What First-Sight Oral Reading Performed by Japanese EFL Learners Implies:
General English Proficiency or Working Memory Capacity?

Takue OHNO
Graduate School, Hokkaido University

Abstract

This research is an attempt to examine the possibility of using oral reading as a diagnostic approach to learners’ general English proficiency. Oral reading, here, is confined to ‘first-sight’ oral reading, and is evaluated in terms of oral reading rate and comprehension tests. In the experiment, participants were required to read passages aloud without any preceding time for comprehension. Two conditions were prepared for the oral reading: one was to read orally focusing on the reading speed, and the other focusing on the comprehension. As this task is deeply related to working memory in that it contains dual-task procedure, the primary purpose of this study was to investigate which of the two variables, learners’ general English proficiency or working memory capacity, would be more reflected in the task performance. Eighty Japanese EFL learners were tested and the data was analyzed firstly based on the group made up by STEP score, and secondly on the group by RST score through 2(High and Low groups)×2(oral reading conditions)×3(difficulty levels of the texts) ANOVA. The result was that the learners’ oral reading has reflected their general English proficiency more than their working memory capacity. Regression analysis was also conducted with STEP score (dependent variable) and oral reading rate, comprehension score (independent variables). The result was that comprehension score of the oral reading focusing on the meaning has reflected learners’ English proficiency best in this study.

1. Introduction

“Oral reading performed by EFL learners can provide fairly good hints as to their general English proficiency.” This is a true feeling shared, though intuitively, by many English teachers, and in
fact it is commonly used as a sort of testing method in everyday English classrooms or even in official English certification tests. However, in most of these cases, oral reading is required only after a period of comprehension and is evaluated in terms of pronunciation, intonation, phrasing, fluency, delivery and the like by examiners’ rather subjective impressions (although based on specific criteria). As the matter of course, the learners’ oral reading should be delivered corresponding to those judgmental factors. In short, the oral reading tests like above could be regarded as an indirect and supplemental assessment after the main tests of pencil-and-paper style rather than a test in its own right.

In contrast to the oral readings described above (hereafter referred to as ‘traditional oral reading’), the first-sight oral reading in this study requires that the readers should both decode letter strings phonologically in order of their appearance and understand the text they have never seen almost simultaneously. Readers are prohibited from retracing or pausing for reflection and they are also supposed to take a small scale comprehension test immediately after the reading. The assessment of the oral reading performance is not impressionistic but more objective by adopting an oral reading rate and comprehension score as two measures. It is quite obvious that this type of oral reading imposes far more of a burden on readers than a traditional one does, because it consists of dual-task procedure, in which participants have to carry out two tasks simultaneously (decoding and understanding in this study), and consequently, doing the task successfully requires higher English proficiency. However, owing to the nature of the task, influence of individual differences of working memory capacity cannot be disregarded. In order for the oral reading in this study to be a possible approach to know learners’ English proficiency, it should be verified that the task is not so much related to working memory capacity but to English proficiency. If the oral reading in this study reflects learners’ working memory capacity rather than their English proficiency, the task has to be regarded as a test for working memory capacity, dismissing the possibility of using it as an English proficiency test. Otherwise, it would survive as one of the candidates for representing learners’ English proficiency. Thus, this is a kind of pilot study to open the way for applying oral reading to evaluating directly and objectively the learners’ English proficiency.

2. Background

2.1 Working memory

Working memory represents a human memory system that is assumed to be responsible for both processing and storage of the information at the same time within a rather limited short period of time (e.g. Baddeley, 1986). Although working memory is related to most of human cognitive activities in every day life, its function can be easily exemplified, say, when thinking of reading. In reading, “incoming information is decoded perceptually, recognized, and integrated with the textual interpretation through the use of syntax, pragmatics, and semantic information while the products of these processes are being stored for a short period (Osaka & Osaka, 1992: 287).”
For example, it is working memory that enables us to interpret what a certain deictic indicates by retaining old information activated and using it on immediate processing. Or, when a reader encounters an ambiguous word in a sentence, (s)he may be able to process continued sentences, holding some possible interpretations of the ambiguous word, until incoming new information settles the most plausible interpretation. In short, thanks to working memory, we are able to do quite complicated tasks with the storing and processing being managed in a parallel manner.

Though there is much to say about working memory, two things should be emphasized here, which seems to be relevant to this study. Firstly, working memory capacity is thought to be limited. Therefore, working memory operates within the relationship of the trade-off between processing and storage functions. Imagine when you are reading very difficult passages. The large part of the working memory capacity is occupied for processing, so you will find it difficult to trace the story line or remember the gist of the story. Consequently, forgetting, regression of eye movement, and/or reduction of reading rate may occur. If this trade-off concept is applied to the oral reading in this study, and if a reader has high English proficiency (this entails that (s)he can take advantage of rather automatized bottom-up processing), (s)he can afford to use the more working memory capacity for storage. Secondly, researches have revealed that the trade-off between processing and storage can be thought as a potential source of individual differences in reading comprehension. In connection with this, reading span tests (RST) have been developed to measure working memory capacity, based on the trade-off in working memory system.

2.2 Reading Span Test (RST)

RST was developed by Daneman and Carpenter (1980) in order to measure working memory capacity. The original version was in English for the subjects whose first language (L1) was English. This test has been said to show much closer relation to reading comprehension ability than the traditional digit span or word span measures do.

The original version of RST is composed of 60 unrelated sentences, with 13 to 16 words each. Each sentence, which ends in a different word, is typed on an index card. The cards are arranged as follows: if it is a two-sentence testing condition, there are three sets of two sentences. Under the three-sentence condition, there are also three sets of three sentences. Likewise, under the four-sentence condition, there are three sets of four sentences. This continues until a six-sentence condition, so it makes 60 sentences in total. A blank card is inserted to mark the end of each set. The subject is shown one card at a time and asked to read the sentence aloud. The next card is shown as soon as the subject reads the sentence aloud. When a blank card appears, the subject has to recall the last word of each sentence. For example, under the five-sentence condition, the subject reads, per set, five sentences one by one, and then recall the last words of each sentence (s)he has just read (in this case 5 words). Since each sentence-condition has three sets of trials, this subject is expected to recall 15 words in all through three sets of trials under a five-sentence condition. The level of sentence-condition under which a subject is correct on two out of three sets is re-
gadoed as his/her score of reading span. During RST, reading aloud functions as an articulatory suppression. This means that the subject cannot rely on the subvocal rehearsal system in trying to remember the target words, for reading aloud interferes with subvocal rehearsal.

There have been some attempts to investigate the relationship between working memory and reading comprehension ability in a second (or foreign) language (L2) (e.g. Harrington & Sawyer, 1992; Osaka & Osaka, 1992; Osaka, Osaka & Groner, 2000). The findings relevant to this study is that working memory capacity and reading comprehension ability have close correlation to each other even in L2 and that working memory for reading appears to be independent of the language used in the test as long as the subject is an advanced level learner of the language.

3. Research Questions

As mentioned so far, first-sight oral reading tasks in this study, which are to examine how much comprehension will accompany oral reading, will inevitably be influenced by the readers’ English proficiency. At the same time, however, because of the nature of the tasks, it is possible that the influence of individual differences in working memory capacity will exceed that of English proficiency on the tasks. In terms of whether oral reading in this study can be an indicator of learners’ English proficiency, the comparison between the influence of English proficiency and working memory capacity on the task performance should be made. In terms of the trade-off in working memory capacity, two types of oral reading (focusing on reading speed or reading comprehension) are to be compared. Thus the research questions are the following:

1. Can first-sight oral reading be an indicator of learners’ English proficiency rather than an indicator of working memory capacity?
2. If it is the case, which type of oral reading, and which of the two, reading rate or comprehension score, will best represent learners’ English proficiency?

Here ‘English proficiency’ refers to the general English proficiency which can be measured by widely accepted English certification tests such as the one administered by the Society of Testing English Proficiency Inc. (hereafter referred to as ‘STEP test’).

4. Method

4.1 Participants

Eighty second-year students in a 5-year national college of art participated in the experiment. Although they are the same age as the second-year students in regular senior high schools, they are quite different in terms of English learning situations. For example, they have only one 90-minute English class and two 45-minute English conversation classes per week. In addition to this, they seldom need to prepare for the entrance examination for the universities.
4.2 Materials
(1) Working memory capacity: Japanese version\(^2\) of Reading Span Test (Osaka, 2002) was used. This was basically the same as the one described in 2.2 except for three points. First, it was in Japanese. Second, target words were not always placed at the end but were found at any position within each sentence, and they are underlined. Third, there were 70 sentences in all, distributed into five sets of two-, three-, four-, and five-sentence condition. Consequently, the level of sentence-condition under which a subject was correct on three out of five sets was regarded as his/her score of reading span. If a subject was correct on only two out of five sets, 0.5 point was added.
(2) General English proficiency: Pre-2-grade STEP test was adopted. It consisted of four sections: grammar and vocabulary (full mark 25), reading comprehension (20), composition (5), and listening comprehension (20), totaling to 70 in a full mark. All the questions were multiple-choice.
(3) Oral reading competence: Test materials were invented by the author in the form of two separate booklets, for the test was to be administered on two separate occasions. Both booklets consisted of six passages, each of which was immediately followed by a comprehension test composed of 6 questions. Participants were required to read these passages aloud under two conditions: three passages were for oral reading focusing on the reading rate (hereafter referred to as ‘or-R’) and the other three passages for oral reading focusing on the comprehension of the texts (hereafter referred to as ‘or-C’). Additionally, these three each texts corresponded to three difficulty levels by choosing them from the exercise books for STEP 3-grade test, pre-2-grade and 2-grade test. They were modified by the author such that each passage consisted of approximately 100 words and had similar readability indices according to the difficulty levels, based mainly on the Dale-Chall Grade Level.

4.3 Procedure
Three kinds of tests were delivered on five separate days using regular English class times in the following order: RST\(^1\)→ STEP test→ RST\(^2\)→ oral reading test\(^1\)→ oral reading test\(^2\).
(1) RST was conducted in pairs though this type of test is usually administered in a one-to-one (examiner and examinee) style. Test materials were given to each pair in the form of a booklet. Within a pair, one was to take the test, and the other was to assist turning the pages as soon as the partner read each sentence. After sufficient instructions and practice trials, the test was started at the same time for every participant. The same test was given reversing the roles within the pairs. In order to avoid task familiarity, one week of interval was set in between. All these performances were tape recorded individually in the language laboratory. The scores were calculated by the author afterward.
(2) STEP pre-2-grade test was delivered as formally done by STEP Inc., with 65-minute writing section and 20-minute listening section. The scores were calculated by the author afterward.
(3) Considering participants’ fatigue and, of course, time constraint, the oral reading test was divided into two tests with one week of interval. In each test, participants were asked to read the
passages orally under the two conditions. One was or-R, in which participants were supposed to read as fast as possible at the cost of comprehension, and the other was or-C, in which participants were supposed to comprehend as much as possible at the cost of reading speed. The test was done altogether individually on the language lab booth with all of their performances tape recorded. Every time they read aloud one passage, they turned the page and answered the comprehension test, and again moved on to the next page for the new passage. This pattern was repeated three times under each of the condition. Participants were allowed neither to retrace the passages nor to pause to ponder under both of the conditions. The order of reading condition was counterbalanced within subject, and the order of the passages was also counterbalanced between subjects. After the test, the author collected the individual reading rate on each passage by the number of the syllable uttered per minute, and checked the comprehension test, with full mark 6 per each passage.

4.4 Data analysis

The data of the participants who did not take some of the tests, which were carried out on five separate days, were excluded. The data with some defaults were also eliminated. Thus, the final number of the participants available was 77, 12 males and 65 females. Age range was 17 to 18 with one exception of 19 years old. First, STEP group was made up according to the STEP test score. It consisted of 22 high-scored (H-STEP subgroup) and 22 low-scored samples (L-STEP subgroup) chosen from the highest and the lowest, respectively. Then, RST group was made up according to the RST score. It consisted of 22 high-scored (H-RST subgroup) and 22 low-scored samples (L-RST subgroup) chosen from the highest and the lowest, respectively. The data of reading rate and comprehension score were analyzed through 2(High and Low subgroups)×2(oral reading conditions)×3(difficulty levels of the texts) ANOVA for both STEP group and RST group in order to compare the magnitude of influence of English proficiency and working memory capacity. Next, based on the entire data of 77 participants', stepwise regression analysis was conducted with STEP score being a dependent variable, and reading rate and comprehension score being independent variables in order to detect the largest explanatory variate.

5. Results and Discussion

5.1 Results of analysis of variance

The results of ANOVA are shown in Figure 1 on the next page.

**Oral reading rate:** For the group made up by the STEP score, the main effect of the H-L group, the main effect of difficulty level of the texts, and the main effect of the reading condition were significant (F(1, 42) = 21.11, p<.01; F(2, 84) = 140.81, p<.01; F(1, 42) = 60.71, p<.01, respectively). For the group made up by the RST score, the main effect of difficulty level of the texts, and the main effect of the reading condition were significant (F(2, 84) = 147.86, p<.01; F(1, 42) = 58.34, p<.01, respectively), whereas the main effect of H-L group was not significant. There were
Figure 1. Oral reading rate and Comprehension score under each reading condition.
no interactions between any factors.

**Comprehension score:** For the group made up by the STEP score, the main effect of H-L group, the main effect of difficulty level of the texts were significant \((F(1, 42) = 56.70, p<.01; F(2, 84) = 153.64, p<.01\), respectively). For the group made up by the RST score, the main effect of H-L group, the main effect of difficulty level of the texts, and the main effect of the reading conditions were significant \((F(1, 42) = 8.94, p<.01; F(2, 84) = 169.39, p<.01; F(1, 42) = 24.68, p<.01\), respectively). There were no interactions between any factors.

5.2 Results of regression analysis

As Table 1 shows, the largest explanatory variate was Comprehension score under or-C, with the second largest Reading rate under or-R and the third Comprehension score under or-R. The multiple regression equation was significant \((F(3, 73) = 29.74, p<.01\). Model collection by stepwise regression provided three significant models. According to model [1], 40.4% of English proficiency can be explained by Comprehension score under or-C. According to the model [2], 50.0% can be explained by Comprehension score under or-C and Reading rate under or-R.

### Table 1. Significant explanatory variates for English proficiency

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>B</th>
<th>(\beta)</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension score in or-C</td>
<td>.689</td>
<td>.344</td>
<td>3.338**</td>
<td>.001</td>
</tr>
<tr>
<td>Reading rate in or-R</td>
<td>.116</td>
<td>.313</td>
<td>3.688**</td>
<td>.000</td>
</tr>
<tr>
<td>Comprehension score in or-R</td>
<td>.579</td>
<td>.284</td>
<td>2.849**</td>
<td>.006</td>
</tr>
</tbody>
</table>

\[Y = 6.00 + 0.689 \times \text{Comprehension score in or-C} + 0.116 \times \text{Reading rate in or-R} + 0.579 \times \text{Comprehension score in or-R}, \quad R^2 = .550, \quad F(3,73)= 29.741** \]

**Table 2. Model Collection by stepwise regression analysis (Independent variable: English proficiency)**

<table>
<thead>
<tr>
<th>Model</th>
<th>Multiple correlation coefficient</th>
<th>Coefficient of multiple determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1] Comprehension score in or-C</td>
<td>.636 (.636)</td>
<td>.404 (.404)</td>
</tr>
<tr>
<td>[2] Comprehension score in or-C</td>
<td>.707 (.701)</td>
<td>.500 (.492)</td>
</tr>
<tr>
<td>Reading rate in or-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[3] Comprehension score in or-C</td>
<td>.742 (.740)</td>
<td>.550 (.548)</td>
</tr>
<tr>
<td>Reading rate in or-R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension score in or-R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The numbers in parentheses are the results when ‘Reading rate in or-R’ is substituted by ‘Reading rate in or-C’.

5.3 Discussion

**For Research Question 1:** From the results of ANOVA, STEP group showed a clearer distinction
between H-L subgroups than RST group did both on reading rate and comprehension score. This implies that oral reading rate and comprehension score in this study represent learners’ general English proficiency rather than their working memory capacity.

In terms of reading rate, the main effect of H-L group was significant only on the part of STEP group. This supports that the task reflected not working memory capacity but English proficiency, which is quite understandable because the automatized phonological coding is a part of English proficiency.

In terms of reading comprehension score, the fact that the main effect of reading condition was marginally significant on the part of STEP group (F(1, 42) = 3.50, p < .10) indicated that participants had read aloud with a similar degree of understanding regardless of the reading conditions. One interpretation of this result is that it might be English proficiency that yielded a similar degree of understanding. In other words, for instance, high English proficiency did not allow the readers to ignore the meaning of the passages even under or-R, where they were instructed not to be worried about the upcoming comprehension test. On the other hand, the main effect of reading condition on comprehension score was significant on the part of RST group as reported in 5.1. This means each oral reading condition was well reflected following the conditional instructions. Perhaps this indicates that working memory capacity might control the mode of oral readings following each condition required, that is to say, working memory contributes to differentiate the mode of oral reading.

To sum up, in the first-sight oral reading by the beginners or beginners to intermediate learners like in this study, the influence of working memory capacity does not exceed that of general English proficiency. Thus it can be said that the oral reading in this study has a possibility to be an indicator of learners’ English proficiency.

**For Research Question 2:** Based on the possibility that oral reading in this study can be an indicator of English proficiency, regression analysis was conducted to identify the largest explanatory variate. From the results, comprehension score under or-C was the one. It seems quite natural that comprehension score was larger explanatory variate than reading rate because the latter was only the reading speed attributed to learners’ automatized phonological coding in English, whereas the former was accompanied by understanding as well. However, in a practical sense, if these tasks are to be applied in every day classrooms, oral reading under or-C seems enough, both for reading rate and comprehension score, instead of the combination of comprehension score under or-C and reading rate under or-R shown as model [2] in Table 2.

6. Conclusion

In conclusion, the oral readings in this study more reflected learners’ general English proficiency than their working memory capacity. As the participants were still on the way of learning the language and not sufficient in English proficiency, it may be surmised that the influence of
English proficiency was much more illuminated in the oral reading performance. Thus the possibility for oral reading in this study to be a testing approach to the learner’s English proficiency was supported.

For future study, there seems to be a developmental relationship between English proficiency and working memory. When the same task was done by L1 English speakers, their performances were distinct under two reading conditions. Perhaps, exceedingly high English proficiency entails being able to differentiate the mode of oral reading thoroughly according to the reading conditions. This will be accompanied, then, by a more salient influence of working memory capacity, which will exceed and blur the influence of English proficiency. Reversely speaking, the participants in this study have not reached this stage yet, allowing an oral reading task to be a possible indicator of their English proficiency rather than of working memory capacity. The further experiment on advanced L2 English learners should be of great interest in the future.

Notes

1 The term ‘working memory capacity’ seems rather problematic in that it might cause the arguments about which counts toward human cognitive activities, ‘capacity’ or ‘efficiency’ of working memory. At this moment it cannot be clearly answered, yet the author chose the term ‘capacity’ following Daneman and Carpenter (1980).

2 The reason for adopting the Japanese version was deeply related to what was mentioned in 2.2. Thinking of the participants’ English proficiency level, the English version seems to be too much for them. Thus, in order to minimize the impeding factors and to measure each participant’s fundamental working memory capacity, the Japanese version was selected.

3 Since the correlation coefficient between Reading rates under or-C and or-R was very high (r = .960), one of these two data should be discarded in the analysis to avoid multicollinearity. Even if the data of Reading rate under or-C was adopted, instead of or-R, the value would change little (R^2 = .548).

References


芳阪直行(2000).『脳とワーキングメモリ』. 京都：京都大学学術出版会.

芳阪満里子(2002).『脳のメモ帳ワーキングメモリ』. 東京：新曜社.