Effects of Word Recognition Speed, Accuracy, and Automaticity on Reading Ability

Masumi KOJIMA
Graduate School, Nagoya University

Abstract

This study investigated the significance of the effects of word-recognition speed, accuracy, and automaticity on three different English-reading proficiency groups: two groups of English as a foreign language (EFL) learners and one group of native English speakers. In addition, whether the effects of word-recognition speed, accuracy, and automaticity varied according to word frequency levels was examined. Automaticity of word recognition was measured in terms of the coefficient of variation of reaction times (CVRT), proposed by Segalowitz and Segalowitz (1993). Overall results showed that word-recognition accuracy and speed had significant effects in differentiating among the three reading proficiency groups, while moderate effects were observed for word-recognition automaticity on them. In other words, the more proficient a person becomes in reading, the more quickly, accurately, and automatically he or she can recognize words. Interestingly, the effects of word-recognition accuracy and speed grew larger when word frequency decreased, while such changes were not observed for word-recognition automaticity. These results suggest that learners need to recognize words with a wider range of frequency—not only accurately but also rapidly and automatically.

1. Literature Review

In current theories of reading comprehension, automaticity is seen as a critical way for readers to engage in multiple processes, simultaneously or in a parallel fashion (Grabe, 2009). The reason why automaticity is important is because automatic processes do not place great demands on processing resources in working memory and therefore can be carried out while a person focuses on other tasks (Pressley, 1998).

Grabe and Stoller (2002) argue that the most fundamental requirement for fluent reading comprehension is rapid and automatic word recognition. In this paper, Koda’s (2005) definition of word recognition is employed: Word recognition refers to the processes of obtaining not only words’ sounds but also words’ meanings. Koda (2005) points out that inefficient word recognition is common among low-proficiency readers, irrespective of their first language (L1) backgrounds. She argues that with seriously restricted word-recognition skills, low-proficiency readers expend
considerable time and energy in their visual-information sampling, which severely limits their ability to use multiple information sources.

Grabe (2009) points out that it is possible for certain contexts to initiate word recognition, yet lead to no or very little lexical access. In a second-language (L2) reading context, for example, learners may be able to recognize that a word is a real word in the L2 because they have seen it before, and they even may be able to pronounce it, but they cannot access the meaning of the word simply because there is no lexical entry for the word in their mental lexicon.

The distinction between the speed and automatization of word recognition is also important. According to Grabe (2009), the key characteristics associated with automaticity (aside from speed) are that we cannot stop ourselves from carrying out the process, and we cannot reflect upon on the process. Accuracy, speed, and automaticity are considered to be three important factors of word recognition.

The important question is: How can we quantify the degree of automaticity in the word recognition of individual readers? Segalowitz and his colleagues attempted to answer this question (e.g., Segalowitz, & Segalowitz, 1993; Segalowitz, Watoson and Segalowitz, 1995). They argued that skilled word recognition involves a blend of automatic and controlled processes. Automatic processes operate without interference from other sources of information. Therefore, these processes generally operate very fast and are relatively stable in terms of execution time. Controlled processes, on the other hand, involve decision-making and the evaluation of information. Thus, they generally operate slowly and are relatively more variable in their execution time. The more highly skilled the performance, the more this blend will be characterized by the automatization of those underlying processes, and the execution time, on the whole, becomes faster and less variable. Segalowitz et al. warn us, however, that faster performance may reflect a general speed-up of all underlying mechanisms and not an actual change in the blend of automatic and controlled components. In order to reduce the proportion of controlled processes, some redundant and more variable components need to be eliminated along with recombining some components. These qualitative changes in the underlying processes involved in performing a task are referred to as restructuring. Automaticity in this approach is defined as the qualitative changes resulting from the restructuring of the processes, rather than simply a faster processing time. A statistical index, the coefficient of variation of reaction times ($CV_{RT}$)– which is the standard deviation (SD) of reaction times (RT) divided by the mean RT–was adopted to differentiate between speed-up and restructuring. It is accepted that a change in RT without reduced $CV_{RT}$ is regular speed-up, but a change in RT accompanied by a reduced $CV_{RT}$ is considered an indication of restructuring; thus, in that case, one can claim that automatization has occurred.

Fukkink, Hulstijn, and Simis (2005) investigated automatization of word recognition in L2 with a computer-based training program, using Dutch L1 students with intermediate knowledge of L2 English. In their study, word recognition and the reading-comprehension performance of an
experimental group and a controlled group were compared. After lexical access training, accuracy rate (AC) and RT of word recognition were significantly improved for the experimental group compared to the controlled group, yet the two groups were not differentiated in terms of $C_{RT}$ and reading comprehension. The researchers concluded that there was no transfer of acceleration of lexical access to higher-order text comprehension.

Akamatsu (2008) examined the effects of word-recognition training on the word-recognition processing of EFL learners. One of the main aspects of this study concerned word-frequency effects in the automatization of word recognition. Before training, high-frequency words were accessed more accurately, rapidly, and automatically as compared to low-frequency words. After training, the processing of high-frequency words in terms of RT was improved, but not in terms of AC and $C_{RT}$. On the other hand, AC, $C_{RT}$, and RT were significantly improved for low-frequency words. The researcher concluded that the improvement in the processing of high-frequency words was associated with simple speed-up; the processing of low-frequency words, on the other hand, was associated with automatization.

Although Fukkink et al. (2005) concluded that there was no transfer of acceleration of word recognition to reading comprehension, the lack of improvement in $C_{RT}$ might have caused the lack of improvement in higher-order text comprehension. The effects of word-recognition speed, accuracy, and automaticity on reading comprehension have not been fully examined by previous studies.

This study is basically in response to the need to examine the effects of word-recognition speed, accuracy, and automaticity on reading ability. To this end, we compared the word-recognition processing of two groups of EFL learners with different reading proficiency levels and native speakers of English. The present study is also motivated to investigate the word-frequency effects on the three factors of word recognition. The following research questions are addressed to fulfill these purposes.

1. How large are the effects of word-recognition speed, accuracy, and automaticity on three different English-reading proficiency groups: two groups of EFL learners and one group of native English speakers?
2. Are the effects of word-recognition speed, accuracy, and automaticity on three groups varied according to word frequency levels?

2. Method

2.1 Participants

The participants in the present study were 44 Japanese EFL students and 22 native speakers of English. All the Japanese students were undergraduate or graduate students studying in Japan. They were divided into two groups based on reading scores on the TOEIC test: advanced readers
(365–495) and intermediate readers (210–340). The advanced group consisted of 22 participants whose fields are related to the English language studies, such as linguistics. The intermediate group also comprised 22 participants, but their majors were varied. All native English speakers were university students or English teachers living in Japan. The average age of the Japanese students was 26.4 years and the average age of the native English speakers was 29.6 years.

2.2 Materials

Recognizing the importance of obtaining meanings of words in word recognition, an antonym semantic decision task was employed. In this task, participants are required to judge whether a pair of English words—as a prime and then as a target (e.g., “up-down”)—are antonyms. A total of 128 prime words were taken from the top 2,000 high-frequency words of the 8,000 Basic Word List of Japan Association of College English Teachers (JACET). Among them, approximately 60% of the prime words were taken from the top 1,000 words of JACET and the remaining 40% were taken from the next 1,000 high-frequency words of JACET. All the stimulus words were expected to be familiar to the participants. Thirty-two corresponding antonymous target words were taken from each 1,000 words of the top 4,000 words of JACET; a total of 128 words were selected. Then, half of the target words were replaced by unrelated words to create non-antonym conditions. The other half of the target words were replaced as well, to be used as a counterbalanced list. All unrelated words had the same class of part of speech, the same number of characters, and belonged to the same frequency level of JACET as the original target words.

2.3 Procedures

The participants were tested individually. They received written instructions and 15 practice items before they were tested in the experimental trials. After the main experiment, all EFL learners also completed a questionnaire in which they were asked whether the target items were known to them; they were also asked for their background information. The entire procedure lasted typically 20 minutes for each participant.

The participants’ main task was to decide whether the meaning of the target word was antonymous to the meaning of the prime word. They were required to respond to the task as quickly and accurately as possible. The 128 word pairs were divided into four sets of trials, with each trial consisting of 32 word pairs. The frequency levels of the target words were equally distributed among the four trials. Between each trial, participants were allowed to take a brief rest, if required.

Materials were presented on the screen of a personal computer. Participants were asked to press the right (“Yes”) or left (“No”) button of a mouse as soon as they decided whether a target word was antonymous to a prime word. Each pair of items was presented randomly by the computer program, which was developed by the author for the present study. RT was recorded to the nearest millisecond.

2.4 Analysis

AC ratios, mean RT, and the CVRT of each participant were calculated for each frequency
level of target words. Only anonymous pairs were used for the calculation of AC ratios, and only correct responses of antonyms were used for the calculation of mean RT and CVRT. Items that a participant described as unknown to him or her in the questionnaire were also eliminated from the calculation. An RT outside 2.5 standard deviations in both the high and low range was excluded. Within each frequency level, if an RT was outside of 2.5 standard deviations, it was replaced by the boundaries indicated by 2.5 standard deviations from the individual means of participants.

The result of normality tests revealed that the AC ratios of each participant were not normally distributed, owing to a ceiling effect. In order to deal with this problem, the AC ratios were subjected to inverse sine transformation (Tanaka & Yamagiwa, 1992) before they were submitted to ANOVA.

Two-way ANOVAs, with four frequency levels of target words as a within-participants factor and proficiency level (intermediate, advanced, native speaker) as a between-participants factor, were conducted for AC rates, mean RT, and CVRT. According to JACET, frequency levels were defined as follows: level 1 for the top 1,000 most frequent English words taken from JACET, level 2 for the second top 1,000, and levels 3 and 4 for the third and fourth top 1,000 high-frequency words, respectively. In order to compare the effects of AC, RT, and CVRT on the three different reading proficiency groups, $\eta^2$ was calculated. $\eta^2$ describes the proportion of total variability attributable to a factor (effect size). These effect sizes are an indicator of how large the effect of a variable is upon differentiating between the groups. One advantage of this measure is that even different units of measures can be directly compared to one another.

### 3. Results

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intermediate $(n = 22)$</th>
<th>Advanced $(n = 22)$</th>
<th>Native $(n = 22)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AC $(SD)$</td>
<td>RT $(SD)$</td>
<td>CVRT $(SD)$</td>
</tr>
<tr>
<td>level 1</td>
<td>91.2 (8.6)</td>
<td>929 (179)</td>
<td>0.39 (0.11)</td>
</tr>
<tr>
<td>level 2</td>
<td>82.4 (18.2)</td>
<td>1061 (180)</td>
<td>0.39 (0.09)</td>
</tr>
<tr>
<td>level 3</td>
<td>65.9 (16.5)</td>
<td>1217 (294)</td>
<td>0.39 (0.11)</td>
</tr>
<tr>
<td>level 4</td>
<td>44.9 (13.6)</td>
<td>1345 (335)</td>
<td>0.36 (0.12)</td>
</tr>
</tbody>
</table>

*Note.* Unit: Percentage for AC and milliseconds for RT.
3.1 Accuracy rates

The results of two-way ANOVA for AC rates show that significant main effects for groups, $F(2, 63) = 38.64, p < .01$, and for frequency levels, $F(3, 189) = 133.77, p < .01$. The interaction of frequency levels and proficiency groups was also significant, $F(3, 189) = 19.33, p < .01$. A closer examination of the interaction revealed that the simple main effects of frequency levels were significant within all three groups ($p < .01$). Although the simple main effect of the groups was not significant at frequency level 1, they were significant at all other frequency levels ($p < .01$) (see Figure 1).

Tables 2 and 3 show the results of multiple comparisons by LSD (5% significance level). Overall, the results show that native speakers recognized words more accurately than the learners' groups and that advanced learners responded more accurately than intermediate learners; however, the three groups did not respond differently in regard to the AC of very high-frequency words. Their AC decreased when the frequency of words became lower.

3.2 Reaction times

Regarding the RT, the ANOVA obtained significant main effects for groups, $F(2, 63) = 12.16, p < .01$, and for frequency levels, $F(3, 189) = 63.98, p < .01$. The interaction of frequency levels and proficiency groups was also significant, $F(3, 189) = 5.80, p < .01$. The simple main

![Figure 1. AC for the frequency levels by the groups.](image1)

![Figure 2. RT for the frequency levels by the groups.](image2)
effects of groups and frequency levels were significant at all conditions ($p < .05$ for the groups' simple main effect at frequency level 1, $p < .01$ for all other conditions) (see Figure 2).

Tables 4 and 5 indicate the results of multiple comparisons by LSD (5% significance level).

In general, native speakers responded faster than learners’ groups, and advanced learners reacted more rapidly than intermediate learners, but the difference between the native speakers and the advanced learners was significant only at the frequency level 2. In general, participants’ word recognition became slower when the frequency of words decreased.

### 3.3 Coefficient of variation (CV<sub>RT</sub>)

The ANOVA for the CV<sub>RT</sub> scores produced significant main effects for groups, $F(2, 63) = 6.0, p < .01$, and for frequency levels, $F(3, 189) = 3.09, p < .05$. The interaction of frequency levels and reading proficiency groups was not significant, $F(3, 189) = 0.43, p = n.s.$

The results of the multiple comparisons by LSD (5% significance level) indicate that the CV<sub>RT</sub> scores of intermediate learners were significantly higher than those of advanced learners and English native speakers. There was, however, no significant difference between the CV<sub>RT</sub> scores of advanced learners and native speakers. The CV<sub>RT</sub> scores with words of frequency level 1 were higher than those with words of frequency level 4, while such a statistical significance was not observed for any other pairs of frequency levels.

### 3.4 Effect size

Table 6 indicates the effect sizes ($\text{Eta}^2$) of AC, RT, and CV<sub>RT</sub> which is attributable to the main effect of groups and the simple main effect of groups at each frequency level. Eta$^2$ of CV<sub>RT</sub> by each frequency level was not calculated, because the interaction of frequency levels and

---

**Table 4 Multiple Comparisons of RT for the Frequency Levels by the Groups**

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate</td>
<td>&lt;</td>
<td>&lt;</td>
<td></td>
<td>&lt;</td>
</tr>
<tr>
<td>Advanced</td>
<td>&lt;</td>
<td>=</td>
<td>&lt;</td>
<td>&lt;</td>
</tr>
<tr>
<td>Native Speaker</td>
<td>=</td>
<td>=</td>
<td>&lt;</td>
<td>&lt;</td>
</tr>
</tbody>
</table>

**Table 5 Multiple Comparisons of RT for the Groups by the Frequency Levels**

<table>
<thead>
<tr>
<th></th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>Intermediate</td>
<td>&gt;</td>
<td>Native</td>
</tr>
<tr>
<td>Level 2</td>
<td>Intermediate</td>
<td>&gt; Advanced</td>
<td>Native</td>
</tr>
<tr>
<td>Level 3</td>
<td>Intermediate</td>
<td>&gt; Advanced</td>
<td>= Native</td>
</tr>
<tr>
<td>Level 4</td>
<td>Intermediate</td>
<td>&gt; Advanced</td>
<td>= Native</td>
</tr>
</tbody>
</table>

---

**Figure 3. CV<sub>RT</sub> for the frequency levels by the groups.**
proficiency groups was not significant for CV<sub>RT</sub>. Effect sizes (Eta<sup>2</sup>) between 0.01 and 0.06 are considered to be small, values from 0.06 to 0.14 indicate a medium effect, and values of 0.14 and above indicate large effects (Cohen, 1988).

Large effects were observed for overall AC and RT, and moderate effect was found for CV<sub>RT</sub> on the whole. At each frequency level, the effect size of AC became larger when the frequency of target words decreased. As in the case of RT, the effect sizes for frequency level 1 and 2 were small, but they turned out to be moderate for words with frequency level 3 and 4. The effect of CV<sub>RT</sub> to discriminate among the three reading proficiency groups seems to be constant across the four frequency levels.

**Table 6 Eta<sup>2</sup> for AC, RT and CV<sub>RT</sub> by Frequency Levels**

<table>
<thead>
<tr>
<th>Levels</th>
<th>AC</th>
<th>RT</th>
<th>CV&lt;sub&gt;RT&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.19*</td>
<td>0.20**</td>
<td>0.10*</td>
</tr>
<tr>
<td>Level 1</td>
<td>0.00</td>
<td>0.01*</td>
<td>-</td>
</tr>
<tr>
<td>Level 2</td>
<td>0.03*</td>
<td>0.04*</td>
<td>-</td>
</tr>
<tr>
<td>Level 3</td>
<td>0.11*</td>
<td>0.08*</td>
<td>-</td>
</tr>
<tr>
<td>Level 4</td>
<td>0.15**</td>
<td>0.09*</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note. small, * medium, ** large

4. Discussion

As for RQ 1 ("How large are the effects of word-recognition speed, accuracy, and automaticity on three different English-reading proficiency groups: two groups of EFL learners and one group of native English speakers?"), the results of effect sizes for AC, RT, and CV<sub>RT</sub> revealed that word-recognition speed and accuracy have a large effect on different reading proficiency groups, while automaticity of word recognition has a medium effect on differentiating among them. It is suggested that the more quickly, more accurately, and more automatically a person can recognize words, the better he or she can read, but the effects of speed and accuracy are larger than those of automaticity.

Regarding RQ 2 ("Are the effects of word-recognition speed, accuracy, and automaticity on three groups varied according to word frequency levels?"), the results indicated that the effects of word-recognition speed and accuracy became greater when the frequency levels of words decreased. On the other hand, the effects of word-recognition automaticity were constant across the four frequency levels on three different reading proficiency groups.

Figure 1 visually shows that the difference of AC among three groups becomes greater when the frequency levels of words decreases. The AC of intermediate learners, as compared to advanced learners and native speakers, were most strongly affected by the frequency levels of words, and their AC deteriorated as the frequency of target words became lower. Native speakers’ AC decreased only when they were tested with the level 4 frequency of words.

Figure 2 shows the similar trends of RT of three groups by four frequency levels. The difference of mean RT between the intermediate learners and the advanced learners or native speakers turned out to be larger as the frequency levels of words decreased, while such a
difference was not as obvious between the advanced learners and the native speakers. The RT of intermediate learners were more strongly influenced by the frequency levels of words than were the RT of the advanced learners and native speakers, and the intermediate learners took more time to recognize a word as the frequency of target words decreased. The results of the multiple comparisons show that the native speakers' RT increased only with the level 4 frequency of words. These results suggest that for native speakers, who are relatively highly skilled readers of the language, word recognition is consistently fast and accurate across wider range of frequency levels, as compared to language-learners.

The results of ANOVA for CV_{RT} and the post-hoc analysis imply that the word recognition of the native speakers and the advanced learners was more automatic than that of the intermediate learners. Although CV_{RT} scores of advanced learners were generally higher than those of native speakers, they were not statistically significant. One possible explanation for this result is that their automaticity of word recognition is not differentiated by CV_{RT} measures. Another explanation is that deciding whether a word pair is antonymous or not involves decision-making, and some amount of controlled processes always remain. As Akamatsu (2008) points out, once the automaticity of one's word-recognition mechanism reaches its optimal level, decreases in CV_{RT} may not be observable.

Figure 3 shows that the scores of CV_{RT} were relatively constant across the four frequency levels, regardless of which reading proficiency group was being tested. This result suggests that the restructuring of underlying processes of word recognition takes more time, and the changes of the CV_{RT} are subtler compared to the changes of AC and RT. It was expected that the learners’ CV_{RT} scores would increase as the words’ frequency decreased, because they are more exposed to high-frequency words than low-frequency words, and their word recognition of high-frequency words is supposed to be more automatic than that of low-frequency words. Contradictory to those expectations, however, the CV_{RT} scores of level 4 frequency words were smaller than those of words with frequency level 1. One possible explanation for this is that the AC deteriorated as the words’ frequency decreased, especially at frequency level 4. As a result of eliminating wrong responses, the proportion of SD did not increase more than the correspondingly proportional increase in RT; this in turn led to smaller CV_{RT} with lower frequency words. If this assumption is true, CV_{RT} cannot be directly compared when the AC rates differ among conditions.

5. Conclusions and Implications

The results of this study provide some evidence that the effects of word-recognition speed and accuracy are large on different reading proficiency groups. In other words, the more proficient a person becomes in reading, the more quickly and accurately he or she can recognize words. Those effects grow larger when word frequency decreases.

The current study also suggests that the effects of word-recognition automaticity are
substantial on different reading proficiency groups, although they are not as large as the effects of word-recognition speed and accuracy. However, because some unexpected results were observed with the scores of CVRT, we need to investigate further the contribution of word-recognition automaticity to individual differences in reading proficiency.

Here are some educational implications. In order to improve reading skills, learners need to recognize words with a wider range of frequency, and they must recognize words more quickly, more accurately, and more automatically. Their priority seems to be to increase the breadth of word knowledge and then improve the lexical access to those words, in terms of speed and automaticity. To attain these goals, as Grabe (2009) points out, learners need to practice word recognition continuously with enormous amounts of meaningful input.

References


