Teaching English Language, Vol. 17, No. 2, Summer & Fall 2023, pp. 213-250

Teaching English Language Journal
ISSN: 2538-5488 – E-ISSN: 2538-547X – http://tel.journal.org
© 2023 – Published by Teaching English Language and Literature Society of Iran

Please cite this paper as follows:
Mahshanian, A., Shahnazari, M., & Moinzadeh, A. (2023). The contribution of working memory and word recognition to second language reading across different proficiency levels: An eye-movement study. *Teaching English Language, 17*(2), 213-250. https://doi.org/10.22132/TEL.2023.407277.1492

Research Paper

The Contribution of Working Memory and Word Recognition to Second Language Reading across Different Proficiency Levels: An Eye-Movement Study

Amir Mahshanian
PhD Candidate, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran

Mohammadtaghi Shahnazari[1]
Associate Professor, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran

Ahmad Moinzadeh
Associate Professor, Department of English Language and Literature, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran

Abstract

This study examined the role of working memory (WM) and word recognition in L2 reading across different proficiency levels. 120 Iranian EFL learners were placed in three proficiency groups based on their IELTS scores. The battery of tests used in this study included a reading span task to measure WM, an eye movement word vs. non-word task to measure word recognition ability, an L1-recall task, and a multiple-choice reading test to measure reading comprehension. Correlations were carried out to examine the connections between WM, word recognition skills, and L2 reading performance. Regression analyses were also conducted to test whether WM and word recognition can predict reading performance at different levels of proficiency. The results showed that there were significant correlations

[1] Corresponding author: m.shahnazari@fgn.ui.ac.ir

between WM and L2 reading in the beginner group and between WM and word recognition speed in the beginner and intermediate groups. Regression analyses indicated that WM is a strong and direct predictor of reading performance at a beginner level of proficiency and a predictor of how fast less proficient readers recognize words in context. Highlighting the important role of WM in the word recognition ability of less proficient L2 readers, this study also showed that second language reading is not related to the accuracy or speed of word recognition across proficiency levels.

Keywords: working memory, lower-level processing, word recognition, L2 reading

Received: March 7, 2023
Accepted: October 29, 2023

1. Introduction

Reading is a complex mental activity that exploits multiple cognitive and linguistic resources to process information in written text and obtain an understanding of it (Habók & Magyar, 2019). The complexity of reading calls for the simultaneous recruitment of multiple lower- and higher-level processing skills and several domain-general comprehension mechanisms. According to Perfetti and Hart (2002), lower-level skills involve processing the orthographic, semantic, syntactic, and phonological features of words, which, as Shimono (2019) emphasized, allow for the adoption by the reader of conscious strategies such as inference-making, comprehension monitoring, and goal setting (i.e., higher-level processing skills) that are necessary for text comprehension. Effective integration of these skills, which are essential for information processing at a word-, sentence-, or text level, is an indicator of successful reading comprehension (Spencer, Richmond, & Cutting, 2020).

Drawing on a component skills view of reading proposed by Carr and Levy (1990), Grabe (2009) argued that reading comprehension issues from two levels of processing: lower-level and higher-level. Lower-level processing involves word recognition, syntactic parsing, and semantic encoding. Word recognition is the process of decoding individual words for their pronunciations (Smith Gabig, 2010) and meanings (Grabe, 2009). Although essential for word-level comprehension, a grapheme-to-phoneme transformation is not sufficient for processing larger text units. To interpret phrasal groupings, word order, and clausal relationships, a second lower-level skill, that is, syntactic parsing, is required. Finally, processing at a lower level follows an integration of word meanings and their syntactic structures (i.e., semantic proposition encoding) to make sense of clauses and sentences. On the other hand, higher-level processing describes the reader's intentional use of strategies like goal-setting, comprehension monitoring, and inference-making to draw meaning from the text (Habók & Magyar, 2019).

Rapp et al. (2007) argued that in addition to lower- and higher-level processing abilities, proficient reading is heavily dependent on cognitive processing systems such as working memory. The creation of a coherent text model requires the integration of lower-level abilities like decoding, higher-level abilities (i.e., discourse processes linked to comprehension), and cognitive processing skills like working memory (Arrington, 2014). During reading, our mind relies heavily on working memory capacity (WMC) to store text information to facilitate comprehension (Just & Carpenter, 1992). WM, as a capacity-limited workspace for storing and online processing information (Borella & de Ribaupierre, 2014; Carpenter & Just, 1989), plays an important role in allocating cognitive resources to different aspects of language (Ahmadian, 2020). This role has been extensively investigated in L2 research, where a consensus has been achieved on the contribution of WM to reading in a second language (Chang et al., 2019). However, the independent contributions of word recognition (WR), as a lower-level processing skill, and WM, as a cognitive source of individual difference (ID), to L2 reading have yet to be examined. Moreover, the question as to whether language proficiency can play a role in the link between WM and reading, WM and WR, and reading and WR has remained unanswered in studies of second language acquisition.

Investigating the interaction between ID variables and lower-level processes would improve the current knowledge on the mechanisms and processes involved in the development of reading comprehension in a second language. Therefore, this study investigated: (1) the interaction between WM, WR, and L2 reading; and (2) whether and how L2 proficiency mediates such a relationship.

2. Literature review and theoretical framework

Given the importance of reading in education, it is not surprising that literature is filled with a considerable number of models on mechanisms and processes involved in reading. An important text-comprehension theory highlighting the contribution of WM and word recognition to reading is that of Walczyk (1995). Proposing the compensatory-encoding model, Walczyk (1995, 2000) argued that readers with less-developed lower-level processing skills (e.g., lexical processing) compensate for their inefficient word-recognition ability by resorting to higher-level processing skills and strategies. The conceptualization of reading in Walczyk's model is very much consistent with ideas put forth by Perfetti (1988) in the verbal-efficiency model and those by Stanovich (1980) in the interactive-compensatory model. The compensatory-encoding model is, in fact, a variant on the verbal-efficiency theory (Grabe, 2009) and shares with it the basic assumption that efficient comprehension arises from automatic lower-level processing and that working memory is an important factor in predicting reading performance (Walczyk, 2000; Walczyk et al., 2001). With the significant role of cognitive strategies and lower-level reading processes, it is essential to investigate how reading in a second language could be influenced by working memory and word recognition ability.

2.1 Working memory and reading

Working memory (WM) is a cognitive system that is limited in capacity and responsible for temporarily storing and processing information (Baddeley, 2010). It is a multimodal workspace consisting of three domain-specific subcomponents (i.e., the phonological loop for processing phonological data, the visuospatial sketchpad for processing visual data, and the episodic buffer for integrating and retrieving different sorts of data from long-term memory; Baddeley, 2019) and a central executive that makes decisions about the allocation of available resources (Alptekin & Ercetin, 2009). Research has shown that a large number of language processing skills, such as vocabulary learning, writing, listening, and, most importantly, reading comprehension, are heavily dependent on WMC (Alptekin & Ercetin, 2010).

Second language research has long concentrated on how WM and reading comprehension are related. The important role of WMC in L2 reading has been established in past research. (e.g., Alptekin & Ercetin, 2009; Chang et al., 2019; Harrington & Sawyer, 1992). In the most seminal study, Harrington and Sawyer (1992) found a moderate link between WM and L2 reading among advanced Japanese learners. In another study, Lesser (2007) reported that among beginning Spanish learners, WM had a strong correlation with L2 reading ability. Alptekin and Erçetin (2011) reported similar results and suggested that WM and topic familiarity had major impacts on L2 reading comprehension, especially in inferential understanding of the text. More recently, Jung's (2018) study showed that WM is a key factor in predicting the reading performance of Korean L2 learners. This was also evident in Chang et al.'s (2019) study, which suggested that WMC is a strong predictor of reading comprehension and is highly correlated with L2 grammar knowledge and writing performance.

Despite a vast body of data showing a link between WMC and second language reading ability, some studies (e.g., Chun & Payne, 2004; Georgiou & Das, 2016; Joh & Plakans, 2017) have reported non-significant, weak, or indirect relationships between WM and reading performance in a second language. For instance, Chun and Payne (2004) reported no link between reading proficiency and the WMC of German L2 English learners. Also, the results of Georgiou and Das's (2015) study revealed that WMC cannot not predict reading ability in a second language. Non-significant correlations were also reported by Joh and Plakans (2017), who maintained that WM becomes important in predicting L2 reading ability only when readers benefit from some sort of prior knowledge of the text. Most recently, Shahnazari (2023) argued that while WMC plays a key role in predicting the reading performance at a lower level of proficiency, it is a non-significant factor in predicting the reading performance of proficient learners. Given these conflicting results, Huang, Ouyang, and Jiang (2022) argued that in recognition of the link between WM and reading, attention should be turned to the underlying processes of reading comprehension (e.g., word recognition, comprehension monitoring, syntactic parsing, goal setting, etc.).

2.2 Word Recognition and Reading

Word recognition is a lower-level reading process that is widely described as the capacity to read individual words (Adlof et al., 2006). The centrality of word recognition in text comprehension (Rayner & Reichle, 2010) has convinced researchers to refer to it as the 'reading engine' (Adams, 2004). Cadime et al. (2017), however, argued that there is more to word recognition than simply reading single words and that successful reading rests heavily on the reader's ability to reach the appropriate meaning of the single words both rapidly and accurately. Accuracy in decoding single words is key to text comprehension, so incorrect or incomplete recognition of text components (i.e., single words) can impair understanding of a sentence or even larger discourse units (Hoover & Tunmer, 2022). Word recognition speed, on the other hand, is a key factor in successful reading since it is associated with automatization (Daneman & Carpenter, 1980); that is, non-automatic processing of a single word will make word recognition slower and more demanding. Thus, inaccurate and slow processing of text information at a lower level (i.e., word recognition) will restrict available resources (e.g., working memory), which are necessary for higher-level processing, and will ultimately impede comprehension (Cutting & Scarborough, 2006).

A body of research has supported the major role played by word recognition in reading (e.g., Cutting & Scarborough, 2006; Zinar, 2000). Furthermore, studies on the developmental stages of reading comprehension showed that accurate and fluent word recognition makes important contributions to reading performance (Goldenberg, 2022). For example, in a study on the role of comprehension monitoring and word recognition in reading, Zinar (2000) reported that word recognition is the strongest predictor of variability in reading performance. In another study, Cutting and Scarborough (2006) found that word recognition and listening comprehension each contribute significantly and independently to the prediction of reading performance. This finding is also evident in Ouellette and Beers's (2010) study, suggesting that recognition of irregular words and serial decoding can separately influence reading performance.

Similar findings were also reported by Nobre and Salles (2016), with significant correlations found between word recognition, semantic priming, and reading. More recently, using a longitudinal design, Cadime et al. (2017) found that even after four years of primary school education, word recognition ability is still a significant indicator of reading achievement. It can be taken from these studies, along with others (e.g., Brown et al., 1993), that reading comprehension in developmental readers is heavily dependent on their ability to identify individual words in isolation.

2.3 The Contribution of WM and Word Recognition to Reading

Compared to the extensive work on the contributions of WM and word recognition to reading (e.g., Chang et al., 2019; Cutting & Scarborough, 2006), little is known about how WM interacts with word recognition in reading. In this regard, Goff, Pratt, and Ong (2005) suggested that while orthographic processing is the strongest independent predictor of reading performance, the role of WM in successful reading is little. Evidence from research on children with learning difficulties has also revealed that deficits in WMC are related to inaccurate and slow word recognition (WR), which ultimately results in major problems in text comprehension (e.g., Leather & Henry, 1994; Swanson, 2003; Swanson and Ashbakar, 2000). More recently, El-Mir (2017) reported that in Arabic orthography, WMC is a significant indicator of word recognition ability.

The same results also came from Abu-Rabia's (1995) study on Arabic orthography, Kail and Hall's (2001) study on English orthography, and So and Siegel's (1997) study on Chinese orthography. However, absent from these studies is whether high-span readers, with rapid and accurate word recognition abilities, are also successful in reading comprehension. That is, although it is evident from past research that WMC influences word recognition, whether or not the influence of WM on word recognition can contribute to reading performance is still subject to uncertainty.

2.4 This study

The contradictory findings from second language studies on the interplay between WM and reading and the existence of little knowledge about how word recognition and L2 proficiency moderate this relationship call for a closer examination of the interaction between WM, the lower-level skill of word recognition, and reading comprehension (Chang et al., 2019; El-Mir, 2017). With this in mind, the present study set out to address two important questions:

1. Is there a significant difference between L2 reading performance across proficiency levels?
2. Is there a significant relationship between WM, word recognition ability, and L2 reading performance across different proficiency levels?
3. What effect, if any, do WM and SE have on L2 reading ability, across different levels of proficiency?

3. Method

3.1 Participants

Participants were 120 L1 Persian English learners, selected from a sample of 151 university students majoring in translation studies, ~~linguistics,~~ literature, medicine, nursing, pharmacology, and dentistry. Based on the overall scores (the mean of scores in each skill), participants were placed in three groups: beginner (n = 40), intermediate (n = 40), and advanced (n = 40). Table 1 summarizes descriptive data for scores on the language proficiency measure in each group.

Table 1
Descriptive Statistics for Scores on the Language Proficiency Measure

	N	M	SD	Min	Max
Beginner	40	3.62	0.44	3.00	4.00
Intermediate	40	5.68	0.61	4.50	6.50
Advanced	40	7.51	0.44	7.00	8.50
Total	120	5.60	1.67	3.50	8.00

Note: N=Number; M=Mean; SD=Standard Deviation; Confidence interval: 95%.

Following Cambridge Assessment Scales, the participants were placed in three proficiency groups using the Common European Framework of Reference (CEFR). As such, we transformed the IELST scores into a CEFR scale that is level-based. A score on the IELTS scale between 0 and 4 is equal to CEFR A1 or A2. According to the scale, "beginner users" with limited proficiency in using English in daily life are classified as A1 and A2. A score on the IELTS scale between 4 and 6.5 is similar to CEFR B1 or B2. These two levels are referred to as "intermediate users" since they possess enough interactivity, communication, and understanding abilities in common English. IELTS scores between 7 and 9 are similar to CEFR C1 and C2. CERF C1 and C2 are defined as “advanced users” with a fully operational command of English who can comprehend a wide range of complex, lengthy texts and perceive inferential meaning.

3.2 Materials

3.2.1 Test of English language proficiency

Participants' proficiency was measured using an IELTS mock test (Cambridge University Press and Cambridge Assessment, 2020, 2021, 2022) that had four sections: listening, reading, writing, and speaking. There were band scores from 1 to 9, which were the average of scores in all four components. Sections on reading and listening were scored according to the answer keys provided by Cambridge University Press and Cambridge Assessment (2020, 2021, 2022), and the speaking and writing sections of the tests were scored by two examiners with 10 years of EFL and IELTS teaching experience. The inter-rater reliability was estimated at 91% for the speaking section and 88% for the writing section.

3.2.2 Working memory measure

An L1 reading span test (RST) was used to measure WM in the present study. In this task, the participants were requested to silently read a total of 54 active sentences (8–13 words in length) on a computer screen, decide whether each sentence, shown on a PowerPoint slide for 7 seconds, is acceptable in Persian by orally saying "*dorost*" meaning correct or "*nadorost*" meaning incorrect (the processing aspect of WM), and then recall the final words, which were all verbs (the storage aspect of WM) in sets of 3, 4, 5, and 6. To ensure an understanding of the task requirements, participants were presented with 10 practice sentences in sets of 2 and 3. Half of the sentences in this task were syntactically or semantically impossible, and the other half were semantically and syntactically acceptable in Persian. For scoring processing, one mark was awarded for each correct verification of sentences, and for scoring the storage aspect, one mark was awarded for correct and in-order recall of final words, so each participant received a score of processing and a score of recall from 54. Following Conway et al.'s (2005) guidelines, since all participants' processing capacity was above 85%, recall scores were used in data analysis as indicative of WMC.

3.2.3 Reading comprehension measures

To compare reading performance under timed and untimed conditions, two independent tasks were used to measure L2 reading at different proficiency levels. The first task was a timed reading test. This test was offered in three different formats and consisted of two reading passages with 20 multiple-choice questions. Form A was devised for the beginner group; form B was designed for the intermediate group; and form C was for the advanced group. All passages (with lengths ranging between 53 and 833) were borrowed from Active Skills for Reading 2-4 (Anderson 2008, 2013). Following Day and Park's (2005) taxonomies, the 10 items after each passage included questions about vocabulary, pronoun references, paraphrasing, positive or negative factual information, text purpose, and inferential understanding. To identify the poor items and remove them from the test, two experienced EFL teachers (both lecturers in applied linguistics) were required to check and comment on the test items. The revised version, piloted with 30 EFL learners, showed that irrespective of the level of proficiency, the participants needed around 30 minutes to complete the task. As for the scoring, one mark was assigned to correct responses and none to incorrect responses.

The second task was an L1-recall reading test in which participants were required to reproduce ideas obtained from the text in their first language. As a valid measure (Shohamy, 1984), this task controls for the lack of L2 knowledge, which might affect the way participants show their understanding of the text. In addition, to control for the effects of high- or low-WMC on comprehension (see Grabe & Stoller, 2002; Koda, 2005), during task completion, the participants were at liberty to turn back to the text multiple times. Word length, difficulty, and the number of idea units were different in the task designed for each group. In total, there were 80, 122, and 194 idea units in the beginning, intermediate, and advanced tests, respectively.

In the current study, idea units were defined as a part of a proposition that can be separated from another part with intonational pauses (Johnson, 1970). Therefore, adopting the most reliable method to identify idea units (Hopewell, 2011), we recruited two English native speakers and asked them to read the passages at their own natural speed and mark the text with slashes where they paused. The agreement between the two readers was estimated at 94%, and a third reader was asked to resolve the discrepancies. Before the experiment, the test was piloted with 30 EFL learners. Although the pilot study revealed that task completion may take 40–60 minutes at different proficiency levels, participants were reassured that they were given extra time if necessary. As for scoring, two experienced EFL teachers were recruited to assign one mark to each accurate reproduction of an idea unit and no mark to an inaccurate or incomplete reproduction of these units. It should be noted that the inter-rater reliability between the rates was estimated at over 96%.

3.2.4 Word recognition measure

The total reading time for each target word was recorded in an eye-movement task to measure word recognition abilities. The eye-movement task involved recognizing words vs. non-words. A total of 60 words (Table 2) were used in English sentences and posited randomly (i.e., irrespective of regularity, sensibility, or phonemic features) in 60 slides. The words were of three types: regular, irregular, and non-sense. The 20 words in each category were then re-grouped based on the number of phonemes in such a way that each group contained ten 3-phoneme and ten 5-phoneme words. To control for the length effect, only 5-letter words were selected in this study.

The list of three- and five-phoneme regular and three-phoneme irregular words was borrowed from Rey and Schiller (2005), where a good frequency match was observed between the two groups (0.0071% and 0.0072% for three- and five-phoneme words, respectively). The list of five-phoneme irregular words (enjoying a high-reliability estimate: α =92%) was adopted from Ouellette and Beers (2010), with a few modifications and checked for frequency using WordGen software (Duyck, Desmet, Verbeke, & Brysbaert, 2004), which showed a good match against Rey and Schiller's (2005) list. The list of non-words, including 10 three-phoneme and 10 five-phoneme words, was created by the authors based on the phonological features of regular and irregular word lists. Table 2 shows the list of regular, irregular, and nonsense words used in the word vs. non-word recognition task.

Table 2
List of Regular, Irregular, and Nonsense Words

WT	NF	N	Words
Reg	**3-phon**	10	beech, feign, peach, poach, pouch, baulk, kneel, whirl, churn, niece
	5-phon	10	drops, twist, blunt, spank, crisp, skulk, drift, stunt, cleft, brisk
Irr	**3-phon**	10	route, thief, guise, thyme, mauve, weird, seize, niche, vague, lathe
	5-phon	10	sugar, blind, briar, retry, guitar, waver, choir, react, ninth, ricin
Non	**3-phon**	10	teign, meach, thurn, kreef, raunt, luize, knief, yathe, haint, fient
	5-phon	10	prump, blint, clont, spunt, tlips, iscon, sudam, gorut, sutar, feily

Note. WT= Word Type; NP= Number of Phonemes; N=Number.

Prior to the experiment, participants committed competent received oral instructions in L1 on how they should verify words against non-words. They were instructed to push the control button on the right side of their keyboards if they thought the underlined string of letters was a real word, and the control button on the left side if they thought the sentence included an underlined non-word. Following the instructions, 10 sample sentences were practiced with participants. For completing the task, participants sat between 60 and 67 cm away from the screen, their heads were fixed on a stand. In the calibration process, they were asked to keep track of a dot that was moving on the screen with their eyes. Data elicited from this experiment were 1) participants' scores in the word recognition task (with correct recognitions receiving one mark and incorrect recognitions receiving no mark), 2) first-fixation time (the time spent on the target word for the first time), 3) gaze time (the entire amount of time that eyes are fixed on the target word before moving on), and 4) total reading time (the entire lengths of all word fixations added together).

3.2.5 Eye-tracking device

On a light gray background with black text and a Times New Romans 30-point font, 60 sentences (including underlined target words) were displayed one at a time on a 20-inch Bina 992 eye-tracking monitor with a 200 Hz data rate and a display resolution of 1440 x 900 pixels. A nine-point eye-calibration test was carried out at a medium speed before the task. The participants' eyes were 60–75 cm away from the monitor, and the eye-tracker recorded both eyes' movements with average binocular tracking turned on. As a head-fix device, Bina 992 did not allow for head motions throughout the trials, so chin and forehead rests were employed. Data was recorded and analyzed with Coglad software using the default fixation filter. The length of the window was set at 20 ms and the velocity threshold was fixed at 30 degrees per second. The maximum angle between fixes was at 0.5 (to merge adjacent fixes), and no minimum or maximum fixation duration was considered.

3.3 Procedure

The process of data collection started with an IELTS mock test to measure participants' English proficiency. The participants were given all three parts of the written test (listening, reading, and writing) in one day. Participants took the speaking test one week after the written test. The participants were divided into three proficiency groups—beginning, intermediate, and advanced—based on their IELTS test results. A reading span test was given to participants in the following session to assess their WMC. They received two scores based on their processing and storage capacities. However, following Conway et al. (2005), only storage scores were reported in this study. The third stage consisted of a series of reading tasks separated by a 30-minute rest. The first and second tasks were a timed multiple-choice reading test and an untimed L1-recall test, respectively. Finally, participants completed a 60-item computerized word vs. non-word task. This last session was designed to assess participants' word recognition skills using an eye-tracking device.

3.4 Data Analysis

Following the assessment of participants' results on the tests, as outlined in the preceding sections, the data from the measures underwent statistical analysis. Descriptive statistics, two-way analyses of variance, correlations, and regressions were the four components of the statistical analysis. First, using the scores of 120 participants, descriptive statistics for the study's variables were calculated. Second, a Scheirer-Ray-Hare test was used to compare how the three groups performed on each measure. Third, correlation analyses were performed to see how WM and other variables interacted. Finally, WM and word recognition were used as predictors in regression analyses to determine how each variable contributes to L2 reading comprehension in particular.

4. Results

The results are broken down into four parts. Descriptive statistics for the reading test (RT), the reading span test (RST), the word recognition task, and the reading task (L1-recall task) are reported in the first section. In the second part, the Kruskal-Wallis test results are presented, following a comparison between groups based on performance on the above tasks. The correlations between these variables are described in the third section, and in the final section, regression analyses were carried out to examine the predicting power of WM and WR in reading performance.

4.1 Descriptive Statistics

4.1.1 Descriptive statistics for the WM measure (RST)

The descriptive statistics of the RST (Table 3) showed that the beginner group (M = 16.32) outperformed the intermediate (M = 14.97) and the advanced group (M = 13.50). In addition, the beginner group's RST scores showed more fluctuations (SD = 8.95) than those of the intermediate (SD = 6.71) and advanced (SD = 5.38) groups.

Table 3
Descriptive Statistics for RST

		N	Mean	SD
RST	**Beginner**	40	16.32	8.95
	Intermediate	40	14.97	6.71
	Advanced	40	13.50	5.38
	Total	120	14.93	7.20

Note. RST=Reading Span Test; Confidence interval: 95%.

4.1.2 Descriptive statistics for the reading measures

Table 4 summarizes the descriptive data from the two reading tasks (i.e., RT and L1-recall) performed by the participants. Although the advanced group (M = 16.8) outperformed the other two groups on the reading test (M = 15.25 and 12.42 for the intermediate and beginner groups, respectively), their performance on the L1-recall task was weaker (M = 91.15) than the intermediate (M = 92.34) and the beginner group (97.45).

Table 4
Descriptive Statistics for the Reading Measures

Tasks		RT		L1-Recall	
Proficiency	**N**	**Mean**	**SD**	**Mean**	**SD**
Beginner	40	12.425	3.37	97.45	27.32
Intermediate	40	15.250	2.58	92.34	35.92
Advanced	40	16.80	1.45	91.15	21.77
Total	120	14.825	3.15	93.65	28.81

Note. RT= Reading Test; SD= Standard Deviation; Confidence interval: 95%.

4.1.3 Descriptive statistics for word recognition accuracy

Table 5 summarizes the descriptive data of participants' performance on the word vs. non-word task in terms of accuracy of verifications. The raw statistics suggest that the advanced group (M = 78.54) outperformed the intermediate (M = 75.58) and the beginner group (M = 62.12). However, there is only a slight variation in the raw mean (M) between the advanced and intermediate groups. Figure 1 depicts the mean variations in word recognition accuracy across the three groups.

Table 5
Descriptive Statistics for Word Recognition Accuracy

P-level		Beginner (N=40)		Intermediate (N=40)		Advanced (N=40)		Total (N=120)	
Word type		**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3-ph	**Reg**	62.25	20.18	65.25	18.40	57.50	12.35	61.66	17.45
	Irr	61.75	20.24	70.00	10.38	66.00	14.29	65.91	15.74
	Non	59.50	19.86	73.25	14.92	84.00	12.15	72.25	18.76
	Tot	61.16	19.96	69.50	15.16	69.16	16.98	66.61	17.85
5-ph	**Reg**	68.75	12.65	79.00	12.97	86.50	7.70	78.08	13.42
	Irr	76.50	10.99	93.50	7.700	96.00	6.72	88.66	12.22
	Non	62.00	14.36	72.50	14.98	81.25	8.53	71.91	15.07
	Tot	69.08	13.96	81.66	15.02	87.91	9.77	79.55	15.25
All	**Tot**	65.12	17.64	75.58	16.25	78.54	16.71	73.08	17.81

Note: P-level=Proficiency Level; N=Number;3-5-ph=3-5-phoneme; Reg= Regular; Irr=Irregular; Non=Nonsense; Tot=Total; SD= Standard Deviation

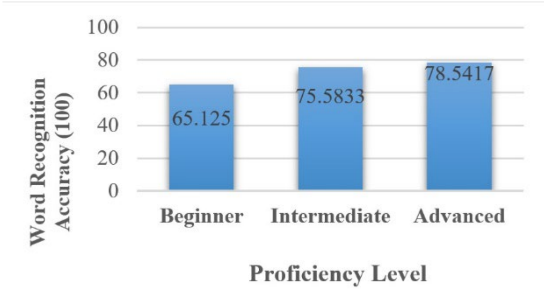

Figure 1. Mean Differences in Word Recognition Accuracy between the Three Groups

Despite the overall advantage of the advanced group over the other two groups, recognition of three-phoneme words followed different patterns across groups. In recognition of three-phoneme words, the intermediate group outperformed the other groups (M = 62.25 for regular words; M = 70 for irregular words; M = 73.25); the advanced group performed better in identifying irregular (M = 65.9) and non-sense words (M = 72.25), compared to the beginner group (M = 61.75 for irregular words; M = 79.5 for nonsense words). The beginner group's performance on regular word detection, however, was only marginally better than that of the advanced group (M = 62.25 for the beginner group; M = 61.6 for the advanced group).

Recognition of five-phoneme words (irrespective of regularity or sensibility) followed the same pattern, with the advanced group outperforming the other two groups (M = 86.5; M = 96; M = 81.25, for regular, irregular, and nonsense words, respectively) and the beginner group receiving the lowest scores (M = 68.75; M = 76.5; M = 62, for regular, irregular, and nonsense words, respectively). Between-group mean differences in word recognition accuracy based on the type of words are displayed in Figure 2.

4.1.4 Descriptive statistics for word recognition speed

The descriptive data on participants' word recognition speed is displayed in Table 6. The raw statistics show that the advanced group was the fastest in word recognition (M = 1.16) among the three groups. While the intermediate group (M = 1.43) was noticeably slower than the advanced group, word recognition in this group was only slightly faster than the beginner group (M = 1.48). Moreover, variation in the speed of word recognition in the beginner group (SD = 0.76) was considerably greater than in the intermediate (SD = 0.45) and advanced groups (SD = 0.40).

Table 6
Descriptive Statistics for Word Recognition Speed

P-level		**Beginner (N=40)**		**Intermediate (N=40)**		**Advanced (N=40)**		**Total (N=120)**	
Word type		**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
3-ph	**Reg**	1.74	0.75	1.62	0.67	1.06	0.35	1.47	0.68
	Irr	1.85	0.92	1.63	0.59	1.40	0.48	1.63	0.71
	Non	1.08	0.61	1.38	0.66	1.22	0.39	1.23	0.57
	Tot	1.56	0.84	1.54	0.64	1.23	0.43	1.44	0.67
5-ph	**Reg**	1.40	0.59	1.37	0.64	1.35	0.59	1.37	0.60
	Irr	1.12	0.61	1.33	0.50	0.83	0.24	1.09	0.51
	Non	1.36	0.68	1.47	0.61	1.14	0.45	1.32	0.60
	Tot	1.29	0.64	1.39	0.59	1.11	0.49	1.26	0.58
All	**Tot**	1.48	0.76	1.43	0.45	1.16	0.40	1.35	0.57

Note: P-level=Proficiency Level; N=Number;3-5-ph=3-5-phoneme; Reg=Regular; Irr=Irregular; Non=Nonsense; Tot=Total; SD= Standard Deviation

The mean differences in word recognition speed between the three groups are shown in Figure 2.

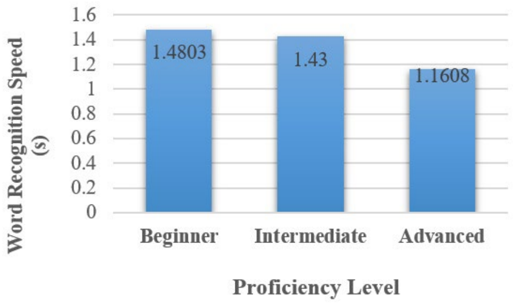

Figure 2. Mean Differences in Word Recognition Speed between the Three Groups

Similar patterns were also observed for the recognition of three- and five-phoneme words, with the advanced group being faster than other groups in identifying both three-phoneme (M = 1.23) and five-phoneme words (M = 1.11) and the intermediate group being slowest in the recognition of three- and five-phoneme words (M = 1.54 for three-phoneme words; M = 1.39 for five-phoneme words). Between-group mean comparisons in the speed of word recognition based on the type of words are shown in Figure 4.

4.2 Between-group Comparisons

Results from the Levene's and Kolmogorov-Smirnov tests indicated that the variables of this study do not follow a normal distribution and that they are not homogenous ($p < 0.05$). Given this result, a Kruskal-Wallis test was carried out to examine whether the three proficiency groups' performance on the measures of this study was significantly different (see Table 7). The results indicated that the three proficiency groups' scores on the RST (H(2) = 106.142, p = 0.430, with the mean ranks of 36.44, 62.65, and 82.41 for the beginner, intermediate, and advanced groups, respectively) and the L1-recall task (H (2) = 1.056, p = 0.590, with the mean ranks of 64.80, 56.90, and 59.80) did not significantly differ from one another. However, there were substantial differences in how well each group performed on the IELTS test of proficiency (H (2) = 108.150, p = 0.000, with the mean ranks of 20.50, 60.50, and 100.50 for the beginner, intermediate, and advanced groups, respectively) and the reading test-RT (H (2) = 35.604, p = 0.000, with the mean ranks of 36.44, 62.65, and 82.41).

Table 7
Kruskal-Wallis One-way Analysis of Variance

	IELTS	RST	RT	L1-Recall
Kruskal-Wallis H	106.142	1.688	35.604	1.056
df	2	2	2	2
Asymp. Sig.	**.000**	.430	**.000**	.590

Considering results from the tests of normality and homogeneity of variances, a Scheirer-Ray-Hare test, a non-parametric equivalence of a two-way ANOVA, was also conducted to compare word recognition performance based on word type across groups. Two main effects (i.e., proficiency and word type) and the interaction effect between these two were considered (Table 8).

Table 8
Scheirer–Ray–Hare Test Results for Word Recognition Accuracy

		Df	Sum Sq	H	P-value
Proficiency	3-ph	2	101761	9.74	**0.007
	5-ph	2	1024730	98.73	**0.000
Word-type Reg-irreg-non	3-ph	2	252886	24.22	**0.000
	5-ph	2	781880	75.33	**0.000
Proficiency □Word-type	3-ph	4	408842	39.17	**0.000
	5-ph	4	56240	5.41	0.246
Residuals	3-ph	351	2983638		
	5-ph	351	1863212		

Note: DV: Score, Observations: 360, D: 0.9637748 (3-phoneme), D: 0.9583565 (5phoneme), MS total: 10830, *p < 0.05, **p < 0.01

The results of Scheirer-Ray-Hare, presented in Table 8, showed that the accuracy of three-phoneme word recognition was significantly influenced by proficiency (p = 0.007), word type (p = 0.00), and the interaction between proficiency and word type (p = 0.00). Overall, participants in the beginner group were the weakest, and those in the intermediate and advanced groups were not significantly different in recognizing three-phoneme words; however, this result does not mean that the beginner group was worse than the other groups in recognition of all types of words (i.e., regular, irregular, and non-sense). While their ability to recognize three-phoneme irregular and non-sense words was significantly lower than the other groups, they outperformed the advanced group in recognition of three-phoneme regular words. Moreover, scores of three-phoneme irregular word recognition in the intermediate group were significantly higher than the advanced and beginner groups ($p < 0.05$ for all). In the case of non-sense word recognition, the advanced group's scores were significantly higher than those of the intermediate and beginner groups (see Figure 3).

The findings for five-phoneme words revealed that the accuracy of word recognition was significantly influenced by proficiency ($p = 0.00$) and word type ($p = 0.00$), but not by the interaction between these two ($p = 0.246$). Overall, the advanced group's five-phoneme word recognition of all word types was much better than the intermediate and beginner groups ($p < 0.05$), and the intermediate group's five-phoneme word recognition was significantly more accurate than the beginner group ($p < 0.05$).

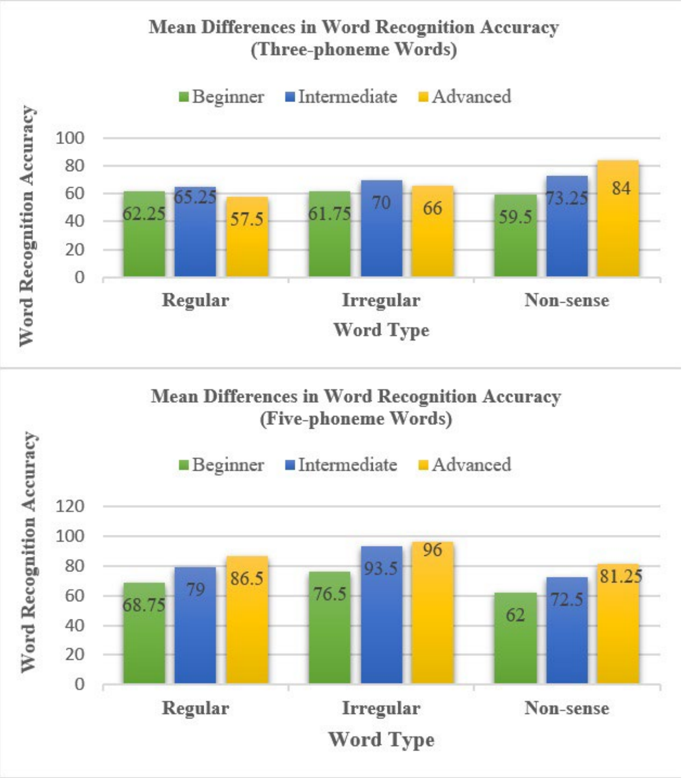

Figure 3. Between-group Mean Differences of Word Recognition Accuracy Based on Word Type

As for word recognition speed, the Scheirer-Ray-Hare test results (see Table 9) indicated that the speed of three-phoneme word recognition was significantly influenced by proficiency (p = 0.0002), word type (p = 0.0001), and the interaction between proficiency and word type (p = 0.002).

Table 9
Scheirer–Ray–Hare Test Results for Word Recognition Speed

		Df	Sum Sq	H	P-value
Proficiency	3-ph	2	180584	16.675	**0.0002
	5-ph	2	171455	15.8326	**0.0003
Word-type	3-ph	2	198531	18.333	**0.0001
Reg-irreg-non	5-ph	2	183102	16.9081	**0.0002
Proficiency □Word-	3-ph	4	179525	16.578	**0.0023
type	5-ph	4	98307	9.0779	0.05918
Residuals	3-ph	351	3329103		
	5-ph	351	3434832		

Note: DV: Score, Observations: 360, D: 0.9999416 (3-phoneme), D: 0.9999297 (5-phoneme), MS total: 10830, *p < 0.05, **p < 0.01

While recognition of regular and irregular three-phoneme words followed a similar pattern, with the advanced group being much faster than the intermediate group ($p < 0.05$) and the intermediate group being considerably faster than the beginner group ($p < 0.05$), a different pattern was observed in non-sense word recognition, with the beginner group detecting words significantly faster than the intermediate and advanced groups (p < 0.05 for both). As regards five-phoneme words, word recognition speed was significantly influenced by proficiency (p = 0.000365) and word type (p = 0.000213) but not by the interaction between these two (0.059181). Overall, the advanced group's five-phoneme word recognition of all word types was significantly faster than the other groups (p < 0.05). Moreover, while the beginner and intermediate groups' performances were almost similar in terms of speed of regular word recognition, the beginner group performed significantly faster in the recognition of irregular and non-sense words (p < 0.05 for both).

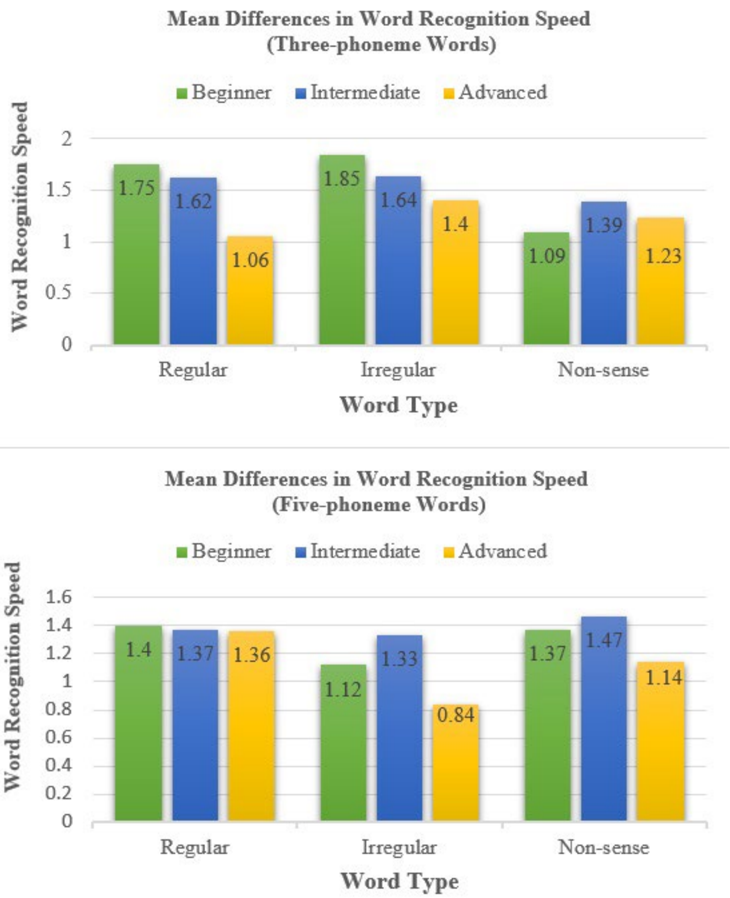

Figure 4. Between-group Mean Differences of Word Recognition Speed Based on Word Type

4.3 Correlations

The sample utilized for this study does not conform to a normal distribution, as demonstrated by the Shapiro-Wilk and Kolmogorov-Smirnov test results. Thus, Spearman's correlation coefficient (ρ) was employed to answer the research questions regarding the link between WM, word recognition, and reading. In the following, the correlation results of this analysis at three levels of proficiency—beginner, intermediate, and advanced, are demonstrated respectively.

4.3.1 Correlations between WM, WR, RT, and L1-recall in the beginner group

Results from correlations between WM, WR, RT, and L1-recall in the beginner group are summarized in Table 10. While WM is shown to be correlated with word recognition speed (p = 0.030) and with reading (p = 0.035 for RT; p = 0.015 for L1-recall), no significant correlations were found between WM and word recognition accuracy or between WM and reading (p > 0.05 for all).

Table 10

Correlations in the Beginner Group

		WR-Score	WR-Time	RST	RT	L1-R
WR-Score	Correlation Coefficient	1.00				
	Sig. (2-tailed)	.				
WR-Time	Correlation Coefficient	-0.07	1.00			
	Sig. (2-tailed)	0.65	.			
RST	Correlation Coefficient	0.04	0.34*	1.00		
	Sig. (2-tailed)	0.78	**0.03**	.		
RT	Correlation Coefficient	0.05	0.29	0.33*	1.00	
	Sig. (2-tailed)	0.71	0.06	**0.03**	.	
L1-R	Correlation Coefficient	0.09	0.26	0.38*	0.32*	1.00
	Sig. (2-tailed)	0.58	0.09	**0.01**	**.04**	.

Note: WR-Score=Word Recognition Accuracy; WR-Time=Word Recognition Speed; RST= Reading Span Test; RT= Multiple-choice Reading Test; L1-R: L1-recall

*. Correlation is significant at the 0.05 level (2-tailed)

Independent significant correlations were also found between WM and the speed of three-phoneme word recognition irrespective of word type (p < 0.01 for regular, irregular, and nonsense words). Similarly, significant correlations were found between WM and the speed of five-phoneme word recognition ($p < 0.05$ for all types of words).

4.3.2 Correlations between WM, WR, RT, and L1-recall in the intermediate group

Working memory (RST), word recognition accuracy (WR-Score), and L2 reading (RT and L1-recall) were not found to be significantly correlated in this group. However, WM was shown to be negatively correlated with the speed of word recognition (WR-Time) in the intermediate group (ρ= -0.39, sig = 0.011). Individual negative correlations were also found between WM and the speed of three- and five-phoneme word recognition, irrespective of word type (p < 0.05 for all regular, irregular, and nonsense words).

Table 11
Correlations in the Intermediate Group

		WR-Score	WR-Time	RST	RT	L1-R
WR-Score	Correlation Coefficient	1.00				
	Sig. (2-tailed)	.				
WR-Time	Correlation Coefficient	-0.08	1.00			
	Sig. (2-tailed)	0.62	.			
RST	Correlation Coefficient	-0.19	-0.39	1.00		
	Sig. (2-tailed)	0.23	**0.01**	.		
RT	Correlation Coefficient	0.22	0.14	-0.22	1.00	
	Sig. (2-tailed)	0.15	0.36	0.16	.	
L1-R	Correlation Coefficient	0.29	0.16	-0.02	0.09	1.00
	Sig. (2-tailed)	0.06	0.29	0.87	0.57	.

Note: WR-Score=Word Recognition Accuracy; WR-Time=Word Recognition Speed; RST= Reading Span Test; RT= Multiple-choice Reading Test; L1-R: L1-recall
*. Correlation is significant at the 0.05 level (2-tailed)

4.3.3 Correlations between WM, WR, RT, and L1-recall in the advanced group

In the advanced group, no correlations were found between working memory (RST), word recognition (WR-Time and Score), and L2 reading (RT and L1-recall). Also, no independent correlations were found between WM and three- or five-phoneme words ($p > 0.05$ for all regular, irregular, and nonsense words). Table 12 summarizes the results of the correlations between the variables in the advanced group.

Table 12
Correlations in the Advanced Group

		WR-Score	WR-Time	RST	RT	L1-R
WR-Score	Correlation Coefficient	1.00				
	Sig. (2-tailed)	.				
WR-Time	Correlation Coefficient	0.09	1.00			
	Sig. (2-tailed)	0.58	.			
RST	Correlation Coefficient	-0.20	0.10	1.00		
	Sig. (2-tailed)	0.19	0.53	.		
RT	Correlation Coefficient	-0.22	-0.03	-0.02	1.00	
	Sig. (2-tailed)	0.15	0.84	0.88	.	
L1-R	Correlation Coefficient	-0.29	-0.20	0.24	0.28	1.00
	Sig. (2-tailed)	0.06	0.21	0.12	0.07	.

Note: WR-Score=Word Recognition Accuracy; WR-Time=Word Recognition Speed; RST= Reading Span Test; RT= Multiple-choice Reading Test; L1-R: L1-recall

4.4 Regression Analysis

Regression analyses were carried out for significant correlations between WM and reading in the beginner group and between WR-time and RT in the beginner and intermediate groups. These analyses sought to identify the unique contributions that WM and word recognition made to reading. The Durbin-Watson Test results (DW =1.623, 1.327, 1.855, and 1.579), the normal-probability plot (forming a diagonal line from bottom left to top right), and the scatterplot (with unsystematic patterns) showed that all four assumptions associated with a linear regression model (i.e., linearity, homoscedasticity, independence, and normality) were met in the following analyses.

4.4.1 Regression between WM and Reading (RT and L1-recall) in the Beginner Group

Regression analysis between WM (RST), RT, and L1-recall in the beginner group ($p = 0.009$) showed that WM accounts for 16.7% of the variance in reading performance on RT, i.e., timed conditions ($R^2 = 0.167$)

and for 17.1% of the variance under untimed conditions, i.e., L1-recall (R^2= 0.171).

Table 13
Regression between WM and Reading (RT and L1-recall) in the Beginner Group

		R	R^2	B	F	t	P
RT	**RST**	0.408	0.167	0.154	7.61	2.76	0.009
L1-recall	**RST**	0.414	0.171	1.26	7.838	2.8	0.008

4.4.2 Regression between WM and WR-time in the Beginner Group

Results from the regression between WM (RST) and WR-time in the beginner group are shown in Table 14. The linear regression (p = 0.019) showed that WM accounted for 13.6% of the variance in word recognition speed, as measured by total fixation duration, in the beginner group (R^2 = 0.136). In addition, the positive relationship between WM and WR-time (B = 0.031) demonstrated that beginner participants with higher WMC spent more time on word recognition, irrespective of the type of the words.

Table 14
Regression between RST and WR-Time in the Beginner Group

	R	R^2	B	F	t	P
RST	0.368	0.136	0.031	5.959	2.441	0.019

4.4.3 Regression between WM and WR-Time in the Intermediate Group

The linear regression model (see Table 15) between WM (RST) and WR-Time (p = 0.009) indicated that WM accounts for 16.8% of the variance in word recognition speed (WR-Time) in the intermediate group (R^2 = 0.168). However, the negative relationship between WM and WR-time in this group (B = -0.028) showed that high-span readers in the intermediate group are faster in word recognition than low-span readers, irrespective of the type of the word.

Table 15
Regression between SE-time and RT in the Beginner Group

	R	R^2	B	F	t	P
RST	0.410	0.168	-0.028	7.681	-2.771	0.009

5. Discussion

This study aimed to examine whether WM and word recognition were direct predictors of L2 reading performance at different levels of proficiency. One important finding of this study was that WM and L2 reading are directly related to each other, albeit only at a low proficiency level. Adding to previous research (e.g., Demir & Erçetin, 2020; Jung, 2018) maintaining that WM is an individual predictor of L2 reading performance, we argue that WMC may also affect underlying processes of reading comprehension (e.g., word recognition) and that the link between WM and reading can be moderated by language proficiency. More precisely, this study suggests that at a beginner level of proficiency, WM is a direct and strong predictor of reading ability both under timed and untimed conditions; however, at higher levels of proficiency (i.e., intermediate and advanced levels), WMC is not an important factor in predicting reading performance.

The results also showed that L2 reading performance is not related to the accuracy of word recognition across proficiency levels, nor is it dependent on its speed. No significant correlations were found between WR (speed and accuracy) and reading under timed vs. untimed conditions (i.e., the multiple-choice reading test vs. the L1-recall reading task). This contradicts findings from L1 studies (e.g., Nobre & Salles, 2016; Zinar, 2000), which assert that word recognition makes an independent contribution to reading comprehension. While it is well established in L1 research that reading depends on accurate and fast word recognition (Hoover & Tunmer, 2022), reading in a second language seems to be associated with other linguistic factors rather than only word recognition skills. Furthermore, this finding casts serious doubts on the generalizability of the simple view of reading (Gough and Tunmer, 1986) to L2 processing. While previous research has extensively supported the SVR's claims that word recognition is a significant predictor of reading performance (e.g., Cadime et al., 2017), the findings of the current study indicated that success in second language reading cannot be accounted for by word recognition skills across different proficiency levels.

Although word recognition did not directly influence reading performance in a second language, WM was shown to be linked with the speed of word recognition at lower levels of proficiency. More precisely, the results indicated that 13.6% of the variation in the speed of word recognition in the beginner group can be accounted for by individual differences in WMC. This meant that high-span readers in the beginner group spent more time recognizing words (regardless of their type) than low-span readers. WMC also explained 16.8% of the variance in the speed of word recognition in the intermediate group. However, the negative relationship between WM and WR-time in this group ($B = -0.028$) showed that participants with high WMC are faster at word recognition than low-span readers, irrespective of the type of word. The relationship between WM and word recognition in the advanced group was not significant, suggesting that advanced learners do not resort to their WM resources for identifying words and that word recognition is an automatized process for advanced learners. In light of this finding and the significant effects of WMC on reading performance at a beginner level of proficiency, it could be argued that less skilled readers devote a large portion of their WM resources to lower-level processing skills such as word recognition, while few resources, if any, are left for higher-level text processing (e.g., inferential understanding and comprehension monitoring), which results in weaker performance on reading tasks.

The relationship between WM, word recognition speed, and reading only among less proficient learners and the significant differences between reading performances on timed conditions across proficiency levels support research by Walczyk (1995, 2000) maintaining that efficient comprehension arises from automatic lower-level processing and that WM plays a key role in L2 reading performance. Walczyk's claims, giving rise to the compensatory-encoding model, suggest that text comprehension is the result of the continuous adoption of compensatory strategies to make up for inefficiencies at lower or higher levels of processing (Grabe, 2009). That is, automatic lower-level processing (e.g., word recognition) of skilled readers will help them save up some workspace in their WM for processing at a higher level (e.g., making inferences, taking reading strategies, monitoring comprehension, etc.), which ultimately ends in better reading performance. However, since word recognition (a lower-level skill) is not an automatic process for less proficient readers, they resort to their WM resources for identifying individual words. This will leave them with no WM resources to be employed for higher-level text processing, which results in weaker reading performance.

A final finding of this study was that reading performances under timed conditions (i.e., multiple-choice tests of reading) were significantly different across proficiency groups, with advanced participants receiving the highest scores and beginner participants receiving the lowest scores on the test. However, no such differences were observed between groups under untimed conditions (i.e., the L1-recall task), suggesting that L2 proficiency is a crucial factor in accounting for the speed of text processing and not general reading ability. Consistent with Walczyk's (1995) model and research on timed reading (e.g., Breznitz, 2006; Breznitz & Share, 1992), this study showed that scores obtained from reading comprehension tests will improve if only a certain amount of time is added to the response time of the task. Thus, it could also be maintained that reading in a second language is a time-sensitive cognitive task that involves the simultaneous processing of text information at both lower and higher levels.

6. Conclusion, Implications, Limitations, and Future Directions

This study examined the roles of WM and word recognition ability in second-language reading at three levels of proficiency. While WM was shown to be an individual predictor of L2 reading achievement and word recognition speed in less proficient learners, results showed that the speed and accuracy of word recognition are not significant factors in predicting L2 reading performance across proficiency levels. This means that less proficient readers benefit from their WMC in rapid recognition of words and text processing more than proficient readers. Furthermore, the results showed that reading performance across different proficiency levels is heavily dependent on time pressure, with significant differences observed between groups only when the reading task is timed. The findings point to the conclusion that WM and word recognition ability play a major role in reading only among less proficient and slow readers.

The findings provide two important implications for teaching English as a second or foreign language. First, in teaching reading materials to students at lower levels of proficiency, attempts should be made by teachers to keep the time-pressure effects on reading tasks to a minimum. This could be done by assuring students that they will be given extra time if they cannot complete reading tasks within the given time limit. This way, reading for comprehension will be promoted over reading for task completion, which will reassure teachers that any future deficits in reading performance could be attributable to variables that are not time-sensitive. Therefore, adding time limits to reading tasks should be postponed until a certain level of automaticity in word recognition (e.g., at higher levels of proficiency) is reached by second or foreign-language learners. Second, materials developers and task designers should provide less proficient learners with reading materials and tasks that facilitate automatic word recognition skills. If word recognition, which is a lower-level processing skill, becomes automatic in less proficient learners, they could use their cognitive resources (e.g., their capacity-limited WM) to process written texts at a higher level (e.g., making inferences) and reach a deeper understanding of them.

Despite its pedagogical implications, an important methodological issue has limited the findings of this study. Word recognition speed was operationally defined as the summed fixation durations of all fixations on the target words (i.e., total reading time) since it reflects both visual identification of the words and semantic activation of their meaning (Dirix, Brysbaert, & Duyck, 2019). Given that different eye tracking results (e.g., first fixation, gaze time, rereading time, and overall reading time) may tap into distinctive phases of word recognition (Boston, Hale, Kliegl, Patil, & Vasishth, 2008), it is very difficult to generalize the results of the current study to all stages of word recognition, so it may be relevant for future eye-movement researchers to investigate whether WMC can predict first fixation, gaze time, and rereading time in the process of word recognition, and more importantly, whether any of these eye-movement data is an individual predictor of reading ability.

Finally, investigating lower-level reading processes was also limited to word recognition ability in this study; however, past studies have established that reading comprehension is also influenced by other lower-level reading skills such as semantic encoding ability and syntactic parsing (Grabe, 2009). Therefore, another suggestion for future research would be a study on how the interaction between WM and other lower-level reading processes, such as syntactic and semantic abilities, can affect reading performance in a second language.

References

Abu-Rabia, S. (1995). Learning to read in Arabic: Reading, syntactic, orthographic, and working memory skills in normally achieving and poor Arabic readers. *Reading Psychology*: *An International Quarterly, 16*(4), 351-394.

Adams, M. J. (2004). Modeling the connections between word recognition and reading. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5^{th} ed., pp. 1219–1243). International Reading Association.

Adlof, S. M., Catts, H. W., & Little, T. (2006). Should the simple view of reading include a fluency component? *Reading and Writing*, 19(9), 933–958.

Ahmadian, M. J. (2020). Explicit and implicit instruction of refusal strategies: Does working memory capacity play a role? *Language Teaching Research*, *24*(2), 163-188.

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Anderson, N. J. (2013). *Active skills for reading 3*. Boston: Cengage Learning.

Anderson, N. (2008). Active skills for reading: *Books 1-4*. Boston: Heinle.

Alptekin, C., & Erçetin, G. (2009). Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System*, *37*(4), 627-639.

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, *33*(2), 206-219.

Alptekin, C., & Erçetin, G. (2011). Effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Quarterly*, *45*(2), 235-266.

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*, *18*(5), 325-346.

Baddeley, A. (2019). The episodic buffer. *Working memories Postmen, Divers and the cognitive revolution (50-60). Routledge*.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136-R140.

Borella, E., & De Ribaupierre, A. (2014). The role of working memory, inhibition, and processing speed in text comprehension in children. *Learning and Individual Differences, 34*, 86-92.

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, *2*(1).

Brown, J. I., Fishco, V. V., & Hanna, G. S. (1993). *Nelson-Denny reading test: Technical report, forms G & H*. Riverside Publishing Company.
Breznitz, Z. (2006). *Fluency in reading: Synchronization of processes*. Routledge.
Breznitz, Z., & Share, D. L. (1992). Effects of accelerated reading rate on memory for text. *Journal of Educational Psychology*, *84*(2), 193.
Cadime, I., Rodrigues, B., Santos, S., Viana, F. L., Chaves-Sousa, S., do Céu Cosme, M., & Ribeiro, I. (2017). The role of word recognition, oral reading fluency and listening comprehension in the simple view of reading: a study in an intermediate depth orthography. *Reading and Writing*, *30*, 591-611.
Carpenter, P. A., & Just, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 31-68). Hillsdale, NJ: Erlbaum.
Carr, T. H., & Levy, B. A. E. (1990). *Reading and its development: Component skills approaches*. Academic Press.
Chang, X., Wang, P., Cai, M. M., & Wang, M. (2019). The predictive power of working memory on Chinese middle school students' English reading comprehension. *Reading & Writing Quarterly*, *35*(5), 458-472.
Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behavior. *System*, *32*(4), 481-503.
Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.
Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277-299.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.
Demir, B., & Ercetin, G. (2020). Exploring the role of proficiency as a mediator of the relationship between working memory and text comprehension in the L2. *Turkish Online Journal of English Language Teaching*, *5*(3), 156–178.
Dirix, N., Brysbaert, M., & Duyck, W. (2019). How well do word recognition measures correlate? Effects of language context and repeated presentations. *Behavior Research Methods*, *51*, 2800-2816.
Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. *Behavior Research Methods, Instruments & Computers, 36*(3), 488–499.

El-Mir, M. (2017). The effect of working memory capacity on word recognition speed in Arabic second-grade readers. *Arab Journal of Psychology*, *3*(1), 149-160.

Georgiou, G. K., & Das, J. P. (2016). What component of executive functions contributes to normal and impaired reading comprehension in young adults? *Research in Developmental Disabilities*, *49*, 118-128.

Goldenberg, C. (2020). Reading wars, reading science, and English learners. *Reading Research Quarterly*, *55*, S131-S144.

Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills, and components of reading decoding in a normal sample. *Reading and Writing*, *18*, 583-616.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6-10.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.

Grabe, W. & Stoller, F. L. (2002). *Teaching and Researching Reading.* Longman.

Habók, A., & Magyar, A. (2019). The effects of EFL reading comprehension and certain learning-related factors on EFL learners' reading strategy use. *Cogent Education*, *6*(1), 1616522.

Huang, L., Ouyang, J., & Jiang, J. (2022). The relationship of word processing with L2 reading comprehension and working memory: Insights from eye-tracking. *Learning and Individual Differences*, *95*, 102143.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25-38.

Hoover, W. A., & Tunmer, W. E. (2022). The primacy of science in communicating advances in the science of reading. *Reading Research Quarterly*, *57*(2), 399-408.

Hopewell, S. (2011). Leveraging bilingualism to accelerate English reading comprehension. *International Journal of Bilingual Education and Bilingualism*, *14*(5), 603-620.

Joh, J., & Plakans, L. (2017). Working memory in L2 reading comprehension: The influence of prior knowledge. *System*, *70*, 107-120.

Johnson, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, *9*(1), 12-20.

Jung, J. (2018). Effects of task complexity and working memory capacity on L2 reading comprehension. *System*, *74*, 21-37.

Kail, R., & Hall, L. K. (2001). Distinguishing short-term memory from working memory. *Memory & Cognition*, *29*(1), 1-9.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology*, *58*(1), 88-111.

Lesser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning*, *57*(2), 229-270.

Nobre, A. D. P., & Salles, J. F. D. (2016). Lexical-semantic processing and reading: Relations between semantic priming, visual word recognition and reading comprehension. *Educational Psychology*, *36*(4), 753-770.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and writing*, *23*, 189-208.

Perfetti, C. A. (1988). Verbal efficiency in reading ability. In M. Dannamen, G. E. Mackinnon & T. G. Waller (Eds.), Reading research: Advances in theory and practice Vol. 6 (pp. 109–143). Academic Press.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of Functional Literacy*, *11*, 67-86.

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*, 289–312.

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 787-799.

Shimono, T. R. (2019). The effects of repeated oral reading and timed reading on l2 oral reading fluency. *The Reading Matrix: An International Online Journal*, *19*(1).

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, *1*(2), 147-170.

Smith Gabig, C. (2010). Phonological awareness and word recognition in reading by children with autism. *Communication Disorders Quarterly*, *31*(2), 67-85.

So, D., & Siegel, L. S. (1997). Learning to read Chinese: Semantic, syntactic, phonological, and working memory skills in normally achieving and poor Chinese readers. *Reading and Writing*, *9*, 1-21.

Spencer, M., Richmond, M. C., & Cutting, L. E. (2020). Considering the role of executive function in reading comprehension: A structural equation modeling approach. *Scientific Studies of Reading*, *24*(3), 179-199.

Swanson, H. L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology*, *85*(1), 1-31.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32-71.

Swanson, H. L., & Ashbaker, M. H. (2000). Working memory, short-term memory, speech rate, word recognition and reading comprehension in learning disabled readers: Does the executive system have a role? *Intelligence, 28*(1), 1-30.

Walczyk, J. J. (1995). Testing a compensatory-encoding model. *Reading Research Quarterly 30,* 396-408.

Walczyk, J. J. (2000). The interplay between automatic and control processes in reading. *Reading Research Quarterly*, *35*(4), 554-566.

Walczyk, J. J., Marsiglia, C. S., Bryan, K. S., & Naquin, P. J. (2001). Overcoming inefficient reading skills. *Journal of Educational Psychology*, *93*(4), 750.

Zinar, S. (2000). The relative contributions of word identification skill and comprehension-monitoring behavior to reading comprehension ability. *Contemporary Educational Psychology*, *25*(4), 363-377.

2023 by the authors. Licensee Journal of Teaching English Language (TEL). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution–Non Commercial 4.0 International (CC BY-NC 4.0 license). (http://creativecommons.org/licenses/by-nc/4.0).