The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

SATORI, Miki
Graduate school, Kumamoto University

Abstract
The present study was to examine if working memory capacity was related to the performance of Japanese EFL learners on listening comprehension sub-skills. More specifically, the study focused on L1 and L2 working memory capacity, and its effect on two dimensions in L2 listening comprehension: literal comprehension and inferential comprehension. The study also investigates the extent to which the role of working memory in the processing of L2 listening varies according to different levels of L2 proficiency. The results of the correlation and regression analysis suggested that the individual differences in working memory capacity were related to inferential comprehension in the higher-level group, and both inferential and literal comprehension in the lower-level group. The results of the study will discuss the implication of these findings and suggest that working memory resources can be allocated to the execution of top-down processing in the case of the higher-level listeners, and both top-down and bottom-up processing in the case of the lower-level listeners.

Keywords: listening, working memory, comprehension sub-skills

1. Introduction
Listening comprehension is not a passive activity but an active process in which the listener must discriminate between sounds, understand vocabulary and grammatical structures, interpret stress and intonation, and interpret them within the sociocultural context. Integrating all of these involves a great deal of complex cognitive activity on the part of the listener (Field, 1998; Vandergrift, 1999). Listening is also a personal process affected by the individual differences in non-linguistic knowledge such as background knowledge about the world (Buck, 2001). Different types of linguistic knowledge and psycholinguistic skills are comprised in listening processing, and affect
The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

the listening performance by interacting with each other (Richards, 1983). As one of such various factors affecting the listening processes, working memory (henceforth, WM) capacity is thought to play a critical role in listening processing by storing the result of the listeners’ comprehension as they deal with the information in a spoken discourse at the same time. According to Just and Carpenter’s (1992) theory, WM capacity constrains comprehension and the total amount of the activation mediating processing and storage varies among individuals. Most of the results of the studies on the contribution of WM capacity to L2 performance suggest that having a larger WM capacity has advantage in using and learning L2 (Miyake & Friedman, 1998).

The recent research in reading has investigated the contribution of WM capacity and background knowledge to inferential processing during reading comprehension (Alptekin & Erçetin, 2010; Andreassen & Bråten, 2010; Rai, Loschky, Harris, Peck, & Cook, 2011). However, there has so far been little research that addressed the role of WM capacity in literal and inferential comprehension in L2 listening. Based on a hypothesis that the limited capacity of WM could affect both literal and inferential comprehension in L2 listening, the present study aims at examining the contribution of WM capacity to both bottom-up and top-down processing in L2 listening.

2. Background

2.1 Top-down and bottom-up processing in L2 listening

Some earlier studies in L2 listening found that both top-down and bottom-up processing play critical roles in L2 listening comprehension (Lynch, 1998; Rubin, 1994; Vandergrift, 2004). Listeners use top-down, or knowledge-based processing when they use pre-existing knowledge to understand the text and construct a conceptual mental model of what the text was about. Listeners also use bottom-up, or a data-driven processing when they decode the incoming linguistic input and build the meaning by combining increasingly larger information from surface-level up to discourse-level features. These two processes interact with each other to help listeners in real-life listening, and the degree to which they need to use both processing to their advantage depends on their purpose for listening (Vandergrift, 2004). Shohamy and Inbar (1991) investigated the effect of text and question types on listening comprehension, and found that ‘(w)hile high-level listeners seemed to process the text in a knowledge-based manner, the low level test takers seemed to process the text in a data-driven manner’ (p. 35). These two processes are in competition for limited space of WM capacity (Just & Carpenter, 1992). Because of limited L2 linguistic knowledge, low-level listeners can
process little of what they hear automatically. As a result, they need to pay conscious attention to individual words and devote more of the limited cognitive resources to bottom-up processing (Vandergrift, 2004). Less-skilled listeners also need plenty of contextual support to compensate the gap of comprehension, due to the lack of the automatized linguistic decoding skills (Goh, 2002; Segalowitz & Segalowitz, 1993). Therefore, automatizing bottom-up processing such as decoding and syntactic parsing has the potential of helping low-level listeners save cognitive resources devoted to bottom-up processing and improve their listening comprehension (Vandergrift, 2004).

2.2 Taxonomy of L2 Listening comprehension sub-skills

Listening comprehension is not a simple process of decoding an incoming input but a complex multidimensional process, and different researchers have attempted to describe them regarding the taxonomy of comprehension underlying the listening process. There has been, however, no general consensus on the adequate definition of listening construct (Buck, 2001).

The taxonomy of L2 comprehension sub-skills used for the present study was based on the one developed by Shohamy and Inbar (1991). Based on the strategies utilized for text processing, they divided listening comprehension questions into two dimensions: global and local. Global comprehension requires ‘the test takers to synthesize information, draw conclusions and focus on cause and effect relationships and on inferences’. Local comprehension requires ‘the test takers to locate details, understand single words which have contextual support, paraphrase and recognize fact’ (p.29). They found that listeners performed better on local comprehension questions than on the global ones across all levels of listeners. The results suggested that comprehension performance differs across the levels of cognitive processing, and deeper-level processing requiring generating, inferring, and synthesizing the information was more cognitively challenging than data-driven processing directly related to the text-base information. According to Kintsch’s (1998) construction-integration model, text comprehension constructs two levels (the text-base and the situation model) of representational architecture. The text-base level corresponds to a semantic representation of the proposition constructed from the linguistic input in the text, and the situation model corresponds to a knowledge-based mental representation of what the text is about. Both text-base representation and deeper-level situation model seem to place cognitive load on WM capacity (Graesser, Singer, & Trabasso, 1994).
2.3 Framework of WM

The current conception of WM emphasizes the dynamic nature of simultaneous processing and storage activity (Miyake & Friedman, 1998). WM is thought to play an important role in various kinds of cognitive operation requiring controlled processing (Engle, 2002; Ortega, 2009). WM is especially important in language comprehension (Daneman & Merikle, 1996; Harrington & Sawyer, 1992; Miyake & Friedman, 1998; Walters & Caplan, 1996; Leeser, 2007).

Among the variety of models of WM, one of the most influential models is Baddeley’s multi-component model, which comprises the central executive and its slave systems. The central executive component of WM has a variety of functions including attention control, coordinating the storage, processing of information, and activating representation within long-term memory (Baddeley, 2007). Engle, Kane, and Tuholski (1999) argue that WM is regarded as a construct separate from that of short-term memory (henceforth, STM), and assume WM capacity is not about storage or memory but about ‘the capacity for controlled, sustained attention in the face of interference and distraction’ (p.104). According to a controlled-attention theory of WM (Engle & Kane, 2004), there is a domain-general component of WM responsible for controlling attention as well as domain-specific components for temporarily representing information. The domain-general controlled attention ability is related to both higher-level cognition such as L2 comprehension and lower-level cognition requiring cognitive control (Colflesh & Conway, 2007). Colflesh and Conway state that greater performance in the WM tasks may reflect greater controlled attention and/or better use of domain-specific skills and strategies to aid maintenance.

Both storage and processing aspect of WM capacity can be measured separately or in combination (Juffs & Harrington, 2011). Whereas the traditional passive STM tasks such as digit span tasks tap only the storage component of the WM systems, the complex WM tasks tap both storage and processing (controlled attention) functions of the WM systems. A widely used complex span test is the sentence span task developed by Daneman and Carpenter (1980). Reading Span Test (henceforth, RST) is commonly used to measure WM capacity, and Listening Span Test (henceforth, LST) is used to measure WM capacity in the case of listening. Engle, Tuholski, Laughlin, and Conway (1999) consider that if the shared variance between STM task and complex WM task reflects storage component of WM, the residual of complex WM capacity should reflect controlled attention.
2.4 Working memory and reading comprehension sub-skills

In addition to numerous studies on contribution of WM to reading comprehension (Harrington & Sawyer, 1992; Miyake & Friedman, 1998; Waters & Caplan, 1996), there have been attempts to examine the relationship between WM capacity and reading comprehension sub-skills, especially inferential comprehension in L1 and L2 reading. In the case of L1 reading, Andreassen and Bråten (2010) investigated the contribution of word recognition skills, strategic text processing, reading motivation, and WM to L1 reading comprehension. The results of the hierarchical multiple regression analysis showed that WM capacity was found to be predictive of comprehension independently of word recognition skills when the text consisted of a longer passage, and contained a large portion of inferential questions. In the case of L2 reading, Alptekin and Erçetin investigated the relationship of both L1 and L2 WM capacity to L2 reading comprehension with proficient L2 readers through RST in L1 (Turkish) and L2 (English). Alptekin and Erçetin separated L2 reading comprehension into its two dimensions: literal comprehension and inferential comprehension. Their study yielded interesting results. Firstly, there was no significant relationship between literal comprehension and any WM capacity measures. Secondly, inferential comprehension correlated positively with L2 RST performance. The results suggest that WM capacity plays a greater role in inferential reading rather than literal reading, because inferential comprehension requires more cognitive resources than literal comprehension. The results of the study should be applicable to L1 users and advanced L2 learners, but the role of WM may give rise to different patterns of influence on literal and inferential comprehension in L2 listening depending on their proficiency levels.

2.5 The role of WM in listening comprehension

Compared to the number of the study on the influence of WM capacity on reading comprehension, relatively little research has focused on the influence of working memory capacity on listening comprehension. Montgomery, Polunenko, and Marinellie (2009) examined the role of L1 WM in children’s comprehension of spoken narrative, and showed that L1 WM measured by L1 LST accounted for a significant 7.9% of unique variance in L1 listening comprehension. The only study on the role of WM in L2 listening for Japanese EFL university students was conducted by Sakuma (2004). The results showed moderate correlations among WM capacity measured by L2 LST, L2 listening comprehension test, and its vocabulary and grammar sections.
Based on the controlled attention theory of WM (Engle & Kane, 2004), the relationship between WM capacity and controlled attention in the dichotic listening was explored by Colflesh and Conway (2007). The results of the study suggested that individuals with greater WM capacity outperformed ones with lower WM capacity not only in selective attention but also in divided attention. Colflesh and Conway also showed that WM capacity significantly affected listening task performance only when the tasks place greater demands on executive attention. Considering these results, the individual differences in WM capacity are thought to be associated with the ability to focus attention and avoid distraction to meet the demanding task goals while listening.

2.6 Relation between L1 and L2 WM capacity

The findings of the research that addressed the parallelism between L1 and L2 WM capacity suggest that cognitive resources underlying WM capacity are shared by the two languages and the relationship is language-independent (Osaka & Osaka, 1992; Osaka, Osaka, & Groner, 1993). Osaka and Osaka found a high correlation ($r=.72$) between L1 (Japanese) RST and L2 (English) RST with highly skilled L2 users. Osaka et al. (1993) also confirmed a high correlation ($r=.85$) between L1 (German) and L2 (French) RST with highly proficient bilinguals. As Miyake and Friedman (1998) state, L2 processing may share the same WM resources as L1 processing at least in bilingual-level L2 users. Other studies suggest that L1 and L2 WM capacity correlated moderately with each other (Harrington & Sawyer, 1992; Juffs, 2004; Miyasako, 2006; Miyake & Friedman, 1998). For example, the correlations between L1 (Japanese) and L2 (English) RST were found to be .39 (Harrington & Sawyer, 1992) and .47 (Miyasako, 2006). Most of the previous studies were conducted with proficient L2 users, but in the case of lower-level L2 users, the relationship between L1 and L2 WM capacity might be affected by their competence in L2 (Sagarra, 2008). In this regard, Van den Noort, Bosch, and Hugdahl (2006) investigated the interaction between WM capacity and the language proficiency level with fluent German (L2) users of Dutch (L1) speakers who learned Norwegian (L3) for six months before the study. The results showed that their performance in L1, L2 and L3 RST increased in relation to their language proficiency level. There was also found to be significant correlations ($r>.75$) among L1, L2 and L3 RST. These findings suggest that substantial amounts of the variance in WM capacity are closely related among the three languages.

Whether WM capacity should be measured in the L1 or L2 would be another important issue in the study of WM capacity and L2 performance (Alptekin & Erçetin, 2006).
2009). Whereas Leeser (2007) found that elementary learners of Spanish with higher span of L1 WM capacity outperformed in the L2 text recall than those with lower span of L1 WM capacity, Chun and Payne (2004) found no relationship between L1 RST and L2 reading comprehension. Walter (2004) investigated the relationship between L1 and L2 RST and L2 reading comprehension for French learners of English. The correlation analysis showed that the correlation between L1 RST and L2 summary completion was comparatively low ($r$=.33). On the other hand, there was a strong relationship ($r$=.73) between L2 RST and L2 summary completion. These findings suggest that L2 reading comprehension is related more to L2 WM capacity than L1 WM capacity. Miyake and Friedman (1998) also examined the relationship between L1 and L2 LST, and L2 syntactic processing. The results of a path analysis indicated that L1 WM capacity was a mediator variable directly related to L2 WM capacity but indirectly related to L2 syntactic comprehension. Considering these research findings, L2 WM capacity, by and large, plays a greater role in L2 reading and listening than L1 WM capacity.

3. Research Questions

Separating L2 listening comprehension into its literal and inferential dimensions, the present study attempted to examine the role of L1 and L2 WM in L2 listening with both advanced and beginner L2 learners through LST in L1 (Japanese) and L2 (English). As Alptekin and Erçetin (2010) suggest, WM capacity might be differently related to these two dimensions of comprehension. Based on the hypothesis that the role of WM might differ across the proficiency levels of the L2 users, more specifically, that there is a meaningful association between WM capacity and inferential listening of the skilled listeners, and literal listening of the less-skilled listeners, the present study attempted to provide answers to the following research questions:

RQ1: Are individual differences in L1 and L2 WM tasks related to literal and inferential comprehension in L2 listening?

RQ2: How does the role of L1 and L2 WM in literal and inferential comprehension differ across L2 proficiency levels?
4. Method
4.1 Participants

Participants in this study were 149 Japanese 1st and 2nd year students from a technical college in Japan. The majority were 19-20 years of age and homogeneous in their social and educational background. Their major was English language and they had three TOEIC classes in a week. They were to take TOEIC IP test 4 times in a year. In terms of their performance on the TOEIC IP test conducted in July, their scores ranged from 205 to 970. (The average of their TOEIC scores =444.7, SD =151.2). Before participating in the research, the participants were asked to read and sign a consent form.

4.2 Instruments
4.2.1 TOEIC listening test

The TOEIC (Test of English for International Communication) is a two-hour multiple-choice test that consists of 200 questions divided into listening section and reading section. The listening section consists of four parts and contains 100 questions, each of which has three or four options. These questions are based on a variety of statements, questions, conversations, and talks. Scores of both sections are determined by the number of correct answers, which is converted to a scaled score ranging from 5-495. Scaled scores are transformed and derived from test takers’ raw scores through a proven statistical procedure. In this study, the participants’ scores on the Listening section of the TOEIC IP test conducted in July was used as an indicator of their L2 listening proficiency levels.

4.2.2 L2 listening comprehension test

The participants’ literal and inferential comprehension in L2 listening was measured with a multiple-choice listening test (Stafford, 2009) which included the same type of questions as Part III and Part IV of the TOEIC listening test. The test contains 48 multiple-choice questions based on 18 extended conversations and talks. For each of the 18 passages, there are two or three questions. These passages encompass a wide range of genres reflecting the circumstance of real world listening and difficulty levels. The aural texts contain features of a normal speech rate and phonological modification. At the same time, to reduce floor and ceiling effects, the test items were carefully chosen to be at the right difficulty level for the population. Most of the lexical items used in the test were familiar words the students recognize in
the written forms of them, but they might not recognize in the spoken forms of them. Based on Shohamy & Inbar’s (1991) taxonomy of listening questions, half of the questions were textually explicit, while the other half were textually implicit. Textually explicit questions were regarded as measuring literal comprehension in that their answers could be derived directly from the text. On the other hand, textually implicit questions mainly tapping elaborative inference were regarded as measuring inferential comprehension requiring extraction of the deeper meaning of the text by going beyond the text-base into constructing a mental model of what the text was about (see Appendix A for examples of the test items).

4.2.3 L1 and L2 Listening Span Test (LST)

A widely administered test of aural WM capacity is the Sentence Listening Span test. The test is recognized to simultaneously tap both the processing and storage functions of WM. There are a number of variations on this test format. In the present study, L1 WM capacity was measured through the Japanese version of LST (Endo & Osaka, 2011). On the other hand, L2 WM capacity was measured through the ESL version of LST developed by Ushiro and Sakuma (2000). Both LSTs were conducted from the two-sentence condition to the five-sentence condition. Three sets of sentences were presented in each sentence condition. The participants listen to the set of sentences. Following each sentence, they are obliged to perform a sentence verification task to indicate their comprehension of the material. After listening to the entire set, the participants are required to recollect the last word of each sentence in the ESL version of LST, and the first word of each sentence in the Japanese version of LST. Scoring was conducted based on the total number of the target words recalled correctly rather than traditional span scores on the basis of Freedman and Miyake’s (2005) findings that the total number of words recalled or the proportion of words per set averaged across all sets had higher reliability and higher correlations with reading comprehension than did traditional span score (see Appendix B for examples of the test items).

4.2.4 L1 and L2 Digit Span Test (henceforth, DST)

In the present study, STM capacity was measured through L1 and L2 random digit span tasks (DST) carefully pilot tested and recoded by Joyce (2008). In the pilot test, the number of digits per string and the number of strings were determined to ensure that there was neither a floor nor a ceiling effect induced by an insufficient number of
appropriately sized strings. Owing to a difference in the difficulty of the tests, the L1 DST ranged from six to nine digits while the L2 DST ranged from five to eight digits. Each of the items only contained the numbers from one to nine, and each digit was used only once. The order of the numbers was randomly selected. The recorded materials were uttered with a falling intonation, and there was a one second interval between each digit. The participants attempted to reproduce the sequence of numbers on the answer sheet immediately after hearing each string. After completing their answers, they turned over the sheet to remove any possible advantage from seeing their previous responses. To receive credit for an answer, the participants were required to both reproduce all the necessary digits, and ensure that they were in the correct order. One point was allocated for each correctly reproduced string. There were 16 items in both measures (see Appendix C for examples of the test items).

4.3 Procedure

The study was conducted in July 2011. The TOEIC IP test was administered to all the participants outside of class time on the same day. Within a week after the TOEIC IP test was conducted, the participants took all the other tests during their TOEIC class time, which were divided into two lessons to reduce their burden.

5. Results

5.1 Results for RQ 1

5.1.1 Descriptive Statistics

The descriptive statistics for all measures are presented in Table 1. As can be seen in Table 1, inferential comprehension questions were slightly more difficult than literal comprehension questions.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>Full marks</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEIC Listening test</td>
<td>100</td>
<td>495</td>
<td>272.85(54.9%)</td>
<td>84.56</td>
<td>149</td>
</tr>
<tr>
<td>Inferential comprehension</td>
<td>24</td>
<td>24</td>
<td>12.91(53.8%)</td>
<td>4.40</td>
<td>149</td>
</tr>
<tr>
<td>Literal comprehension</td>
<td>24</td>
<td>24</td>
<td>14.59(60.8%)</td>
<td>4.04</td>
<td>149</td>
</tr>
<tr>
<td>L1WM</td>
<td>42</td>
<td>42</td>
<td>34.81(82.8%)</td>
<td>6.01</td>
<td>149</td>
</tr>
<tr>
<td>L2WM</td>
<td>42</td>
<td>42</td>
<td>25.95(61.8%)</td>
<td>6.61</td>
<td>149</td>
</tr>
<tr>
<td>L1STM</td>
<td>16</td>
<td>16</td>
<td>5.54(34.6%)</td>
<td>2.80</td>
<td>149</td>
</tr>
<tr>
<td>L2STM</td>
<td>16</td>
<td>16</td>
<td>8.15(50.9%)</td>
<td>3.22</td>
<td>149</td>
</tr>
</tbody>
</table>
The inferential comprehension questions yielded fairly good internal consistency (Cronbach’s $\alpha=.76$), as did the literal comprehension questions (Cronbach’s $\alpha=.73$).

### 5.1.2 Correlation Analysis

The correlations among the variables are shown in Table 2. Firstly, L1 WM capacity significantly yielded weak correlations with L2 listening comprehension, L2 inferential comprehension, and L2 literal comprehension. L2 WM capacity also significantly formed slightly stronger, but moderate correlations with L2 listening comprehension, L2 inferential comprehension, and L2 literal comprehension. On the other hand, L1 STM capacity did not have a correlation with any L2 listening comprehension, whereas L2 STM capacity significantly formed weak correlations with both L2 inferential and literal listening comprehension.

There was also a significant correlation between the L2 WM capacity and the L1 WM capacity ($r=.49$, $p<.01$), and L2 STM capacity and L1 STM capacity ($r=.43$, $p<.01$). L1 WM capacity significantly correlated with both L1 STM capacity ($r=.25$, $p<.01$) and L2 STM capacity ($r=.25$, $p<.01$), whereas L2 WM capacity significantly correlated with only L2 STM capacity ($r=.19$, $p<.05$). The relationship between WM capacity and STM capacity, which was expected given that LST included a storage component, was found to be comparatively weak.

Thirdly, both L2 inferential and literal comprehension significantly formed strong correlations with L2 listening comprehension measured with TOEIC listening test.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Correlations among the variables ($N=149$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>(1) TOEIC Listening test</td>
<td>---</td>
</tr>
<tr>
<td>(2) Inferential comprehension</td>
<td>.758**</td>
</tr>
<tr>
<td>(3) Literal comprehension</td>
<td>.699**</td>
</tr>
<tr>
<td>(4) L1WM</td>
<td>.191*</td>
</tr>
<tr>
<td>(5) L2WM</td>
<td>.428**</td>
</tr>
<tr>
<td>(6) L1STM</td>
<td>.062</td>
</tr>
<tr>
<td>(7) L2STM</td>
<td>.075</td>
</tr>
</tbody>
</table>

*Note.* $^*p<.05$, $^{**}p<.01$
5.1.3 Hierarchical Stepwise Regression Analysis

The results of the hierarchical stepwise regression analysis for predicting the scores on the inferential and literal listening comprehension tests are shown in Table 3 and Table 4, respectively. In this analysis, L1 and L2 STM capacity were entered in Step 1, and L1 and L2 WM capacity were entered in Step 2 with stepwise method. In the case of the inferential listening comprehension, L2 STM capacity ($\beta = .16, p < .05$) was a positive predictor of the model ($F = 3.937, p < .05$, Adjusted $R^2 = .019$) in Step 1. In Step 2, L2 WM capacity, included in the equation model, accounted for a significant additional 14.1% of the variance in the model ($F = 14.609, p < .001$, Adjusted $R^2 = .155$). With the addition of L2 WM capacity in the equation in Step 2, the standard partial regression coefficients showed that L2 WM capacity ($\beta = .38, p < .001$) made the only unique contribution to L2 inferential listening comprehension, as shown in Table 3. In the case of literal listening comprehension, after L2 STM capacity was entered in the equation in Step 1, L2 WM capacity accounted for a significant additional 10.8% of the variance in Step 2. The standard partial regression coefficients show that L2 WM capacity ($\beta = .34, p < .001$) was the only unique predictor of the model ($F = 12.987, p < .001$, Adjusted $R^2 = .139$), as shown in Table 4.

Table 3
Hierarchical stepwise regression predicting inferential listening comprehension (n=149)

<table>
<thead>
<tr>
<th>Block</th>
<th>Predictors</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L2 STM</td>
<td>.161*</td>
<td>.026</td>
<td>.026</td>
<td>3.937*</td>
</tr>
<tr>
<td>2</td>
<td>L2 STM</td>
<td>.087</td>
<td>.167</td>
<td>.141</td>
<td>24.649***</td>
</tr>
<tr>
<td></td>
<td>L2 WM</td>
<td>.382***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *$p < .05$, ***$p < .001$

Table 4
Hierarchical stepwise regression predicting literal listening comprehension (n=149)

<table>
<thead>
<tr>
<th>Block</th>
<th>Predictors</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L2 STM</td>
<td>.207*</td>
<td>.043</td>
<td>.043</td>
<td>6.557*</td>
</tr>
<tr>
<td>2</td>
<td>L2 STM</td>
<td>.142</td>
<td>.151</td>
<td>.108</td>
<td>18.630***</td>
</tr>
<tr>
<td></td>
<td>L2 WM</td>
<td>.335***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *$p < .05$, ***$p < .001$
SATORI, Miki

5.2 Results for RQ 2

5.2.1 Descriptive Statistics

To examine the influence of L2 proficiency on the role of WM in L2 listening, the participants were divided into two groups according to their performance on the scores of TOEIC listening test. The top 50 participants who scored 300 or above were regarded as the higher-level group and the bottom 49 participants who scored 220 or below were regarded as the lower-level group. The descriptive statistics for each group are shown in Table 5 and Table 6, respectively.

Table 5

*Descriptive statistics for the higher-level group*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEIC Listening test</td>
<td>372.50(75.3%)</td>
<td>50.77</td>
<td>50</td>
</tr>
<tr>
<td>Inferential comprehension</td>
<td>17.42(72.6%)</td>
<td>3.15</td>
<td>50</td>
</tr>
<tr>
<td>Literal comprehension</td>
<td>18.22(75.9%)</td>
<td>2.44</td>
<td>50</td>
</tr>
<tr>
<td>L1 WM</td>
<td>36.10(86.1%)</td>
<td>4.40</td>
<td>50</td>
</tr>
<tr>
<td>L2 WM</td>
<td>29.38(70.0%)</td>
<td>5.82</td>
<td>50</td>
</tr>
<tr>
<td>L1 STM</td>
<td>5.76(36.0%)</td>
<td>2.75</td>
<td>50</td>
</tr>
<tr>
<td>L2 STM</td>
<td>8.10(50.6%)</td>
<td>3.01</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 6

*Descriptive statistics for the lower-level group*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEIC Listening test</td>
<td>187.55(37.9%)</td>
<td>29.32</td>
<td>49</td>
</tr>
<tr>
<td>Inferential comprehension</td>
<td>10.10(42.1%)</td>
<td>2.66</td>
<td>49</td>
</tr>
<tr>
<td>Literal comprehension</td>
<td>11.78(49.1%)</td>
<td>3.57</td>
<td>49</td>
</tr>
<tr>
<td>L1 WM</td>
<td>34.10(81.2%)</td>
<td>7.04</td>
<td>49</td>
</tr>
<tr>
<td>L2 WM</td>
<td>23.29(55.5%)</td>
<td>5.21</td>
<td>49</td>
</tr>
<tr>
<td>L1 STM</td>
<td>5.65(35.3%)</td>
<td>3.00</td>
<td>49</td>
</tr>
<tr>
<td>L2 STM</td>
<td>7.65(47.8%)</td>
<td>3.49</td>
<td>49</td>
</tr>
</tbody>
</table>

5.2.2 Correlation Analysis

The correlations among the variables for each proficiency group are shown in Tables 7 and 8. There were found to be some marked differences between the two sets of correlations. Most notably, the correlations between L1 and L2 WM capacity, and L2 listening comprehension in the lower-level group were greater than the corresponding correlations in the higher-level group. Similar patterns emerged for L2
The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

literal comprehension. On the other hand, the correlation between L2 WM capacity and L2 inferential comprehension in the higher-level group was greater than the corresponding correlation in the lower-level group, whereas the correlation between L1 WM capacity and L2 inferential comprehension in the lower-level group was greater than the corresponding correlation in the higher-level group. In the higher-level group, both L1 and L2 STM capacity did not have a significant correlation with any L2 listening comprehension. On the other hand, L2 STM capacity significantly correlated with only L2 literal comprehension in the lower-level group.

Table 7

*Correlations among the variables for the higher-level group (N=50)*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) TOEIC Listening test</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(2) Inferential comprehension</td>
<td>.490**</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(3) Literal comprehension</td>
<td>.320*</td>
<td>.635**</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(4) L1 WM</td>
<td>.030</td>
<td>.132</td>
<td>-.169</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(5) L2 WM</td>
<td>.182</td>
<td>.349*</td>
<td>.003</td>
<td>.333*</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(6) L1 STM</td>
<td>.003</td>
<td>.210</td>
<td>.111</td>
<td>.177</td>
<td>.068</td>
<td>---</td>
</tr>
<tr>
<td>(7) L2 STM</td>
<td>.089</td>
<td>.251</td>
<td>.136</td>
<td>.324*</td>
<td>.195</td>
<td>.357*</td>
</tr>
</tbody>
</table>

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

Table 8

*Correlations among the variables for the lower-level group (N=49)*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) TOEIC listening test</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(2) Inferential comprehension</td>
<td>.278</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(3) Literal comprehension</td>
<td>.349*</td>
<td>.348*</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(4) L1 WM</td>
<td>.408**</td>
<td>.387**</td>
<td>.422**</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(5) L2 WM</td>
<td>.367**</td>
<td>.190</td>
<td>.374**</td>
<td>.354*</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(6) L1 STM</td>
<td>.301*</td>
<td>.219</td>
<td>.281</td>
<td>.292*</td>
<td>.221</td>
<td>---</td>
</tr>
<tr>
<td>(7) L2 STM</td>
<td>.184</td>
<td>.257</td>
<td>.392**</td>
<td>.457**</td>
<td>.274</td>
<td>.748**</td>
</tr>
</tbody>
</table>

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

5.2.3 Hierarchical stepwise regression analysis

In the case of the higher-level group, the results of the hierarchical stepwise regression analysis yielded a significant model ($F = 6.643, p < .05$, Adjusted $R^2 = .103$)
predicting inferential comprehension, as shown in Table 9. As seen in Table 9, L2 WM capacity was the significant predictor of L2 inferential comprehension, and accounted for a significant 12.2% of the variance in the model. Neither L1 nor L2 STM capacity was predictive of L2 inferential comprehension. On the other hand, none of the memory variables were significantly predictive of L2 literal comprehension.

Table 9

Hierarchical stepwise regression predicting inferential listening comprehension for the higher-level group (n=50)

<table>
<thead>
<tr>
<th>Block</th>
<th>Predictors</th>
<th>β</th>
<th>R²</th>
<th>R² change</th>
<th>F change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L2 WM</td>
<td>.349</td>
<td>.122</td>
<td>.122</td>
<td>6.643*</td>
</tr>
</tbody>
</table>

Note. *p < .05

In the case of the lower-level group, the results of the hierarchical stepwise regression analysis are shown in Tables 10 and 11.

Table 10

Hierarchical stepwise regression predicting inferential listening comprehension for the lower-level group (n=49)

<table>
<thead>
<tr>
<th>Block</th>
<th>Predictors</th>
<th>β</th>
<th>R²</th>
<th>R² change</th>
<th>F change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L1 WM</td>
<td>.387**</td>
<td>.150</td>
<td>.150</td>
<td>8.283**</td>
</tr>
</tbody>
</table>

Note. ** p < .001

Table 11

Hierarchical stepwise regression predicting literal listening comprehension for the lower-level group (n=49)

<table>
<thead>
<tr>
<th>Block</th>
<th>Predictors</th>
<th>β</th>
<th>R²</th>
<th>R² change</th>
<th>F change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L2 STM</td>
<td>.392**</td>
<td>.154</td>
<td>.154</td>
<td>8.526**</td>
</tr>
<tr>
<td>2</td>
<td>L2 STM</td>
<td>.313*</td>
<td>.230</td>
<td>.077</td>
<td>4.591*</td>
</tr>
<tr>
<td></td>
<td>L2 WM</td>
<td>.288*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

As seen in Table 10, L1 WM capacity was the only significant predictor of L2 inferential comprehension, and accounted for a significant 15.0% of the variance in the model (F = 8.283, p < .01, Adjusted R² = .132). In the case of literal comprehension, L2
The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

STM capacity was a positive predictor of the performance in Step 1, and accounted for a significant 15.4% of the variance in the model \( F = 8.526, p < .01, \text{Adjusted } R^2 = .136 \), as shown in Table 11. In Step 2, L2 WM capacity, included in the equation model, accounted for a significant additional 7.7% of the variance in the model \( F = 6.884, p < .001, \text{Adjusted } R^2 = .197 \). The standard partial regression coefficients show that L2 STM capacity \( (\beta = .31, p < .05) \) made slightly greater contribution to L2 literal listening comprehension than L2 WM capacity \( (\beta = .29, p < .05) \). In combination, L2 STM capacity and L2 WM capacity accounted for a significant 23.0% of the variance.

6. Discussion

The study examined the role of L1 and L2 WM in literal and inferential comprehension in L2 listening.

The RQ 1 addressed the effects of individual differences in WM capacity on literal and inferential comprehension in L2 listening. Contrary to the expectation that WM capacity would play a greater role in inferential comprehension than in literal comprehension, the results suggested that individual differences in WM capacity influenced both literal and inferential comprehension in L2 listening. High-span listeners, with more working memory capacity, are thought to retain more information and revise prior information as they listen more effectively than the low-span listeners. As a result, the high-span listeners possibly showed significantly greater performance in both literal and inferential comprehension in L2 listening than the low-span listeners. The results of the present study are consistent with the previous study conducted by Alptekin & Erçetin (2010) showing that WM plays an important role in inferential comprehension in L2 reading with highly proficient L2 readers, but contradict their result that literal comprehension in L2 reading did not correlate with WM capacity. In the present study, WM capacity was also found to be related to literal comprehension in L2 listening. Directing the attention to the surface level information and keeping it active in phonological loop, and integrating or revising the information retrieved from long-term memory while listening seem to be both held in central executive resources in WM. One reason why literal listening comprehension also correlated with WM capacity would be that literal comprehension, as a data-driven process, relies on the level of language proficiency and surface level decoding or syntactic parsing skills (Alptekin & Erçetin, 2010).

The discussion now turns to RQ 2 concerning how the role of WM in literal and inferential comprehension in L2 listening differs across L2 proficiency levels. To this
end, the skilled and less-skilled listeners were compared. In the case of the skilled listeners, the results of the correlation and regression analysis showed that L2 WM capacity significantly formed a medium effect on L2 inferential comprehension and accounted for a significant 12.2% of the variance in the model. Literal comprehension, however, did not correlate with any WM capacity measures. The results are in line with the findings of the studies on the contribution of WM to L2 reading comprehension sub-skills (Alptekin & Erçetin, 2010; Van den Noort et al., 2006) and suggested that literal comprehension, a data-driven processing related to the surface level features, may not have been cognitively challenging to their WM capacity. It is quite likely that the skilled listeners rely less on decoding or syntactic parsing processing, because their lower-level processing skills are more automatized. Therefore, literal comprehension, a data-driven processing did not impose on WM capacity a heavy cognitive load. On the other hand, in the case of the less-skilled listeners, both L1 and L2 WM capacity formed a moderate correlation with both inferential and literal comprehension in L2 listening. The regression analysis showed that L1 WM capacity predicted a significant 15.0% of the unique variance in the model of L2 inferential comprehension, and L2 STM capacity and L2 WM capacity accounted for a significant 23.0% of the variance in the model of literal comprehension. The less-skilled listeners possibly focused on scratching surface level information and keeping it active in STM in literal comprehension, because they lack automatized linguistic decoding skills. Individual differences in both storage STM capacity and processing capacity of WM, therefore, may have influenced their L2 literal comprehension performance. The less-skilled listeners are thought to rely on cognitive resources for higher-level processing as well as lower-level processing. In the case of the less-skilled listeners, both text-base representation and deeper-level mental representation seem to have placed cognitive load on WM capacity.

A final finding of the study is that both L1 and L2 WM capacity were related to L2 inferential comprehension performance. L1 and L2 WM capacity seem to share the same or similar cognitive resources utilized in inferential comprehension rather than literal comprehension. According to the controlled attention theory of WM (Engle & Kane, 2004), WM capacity is domain-general, and independent of any one processing task. It is quite likely that high-span individuals outperform low-span individuals in L2 inferential comprehension because of the simple domain-general ability to control attention rather than the domain-specific ability to maintain task-relevant information.
The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

7. Conclusion

In sum, the results of the present study suggested that individual differences in L1 and L2 WM capacity affected both two dimensions of listening abilities: inferential comprehension and literal comprehension. Moreover, the results showed the different patterns of influence of L1 and L2 WM capacity on literal and inferential comprehension in L2 listening. Whereas L2 WM capacity was a significant predictor of L2 inferential listening in the case of the skilled listeners and L2 literal comprehension in the case of the lower level listeners, L1 WM capacity significantly predicted L2 inferential listening in the case of the less-skilled listeners. WM is thought to play an important role in L2 listening comprehension, especially in cognitively demanding listening tasks. The less-skilled listeners are forced to devote more cognitive resources to lower-level processing, because they have limited linguistic knowledge. In the case of the lower-level listeners with lower WM span, even literal listening comprehension may have been cognitively challenging, and required the listeners to allocate more WM resources to decode the surface level information, and leave fewer WM resources for inference-making to integrate the information. In the case of the less-skilled listeners, as the lower-level comprehension sub-skills such as decoding or syntactic parsing get more automatic, it may save the limited capacity of WM, thus more WM resources might be available for higher-level processing sub-skills.

The present study is not without its limitations. One relates to the nature of L1 LST. The performance of L1 LST was quite high overall, including ceiling performance by some of the participants. A restricted range of scores may have affected the pattern and strength of correlations that were observed. A future study may as well include L1 WM measures with greater processing demands. The second limitation concerns the validity of multiple-choice tests. Future work might compare the results on multiple-choice tests with the results on open-ended or short-answer listening comprehension measures. With respect to the role of WM in L2 listening comprehension, future investigation should be conducted in a future study on the potential difference in contribution of WM capacity to performance across multiple choice and open-ended listening comprehension measures.

References


The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening


The Role of L1 and L2 Working Memory in Literal and Inferential Comprehension in L2 Listening

*Psychology-Human Experimental Psychology, 49, 51–79.*

Appendices

Appendix A

Examples of test questions

*M: What did you do with the customer records I gave you an hour ago?*

*They were on my desk earlier this morning.*

*W: Oh, I took them to the copy room and copied them. I put them back on your desk.*

*M: Oh, good. I thought you might have sent them to the customers.*

*I still need to make some changes before I take them to the post office.*

Sample of a textually explicit question

What did the woman do with the papers?

A. She sent them to customers.
B. She put them in the trash.
C. She copied them.
D. She read them.

Sample of a textually implicit question

How does the man feel?

A. Very tired.
B. Disappointed.
C. Angry.
D. Relieved.

Appendix B

Examples of Japanese version of LST

2-sentence condition

2-1 右手でチョキを作り、左手でバーを作ると、折られている手の指は2本となる (F)
2-2 電話は音声を信号に変換して離れた相手に伝えるもので、携帯型もある (T)
Target words: 右手、電話

Examples of ESL version of LST

2-sentence condition
2-1 March is the month that is the third of the year. (T)
2-2 Mathematics is a subject in which students take lessons at a gym. (F)
Target words: year, gym

Appendix C

Examples of ESL version of DST

Stage 1
1-8-4-6-3

Stage 2
2-9-7-3-8-5

Examples of Japanese version of DST

Stage 1
6-2-9-4-8-1

Stage 2
9-3-5-2-8-6-1