The Impact of Automaticity on the Relationship between Working Memory Capacity and Second Language Reading Comprehension

Hikaru HOTTA

Abstract

本研究では、まず先行研究におけるワーキングメモリと第二言語読解における相関関係の「ばらつき」に注目し、その要因の一つとして自動化の概念を提示した。そして、自動化がワーキングメモリと相互作用を起こすことで、ワーキングメモリと第二言語読解の関係性に変化を生じさせるという仮説を展開した。また、相互作用のメカニズムとして 3 つのモデルを紹介し、実験の結果がどのモデルを支持するか検証した。実験の結果として、まず、ワーキングメモリと第二言語読解において有意な相関関係が観測された。また、自動化グループごとに相関関係を調べたところ、高自動化グループでは有意な相関が現れたが、低自動化グループでは観測できなかった。これは、自動化が先行研究における「ばらつき」に影響を与えている可能性を示唆している。さらに、第二言語読解における自動化とワーキングメモリの相互作用を調べたところ、自動化は特にワーキングメモリのスコアが高い被験者に有利にはたらいたことが分かった。この結果は *the-rich-get-richer* モデルを支持し、高自動化グループにおける強い相関関係に寄与していた。

Key Words: working memory, automaticity, L2 reading comprehension

1. Introduction

Humans complete various tasks with great ease, even when the tasks seem to be cognitively demanding. With a sufficient amount of practice, humans learn to type, drive a car, and use language effortlessly and almost unconsciously. This miraculous achievement by humans is often referred to as "automaticity" and has attracted research attention in cognitive psychology. Nonetheless, despite the increasing number of studies, how exactly automaticity interacts with our cognition and how it influences our language use has not been addressed adequately.

To address these issues, the present study focuses on the relationship between Working Memory (WM) capacity and Second Language (L2) reading comprehension. Regarding this relationship, many studies have reported positive correlations between WM capacity and L2 reading comprehension (e.g., Osaka 2002; Walter 2004). However, the level of correlation has varied widely among studies, which may imply that there is another variable impacting the relationship. The present study posits that automaticity plays a significant role in the relationship. More specifically, it is argued that automaticity interacts with WM capacity, which ultimately influences how much WM capacity correlates with L2 reading comprehension. With regard to the interaction effect, the study compared the results with three hypothetical models. The results of the experiment do not only contribute to a better understanding of the WM capacity and L2 reading relationship but also give important insights into how automaticity interacts with our cognition to influence our language use.

2. Working Memory and L2 reading

2.1 Working Memory

Perhaps the most influential model of WM is the one first proposed by Baddeley and Hitch in 1974 (Figure 1). They have identified three components of WM: the central executive, the phonological loop, and the visual-spatial sketchpad. Roughly summarized, the central executive is a supervisory system that serves as a control tower; it regulates information flow coming from the sensory register, and controls where to allocate attentional resources within WM. The phonological loop processes aural information and consists of two components: phonological store and articulatory loop. Through these components, although information stored in the phonological loop is susceptible to rapid decay, rehearsing this information by the articulatory loop allows information to be maintained. To elucidate the function of the phonological loop, Ellis (2001) gives an example of memorizing a phone number; without rehearsing, the number will rapidly decay but repeating the number string contributes to retaining the information while dialing. Lastly, the visual-spatial sketchpad deals with visual and spatial information. For example, remembering the route from one place to another would heavily depend on this component.

Since Baddeley and Hitch (1974), other theoretical models of WM have been proposed (cf. Baddeley 2000; Cowan 2016). Nevertheless, one of the consistent characteristics of WM that is generally agreed upon is its capacity-constrained nature. In contrast to the long-term memory (LTM), WM is known to be limited in its capacity, both in terms of storage and processing. To illustrate this, Miller (1956) conducted a famous experiment which revealed that the average number of digits humans can hold in their WM is seven plus or minus two. Although this experiment was mostly concerned with the storage function, the processing function of WM is also

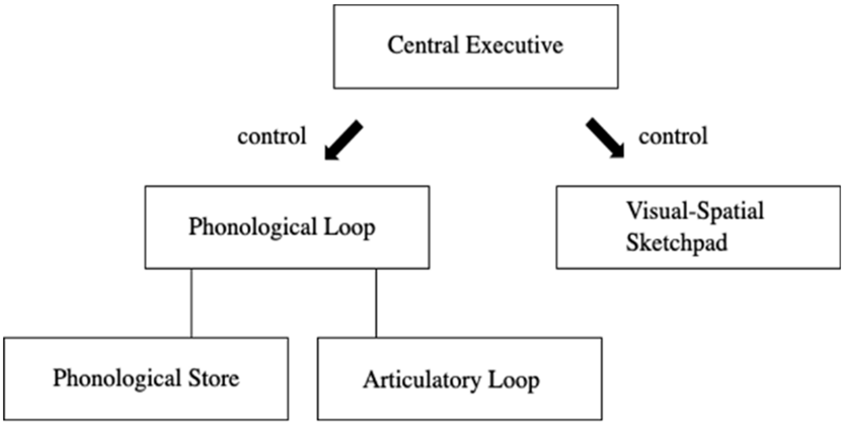

Figure 1. The model of Working Memory proposed by Baddeley and Hitch (1974)

known to be constrained. Consider, for instance, a student reading a text in a classroom while listening to a lecture. Such simultaneous tasks would be most likely difficult as both tasks (reading and listening) are either partially or fully dependent on the same function in WM, the phonological loop. As this example shows, this capacity-constrained characteristic of WM plays a crucial role in reading. This is because reading requires a variety of processes to operate simultaneously. In the next section, how WM plays a role in reading, particularly for L2 reading, is discussed.

2.2 Working Memory and L2 reading comprehension

Since the boom of WM research in cognitive psychology, reading research has seen an increase in studies investigating the relationship between WM and reading comprehension. For instance, many studies empirically examined the correlational relationship between WM and reading comprehension. In the previous literature, the consensus is that WM seems to be positively associated with reading comprehension; that is, the bigger the WM capacity, the better the reading comprehension performance. In L2 reading literature as well, studies are relatively consistent in showing that WM does seem to positively correlate with L2 reading comprehension (e.g., Shin, Dronjic & Park 2019; Walter 2004). However, just like in L1 reading, the level of correlation varies across research from no correlation (e.g., $r = -.068$ in Shahnazari-Dorcheh & Adams 2014) to strong correlation (e.g., $r = .79$ in Walter 2004).

Although the cause of such variation may be methodological (cf. Koda 2005), the present study suggests that there is another variable impacting the relationship between WM and L2 reading comprehension. On this issue, the theoretical model from Perfetti (1985) affords an important insight. According to *the Verbal Efficiency Model*, automatic lower-level processes of reading are essential for cognitive resources to be allocated for comprehension processes. The model assumes the trade-off relationships among various cognitive processes within WM and puts a particular emphasis on the importance of automaticity. In other words, when learners have attained a certain level of automaticity in reading, it changes how WM capacity impacts reading comprehension. Therefore, automaticity may bring important insight into the seemingly inconsistent relationship between WM and L2 reading comprehension.

2.3 Automaticity and reading

First, the distinction between automaticity and automatization should be made to avoid potential confusion. Although the two terminologies are sometimes used interchangeably, researchers often refer to automaticity as "the end result of a process of automatization" (Dekeyser 2001, p.130). Since the interest of the present study is mainly toward the result of automatization rather than the process itself, the term automaticity will be used unless specified otherwise.

In relation to reading, automaticity contributes greatly to reading fluency (Fukkink, Hulstijn & Simis 2005; Koda 2005). Although automaticity is important for every aspect of reading, previous studies often examine the role of automated word recognition in reading. For example, Speciale, Ellis, and Bywater (2004) suggest that for skillful readers of English, phonologically processing frequent sequences of letters such as /th/ is well automated. This is because native speakers store knowledge in the long-term memory that /th/ usually corresponds to the sound of [θ] or [ð]. Consequently, automaticity makes reading less cognitively demanding, which contributes to enhancing reading fluency.

As the example by Speciale et al. (2004) shows, it is clear that automaticity reduces the cognitive burden on the memory system. In fact, Dekeyser (2001) observes that the general shift in the literature on automaticity is "from theories which present automaticity as an issue of how much attention is given to a task to theories that present it as an issue of how memory is used" (p.130). Under this logic, it can be argued that people who have attained high levels of automaticity are likely to have more WM capacity. This assumption is consistent with the above-mentioned *Verbal Efficiency Model* (Perfetti 1985), which suggests that automatized lower-level processes allow cognitive resources to be used for other comprehension processes. Then, what does it say about the variation of the relationship between WM capacity and L2 reading comprehension? This issue is addressed in the next section.

2.4 Interaction between automaticity and Working Memory capacity in L2 reading comprehension

The present study argues that automaticity interacts with WM capacity in L2 reading comprehension, thus impacting the overall relationship between WM capacity and L2 reading comprehension. However, just as asking *whether* automaticity impacts the relationship is important, it is also informative to ask *how* it impacts the relationship. On this issue, the present study compared the results with three hypothetical models: the *compensation model*, the *independent-influences model*, and the *the-rich-get-richer model.*

The three models were taken from Hambrick and Engle (2002), which examined the interplay between background knowledge and WM capacity in L2 reading. Although the topic is partially different from the present study, the models are useful since the overall structure is the same. The three models for the present study are represented in Figure 2. The figure presupposes that the participants were divided into two groups according to their automaticity levels: the high automaticity group and the low automaticity group. The three models are represented by the slopes for the high automaticity group, while the slope for low automaticity group is drawn as a benchmark for all three models.

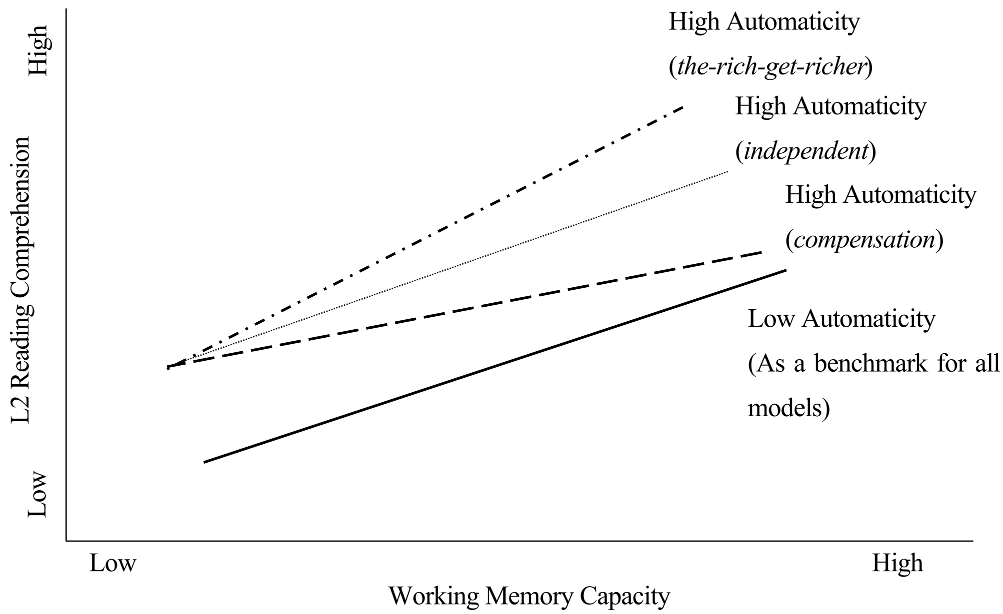

Figure 2. The hypothesized relationship between WM Capacity and L2 reading comprehension inspired by the three models proposed by Hambrick & Engle (2002)

First, the *compensation model* suggests that high levels of automaticity compensate for low levels of WM. This model predicts that learners with low levels of WM benefit more from their automaticity than their high WM counterparts. Consequently, the correlation between WM capacity and L2 reading comprehension should be weaker in the high automaticity group, resulting in the shallow slope in Figure 2. Secondly, the *independent-influences model* posits that automaticity and WM have additive yet independent effects on reading comprehension. According to this model, the correlation does not or should only slightly change between the high and low automaticity groups (note that the slope of the lines for the high and low automaticity are identical according to the *independent-influences model*.) Finally, the *the-rich-get-richer model* suggests that learners with high levels of WM benefit the most from their automated reading process. As a result, the correlation in the high automaticity group will be the stronger than their low WM counterparts.

All in all, the *compensation model* or the *the-rich-get-richer model* could account for the seemingly inconsistent result on the relationship between WM capacity and L2 reading comprehension in the previous literature. In contrast, the *independent-influences model* would fail to explain the variation since the correlation is expected to be more or less the same between the high and low automaticity groups.

2.5 Research Questions

The research questions (RQs) for the present study are as follows:

RQ1: Does WM capacity generally correlate with L2 reading comprehension?

RQ2: Does the correlation between WM capacity and L2 reading comprehension change between high automaticity group and low automaticity group?

RQ3: Which model (the *compensation model*, the *independent-influences model*, and the *the-rich-get-richer model*) does the finding support?

RQ1 is the most general question of three, followed by RQ2 and RQ3 which raise more specific issues. RQ1 is concerned with the overall correlation between WM and L2 reading comprehension regardless of the automaticity levels among participants. This question is to confirm the previous finding that WM capacity correlates with L2 reading comprehension. RQ2 is concerned with *whether* automaticity could, at least partially, explain the variation in correlation between WM and L2 reading comprehension in the previous literature. The grouping method into high and low automaticity groups is more elaborated in the next section. RQ3 is aimed to find *how* automaticity impacts the relationship between WM and L2 reading. For RQ3, the results will be analyzed based on the three hypothetical models from Figure 2.

3. Methodology

3.1 Participants

A total of 30 participants volunteered to participate in the study. All participants were undergraduate students from International Christian University, Japan. Only students whose native language was Japanese were eligible to participate since the task for measuring WM was written in Japanese. However, two participants grew up bilingual in both Japanese and English. They were included in the study because they had no problem with reading Japanese characters. Participants' age ranged from 18 to 23 years old (*M* = 20.47).

3.2 Materials & procedures

3.2.1 Reading Span Task

To measure WM capacity, the present study employed the Reading Span Task (RST). The task was employed in L1 (Japanese) to avoid L2 reading skills from impacting the score (cf. Shin, Dronjic & Park 2019). The Japanese RST employed in this study was based on Osaka (2002). Below, two example sentences are presented with English gloss.[2]

1)	野球	が	初めて	日本	に	伝えられた	の
	Yakyuu	*ga*	*hazimete*	*nihon*	*ni*	*tutae.rare.ta*	*no*
	Baseball	NOM	first	Japan	to	bring.PASSIVE.PAST	NMLZ

は	明治 5 年	ごろ	である。
wa	Meiji-gonen	goro	de-aru
TOP	Meiji 5	around	COPL

(It was around Meiji 5 when baseball was first brought to Japan.)

2)	その	技術	の	レベル	は	しろうと	の	域
	sono	*gizyutu*	*no*	*reberu*	*wa*	*sirooto*	*no*	*iki*
	that	technique	GEN	level	TOP	amateur	GEN	scope

を	はるかに	超えている。
o	harukani	koete-iru
ACC	far	go.beyond-ASP

(Lit. The level of that technique goes far beyond the scope of amateurs.)

Complete details of the procedure can be found in Osaka (2002). Participants read sets of sentences aloud while memorizing the bold and underlined target words. The task consisted of four sections: 2-sentence section, 3-sentence section, 4-sentence section, and 5-sentence section. Each section further consisted of 5 sets, which in total make up 70 sentences throughout the task ($2 \times 5 + 3 \times 5 + 4 \times 5 + 5 \times 5 = 70$) When a participant finished reading a sentence, the experimenter pushed "enter" for the next sentence to appear. The participants were instructed to read sentences as soon as they appear, with a constant speed throughout the task. When a participant finished reading sentences, a white blank screen appeared in which they needed to report the target words in a set. The participants were free to report the target words in any order, but they were not allowed to report from the final words to prevent recency effect. If a participant got more than 3 out of 5 sets correct, they "passed" the section to move on to the next. When they failed to do so, the experiment ended at that point. However, 2 out of 5 sets correct was counted as 0.5 points. The number of sentences included in the last section participants were able to pass was their final score. For example, a participant who made it to the 3-sentence section but failed at the 4-sentence section with 2 out of 5 sets correct received a score of 3.5 (3 + 0.5 = 3.5). Before the task, participants were able to practice with practice sentences until they were comfortable enough with proceeding.

3.2.2 C-test

To the best of the author's knowledge, there has not been any attempt to measure automaticity for reading comprehension in a holistic way. Some studies measure automaticity for word reading using the Stroop task (Stroop, 1935), and others measure automatization for word recognition (Fukkink et al. 2005). However, none of these studies attempt to measure automaticity for reading comprehension.

To overcome this methodological challenge, another test usually employed to measure L2 proficiency appeared as an appropriate alternative: C-test. C-test has been employed as an economic and reliable method of measuring proficiency in the previous literature (Lamb, 2012). The test usually consists of several texts in which every second half of roughly every second word was removed (see the Appendix for the present study's version of the C-test). The task makes use of the *reduced redundancy principle* (RRP) in the sense that speakers can fill in the missing elements in the distorted text because the structure of a natural language is often redundant for speakers of a language (Lamb 2012). However, there has been a concern in the literature regarding what C-test really measures. To address this, Babaii and Ansary (2001) have conducted verbal protocols to find what "cues" participants employ to fill in the missing letters in the C-test. After the analysis, four types of cues emerged: "automatic processing" (16.6%), "lexical adjacency" (54.9%), "sentential cues" (22.4%), and "top-down cues" (6.1%). The category of "automatic processing" is extremely similar if not identical to the concept of automaticity discussed in Section 2. Other cues are broadly concerned with reliance on lexical clues, grammatical features, and background knowledge respectively (Babaii & Ansary 2001).

To apply C-test, which is generally used to measure proficiency, to the present study, which is aimed to measure automaticity for reading comprehension, increasing the proportion of the first cue seems crucial: "automatic processing". Babaii and Ansary (2001) have claimed that when the task is relatively easy, the performance becomes more automated and less subject to conscious control. Moreover, if the task is conducted under time pressure, it is unlikely that participants will rely on "top-down cues", because they generally require participants to go back and forth within the text to find out the cues. Considering these two points, the present study employed C-test at the beginning level to make the reading process more prone to automaticity. Two C-test texts from Gilmore (2011) and two texts from Lamb (2012) were combined to make the C-test for the present study as simple as possible. Second, participants were required to read the texts aloud so that the experimenter could follow and were supplied the correct answer when they were stuck for more than 5 seconds. This procedure is to prevent the participants from taking too much time to fill in the missing items and encourage them to make use of their automatic reading capability to the greatest extent. When the participants made mistakes, the correct answer was supplied by the experimenter. However, the participants were encouraged to minimize the mistakes as much as possible so that their overall time is not measured faster than it should be. Following the tradition to measure automaticity level by the reaction times, the time to complete each text was recorded by the experimenter as the level of automaticity. The shorter the length of time, the greater the automaticity level. Finally, unlike the RST for WM, the texts were written in English for the C-test. This is to reflect the skill-specific nature of automaticity; that is, how much L2 reading skill has been automated can be measured only by the use of L2 texts (Dekeyser 2001).

3.2.3 L2 reading comprehension

To measure L2 reading comprehension skill, one practice TOEFL reading passage, "LOIE FULLER", published by Educational Testing Service (ETS), was employed in this study. There were 14 multiple-choice questions within the test, and 1 point was given to each question except for the Summary Completion Question with 2 points. However, 4 vocabulary questions were excluded in calculating the final score because they were solely made to check the vocabulary knowledge, not the comprehension skills in reading. This made the maximum score of 11. Participants had 20 minutes to complete the test.

3.2.4 Procedure of the entire experiment

The three tasks were all computer-based using the experimenter's laptop. Prior to the experiment, all participants filled out the consent form. Participants completed the tasks in the order of RST, C-test, and the L2 reading comprehension task. Each task was preceded by a practice session, in which participants tried out the short practice samples until they were comfortable enough to move on to the real task. The whole experiment took about 45 minutes. While employing various preventative measures to minimize the chance of spreading the novel coronavirus (wearing masks, social distancing, sanitizing equipment, etc.), the experiment was conducted face-to-face in a classroom at International Christian University.

4. Results

4.1 RQ1: Does WM capacity generally correlate with L2 reading comprehension?

To see the relationship between WM capacity and L2 reading comprehension, Pearson's correlation was calculated. The correlation was calculated between all the variables (RST, C-test, L2 reading comprehension task) to make sure there was no multicollinearity involved. Table 1 shows Pearson's correlation coefficients between all the variables. Consistent with the previous literature, there was a moderate and significant correlation between RST score and L2 reading performance (r = .40, p < .05). However, the stronger correlation was found between the automaticity and L2 reading scores (r = .61, p < .01). Nonetheless, only the moderate level of correlation indicates that the two variables were reasonably distinct constructs. On the other hand, RST and C-test scores were not correlated at all (r = .08, p >.05), implying that they were independent variables.

Table 1. Pearson's correlation coefficients between the three variables (*N* = 30)

	RST	C-test	L2 Reading
RST	—		
C-test	.08	—	
L2 Reading	.40*	.61**	—

*p < .05, **p < .01, df = 28

Table 2. Descriptive statistics of the automaticity variable for the high and low automaticity groups

	Low Automaticity (*N* = 15)			High Automaticity (*N* = 15)		
	M	*SD*	*Range*	*M*	*SD*	*Range*
Automaticity	98.0	15.84	65.26	66.31	9.05	29.0

4.2 RQ2: Does the correlation between WM capacity and L2 reading comprehension change between high automaticity group and low automaticity group?

To find out the change in correlation between WM capacity and L2 reading comprehension depending on automaticity levels (RQ2), participants were divided into two groups; the upper half of the participants belonged to the high automaticity group (N = 15), and the below half of participants into the low automaticity group (N = 15). Table 2 shows descriptive statistics of the automaticity variable, for the high and low automaticity groups. There was significant difference in the automaticity score between the two groups (t (28) = 6.72, p = < .001). Since automaticity is measured by the time taken for the task, the greater number indicated weaker levels of automaticity, unlike other variables. Therefore, the centered scores were generated by subtracting the mean value from the raw scores, and the polarity was reversed to ensure linearity across the three variables.

Since the data in the high and low automaticity group was not normally distributed due to the small sample sizes, the non-parametric Spearman's correlation was calculated for the low and high automaticity group. Table 3 summarizes the correlation coefficients. For the low automaticity group, there was only a moderate and non-significant correlation between RST and L2 reading comprehension (rs = .43, p > .05). However, in the high automaticity group, there was a strong and highly significant correlation (rs = .76, p < .01). The results show that the correlation between WM capacity and L2 reading comprehension changes between the high and low automaticity groups.

Table 3. Spearman's correlation coefficients in the high and low automaticity groups

	Low Automaticity (N = 15)			High Automaticity (N = 15)		
	RST	C-test	L2 reading	RST	C-test	L2 reading
RST	—			—		
C-test	-.06	—		.55*	—	
L2 reading	.43	.09	—	.76**	.45	—

*rs < .05, **rs< .01

However, other correlation coefficients shown in Table 3 may suggest that the interpretation requires a more careful approach. In the high automaticity group, automaticity and WM significantly correlated (rs = .55, p < .05), meaning that the highly significant correlation between WM and L2 reading in the high automaticity group may have been partially confounded by the automaticity variable. However, WM correlated with L2 reading more significantly than automaticity. Moreover, automaticity did not significantly correlate with L2 reading in the high automaticity group (rs = .45, p > .05). Therefore, it should be reasonable to conclude that there was a more significant level of association between WM capacity and L2 reading comprehension in the high automaticity group than the low automaticity group.

4.3 RQ3: Which model (the *compensation model*, the *independent-influences model*, and the *the-rich-get-richer model*) does the finding support?

To see how automaticity impacts the relationship between WM capacity and L2 reading comprehension, the regression lines were drawn on the scatterplot for the high automaticity group and the low automaticity group (Figure 3). The dots represent each data point and were randomly jittered slightly so that overlapping points are reflected in the figure. From the graph, it is quite interesting that for every WM value (x-axis), participants in the high automaticity group scored better than those in the low automaticity group with only a few exceptions. Moreover, as the steeper slope for the high automaticity group shows, they tended to outperform low-automaticity counterparts with the same WM score more when they have achieved high WM scores. Comparing the graph with the three models in Figure 2, the *the-rich-get-richer model* seems to be most in line with the data. The model suggests that participants with high WM scores benefit more from automaticity in their L2 reading comprehension.

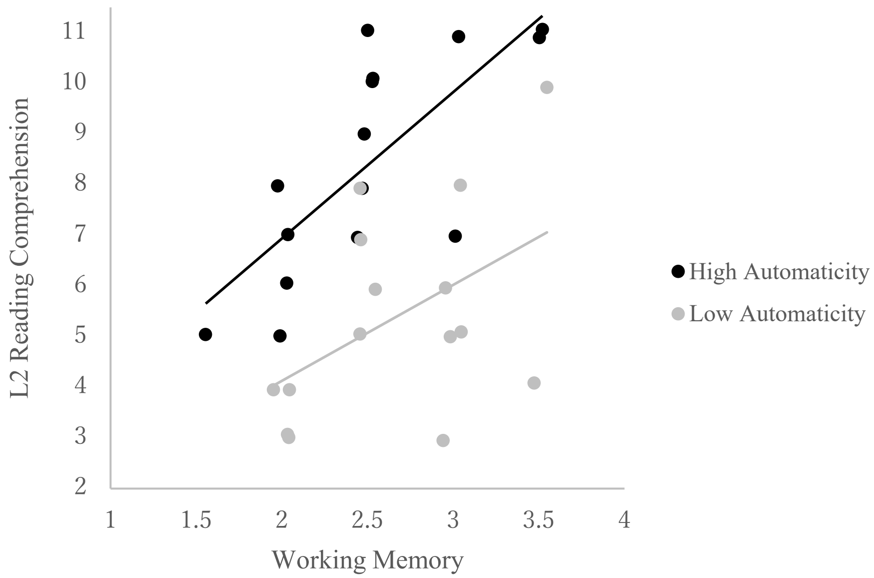

Figure 3. The scatterplot of the relationship between WM and L2 reading comprehension

However, Figure 3 has its own limitations. First, as the sample size for each automaticity group was relatively small ($N = 15$), drawing the regression line for each group could be argued as an inappropriate operation. Second, the grouping method fails to capture the spectrum within groups: in this case, the variation of automaticity. One way to overcome this issue may be to divide up participants into smaller groups to reflect more of the spectrum. However, this reduces the sample sizes even further, which was deemed inappropriate for the correlational analysis[1].

To conclude, WM capacity significantly correlated with L2 reading comprehension. However, the significant correlation was observed in the high automaticity group, but not in the low automaticity group. Regarding how automaticity impacts the relationship between WM capacity and L2 reading comprehension, the *the-rich-get-richer model* seems to best fit the obtained data.

5. Discussion

5.1 Working Memory capacity and L2 reading comprehension

Consistent with the previous literature, the present study has found a significant correlation between WM capacity and L2 reading comprehension ($r = .40, p < .05$). The unique finding of this study is that even when WM was measured in first language (Japanese), there was a significant correlation with L2 (English) reading comprehension. The present study reaffirmed the significant association between WM capacity and L2 reading comprehension even when the task is conducted in a language other than English.

However, it is important not to overplay the implication of the result. In contrast to the present study, the previously mentioned study by Shin (2020) had found that the correlation was weak and non-significant when the RST was conducted in a first language ($r = .17, p > .05$). Since the previous study was a meta-analysis, the sample size was quite large (N = 1337). However, the weak correlation found by Shin (2020) reflects the average value out of a wide range of correlation coefficients. Therefore, the issue ultimately comes down to the initial question, "why has the correlation been so inconsistent?", to which the present study has provided important insights by introducing the concept of automaticity.

5.2 The role of automaticity in the relationship between WM & L2 reading comprehension

In answering RQ 2, the current results found that the correlation was strong and significant in the high automaticity group but non-significant in the low automaticity group. Moreover, in answering RQ 3, the results best supported the *the-rich-get-richer model*, suggesting that people with high WM capacity seem to benefit more from their automated reading abilities.

The result may be best interpreted by *the Verbal Efficiency Model* (Perfetti 1985) introduced in Section 2. According to the model, automaticity of the lower-level processes such as word recognition is essential in the higher-level comprehension processes. The theory claims that one of the significant hurdles for skilled reading is that when the word recognition skills are not well automated, the WM capacity cannot be effectively used for comprehension. In the present study, the low automaticity participants may have failed to use their cognitive resources for comprehension regardless of their WM capacity, which ultimately led to the non-significant correlation between WM capacity and L2 reading comprehension. This interpretation is reflected in Figure 3, in which the low automaticity group struggled to perform well in L2 reading comprehension, even when they have achieved high WM capacity.

In contrast, the high automaticity group seemed to be able to use their WM capacity to the fuller potential. This advantage resulted in the high automaticity group constantly outperforming the low automaticity counterparts with the same WM levels. Moreover, the high WM participants in the high automaticity group tended to score better than any other counterparts, which ultimately hiked up the correlation coefficients between WM capacity and L2 reading comprehension. The finding is in agreement with the *the-rich-get-richer model,* reflected by the steeper slope for the high automaticity group in Figure 3.

5.3 Methodological limitations

There were several methodological limitations in this study that should be addressed. First of all, there were concerns with regard to the automaticity task. As described in 3.2.2, the task only measured speed but not the accuracy of the performance. The experimenter encouraged participants to minimize the mistakes by asking them to simply pause reading when they were unsure of the words to fill in. However, most participants did make a few mistakes by accidentally reading the incorrect words. The unexpected errors have ultimately increased the speed of their performance than it should be. Although the experimenter lacked sufficient knowledge to overcome this methodological challenge, it should have been the case that at least the number of mistakes were counted during the experiment so that it could be reported in the paper for the sake of transparency.

Moreover, performance on the automaticity task may have been impacted by other unexpected variables. When the experimenter briefly interviewed participants at the end of the experiments, some participants have told that they had trouble filling in the blanks simply because they did not know the target words. In this case, vocabulary knowledge, rather than their automated reading skills, impacted their performance. In addition, since C-test is usually used for measuring proficiency, it could be argued that automaticity score only reflects proficiency. Under this assumption, participants may have needed to reach a certain level of proficiency to utilize WM capacity in L2 reading comprehension.

However, this argument is somewhat at odds with some previous findings. For instance, Shahnazari-Dorcheh and Adams (2014) found that the significant correlation between WM score and L2 reading comprehension was only observed at the beginning levels of proficiency but gradually disappeared as the participants' proficiency increased. Moreover, Walter (2004) involved participants with both lower-intermediate proficiency and upper-intermediate proficiency and the correlation was stronger for the lower-intermediate participants ($r = .72$, $p < .0001$) than the upper-intermediate participants ($r = .46$, $p < .01$). If automaticity was a mere reflection of proficiency, the present study should also show weaker correlation for the high automaticity group – but it did not. Therefore, it should be reasonable to conclude that the present study was able to measure the unique feature of automaticity, which was reasonably independent of L2 proficiency.

6. Conclusion

The present study was set out to address the issue in the previous literature, namely the wide variation in the correlation between WM capacity and L2 reading comprehension. The study hypothesized that automaticity may play a role in the variation. The study further hypothesized that automaticity may interact with WM in L2 reading comprehension, which ultimately influences how much WM capacity correlates with L2 reading comprehension. With regard to this interaction effect, the study compared the results to the three models proposed by Hambrick and Engle (2002).

The results of the experiment showed that there was a significant correlation between WM capacity and L2 reading comprehension. However, when the correlation was calculated for each automaticity group, a significant correlation was found among the high automaticity learners but not among the low automaticity learners. This result was significant in revealing how the correlation between WM capacity and L2 reading comprehension can change depending on the automaticity levels. Moreover, with regard to the interaction effect, automaticity seems to benefit high WM participants to a greater extent than the low WM counterparts. This resulted in the strong correlation in the high automaticity group, best supporting the *the-rich-get-richer model* out of the three hypothetical models.

For further research, more studies on the role of automaticity in language use and how it interacts with our cognition are to be awaited. As the present study also struggled, automaticity is extremely hard to measure as it is intertwined with other variables. Nonetheless, more research attention should contribute to uncovering the elusive yet remarkable ability of language learner.

Acknowledgement

This paper grew out of my senior thesis. I am grateful to my senior thesis advisor, Dr. Akiko Fujii, for the continuous support. I would also like to thank my current advisor, Dr. Seiko Fujii for helping me with the English gloss. Lastly, comments from two anonymous reviewers have greatly improved the manuscript.

Notes

1 It was pointed out by an anonymous reviewer that the automaticity score could rather remain as a continuous variable for the statistical analysis. Although the author fully acknowledges this possibility, it was concluded that the operation is at odds with the hypothetical models necessitating the grouping method. The possibility will be explored for my future studies.

2 The abbreviations used in the gloss are as follows: ACC=accusative, ASP=aspectual marker, COPL=copula, GEN=genitive, NOM=nominative, NMLZ=nominalizer, TOP=topic marker

References

Babaii, E., & Ansary, H. (2001). The C-test: A valid operationalization of reduced redundancy principle? *System* (Linköping), *29*(2), 209–219.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation*, *6*, 47–89.

Cowan, N. (2016). *Working memory capacity* (Classic Ed.). New York: Routledge.

Dekeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (Cambridge Applied Linguistics, 125–151). Cambridge: Cambridge University Press.

Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (Cambridge Applied Linguistics, 33–68). Cambridge: Cambridge University Press.

Fukkink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal*, *89*(1), 54–75.

Gilmore, A. (2011). "I prefer not text": Developing Japanese learners' communicative competence with authentic materials. *Language Learning, 61*(3), 786–819.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*(4), 339–387.

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

Lamb, M. (2012). A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning*, *62*(4), 997–1023.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Osaka, M. (2002). *Nou no memochou wākingu memori = Working memory: The sketchpad in the brain*. Tōkyō: Shinyōsha.

Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Shahnazari-Dorcheh, M., & Adams, R. (2014). The relationship between working memory and L2 reading comprehension. *Applied Research on English Language*, *3*(6), 19–34.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics, 41*(4), 873–900.

Shin, J., Dronjic, V., & Park, B. (2019). The interplay between working memory and background knowledge in L2 reading comprehension. *TESOL Quarterly*, *53*(2), 320–347.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity

determine second language vocabulary acquisition. *Applied Psycholinguistics*, *25*(2), 293–321.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.
Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics, 25*(3), 315–339.

Appendix. C-Test

TEXT 1

"Hi! My name's Ben. I'm ten ye--- old a-- I'm fr-- Australia. I've g-- brown ha-- and bl-- eyes. I li-- with m- parents a-- my sis--- in a sm--- house i- Sydney. I- my fr-- time I li-- to sw-- in t-- sea a-- to ri-- my bi--. I al-- enjoy pla---- games o- my comp----."(Lamb 2012, Appendix S2)

TEXT 2

Sarah,

We're in London, having a great time. Our English friends, Alice and Becky, m-- us a- the air---- and th-- we to-- a b-- into t-- city cen---. We're sta---- in a ho--- near Buckingham Pal---, b-- we ha-- not se-- the qu--- yet! Th--- are s- many

thi--- to d- here, b-- it is ve-- expe-----, and t-- weather i- quite co--. I th--- my Eng---- is impr----- fast!

Best wishes,

Etty & family (Lamb 2012, Appendix S2)

TEXT 3

Traditional English breakfast is a very big meal; sausages, bacon, eggs, tomatoes, mushroom. But nowa---- many peo--- just ha-- cer--- with mi-- and su---, or

to--- with marm-----, jam, o- honey. Marmalade a-- jam a-- not t-- same!

Marm----- is ma-- from ora--- and j-- is ma-- from ot--- fruit. T-- traditional brea----- drink i- tea, wh--- people ha-- with co-- milk. So-- people ha-- coffee, of--- instant cof---, which i- made wi-- just h-- water. Many visitors to Britain find this coffee disgusting! (Gilmore 2011, Appendix S1)

TEXT 4

Every morning billionaire Milton Petrie walked from his New York apartment and bought a newspaper from the ragged old man on the street corner. One mor---- the m-- wasn't th---. Petrie lea---- that h- was ve-- ill i- the ci-- hosp----. He pa-- his hosp---- bill a-- later, wh-- the m-- died, pa-- for h-- fun----. The old man was just one of many people that Milton Petrie helped with his money. (Gilmore 2011, Appendix S1)