Factors Related to Reading Span Test Results for Hearing-Impaired Students

Hirohito Chonan¹, * and Takashi Sawa²

¹ Research and Support Center on Higher Education for the Hearing and Visually Impaired, Tsukuba University of Technology, Japan
² Department of Special Needs Education, Tokyo Gakugei University, Japan

Working memory is crucial for a range of cognitive tasks, and the Reading Span Test is widely used to measure linguistic working memory. The present study examined links between linguistic memory scores from the Reading Span Test and results of subsystem tasks testing the phonological loop, central executive, and visuospatial sketchpad ability in 46 hearing-impaired university students. The Reading Span Test, Nonword Repetition Task, WMS-R Visual Memory Span Task, and New Stroop Test II were administered individually to participants. Reading Span Test performance was significantly associated with the Nonword Repetition Task score used to measure phonological loop capability, and with the Stroop interference ratio used to measure central executive capability. Results regarding phonological loop differed from those from studies conducted with normal-hearing students. Our findings suggest the importance of paying attention to phonological encoding when teaching reading to hearing-impaired children.

Key Words: hearing-impaired students, working memory, phonological loop, central executive

Introduction

Working memory refers to the memory's function of temporary storage and manipulation of information necessary for carrying out a range of cognitive tasks and the mechanisms and processes that allow implementation of this function (Baddeley, 1992). Working memory is assumed to comprise four subsystems: the phonological loop, which stores and processes linguistic information; the visuospatial sketchpad, which does the same with visual information; the episodic buffer, which integrates these two types of information, stores them, and compares them to long-term memory; and the central executive, which is responsible for securing and distributing psychological processing resources for performing tasks, judging old and new information, attention and control, changes in processing plans, and supervision of the status of task-processing (Baddeley, 2000).

Among the various tests of working memory, the Reading Span Test has been widely used to measure linguistic working memory. In this test, subjects are required to read aloud a presented passage and memorize target words that are underlined in the text. Texts are presented with two to five sentences, with five trials for each number of sentences. The test is a dual task, comprising the processing task of reading the text aloud and encoding the target words, and the storage task of memorizing the target words.

Individual differences in Reading Span Test performance in studies of participants with normal hearing have been taken to reflect differences in the capabilities of these subsystems and have been assumed to be related to functional differences between the phonological loop, which is involved in encoding and storing target words, and the central executive, which directs attention toward target words and suppresses information other than the target words (Miyake & Saito, 2001; Osaka, 2002). In a study of participants with normal hearing, Jinchou (2010) discovered that individual differences in Reading Span Test performance show little connection to differences in ability of the phonological loop, and instead reflect the abil-
ity of the central executive.

Investigation of the linguistic working memory of hearing-impaired students, who show a delay in the development of reading and writing compared to those with normal hearing, represents an issue in the education of the hearing impaired (Sawa, 2003). Research conducted with hearing-impaired individuals using the same Reading Span Test as with normal-hearing people have suggested that gaps in scores are related to reading scores (Daneman, Nemeth, Stainton, & Huelsmann, 1995; Garrison, Long, & Dowaliby, 1997; Hanson, Liberman, & Shankweiler, 1984; Nakamura, 2000; Sawa, 2011).

As far as studies of the phonological loop as a subsystem among hearing-impaired individuals are concerned, a study showing potential for storage and processing in the phonological loop of sign language words in memory tasks has been reported (Wilson & Emmorey, 1997, 1998), showing that function of the phonological loop among hearing-impaired individuals differs from that in people with normal hearing. Flaherty and Moran (2001) also conducted a study using kanji memory tasks. Based on results showing no greater articulatory suppression effect compared to students with normal hearing, the study suggested that hearing-impaired students may use visual information, applying the phonological loop less than students with normal hearing. In terms of research on the central executive, Flaherty and Moran (2007) conducted a Stroop test with hearing-impaired and normal-hearing university students, reporting from an analysis of reaction times that the Stroop interference ratio of hearing-impaired students was higher than that of students with normal hearing. In addition, Figueras, Edwards, and Langdon (2008) revealed a correlation between language ability and central executive in hearing-impaired students.

In this way, studies related to linguistic working memory among hearing-impaired individuals have reported findings related to links between Reading Span Test scores and reading scores, and the situation concerning subsystem abilities. However, the link between Reading Span Test scores and subsystem abilities has remained unexamined.

Within this context, the aim of the present study was to explain the links between linguistic memory scores as measured by the Reading Span Test and the results of subsystem tasks testing phonological loop, central executive, and visuospatial sketchpad abilities among hearing-impaired students. No correlation was found between phonological loop and Reading Span Test scores in research with people with normal hearing, but an investigation among hearing-impaired people seems warranted, as differences from people with normal hearing in the function of the phonological loop are evident, as described above. Furthermore, as the results of Flaherty and Moran (2001) suggest, hearing-impaired individuals may visually encode target words in the Reading Span Test, so this study also examined the visuospatial sketchpad as a subsystem. By clarifying these points, analysis of reading problems of hearing-impaired students in the process of storage and management of textual information becomes possible, making it possible to obtain data offering suggestions for methods of education. In addition, the present study included a Reading Span Test response error analysis as described by Osaka and Nishizaki (2000), and the results were used to investigate internal processes when carrying out the Reading Span Test.

**Methods**

**Participants**

A summary of the participants is given in Table 1. Participants comprised 46 university students with hearing impairments. Mean hearing ability of better hearing ear was 92.5 dB (HL). All participants were diagnosed with hearing impairments before entering school, 3 participants used cochlear implants, and all

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Subjects’ Attributes and Japanese Language Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of people</td>
<td>Sex</td>
</tr>
<tr>
<td>46</td>
<td>M 29</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

( ) shows standard deviation.
others used hearing aids on a daily basis. Everyday forms of communication were speech and sign language. Japanese language ability was measured using the Item Response Theory (IRT) Japanese Diagnostic Test (NTS, Tokyo, Japan). This test was developed for the purpose of testing Japanese language ability of students in higher education institutions. Japanese language proficiency of the candidate was calculated on the basis of the test score, and expressed as one of six levels from junior high school year 1 level to senior high school year 3 level, corresponding to the normal level of Japanese proficiency for students at that academic level. In the present study, mean Japanese language ability was senior high school year 1 level, which was not considered problematic for executing the Japanese Reading Span Test. The possibility exists that in hearing-impaired students, target words may include words that have not yet been learned, which would affect scores and errors from the Reading Span Test. For this reason, at the end of the test the researcher presented participants with a printout of the PowerPoint slides used in the Reading Span Test, and instructed the participants to underline any words that they were not sure how to read or what the meaning was, or words they had difficulty with. Participants underlined the corresponding words with a red ballpoint pen. Forty-six people reported no unknown words, and were included as participants for analysis in the present study. In addition, participants were asked verbally and by sign language if they had any visual impairment that would hinder execution of the test, and all participants confirmed that no such impairments were present.

Materials

The Reading Span Test, the Nonword Repetition Task, the Wechsler Memory Scale-Revised (WMS-R) Visual Memory Span Task, and the New Stroop Test II were administered individually to participants. Instructions were given to participants by sign language and verbally.

Reading Span Test. The Japanese version of the Reading Span Test of Osaka (2002) was used. The test comprised a practice task with four stimulus sentences and test tasks of two to five sentences, with five trials for each number of sentences. The method of presentation was the same as that described by Osaka and Osaka (1994), with the researcher presenting participants with stimulus sentences one at a time as PowerPoint slides on a computer monitor in front of the subject. The participant was instructed to read the stimulus sentences out loud, one sentence at a time, and to remember the underlined target word in the sentence. For each number of sentences, at the end of the five trials the computer screen showed a question mark, and the participant was then required to write from memory on an answer sheet the target words they remembered. With regard to the order of writing the words from memory, in accordance with Kondo and Osaka (2000), participants were not allowed to write the target word from the final stimulus sentence of the trial as the first word on the answer sheet, in order to avoid recency effect. Other than this, participants were free to write the words in any order they chose. No time limit was set for recall.

Nonword Repetition Task. Phonological loop function was measured using the Nonword Repetition Task, which has a strong positive correlation with Digit Span Test (Gathercole, Willis, and Baddeley, 1991) performance and has been widely used in recent years for measurement of the phonological loop (Yuzawa, 2008). In the present study, the 15 nonword tasks of Okumura (2011) were used. All nonword tasks were made up of four syllables, and the researcher presented each nonword verbally and by fingerspelling for approximately 1 s.

WMS-R Visual Memory Span Task. The Visual Memory Span Task of the WMS-R was used to measure the visuospatial sketchpad. Masunaga, Hamada, Kubota, Yotsumoto, Umemoto, and Yamasita (2006) confirmed that this test reflects the storage and processing of visual information.

In this test, participants were shown a card with eight squares drawn on it. The researcher pointed at the squares in order. In the same-order task, the participants were required to point at the squares in the same order as the researcher. In the reverse-order task, the participants pointed at the squares in the reverse order to that shown by the researcher. The number of squares pointed to was from two to eight in the same-order test and from two to seven in the reverse-order test, and two trials were carried out for each number of squares. The procedure of the same-order task thus complied with instructions for the test.

New Stroop Test II. The Stroop test was used for measurement of central executive function. In the Stroop test, the color and meaning of the stimulus
word sometimes coincide and sometimes do not, and reactions are faster when they coincide. This is believed to be because attention needs to be directed toward one of two attributes, and the other attribute needs to be controlled. The Stroop test thus measures the ability to control attention and interference, and this ability is regarded as equivalent to the executive function ability of attention and control. In the present study, the New Stroop Test II (Hakoda and Watanabe, 2005) was used to measure central executive function. This test comprised four tasks. In Task 1, participants were required to select the color patch corresponding to a color name written in black. In Task 2, color names were written in different colors to the name, and participants selected the color patch corresponding to the name. In Task 3, participants selected the name corresponding to the color patch. In Task 4, color names were written in different colors and participants selected the name corresponding to the color. Task 2 was a reverse Stroop interference condition, Task 4 was a Stroop interference condition, and Tasks 1 and 3 were control conditions.

**Method of Analysis**

*Reading Span Test scoring.* For scoring the Reading Span Test, total words, proportion words, correct sets words, and truncated span methods have all been proposed. Of these, scoring according to total words appears to be the most appropriate from the perspectives of reliability and normality (Friedman and Miyake, 2005). The total words method was therefore adopted for the present study, with 1 point given for one target word remembered, to a possible maximum of 70 points.

For classification of Reading Span Test errors, with reference to Osaka and Nishizaki (2000), Reading Span Test errors were classified into the following five categories: 1) no response, when the response form is left blank; 2) trial internal error, when a non-target word is mistakenly recalled from the stimulus sentence; 3) trial external error, when the participant mistakenly recalls a word that is present in a stimulus sentence shown either before or after the sentence being processed; or 4) insertion error, when a word that is not present in any stimulus sentence is recalled.

Classification of errors was carried out by taking into account the characteristics of the error that was written down and the participant’s self-report of the error. To check the reliability of the classification, evaluations were carried out independently by the author and by an educator specializing in the education of hearing-impaired children. The concordance rate was determined by calculating as a percentage proportion of the total number of errors (1,012 words) for which there was concordance in the classification. The concordance rate was 92.5%. Evaluators discussed the errors for which their classifications differed (76 words), and eventually two words over which judgments did not come to agreement were excluded from the analysis, leaving 1,010 words for analysis. In addition, insertion errors were classified on the basis of the features of the response word and the participant’s self-report into four subtypes: semantic errors, phonological errors, written character errors, and other errors.

*New Stroop Test II.* According to Hakoda and Watanabe (2005), the Stroop interference ratio indicates resistance to linguistic interference, and a lower value indicates less receptivity to interference. The reverse Stroop interference ratio indicates resistance to visual obstruction, and a lower value indicates...
Factors Related to Reading Span Test Results for Hearing-Impaired Students

less receptivity to interference. Osaka and Nishizaki (2000) used the Stroop interference ratio as an index of central executive function when students with normal hearing performed the Reading Span Test. The Stroop interference ratio was therefore also used in the present study. This value is calculated using the following formula, in accordance with the implementation guidelines for the Stroop interference Test:

\[
\text{Interference ratio} = \left( \frac{\text{no. of correct responses in task 3, control condition} - \text{no. of correct responses in task 4, interference condition}}{\text{no. of correct responses in task 3, control condition}} \right) \times 100
\]

Statistical analysis. Mean scores of the Reading Span Test, Nonword Repetition Task and WMS-R Visual Memory Span Task were calculated. Correlations between all variables were checked. Furthermore, multiple regression analyses were used to examine the effects of these variables on Reading Span Test scores.

Results

Table 2 shows mean score (M) and standard deviations (SD) for the Reading Span Test, Nonword Repetition Task, and WMS-R Visual Memory Span Task, mean Stroop interference ratio and standard deviation for the New Stroop Test II, and mean scores and standard deviation for the Japanese Diagnostic Test.

Results for each test were as follows: Reading Span Test, M=48.0 (SD=10.3); Nonword Repetition Task, M=11.8 (SD=2.1); WMS-R Visual Memory Span Task, M=20.3 (SD=2.8); Stroop interference ratio, M=7.1 (SD=10.4); Japanese Diagnostic Test, M=560.5 (SD=59.2).

Furthermore, no significant correlations were identified between Reading Span Test score and hearing ability of participants (r=.14). Next, a Shapiro–Wilk test of normality was conducted for the Reading Span Test scores. Consequently, the participants had p values of .96 of the Reading Span Test. Thus, the normality of participants was confirmed.

Table 3 shows Pearson correlation coefficients and confidence intervals using the Reading Span Test, Nonword Repetition Task, WMS-R Visual Memory Span Task interference ratio angular transformation scores, and Japanese Diagnostic Test scores. For Reading Span Test scores, significant correlations were evident except for IRT Japanese Diagnostic Test scores. A significant correlation between the Nonword Repetition Task and IRT Japanese Diagnostic Test scores was also identified.

Table 4 shows the results of multiple regression

<table>
<thead>
<tr>
<th>Table 2 Mean (SD) Score on Each Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Span Test</td>
</tr>
<tr>
<td>Nonword Repetition Task</td>
</tr>
<tr>
<td>Visual Memory Span Task</td>
</tr>
<tr>
<td>Stroop interference ratio</td>
</tr>
<tr>
<td>Japanese Diagnostic Test</td>
</tr>
</tbody>
</table>

| Table 3 Correlation Coefficients and 95% Confidence Intervals on Each Test |
|------------------------|-----------------|-------------|
| Reading Span Test      | Nonword Repetition Task | Visual Memory Span Task |
|                        | Stroop interference ratio |
| Reading Span Test      | 0.65**           | 0.31*       |
|                        | 0.31 ≤ ρ ≤ 0.72  | 0.02 ≤ ρ ≤ 0.55 |
| Nonword Repetition Task| 0.31*           | 0.32*       |
|                        | 0.04 ≤ ρ ≤ 0.56  | −0.38**     |
| Visual Memory Span Task| 0.31*           | 0.31*       |
|                        | −0.66 ≤ ρ ≤ −0.21| −0.6 ≤ ρ ≤ −0.1 |
| Stroop interference ratio| −0.47**        | −0.20       |
|                        | −0.66 ≤ ρ ≤ −0.21| −0.46 ≤ ρ ≤ 0.095 |
| Japanese Diagnostic Test| 0.22           | 0.30*       |
|                        | 0.08            | 0.02       |
|                        | 0.08 ≤ ρ ≤ 0.36  | −0.21 ≤ ρ ≤ 0.36 |
|                        | −0.32 ≤ ρ ≤ 0.26 |

* p<.05, ** p<.01.
The upper row: correlation coefficients.
The lower row: confidence intervals.
analysis using Reading Span Test scores as the objective variable and Nonword Repetition Task scores, WMS-R Visual Memory Span Task scores and Stroop interference angular transformation scores, which all showed significant correlations with Reading Span Test scores, as explanatory variables, as well as Japanese Diagnostic Test scores, which showed significant correlation with Nonword Repetition Task scores. Results of analysis showed the explanatory ratio to be 44.6%, with the explanatory ratio verified to be significant at the 1% level \( (F(4.41)=10.1, p<.01) \).

The standard partial regression coefficient showed a significant coefficient between the Nonword Repetition Task and Stroop interference ratio. Accordingly, this shows that participants with high scores on the Reading Span Test had high scores on the Nonword Repetition Task and low Stroop interference ratios.

Table 5 shows the number of response errors, proportion of total errors, \( M \) and \( SD \).

The greatest number of errors was seen for non-response errors, totaling 684, constituting 67.6% of all errors. Next came 99 trial errors (9.8%), followed by 76 insertion errors (7.5%), 66 non-trial errors (6.5%), and 85 other errors (8.4%).

Insertion errors were classified on the basis of the features of the response word and the participant’s self-report into four subtypes: 40 semantic errors (52.6%), 24 phonological errors (31.5%), 8 written character errors (10.5%), and 4 other errors (5.2%). Semantic errors are errors in which the target word and inserted error show a semantic similarity. An example of such an error that was seen was mistakenly recalling “letter” when the target word was “postcard.” Phonological errors are errors in which the target word and the mistakenly recalled word show phonological similarity. For example, the error was seen in which “shikkuri to” [nicely] was mistakenly recalled in place of the target word “shikkari to” [firmly]. With regard to phonological-type insertion errors, participants self-reported forgetting the target word and only being able to recall part of it, and responding with an already known word containing the remembered part (for example, the subject mistakenly responded “happa” [leaf] for the target word “hagaki” [postcard]). Written character errors are errors in which the target word and the mistakenly recalled word are written with similar characters. For example, an error was seen in which the target word “suigin” [水銀, mercury] was mistakenly recalled as “kiroku.” [木録]. In the self-report, the participant had forgotten the target word and instead wrote a word with characters that showed a similar overall shape to the part of the characters of the target word remaining in the participant’s memory and to a vaguely remembered word.

### Discussion

This study investigated the relationship between the results of working memory subsystems tests and Reading Span Test performance among hearing-impaired students. The results show that Reading Span Test performance was significantly associated with the Nonword Repetition Task score used to measure phonological loop capability, and with the Stroop interference ratio used to measure central executive capability. From this, higher phonological...
loop and central executive capabilities appear associated with higher Reading Span Test performance. Visuospatial sketchpad capability was not involved in Reading Span Test performance.

These results differ from those of a prior study of students with normal hearing, which found no association between Reading Span Test performance and phonological loop capability (Jinchou, 2010). One possible reason is that encoding methods used by hearing-impaired individuals are more diverse than those of normal-hearing individuals.

This study found cases of textual errors in insertion errors, such as mistakenly recalling the target word “suigin” (水銀, mercury) as “kiroku” (木録), whereas a study of normal-hearing students by Osaka and Nishizaki (2000) did not report any examples of mistakes like “kiroku”, where the phonology was completely different, meanings were unrelated, and only the shapes of characters were similar. Reflective reports on these incorrect responses reveal the existence of students who demonstrated insufficient phonological encoding of the target word, relying instead on visual similarities of the text in their responses. This is similar to the research findings of Flaherty and Moran (2001), who pointed out that hearing-impaired children do not use the phonological loop as much as normal-hearing children, instead using visual information in linguistic working memory tasks. In addition, this study did not show a connection between Reading Span Test scores and visuospatial sketchpad capability, suggesting that skills in the storage and processing of visual information do not promote memory of the target word under this experimental condition. Accordingly, phonological loop capability seems to demonstrate a link with Reading Span Test scores. However, in this study, we could not define the link between the diversity of encoding seen among hearing-impaired students and phonological loop capability. This represents an issue that needs to be researched in greater depth after clarifying memorization strategies used while undertaking the Reading Span Test.

In addition, tasks in the Nonword Repetition Task, which was used in the present study to measure the phonological loop, were presented using fingerspelling. This means that participants may have used fingerspelling in processing the task. A link is also known to exist between the hearing ability of hearing-impaired individuals and smoothness of articulatory movement (Yoshino, 1988), and individual differences in articulatory movement ability from the perspective of hearing ability level can be conjectured to have also manifested in this study. Nonword repetition task scores may thus be affected by articulation ability. Given this situation, careful discussion is needed regarding whether this task measures only phonological loop capability. Further studies that divide hearing-impaired students into groups according to factors associated with phonological loop function are warranted.

The Stroop test interference ratio may involve the central executive selective focus and suppression capability required to select target words in the Reading Span Test. The results of this study were consistent with those of Jinchou (2010), whose study involved normal-hearing individuals. In addition, the fact that the Stroop interference ratio was generally similar to that in normal-hearing individuals suggests that Stroop test results measure central executive selective attention and inhibition function. However, the central executive psychological processes of hearing-impaired individuals cannot be claimed to have been sufficiently investigated in this study, and further examination of similarities and differences between normal-hearing and hearing-impaired individuals in this regard is necessary.

Analysis of types of errors showed the same categories in the responses of hearing-impaired and normal-hearing students, with the exception of textual errors, seen as insertion errors in the responses of hearing-impaired students. These results show the need for further investigation of similarities and differences between normal-hearing and hearing-impaired individuals in terms of the psychological processes underlying execution of the Reading Span Test. In terms of each error, the most frequent type of error in the present study was “no response”. The reason for this is unclear. In a study of normal-hearing participants, Jinchou (2010) found that many different types of error in the Reading Span Test showed correlations to language ability, and whether this remains true for individuals with hearing impairment merits examination. Specifically, there is a need to divide hearing-impaired individuals into groups with high and low linguistic ability, compare errors between these groups, and thereby clarify differences among hearing-impaired people in the psychological process of executing the Reading Span Test, and simi-
larities and differences compared to normal-hearing individuals. In addition, the hearing-impaired students who participated in the present study can be regarded as a group with a relatively higher level of Japanese language ability for hearing-impaired students, as shown by the results of the IRT Japanese Diagnostic Test. According to the study by Jinchou (2010) into the relationship between linguistic ability and Reading Span Test scores among normal-hearing individuals, those with low linguistic ability consume substantial mental resources in processing trial sentences, leading to lower Reading Span Test scores. Whether the same applies to hearing-impaired individuals may need to be examined. Investigation may also be required into whether mental processing during the Reading Span Test is the same between normal-hearing individuals and those with low linguistic ability, including participants in this study.

In summary, this study investigated factors linked to Reading Span Test scores among hearing-impaired students. The results demonstrated that higher phonological loop and central executive capabilities were associated with higher Reading Span Test performance, and that results regarding the phonological loop differed from those of studies conducted in normal-hearing individuals. Our findings suggest the importance of emphasizing the phonological encoding of text when teaching reading to hearing-impaired children. In terms of phonological loop encoding, research to clarify memory strategies has been conducted in normal-hearing students (Endo & Osaka, 2011), and the same kind of study needs to be conducted with hearing-impaired students in the future.

In this study, no link between visuospatial sketchpad abilities and Reading Span Test scores was evident, but errors were observed in items with no phonological or semantic resemblance, but with similarities only in character shapes, suggesting the need for further studies to examine this conclusion, using diverse tasks such as the mental rotation task to measure visuospatial sketchpad abilities.

Consequently, there is a need to clarify whether the results of this study are valid for hearing-impaired individuals in general by carrying out studies that cover participants with a wider range of Japanese language abilities.

References

Factors Related to Reading Span Test Results for Hearing-Impaired Students


