义被动回忆方面效果最好。然而，当学习时间延长后，图像注释可能成为冗余信息。此外，研究结果还表明，具有较高语音工作记忆的学习者可能会储存更多的冗余语音信息，从而干扰信息处理。然而，如果学习者具有更好的大脑执行功能，这种负面影响可能会减轻。

Keywords: Cognitive process, passive vocabulary learning, working memory capacity, vocabulary encoding strategy, multimedia glosses

Introduction

Intensive and rapid interaction may appear in different digital learning tools such as eBook (Xu et al., 2021), videos (Zhu et al., 2022), 3D images (Hassan Ja'ashan et al., 2022), AR (Chen et al., 2022), digital games (Chen & Liu, 2021; Kohnke et al., 2019) in the instruction of language acquisition. Different from traditional learning materials (e.g., printed books), the demonstrated information along with the interaction (e,g., pop-up messages as glosses for novel information or corrective feedback for the wrong answer; Kohnke et al., 2021) are usually not designed for presenting in a long learning period. As these messages are widely considered as a scaffolding method to provide external information and support knowledge acquisition (e.g., Blom et al., 2018; Chen & Yen, 2013), it is essential to explore how learners process the provided information within a short learning period and how to present information to strengthen the memory of information. As limited studies controlled the learning duration strictly within a short learning period, the current study aims to use intensive and rapid vocabulary acquisition tasks to embody learners' information processes and retention, and explore how to design information presentation for better learning results.

Multimedia gloss is a type of annotation that is used to present supplementary information in the digital learning environment with its modality affordance (Mohsen & Balakumar, 2011). The current study adopted multimedia glosses to present the meaning of vocabulary. The design of the multimedia glosses referred to previous studies (e.g., Boers et al., 2017; Kim & Gilman, 2008) and controlled the existence of pictures and sounds. To furtherly analyze how learners process the glosses within a short learning period, the current study adopted the survey of vocabulary encoding strategies and working memory capacity (WMC) tests. As the learning processes entail learners to select, store and organize useful information (Mayer, 2021), it was hypothesized that the survey of vocabulary encoding strategies and WMC tests can assist in embodying the memorizing process and explore the effects of personal characteristics in vocabulary acquisition.

In short, to enhance the learning experience and performance in digitalized learning environments, how learners process information with different information presentation elements within a short learning duration requires more definitude. The current study controlled the task learning duration from 15-35 seconds with different types of multimedia glosses. Through collecting participants' WMC and vocabulary encoding strategies, the factors that affect the processing of vocabulary learning materials can be furtherly analyzed.

Based on the considerations above, the following research questions are addressed.

1. What is the impact of different types of multimedia glosses and learning duration on vocabulary retention?
2. What is the impact of vocabulary encoding strategies on vocabulary retention?
3. What is the impact of working memory capacity (WMC) on vocabulary retention with different types of glosses and learning duration?

Vol. 20, No. 1, Spring 2023, 64-88

Background

Dual Coding Theory

Dual coding theory (DCT) is established based on empirical evidence of mental processes describing the mechanisms in process nonverbal (e.g., image) and verbal information (e.g., linguistic sounds) (Clark & Paivio, 1991). The fundamental assumption behind DCT is that the input information goes through different channels depending on its type (verbal or nonverbal) and task demands (Paivio, 2010). The task, for example, requires learners to name an object in terms of a picture can activate both verbal and nonverbal channels. Depending on the types of input stimuli and task types, verbal and nonverbal information can be processed separately or simultaneously (Paivio, 2010).

More explicitly, DCT emphasized that when learners receive external information, the mental representation is modality-specific (Sadoski & Paivio, 2013). Information can be received from several senses such as visual, auditory, and haptic, and different mental representations can be evoked because of the sensory type. For instance, as for auditory sensory modality, environmental sounds are nonverbal and spoken words are verbal. Regarding visual modality, pictures can be encoded as nonverbal information, and text is encoded as verbal information. In the current study, corresponding pictures of words are the nonverbal information and the corresponding sounds of words are verbal information and provided as glosses in facilitating the information processing.

Cognitive Theory of Multimedia Learning

The cognitive theory of multimedia learning (CTML) contributes to the digital learning resources design with several principles that are constructed considering three underpinning theories: “active processing assumption”, “limited-capacity assumption” and “dual-channel assumption” (Mayer, 2014). The active processing assumption proposed that individuals actively construct and process mental representation (Mayer, 2021). Learners will process information by paying attention to the presented information, and the effort in processing information by selecting and organizing information leads to meaningful learning (Mayer, 2021). To facilitate the processing of information, the coordination of information presentation is emphasized (Jiang et al., 2017; Mayer, 2014).

Limited-capacity assumption indicates that within a certain time, the volume of information that can be processed is constrained by WMC, which leads to the consideration of cognitive load (Baddeley, 2006; Mayer, 2014). There are three types of cognitive load: “intrinsic cognitive load”, “extraneous cognitive load”, and “germane cognitive load” (Sweller, 2011). Intrinsic cognitive load is fixed during the learning period and related to the prior knowledge that learners have handled (Sweller, 2011). Learners will have a lower intrinsic cognitive load when the presented information is what they are familiar with (Sweller & Sweller, 2006). Extraneous cognitive load can be controlled by modifying the information presentation, for example, reduce the demonstration of irrelevant information (Sweller, 2011). The task content that activates the interaction of different elements simultaneously in the brain can affect intrinsic and extraneous cognitive load (Sweller, 2010). Germane cognitive load indicates the capacity that extraneous cognitive load reserved to process intrinsic cognitive load (Sweller et al., 2019).

Dual-channel assumption based on the model of working memory from Baddeley (see Baddeley, 2006) and dual coding theory (see Paivio, 2010) suggested that information will be processed in terms of its types (visual/auditory or verbal/nonverbal) in working memory with two channels (Mayer, 2021). Different types of information will be stored in sensory memory from different sources of sensory input (e.g., ears or eyes) first and be extracted and processed in working memory with two channels in terms of the embedded information (verbal or nonverbal) (Mayer, 2021). The modality principle further proposed to enrich the dual-channel assumption with practical guidelines in education that use auditory information instead of text information to reduce visual input and increase the volume of information that can be digested within a limited time.

Working Memory and Vocabulary Encoding Strategy

Working memory is a mechanism that processes information in certain seconds with limited capacity (Baddeley, 2006). The capacity of working memory constrains learners' information processing when the cognitive load is high (Sweller & Sweller, 2006; Sweller, 2011). While the advantages of the application of multimedia learning resources in supporting information processing and memory have been proposed in plenty of research before, the increased elements in the digital learning environment challenge the capacity of working memory (Anmarkrud et al., 2019). Subsequently, Sweller et al. (2019) contended that the previous studies have not paid enough attention to working memory in multimedia learning resource design. Therefore, the current study took working memory capacity (WMC) into consideration. To further evaluate the impact of WMC in the use of multimedia glosses, the current study adopted the measurements of WMC using Baddeley's model of working memory as a framework.

Baddeley's model of working memory consists of three components: "central executive" and its two slave subsystems "phonological loop" and "visuospatial sketchpad" (Baddeley & Hitch, 1974). The central executive is responsible for coordinating the human cognitive process, and the phonological loop and visuospatial sketchpad are supervised and regulated by it (Baddeley & Hitch, 1974; Henry, 2012). The phonological loop involves the process of temporarily storing phonological information and retrieving it through consecutive articulatory rehearsals (Baddeley, 2006). Visuospatial sketchpad is a system that is assumed to assist visual and spatial information retention (Baddeley, 2006). The current study conducted phonological working memory capacity (PWMC) tests and visuospatial working memory (VSWMC) capacity tests to measure them accordingly.

While WMC limits the amount of information that learners can remember and process, the performance of memory is also affected by the strategy learner used (Kirchhoff & Buckner, 2006). Kirchhoff and Buckner (2006) designed the vocabulary encoding strategy questionnaire with ten types of encoding strategies. They explored the relationship between WMC and encoding strategy proposing how the strategies occupy the cognitive resources in four aspects: "verbal elaboration", "mental imagery", "visual inspection" and "memory retrieval" (see Appendix A). The current study adopted the vocabulary encoding strategy questionnaire to figure out the impact of the use of strategy in the memory of vocabulary with multimedia glosses.

Research Design

Sampling Methods

The current study conducted purposive sampling in recruiting participants with the following criteria: 1) participants' first language is Chinese, 2) participants have learned English before, 3) participants are high school students, 4) participants know how to type words using their learning devices, 5) participants have known less than 10% of words presented in the current study.

Because the treatment entailed learners remembering 60 English words, it is critical to confirm that learners have not learned the words before. All the vocabulary included in the current study was collected from the book "Collins Cobuild Key Words for IELTS: Book 3 Advanced". As the book is devised for learners to achieve IELTS 7 score (Moore, 2011), the vocabulary in this book is assumed to be harder than the vocabulary that high school students in mainland China need to learn. However, as it is difficult to ensure that every word for the participants is unknown, the current study set the criteria that participants were familiar with less than 10% of the words. To achieve the criteria, a vocabulary size test is set before learners participate in the experiment.

Vol. 20, No. 1, Spring 2023, 64-88

Participants

There are 63 high school students from mainland China (35 males and 28 females) were involved after excluding 2 students who learned over 10% of the vocabulary presented in the current study. The participants' years of learning English ranged from 8-12 years with the English lower intermediate level. The participants’ age ranged from 15 to 19 ($\overline{M}$= 16.95, SD = 1.35). As the study is conducted fully online, the learning devices have not been strictly controlled. Because some participants do not have a computer, they used smartphones instead to access the learning content. More explicitly, 71% of participants use smartphones, and 29% use computers to complete the learning task. The learning content and tests have been presented through the platform “QuestionPro” (https://www.questionpro.com/). The platform adopts an adaptive design in information presentation so that the font presentation would be automatically changed in terms of the devices that learners used. Besides, as the learning environment cannot be controlled as well, the current study let participants rate the environmental sound level in a questionnaire.

As the whole study lasted for two weeks, the current study provided several time slots for participants to choose from each day. The participants registered for the study by choosing the provided time slots. At the same time, only one participant attended the intervention with one researcher’s supervision. After the end of the intervention, participants cannot access the learning tasks and tests again. Therefore, it was assumed that during the intervention, participants do not share information with one another.

Participants were required to complete four types of items as Table 1 shows. However, as the current study encountered the technical issue that the survey platform has not saved all of the data correctly and some sets of data do not have enough quality (e.g., the low audio quality in nonword repetition tests), not all the collected data is available for analysis (see Table 1).

TABLE 1
Descriptive Data for the Tested Items

Items	Description	Time assignment (min)	Number of available data
Basic information questionnaire	Including age, gender, and the used devices	1	63
Short-term memory test	Nonword Repetition test	1.5	44
	Digit Span Test	1.5	63
	Visual Patterns Test	2	18
Vocabulary remembering test	Vocabulary size test	2.5	63
	Word acquisition tasks	15	63
	Word retention tests	5	63
Questionnaire	Encoding Strategy Questionnaire	3	49

Measures

A. Nonword repetition test: A PWMC test

The present study examined participants’ PWMC using a nonword repetition test (sometimes called as Children's Test of Nonword Repetition). The test can indicate learners' prior knowledge of English and English acquisition ability (Duncan & Paradis, 2016; Gathercole & Baddeley, 1989, 1996; Gathercole et al., 1994). Compared with the digit span test, though there is a strong correlation between the digit span test score and the nonword repetition test score, the result of the nonword repetition test is a stronger indicator of learners’ vocabulary knowledge (Gathercole et al., 1994).

The adopted nonword in this study was devised by Howell et al. (2017). Four nonwords were chosen at random for each syllable length. After hearing the sounds of the nonwords with 2 to 5 syllables (e.g., two syllables: blomplin /ˈblɔm.plɪn /; five syllables: sembumpiklempet /sɛm.bɵm.pɪ.ˈklɛm.pɛt /), learners need to repeat the sounds (Henry, 2012; Howell et al., 2017). There are 16 nonwords in the Nonword repetition test for the current study.

The measure of testing result is adopted considering the aspects 1) the length of syllables that can be recalled; 2) repetition accuracy (Howell et al., 2017), and the score was recorded as "nonword repetition level" and "nonword repetition score" respectively in this study. Meanwhile, repetition accuracy centered on the exact pronunciation of the phonemes and disregarded the stress and syllable length. (Guiberson & Rodriguez, 2013; Howell et al., 2017).

B. Digit span test: A PWMC test

The digit span test is a typical measurement of individuals' PWMC and indicates central executive function as well (Baddeley, 1992; Gathercole, 1999; Kaushanskaya et al., 2011). It involves two directions in repeating: forward and backward. Both the two types of digit span tests measure individuals' PWMC, and apart from that, the backward digit span test presumably indicates individuals' executive function as well (Gathercole, 1999). Additionally, it was suggested that the results of the digit span test might anticipate the passive learning result of vocabulary (Majerus et al., 2006). The present study conducted the two types of digit span tests with forward and backward directions. The digit length was between 3 and 9 and each length had two sets of digits. After hearing the sounds of the digits, participants were required to input their answers on the survey.

TABLE 2
The Description of PWMC Tests

PWMC tests	Description	Result indicates
Nonword repetition test	After hearing the nonword, repeat it by speaking out	1) PWMC 2) English prior knowledge (e.g., vocabulary size, phonological awareness) 3) English acquisition ability
Forward digit span test	Type in the digits that hear before in positive order	1) PWMC
Backward digit span test	Type in the digits that hear before in reverse order	1) PWMC 2) Central executive function (information executive process)

C. Visual pattern test: A VSWMC test

Visual pattern test measures individuals' VSWMC. It examines the maintenance and processing of visuospatial information (Della Sala et al., 1997). The current study designed four sizes of grids for memorization (i.e., 3*3, 3*4, 4*4, and 5*5). Some grids were painted in black (see figure 1). In accordance with past research, participants were asked to choose the painted cells in a blank grid after displaying for 3s and the preceding grid had vanished (Brown et al., 2006; Levin et al., 2010).

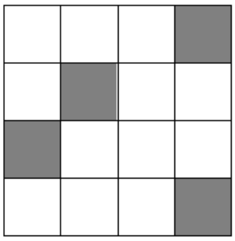

Figure 1. 4*4 grid.

D. Vocabulary acquisition tasks: The design of learning duration & multimedia glosses

Vocabulary acquisition tasks were designed in terms of gloss types and learning duration and were consistent in one acquisition task. For example, if the first word has a 5s learning duration, the subsequent words in the learning task have a 5s learning duration respectively. Each page contained a single word and its corresponding multimedia glosses (see Table 3). The present study designed the learning duration of vocabulary acquisition tasks in terms of individuals' WMC. It was proposed that working memory can retain information for 15s to 30s (Ciccarelli & White, 2015, p.261; Prisko, 1963). As WMC can store around 7 chunks of information (Miller, 1956), each learning task consisted of 5 words in the current study. As a result, the learning duration per word is between 3s and 7s, and the overall learning time is between 15s and 35s (see Figure 2).

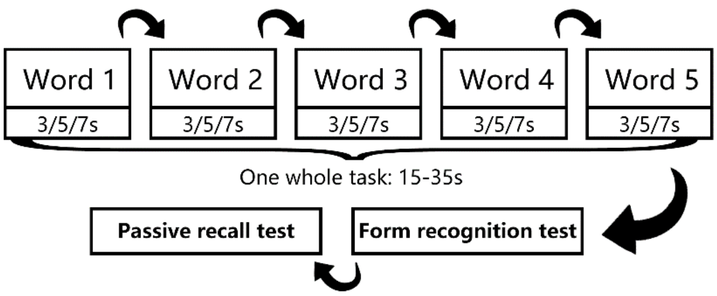

Figure 2. Learning duration for the tasks.

Besides, the current study designed four types of glosses 1) Translation only (T); 2) Translation + corresponding picture (TP); 3) Translation + corresponding sound (TS); 4) Translation + corresponding picture + sound (TPS). Combined with the learning duration, for the task that each word has a 3s learning duration, the glosses were denoted as T3, TP3, TS3, and TPS3 to present. For the 5s learning duration, the glosses were denoted as T5, TP5, TS5, and TPS5. For the 7s learning duration, the glosses were denoted as T7, TP7, TS7, and TPS7.

TABLE 3
Examples of Multimedia Glosses

T	TP	TS	TPS
			

E. Form recognition test & passive recall test: Vocabulary retention test

The form recognition test and passive recall tests are designed in terms of Türk and Erçetin's work (2014). The form recognition test examines participants' *recognition* of words and the passive recall test entails participants recalling the *translation* of the word. As for the form recognition tests, the current study entailed learners selecting 5 learned words after each vocabulary acquisition task from 10 presented words. As for the passive recall tests, the current study entailed learners typing in the translation of learned words.

F. Encoding strategy questionnaire

The current study explored learners' vocabulary encoding strategies in terms of Kirchhoff and Buckner's (2006) strategy questionnaire. Though the strategy questionnaire focuses on cognitive encoding strategy, Miller, Donovan, Bennett, Aminoff, and Mayer (2012) rephrase some descriptions in the questionnaire to fit the questionnaire in measuring the vocabulary memory process. The current study utilized the rephrased version as an encoding strategy questionnaire. In the current study, students entailed rating how often they used each strategy on a 5-point scale from 1 (never) to 5 (always) after the whole vocabulary acquisition task.

Results

Potential Disturbance Variable

As the current study was conducted online, the learning environment cannot be controlled strictly. To enhance the reliability of the study, participants rated the environmental sounds level by themselves for the WMC tests, vocabulary acquisition tasks, and vocabulary retention tests respectively. Pearson correlation analysis was conducted in analyzing the relationship between environmental sounds, WMC test results, and the result of vocabulary retention tests. The result shows that environmental sounds significantly correlated with the result of nonword repetition level in the PWMC tests ($r = .383$, $p < 0.05$), and the result of form recognition tests with TP7 glosses ($r = -.421$, $p < 0.05$). Therefore, the result of nonword repetition level and form recognition tests with TP7 glosses may be affected by environmental sounds.

Q1: The Impact of Different Types of Multimedia Glosses and Learning Duration on Vocabulary Retention

Figure 3 illustrates that, in terms of the mean score of the form recognition test in each learning duration, the outcome of the form recognition test declined when the learning duration was increased. With the three types of learning duration (i.e., 3s, 5s, 7s), the text-only glosses exhibited their advantage in facilitating word recognition in the form recognition test. Compared with the learning result with TS and TP glosses, participants received a higher score with TP glosses when the learning duration is 3s and 5s per word. However, once the learning duration was increased to 7s, participants received the lowest score with TP glosses. Moreover, when the learning duration was extended to 7s per word, learners with T and TPS glosses received the highest score in word recognition.

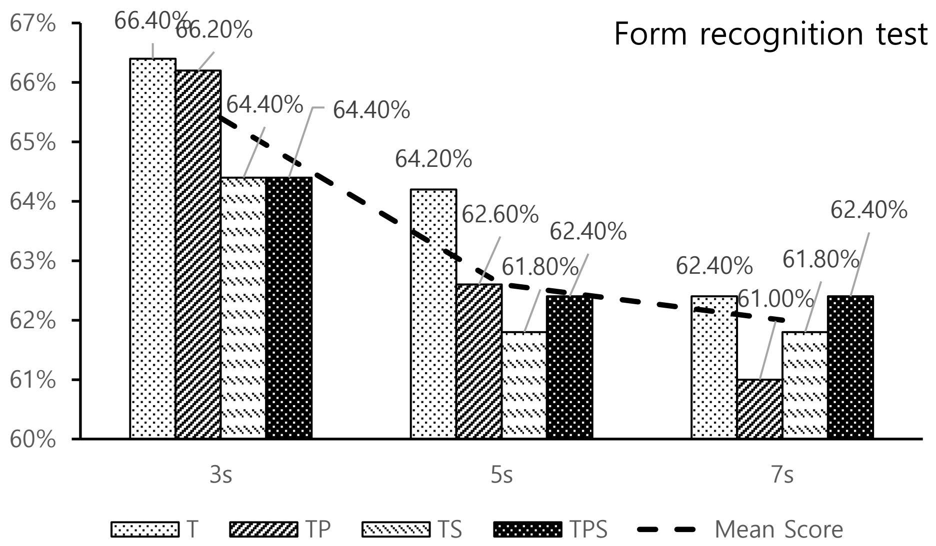

Figure 3. Percentage of correct answers for the form recognition test. x-axis: learning duration per word; y-axis: the correctness of the form recognition test.

In contrast to the results of form recognition tests, participants exhibited the highest mean score in recalling translation in terms of the result of the passive recall test when the learning duration was 5s per word (see Figure 4). The participants' vocabulary learning with T glosses got a relatively stable percentage of correctness when the learning duration was 3s (41%), 5s (42%), and 7s (42%). Besides, though the glosses with pictures (TP and TPS glosses) demonstrated their superiority in supporting translation recall with 3s and 5s learning duration per word, text-only tasks led to the highest learning result when the learning duration was 7s. Regarding the impact of TS glosses, participants scored the lowest on all passive recall tests with 3s, 5s, and 7s learning durations when the glosses were TS.

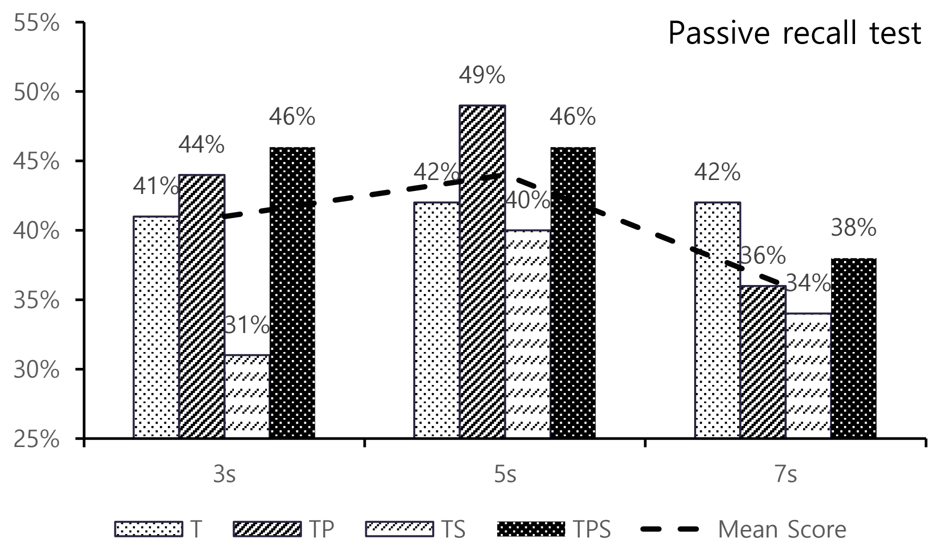

Figure 4. Percentage of correct answers for passive recall test. x-axis: learning duration per word; y-axis: the correctness of the passive recall test.

Q2: The Impact of Vocabulary Encoding Strategies on Vocabulary Retention

Participants rated their use of vocabulary encoding strategy at the end of all of the vocabulary acquisition tasks. In light of Table 4, participants frequently used the strategy of "repeat the words" (Strategy 3), "imagine a typed list of the words" (Strategy 9), and "use prior memories associated with the objects" (Strategy 2) to memorize the words.

TABLE 4
Descriptive Analysis of the Rating of Vocabulary Encoding Strategy

Vocabulary encoding strategies*	Mean	SD
1. Considered the living/nonliving status of the objects	2.27	1.07
2. Used prior personal memories associated with the objects	2.96	1.09
3. Repeated the words to yourself	3.38	1.18
4. Used the starting letter of the words to group them	3.21	1.03
5. Constructed sentences using the word you studied	2.81	1.16
6. Constructed weird/silly/nonsensical sentences	2.79	1.13
7. Constructed a story from the words as they appeared	2.88	1.14
8. Formed a picture of each word in your mind	2.85	1.22
9. Tried to imagine a typed list of the words	3.17	1.08
10. Mentally put words together into categories	2.79	1.11

The current study adopted Pearson correlation analysis for participants' rating of vocabulary encoding strategies and the two vocabulary retention tests (see Appendix B). The strategies that "repeated the words to yourself" (Strategy 3) and "constructed weird/silly/nonsensical sentences" (Strategy 6) have no significant correlation with the two vocabulary retention tests (form recognition test and passive recall test).

*"Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals," by M. B. Miller, C. L. Donovan, C. M. Bennett, E. M. Aminoff and R. E. Mayer, 2012, Neuroimage, 59(1), pp. 83-93 (https://doi.org/10.1016/j.neuroimage.2011.05.060). Copyright 2011 by Elsevier Inc.

Regarding the form recognition test, significant negative effects have been found when using strategies 1 ($r = -.412$, $p < .05$, TP3), 5 ($r = -.357$, $p < .05$, T3; $r = -.394$, $p < .05$, TP3; $r = -.384$, $p < .05$, T5), 6 ($r = -.389$, $p < .05$, TPS7), 7 ($r = -.358$, $p < .05$, T3; $r = -.375$, $p < .05$, T5; $r = -.530$, $p < .01$, TPS5; $r = -.644$, $p < .01$, TPS7), 8 ($r = -.418$, $p < .05$, TS3), and 10 ($r = -.402$, $p < .05$, TPS5). Meanwhile, there are some significant positive relationships between the use of strategy 4 ($r = .554$, $p < .01$) and 9 ($r = .488$, $p < .01$) with TPS3 glosses (see Appendix B).

Regarding the passive recall test, all significant correlations between the use of strategies and translations recalling is positive (see Appendix C). Specifically, significant positive effects have been found when using strategies 1 ($r = .136$, $p < .05$, T3), 2 ($r = .355$, $p < .05$, TPS3; $r = .402$, $p < .05$, T5; $r = .371$, $p < .05$, TS5; $r = .362$, $p < .05$, TPS5; $r = .380$, $p < .05$, TP7), 5 ($r = .417$, $p < .05$, T3), and 8 ($r = .400$, $p < .05$, T3; $r = .492$, $p < .01$, TS3; $r = .378$, $p < .05$, TPS5).

Q3: The Impact of Working Memory Capacity (WMC) on Vocabulary Retention with Different Types of Glosses and Learning Duration

The Pearson correlation analysis among WMC tests shows that there is no significant relationship between nonword repetition tests and digit span tests (see Appendix D). Meanwhile, the nonword repetition score has a strong positive correlation with the nonword repetition score ($r = .906$, $p < .01$), and the result of the forward digit span test is positively related to the backward digit span test ($r = .275$, $p < .05$). As there was a high positive correlation between the result of nonword repetition level and nonword repetition score, the current study only used nonword repetition level as the nonword repetition test result for descriptive analysis.

To analyze the relationship between PWMC and vocabulary retention tests (form recognition tests and passive recall tests), descriptive analysis (see Figure 5, 8) and correlation analysis (see Appendix E, F) were adopted. As the forward digit span test entails only participants' PWMC (see Table 2), the current study mainly used the result of the forward digit span test as the indicator of PWMC. Besides, high and low PWMC groups were arranged in terms of the median of the result of the forward digit span test. The high and low VSWMC groups were arranged in terms of the median of the result of the visual pattern test.

The correlation analysis for the WMC test result and the two vocabulary retention tests were conducted using the Pearson correlation method. As the sample size of the visual pattern test is lower than 30, instead of using a p-value to evaluate the significance of the result, the current study used the Bayesian factor to define the significance (Schisterman et al., 2003).

Form Recognition Test Result

Different from the expectation, high PWMC and VSWMC groups did not always receive higher scores than low PWMC and VSWMC groups (see Figure 5). The high PWMC group received slightly lower scores in TPS3 and TPS7 glosses-assisted tasks. The high VSWMC group received lower scores in TPS3, T5, TP5, TPS5, T7, TS7, and TPS7 glosses-assisted tasks.

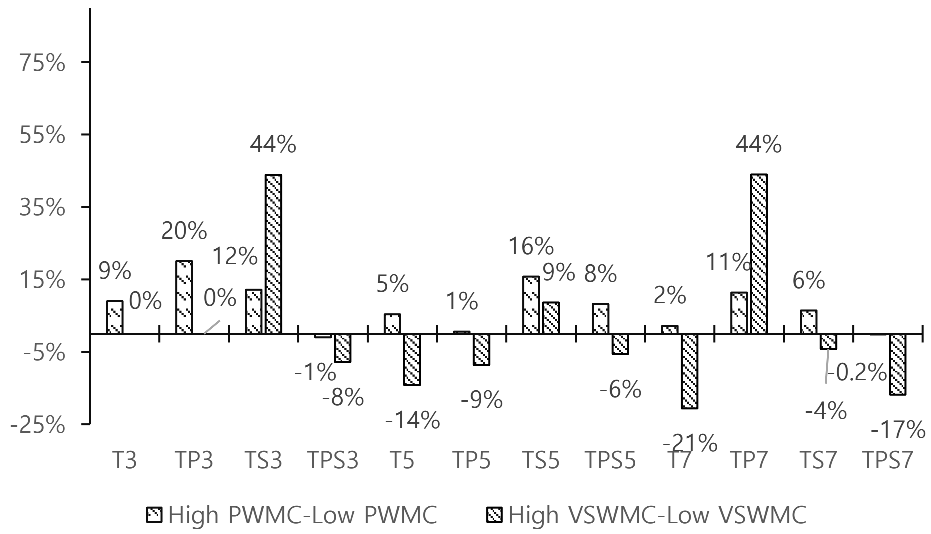

Figure 5. The difference in form recognition test results between the low WMC group and the high WMC group. x-axis: the glosses type; y-axis: the differences between high and low WMC groups in the correctness of the form recognition test.

To figure out the impact of learning duration on the result of form recognition tests, the current study used the mean value of form recognition tests (see Figures 6 and 7). The groups with higher PWMC and VSWMC achieved better learning outcomes in word recognition. When the learning duration was increased, the discrepancy between high and low PWMC groups, and high and low VSWMC groups was reduced.

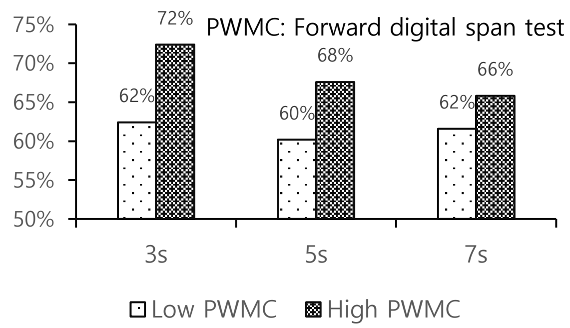

Figure 6. The percentage of the correct answer for the form recognition test with Low PWMC and high PWMC groups. x-axis: learning duration; y-axis: percentage of correctness.

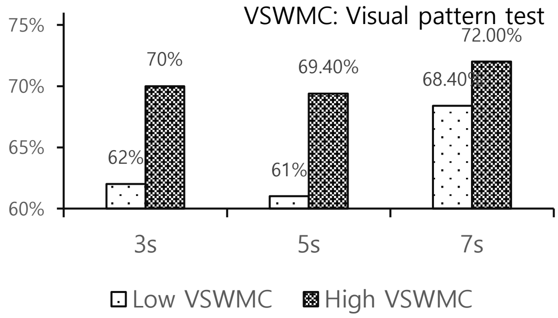

Figure 7. The percentage of the correct answer for the form recognition test with low VSWMC and high VSWMC groups. x-axis: learning duration; y-axis: percentage of correctness.

The correlation analysis provides a clearer analysis of the impact of WMC on the recognition of learned words (see Appendix E). The result of the nonword repetition score indicates better performance in word recognition when the glosses are T3 ($r = .517$, $p < .01$), TP5 ($r = .324$, $p < .05$), TPS5 ($r = .322$, $p < .05$), and T7 ($r = .419$, $p < .05$). The result of the forward digit span test shows that PWMC is closely related with almost types of learning tasks with significant correlation, except for the learning task that adopted picture glosses (TPS3, TP5, TP7, and TPS7). The result of the backward digit span test is positively correlated with the learning result of the tasks that assisted with the T3 ($r = .321$, $p < .05$), T5 ($r = .294$, $p < .05$), TS5 ($r = .384$, $p < .01$), and T7 ($r = .326$, $p < .05$) glosses. Regarding the role of VSWMC in word recognition, as the Bayes factor for all tasks is larger than 1, there is no significant relationship between visual pattern score and the recognition of learned words.

Passive Recall Test Result

The participants were assigned to high PWMC in terms of the three PWMC tests respectively (see Figure 8). When the learning duration was less than 7s, the participants who were assigned to the high PWMC group in terms of the result of the nonword repetition test and forward digit span test, received lower scores in translation remembering than in the low PWMC group. Reversely, participants who received a higher score in the backward digit span test performed better in translation remembering (except for the T3 glosses-assisted task) when the learning duration was less than 7s. Regarding the result of the visual pattern test for VSWMC, the learner who received a higher score in VSWMC, performed better in translation remembering in almost types of glosses (except for TS5 and T7) assisted tasks.

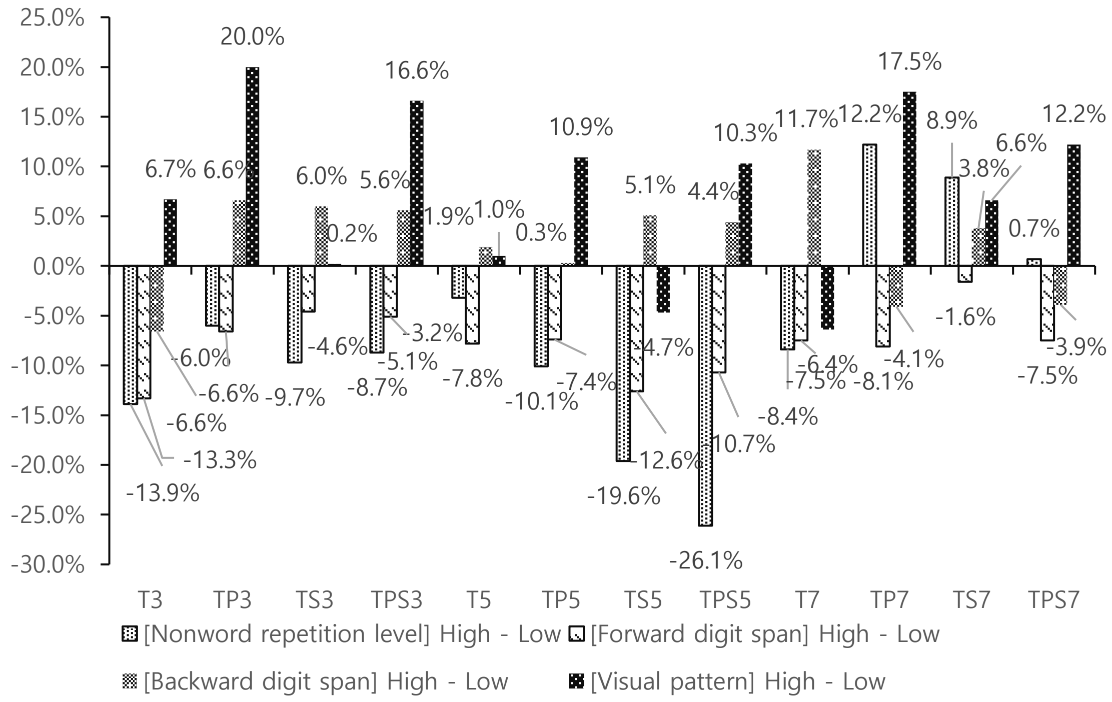

Figure 8. The difference in passive recall test results between the low and high WMC groups. x-axis: the glosses type; y-axis: the differences between high and low WMC groups in the correctness of the passive recall test.

To explore the impact of WMC with different learning duration, the current study used the mean result of the translation remembering score. In terms of Figure 9, learners with higher scores on the forward digit span test received a lower score with 3s, 5s, and 7s learning duration. Besides, the participants with higher scores in nonword repetition tests received lower scores when the learning duration was 3s and 5s. Different from the result in the forward digit span test and nonword repetition test, the participants who achieved a higher score in the backward digit span test received higher scores in the passive recall tests when the learning duration was 3s, 5s, and 7s. Regarding the impact of VSWMC, the result shows that participants who received a higher score in the visual pattern test performed better in passive recall tests.

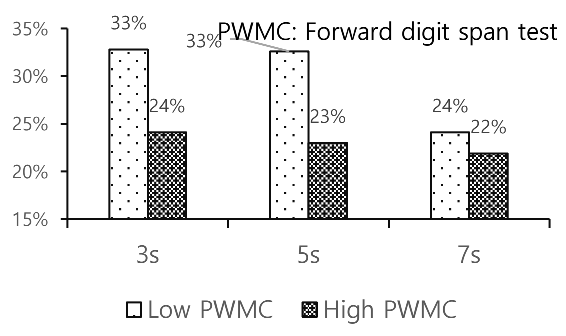

Figure 9. The percentage of the correct answer for the passive recall test with low and high PWMC groups in the forward digit span test. x-axis: learning duration; y-axis: percentage of correctness.

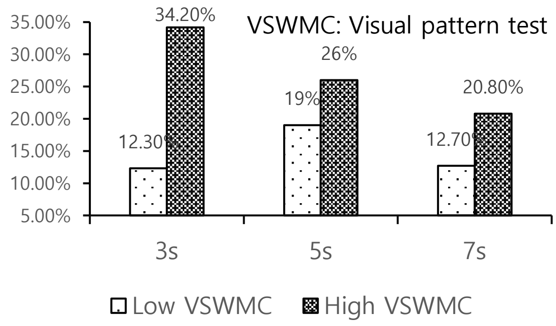

Figure 10. The percentage of the correct answer for the passive recall test with low and high VSWMC groups in the visual pattern test. x-axis: learning duration; y-axis: percentage of correctness.

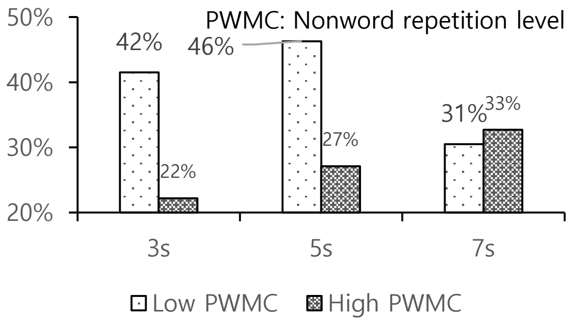

Figure 11. The percentage of the correct answer for the passive recall test with low and high PWMC groups in the nonword repetition test. x-axis: learning duration; y-axis: percentage of correctness.

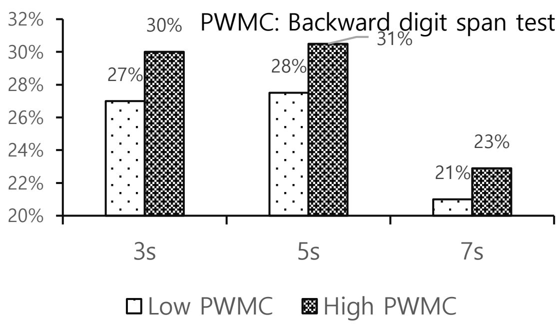

Figure 12. The percentage of the correct answer for the passive recall test with low and high PWMC groups in the backward digit span test. x-axis: learning duration; y-axis: percentage of correctness.

Furthermore, the correlation analysis (see Appendix F) revealed that when the glosses were TS3, while the nonword repetition score predicts a negative impact of PWMC on translation remembering in passive recall tests ($r = -.382$, $p < .05$), the result of backward digit span is positively correlated with it ($r = .354$, $p < .05$). Regarding the learning tasks assisted with TS5 glosses, while both the nonword repetition level ($r = -.361$, $p < .05$) and nonword repetition score ($r = -.405$, $p < .05$) indicate a negative relationship with the result of the passive recall test, the result of the backward digit span test (r = .374, p <.05) is positively correlated with it.

Regarding the correlation between the passive recall test and VSWMC, similar to the correlation analysis in the form recognition test, the current study used the Bayes factor to indicate the significance. Results indicate that when the glosses embed with corresponding pictures (TPS3, r =.589; TP5, r =.489; TP7, r =.510), there is anecdotal evidence to define the significant positive correlation.

Discussion

The present study examined the impact of learning duration and WMC on vocabulary acquisition with different types of multimedia glosses. Two types of vocabulary retention tests were adopted: 1) form recognition test: recognizing the learned 5 words from a 10 words list; and 2) passive recall test: recalling the translation of words. While the form recognition test only requires learners to store and retrieve the information, the passive recall test entails a more cognitive process for remembering the translation and connecting the translation with the presented word.

The Process of Visual and Auditory Information

In terms of the result of vocabulary encoding strategies, while the used strategies lead to better learning performances in passive recall of vocabulary meaning, almost all strategies lead to negative performance in word recognition tasks. It indicates learners' tendency in remembering vocabulary meaning instead of word recognition. As learners could only obtain the meaning of words by looking at the pictures and the meanings, visual information may have a higher priority in reception and processing. Meanwhile, when providing the sounds of words, learners achieved lower learning outcomes in passive vocabulary recall. Though Mayer's dual-channel assumption assumed that learners can process verbal/nonverbal and pictorial/visual information separately (Mayer, 2021) and plenty of studies adopted the modality information to enhance knowledge acquisition (e.g., Inan et al., 2015; Kozan et al., 2015; Syodorenko, 2010), the results indicate a relatively reverse result – when the learning duration is short, auditory information may disturb the processing of visual information and lead to lower performance in recalling the meaning of vocabulary.

Furthermore, the executive function is assumed to support information selection and processes during the learning process – when the corresponding pictures or sounds were added, learners with high backward digit span scores performed better in translation recall tasks compared with the learners with high forward digit span scores (see Figure 8). While the forward digit span test measures individuals' PWMC, the backward digit span test evaluates both PWMC and executive function (Baddeley, 1992; Gathercole, 1999). As learners' executive function is responsible for allocating cognitive resources in attentional control and strategy use (Anderson, 2002), it was assumed that when presenting various types of information simultaneously, learners entail selecting the specific types of information for processing.

Working Memory and Learning Duration

Rapid learning tasks have been conducted as Figure 2 shows. The short learning duration consolidated the impact of working memory. When the learning duration of the whole learning task extended from 15s to 35s (3s to 5s per word), learners' word recognition results decreased (see Figure 3), while learners achieved better scores in translation recall tasks when the learning duration was 5s per word (25s for the whole tasks). As working memory can store information within a limited period (Wilhelm et al., 2013), learners are assumed to have the tendency to forget the information when the duration of the whole learning task was extended. However, as pictorial and text information entails time to recognize and process (Fisher et al., 2019; Van der Zanden et al., 2022), learners achieved the best learning performance in remembering the meaning of words when the learning duration extended to 5s per word.

Besides, the current study found that corresponding pictorial glosses can be redundant information if the learning duration is extended – though learners can recall more translation when providing corresponding pictorial glosses with 3s and 5s learning duration per word, adding pictorial glosses led to a lower learning performance when the learning duration was extended to 7s per word (see Figure 4). As add pictorial glosses increase the information needs to be stored and the extended learning duration entails learners retaining information for a longer time, learners need to keep additional pictorial information for a longer time. Therefore, the current study hypothesizes that because learners' working memory has been challenged, pictorial information may become redundant information and impede the retrieval of vocabulary translation, and leads to lower scores in translation recall.

Furthermore, van den Broek and her colleagues' (2021) discovery may support to explain why pictorial information led to lower learning performance when the learning duration was extended from another perspective. It was suggested that pictorial information may be overestimated in supporting knowledge learning and retrieving as learners are assumed to spend less effort processing information compared with text-only glosses (van den Broek et al., 2021). As process text-only information without a picture as the cue in arousing prior knowledge entails more effort in information processing (van den Broek et al., 2021), when the learning duration is enough for learners to process text-only information, learners who were demonstrated with text-only glosses may spend more effort in knowledge processing and achieved better learning result, which is aligned with the result in the current study when the learning duration extended to 7s per word (see Figure 4).

Pedagogical Implications

Learning opportunities in English are a persistent issue in China – though some wealthier places can afford immersive programs, there are some places that cannot guarantee teachers receive enough training before teaching and quality enough learning materials (Butler, 2015; Wang et al., 2022). As digital learning resources can to some extent relieve the resource allocation issues using the internet, it is essential to consider how to impart knowledge in a digitalized learning environment, instead of simply transferring the content of printed books to digital learning materials.

When one learning tasks include several learning objectives, with intensive and rapid learning tasks, learners will have their priority in selecting the information they would like to process and store quickly. Therefore, instructional designers need to consider the consistency of the learning objectives, to ensure the supplementary resources will not be redundant information for the prime learning objective. For example, regarding the design of second language learning, as information connecting and retrieving entail time, instructional designers need to consider the expected time learners will stay when presenting specific information. If learners are assumed to stay on certain pages for only a short period, it would be better to reduce the learning objectives and present the glosses that specifically support the learning objectives.

Meanwhile, with different encoding strategies, glosses may not be as effective as expected. Therefore, when designing multimedia learning resources, it is important to consider how to scaffold learners to allocate their cognitive resources in information processing. For example, in terms of learners' PWMC and VSWMC, learners can be taught to use different learning strategies, and multimedia learning resources can demonstrate suitable glosses for them accordingly.

Besides, considering working memory capacity, the interval between the learning of novel knowledge and review is recommended less than 30 seconds. In the current study, when the learning duration of the whole task was more than 30 seconds, learners' performance in vocabulary recognition and translation recall was reduced.

Conclusion

The current study reveals the impact of learning duration, working memory capacity, and vocabulary encoding strategies with pictorial and auditory glosses on vocabulary acquisition. The results indicate how individuals allocate cognitive resources within a short knowledge acquisition duration and the types of glosses can scaffold more knowledge learning. Text-only glosses with 3s per word (15s whole learning task) learning duration led to the highest learning result in supporting learners' word recognition and text + picture glosses with 5s per word (25s whole learning task) lead to the best learning performance in learners' passive recall of vocabulary meaning. Different from Mayer's dual channel assumption, when the learning duration is short, the auditory glosses process may impede the process of visual information. Meanwhile, image glosses may become redundant information when the learning duration is extended. Besides, better executive function facilitates information organization and selection which relieves the negative effect caused by redundant information.

Limitations & Further Research Direction

As technical issues caused some WMC test data to be lost, the current study may have reliability issues. It is recommended to carry out further research to determine whether the negative impact of high PWMC on vocabulary translation remembering exists and how the encoding strategies with different capacities in visuospatial and phonological working memory affect the cognitive load. Besides, as the study was conducted online, environment sounds and learning devices (i.e., computers and smartphones) may affect the learning results. Further studies can experiment in a controllable environment, such as a laboratory or computer classroom to reduce the disturbance variables.

As the correctness of all translation recall tasks is less than 50%, it was assumed that the 3s, 5s, or 7s learning duration per word may be too short for learners. For the vocabulary rapid learning tasks, further study can consider longer learning duration per word. Meanwhile, when the learning duration is short, learners' reaction time toward different types of glosses need to be considered as well. As the current study lacks the equipment to measure the data, learners’ reaction time has not been discussed enough in the current study. Further study can explore whether adding different glosses affect learners’ reaction time and whether different reaction time affects the learning result of rapid learning tasks.

The Authors

Meilin Jiang is a research assistant at Hong Kong Metropolitan University. She obtained her bachelor degree from the Education University of Hong Kong and master degree from the University of Hong Kong. Her research interest is technology-enhanced learning.

School of Science and Technology
Hong Kong Metropolitan University
Homantin, Kowloon, Hong Kong, SAR China
Email: mljiang@hkmu.edu.hk

Di Zou is an Assistant Professor at The Education University of Hong Kong. Her research interests include second-language acquisition, technology-enhanced language learning, game-based language learning and flipped classroom. She has published more than 100 research papers in international journals and books, including Computers & Education, Computer Assisted Language Learning, Language Teaching Research, and British Journal of Educational Technology.

Department of English Language Education

The Education University of Hong Kong
10 Lo Ping Rd, Tai Po,
New Territories, Hong Kong, SAR China
Email: dzou@eduhk.hk

Lap-Kei Lee is an Assistant Professor at the School of Science and Technology of Hong Kong Metropolitan University. He obtained his BEng (with First Class Honours) in Computer Engineering and PhD (Computer Science) from the University of Hong Kong. His research interests include the design and analysis of algorithms (especially in online job scheduling and data stream algorithms), natural language processing, educational technology, and bioinformatics.

School of Science and Technology
Hong Kong Metropolitan University
Homantin, Kowloon, Hong Kong, SAR China
Email: lklee@hkmu.edu.hk

Fu Lee Wang is a Professor at Hong Kong Metropolitan University and Dean of School of Science and Technology. He has published over one hundred and fifty articles in refereed journals and conference proceedings, including *Decision Support Systems, IEEE Transactions on Cybernetics, IEEE Intelligent Systems, IEEE Multimedia, IEEE Transactions on Affective Computing, Information & Management, Information Processing & Management, Information Sciences, Neurocomputing,* and *Optics and Lasers in Engineering* among others. He was also the Chair of ACM Hong Kong Chapter and IEEE Hong Kong Section Computer Society.

School of Science and Technology
Hong Kong Metropolitan University
Homantin, Kowloon, Hong Kong, SAR China
Email: pwang@hkmu.edu.hk

Nga-In Wu is a Lecturer at the College of Professional and Continuing Education of the Hong Kong Polytechnic University. She obtained her Bachelor of Arts in Chinese Language and Literature and Linguistics (double major) and Master of Arts in Linguistics from the University of Hong Kong. Her research interests lie mainly in syntax and morphology of Chinese and English, and phonetics of Huizhou dialect.

College of Professional and Continuing Education
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong, SAR China
Email: ngain.wu@cpce-polyu.edu.hk

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82. https://doi.org/10.1076/chin.8.2.71.8724

Anmarkrud, Ø., Andresen, A., & Bråten, I. (2019). Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. *Educational Psychologist, 54*(2), 61-83. https://doi.org/10.1080/00461520.2018.1554484

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Baddeley, A. (2006). Working memory: An overview. *Working memory and education*, 1-31. https://doi.org/10.1016/B978-012554465-8/50003-X

Baddeley, A. D., & Hitch, G. (1974). Working memory. In I. G. A. Bower (Ed.), *The Psychology of learning and motivation.* (pp. 47-90). https://doi.org/10.1016/S0079-7421(08)60452-1

Blom, H., Segers, E., Knoors, H., Hermans, D., & Verhoeven, L. (2018). Comprehension and navigation of networked hypertexts. *Journal of Computer Assisted Learning, 34*(3), 306-314. https://doi.org/10.1111/jcal.12243

Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning, 30*(7), 709-725. https://doi.org/10.1080/09588221.2017.1356335

Brown, L. A., Forbes, D., & McConnell, J. (2006). Limiting the use of verbal coding in the Visual Patterns Test. *The Quarterly Journal of Experimental Psychology, 59(*7), 1169-1176. https://doi.org/10.1080/17470210600665954

Butler, Y. G. (2015). English language education among young learners in East Asia: A review of current research (2004-2014). *Language teaching*, *48*(3), 303-342. https://doi.org/10.1017/S0261444815000105

Chen, I. J., & Yen, J. C. (2013). Hypertext annotation: Effects of presentation formats and learner proficiency on reading comprehension and vocabulary learning in foreign languages. *Computers & Education, 63*, 416-423. https://doi.org/10.1016/j.compedu.2013.01.005

Chen, M. P., Wang, L. C., Zou, D., Lin, S. Y., Xie, H., & Tsai, C. C. (2022). Effects of captions and English proficiency on learning effectiveness, motivation and attitude in augmented-reality-enhanced theme-based contextualized EFL learning. Computer *Assisted Language Learning, 35*(3), 381-411. https://doi.org/10.1080/09588221.2019.1704787

Chen, S. L., & Liu, Y. T. (2021). Learning by designing or learning by playing? A comparative study of the effects of game-based learning on learning motivation and on short-term and long-term conversational gains. *Interactive Learning Environments*, *2021*, 1-15. https://doi.org/10.1080/10494820.2021.1961159

Ciccarelli, S. K., & White, J. N. (2015). *Psychology* (4th ed.) Pearson Education Limited.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210. https://doi.org/10.1007/BF01320076

Della Sala, S., Gray, C., Baddeley, A., & Wilson, L. (1997). *Visual patterns test: A test of short-term visual recall*. Thames Valley Test Company.

Duncan, T. S., & Paradis, J. (2016). English language learners' nonword repetition performance: The influence of age, L2 vocabulary size, length of L2 exposure, and L1 phonology. *Journal of Speech, Language, and Hearing Research, 59*(1), 39-48. https://doi.org/10.1044/2015_jslhr-l-14-0020

Fisher, J. T., Hopp, F. R., & Weber, R. (2019). Modality-specific effects of perceptual load in multimedia processing. *Media and Communication, 7*(4), 149-165. https://doi.org/10.17645/mac.v7i4.2388

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419. https://doi.org/10.1016/s1364-6613(99)01388-1

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

Gathercole, S. E., & Baddeley, A. D. (1996). *The Children's Test of Nonword Repetition.* Psychological Corporation.

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

Guiberson, M., & Rodríguez, B. L. (2013). Classification accuracy of nonword repetition when used with preschool-age Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 44*(2), 121-132. https://doi.org/10.1044/0161-1461

Hassan Ja'ashan, M. N., Abdulaziz Alfadda, H., & Saleh Mahdi, H. (2022). Using a holographic application in learning medical terminology for English as a foreign language students. *Interactive Learning Environments.* Advance online publication. https://doi.org/10.1080/10494820.2022.2093913

Henry, L. (2012). *The development of working memory in children*. https://dx.doi.org/10.4135/9781446251348

Howell, P., Tang, K., Tuomainen, O., Chan, S. K., Beltran, K., Mirawdeli, A., & Harris, J. (2017). Identification of fluency and word-finding difficulty in samples of children with diverse language backgrounds. *International Journal of Language & Communication Disorders, 52*(5), 595-611. https://doi.org/10.1111/1460-6984.12305

Inan, F. A., Crooks, S. M., Cheon, J., Ari, F., Flores, R., Kurucay, M., & Paniukov, D. (2015). The reverse modality effect: Examining student learning from interactive computer-based instruction. *British Journal of Educational Technology, 46*(1), 123-130. https://doi.org/10.1111/bjet.12129

Jiang, D., Renandya, W. A., & Zhang, L. J. (2017). Evaluating ELT multimedia courseware from the perspective of cognitive theory of multimedia learning. *Computer Assisted Language Learning, 30*(7), 726-744. https://doi.org/10.1080/09588221.2017.1359187

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2011). The relationship between vocabulary and short-term memory measures in monolingual and bilingual speakers. *International Journal of Bilingualism, 15*(4), 408-425. https://doi.org/10.1177/1367006911403201

Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Journal of Educational Technology & Society*, *11*(3), 114-126.

Kirchhoff, B. A., & Buckner, R. L. (2006). Functional-anatomic correlates of individual differences in memory. *Neuron, 51*(2), 263-274. https://doi.org/10.1016/j.neuron.2006.06.006

Kohnke, L., Zhang, R., & Zou, D. (2019). Using mobile vocabulary learning apps as aids to knowledge retention: Business vocabulary acquisition. *Journal of Asia TEFL, 16*(2), 683-690.

Kohnke, L., Zou, D., & Ho, E. (2021). Exploring the potential of push notification in ESL: A qualitative inquiry. *Journal of Asia TEFL, 18*(3), 985-994.

Kozan, K., Erçetin, G., & Richardson, J. C. (2015). Input modality and working memory: Effects on second language text comprehension in a multimedia learning environment. *System, 55*, 63-73. https://doi.org/10.1016/j.system.2015.09.001

Levin, D. S., Thurman, S. K., & Kiepert, M. H. (2010). *More than just a memory: The nature and validity of working memory in educational settings*. Psychology Press.

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

Mayer, R. (2014). Cognitive Theory of Multimedia Learning. *In The Cambridge Handbook of Multimedia Learning* (pp. 43-71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. (2021). Cognitive Theory of Multimedia Learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (Cambridge Handbooks in Psychology, pp. 57-72). https://doi.org/10.1017/9781108894333.008

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81. https://doi.org/10.1037/h0043158

Miller, M. B., Donovan, C. L., Bennett, C. M., Aminoff, E. M., & Mayer, R. E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage, 59*(1), 83-93. https://doi.org/10.1016/j.neuroimage.2011.05.060

Mohsen, M. A., & Balakumar, M. (2011). A review of multimedia glosses and their effects on L2 vocabulary acquisition in CALL literature. *ReCALL, 23*(2), 135-159. https://doi.org/10.1017/S095834401100005X

Moore, J. (Ed.). (2011). *Collins Cobuild key words for IELTS: Book 3 advanced*. Collins Cobuild.

Paivio, A. (2010). Dual coding theory and the mental lexicon. *The Mental Lexicon, 5*(2), 205-230. https://doi.org/10.1075/ml.5.2.04pai

Prisko, L. H. (1963). *Short-term memory in focal cerebral damage*. [Unpublished doctoral dissertation]. McGill University.

Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing.* Routledge.

Schisterman, E. F., Moysich, K. B., England, L. J., & Rao, M. (2003). Estimation of the correlation coefficient using the Bayesian Approach and its applications for epidemiologic research. *BMC medical research methodology, 3*(1), 1-4. https://doi.org/10.1186/1471-2288-3-5

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review, 22*(2), 123-138. https://doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76).

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology, 4*(1), https://doi.org/10.1177/147470490600400135

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261-292. https://doi.org/10.1007/s10648-019-09465-5

Syodorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 14*(2), 50-73. http://dx.doi.org/10125/44214

Türk, E., & Erçetin, G. (2014). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning, 27*(1), 1-25. https://doi.org/10.1080/09588221.2012.692384

van den Broek, G. S., van Gog, T., Jansen, E., Pleijsant, M., & Kester, L. (2021). Multimedia effects during retrieval practice: Images that reveal the answer reduce vocabulary learning. *Journal of Educational Psychology, 113*(8), 1587-1608. https://doi.org/10.1037/edu0000499

van der Ven, F., Takashima, A., Segers, E., & Verhoeven, L. (2017). Semantic priming in Dutch children: word meaning integration and study modality effects. *Language Learning, 67*(3), 546-568. https://doi.org/10.1111/lang.12235

Van der Zanden, T., Mos, M. B., Schouten, A. P., & Krahmer, E. J. (2022). What people look at in multimodal online dating profiles: How pictorial and textual cues affect impression formation. *Communication Research*, *49*(6), 863-890. https://doi.org/10.1177/0093650221995316

Wang, Z., Kohnke, L., & Zou, D. (2022). Chinese students' identity construction and negotiation during emergency remote teaching: A case study from Wuhan. *The Journal of AsiaTEFL, 19*(3), 1044-1052.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it?. *Frontiers in psychology, 4*, 433. https://doi.org/10.3389/fpsyg.2013.00433

Xu, Y., Yau, J. C., & Reich, S. M. (2021). Press, swipe and read: Do interactive features facilitate engagement and learning with e-Books?. *Journal of Computer Assisted Learning, 37*(1), 212-225. https://doi.org/10.1111/jcal.12480

Zhu, F., Yang, J., & Pi, Z. (2022). The interaction effects of an instructor's emotions in instructional videos and students' emotional intelligence on L2 vocabulary learning. *Educational Technology Research and Development, 70*(5), 1695-1718. https://doi.org/10.1007/s11423-022-10148-2

(Received July 28, 2022; Revised January 17, 2023; Accepted March 18, 2023)

Vol. 20, No. 1, Spring 2023, 64-88

Appendix A

Encoding Strategy Questionnaire and the high loading components (Kirchhoff & Buckner, 2006)

Strategies	High loading components
1. Considered the living/nonliving status of the objects	Memory Retrieval
2. Used prior personal memories associated with the objects	Memory Retrieval
3. Repeated the words to yourself	No specific
4. Used the starting letter of the words to group them	Verbal Elaboration
5. Constructed sentences using the word you studied	Verbal Elaboration
6. Constructed weird/silly/nonsensical sentences	No specific
7. Constructed a story from the words as they appeared	No specific
8. Formed a picture of each word in your mind	Mental Imagery
9. Tried to imagine a typed list of the words	Verbal Elaboration
10. Mentally put words together into categories	Visual inspection

Appendix B

Correlation matrix for vocabulary encoding strategy and the learning result with different types of glosses in form recognition test

	T3	TP3	TS3	TPS3	T5	TP5	TS5	TPS5	T7	TP7	TS7	TPS7
Strategy 1	-0.215	**-.412***	-0.266	0.034	-.421	0.075	-0.237	-0.276	-0.160	0.087	-0.336	-0.090
Strategy 2	0.023	0.297	0.188	0.258	0.090	0.171	0.114	0.022	0.247	-0.054	0.106	0.033
Strategy 3	0.127	0.035	0.163	0.220	-0.072	0.216	0.231	0.258	0.183	-0.093	0.139	-0.073
Strategy 4	0.147	-0.053	0.311	**.554****	0.138	0.139	0.206	0.068	-0.005	0.119	-0.097	0.161
Strategy 5	**-.357***	**-.394***	-0.288	-0.286	**-.384***	-0.207	-0.304	-0.321	-0.231	0.171	-0.081	-0.230
Strategy 6	-0.106	-0.067	-0.218	-0.081	-0.171	-0.006	-0.083	-0.108	-0.192	0.316	-0.052	**-.389***
Strategy 7	**-.358***	-0.195	-0.256	-0.019	**-.375***	-0.248	-0.351	**-.530****	-0.147	0.154	-0.239	**-.644****
Strategy 8	0.007	-0.108	**-.418***	0.179	0.008	-0.089	-0.133	**-.518****	-0.207	0.314	-0.195	-0.069
Strategy 9	0.296	0.180	0.309	**.488****	0.046	0.300	0.240	-0.085	0.072	0.075	0.182	-0.083
Strategy 10	-0.045	-0.160	-0.219	0.163	-0.113	-0.151	-0.195	**-.402***	-0.117	0.168	-0.144	-0.129

* $p < .05$ ** $p < .01$

Appendix C

Correlation matrix for vocabulary encoding strategy and the learning result with different types of glosses in passive recall test

	T3	TP3	TS3	TPS3	T5	TP5	TS5	TPS5	T7	TP7	TS7	TPS7
Strategy 1	**0.136***	0.212	-0.156	0.045	0.000	0.024	-0.042	-0.037	-0.143	0.305	-0.097	-0.122
Strategy 2	0.072	0.195	0.343	**.355***	**.402***	0.244	**.371***	**.362***	0.212	**.380***	0.225	0.300
Strategy 3	0.113	-0.082	0.089	0.305	0.244	0.178	0.301	0.099	0.074	0.266	-0.031	0.151
Strategy 4	0.165	0.007	0.224	-0.085	0.159	-0.022	0.180	0.008	0.022	0.202	-0.071	-0.106
Strategy 5	**.417***	0.109	0.091	0.300	0.153	0.342	0.332	0.339	0.167	0.313	0.244	0.254
Strategy 6	-0.162	-0.189	-0.332	0.092	-0.147	-0.069	-0.313	-0.098	-0.214	-0.207	-0.243	-0.133
Strategy 7	0.159	-0.236	0.022	-0.035	-0.067	0.008	-0.035	0.058	-0.250	0.207	-0.059	-0.067
Strategy 8	**.400***	0.138	**.492****	0.317	0.161	0.205	0.094	**.378***	-0.021	0.173	0.110	0.297
Strategy 9	0.128	0.025	0.059	-0.022	0.036	-0.157	-0.104	-0.183	-0.274	0.094	0.266	-0.107
Strategy 10	0.193	-0.072	0.085	0.011	-0.026	-0.040	-0.088	0.183	-0.058	0.218	0.248	0.073

* $p < .05$ ** $p < .01$

Appendix D

Correlation matrix of WMC test

	Nonword repetition Score	Forward digit span test	Backward digit span test	Visual pattern test
Nonword repetition Level	**.906****	0.195	0.080	-0.224
Nonword repetition Score		0.271	0.125	-0.090
Forward digit span test			**.275***	0.193
Backward digit span test				-0.192

* $p < .05$ ** $p < .01$

Appendix E

Correlation matrix for WMC and the learning result with different types of glosses in form recognition test

	T3	TP3	TS3	TPS3	T5	TP5	TS5	TPS5	T7	TP7	TS7	TPS7
Nonword repetition Level	**.393***	0.196	0.106	-0.071	0.033	0.280	-0.024	0.153	0.299	0.000	0.032	-0.189
Nonword repetition Score	**.517****	0.305	0.290	0.090	0.176	**.324***	0.164	**.322***	**.419***	0.128	0.162	-0.069
Forward digit span	**.461****	**.474****	**.473****	0.143	**.302***	0.221	**.408****	**.360****	**.386***	0.292	**.332***	0.220
Backward digit span	**.321***	0.233	0.115	0.107	**.294***	0.164	**.384****	0.256	**.326***	0.292	0.124	0.248
Visual pattern	0.132	0.211	0.126	0.093	-0.161	-0.038	0.283	0.032	-0.115	0.205	0.103	-0.076
Bayes Factor	4.503	3.829	4.548	4.741	4.285	4.946	3.076	4.958	4.496	3.559	4.281	4.034

* $p < .05$ ** $p < .01$

Vol. 20, No. 1, Spring 2023, 64-88

Appendix F

Correlation matrix for WMC and the learning result with different types of glosses in passive recall test

	T3	TP3	TS3	TPS3	T5	TP5	TS5	TPS5	T7	TP7	TS7	TPS7
Nonword repetition level	-0.238	-0.110	-0.329	-0.179	-0.073	-0.324	**-.361***	**-.395***	-0.177	0.119	0.219	0.022
Nonword repetition score	-0.273	-0.098	**-.382***	-0.171	-0.107	**-.387***	**-.405***	**-.378***	-0.176	0.125	0.228	-0.089
Forward digit span	-0.177	. 0.024	-0.003	0.037	-0.064	0.124	-0.136	-0.191	0.089	-0.214	0.151	-0.158
Backward digit span	0.183	0.264	**.354***	0.182	0.305	0.103	**.374***	0.191	0.241	0.081	0.264	0.149
Visual pattern	0.459	0.469	0.359	**0.589**	0.283	**0.489**	0.112	0.428	0.368	**0.510**	0.384	0.219
Bayes Factor	1.168	1.091	2.162	0.360	3.044	0.933	4.751	1.451	2.068	0.786	1.896	3.779

* p < .05 ** p < .01