Investigating EFL Learners’ Mental Representations with the Verb-Clustering Test

Shuichi TAKAKI
Graduate School, University of Tsukuba

Abstract

Although many studies assume that the construction of situation models is essential for reading comprehension, few of them have examined EFL learners’ reading process from the perspective of situation models. This study investigated the situation models of EFL readers with the verb-clustering test which is used in the event-indexing model paradigm. (Zwaan, Langston, & Graesser, 1995). The event-indexing model paradigm is highly recommended for classroom instruction because it refers to elements for the development of elaborate situation models.

The verb-clustering test is one way which examines situation models of readers. It is thought to be useful because one can investigate learners’ situation models without disturbing them (Iseki & Kawasaki, 2006). Furthermore this test is effective for classrooms because it does not take much time to administer and it allows teachers to examine the reading processes of L2 learners. This study investigated the situation models of L2 readers with the test.

A total of 122 high school students participated in this study. They read a short narrative text and took the verb-clustering test. They were divided into two groups based on their language proficiency. The result of this study showed that the construction of L2 situation models that are comparable to those of L1 demands a high degree of language proficiency, which is supported by Zwaan and Brown (1996). This study also suggested that the pre-reading instructions had different effects on situation models depending on learners’ language proficiency. Given instructions, good readers could construct more elaborate situation models in contrast to poor readers, who could not.

1. Introduction
1.1 Studies on Mental Representations

Learners’ success in reading depends upon the construction of mental representations of texts (Iseki & Kawasaki, 2006; Zwaan & Radvansky, 1998). Mental representations are memories learners construct about information in a text. According

— 163 —
to van Dijk and Kintsch (1983), mental representations consist of three levels of representation: the surface structure, the text-base, and the situation model. The surface structure level is made up of memory of the text itself. It consists of accurate sequences and running words. The text-base level consists of memories about the relationships of ideas in a text. It shows how the propositions are interconnected to each other. Finally, the situation model level indicates memory of the reader’s mental situation derived from interaction with a text, including actions and states. It represents the reader’s image of a text and not a text as it objectively is. While the surface structure and text-base are constructed from text descriptions alone, the situation model is constructed from an interaction between text descriptions and the reader. Although each level of representation is essential for reading, the situation model is recognized as the most important for mental representations. L1 studies of reading comprehension have maintained that the construction of accurate situation models plays a significant role in reading processes (Zwaan & Radvansky, 1998).

Building upon the implications of those L1 studies, situation models were recognized as one of the important factors for reading success in L2 reading instruction. However, to date, there have been few L2 studies on situation models. One reason could be that such studies have provided few implications for classroom instruction. Many L1 studies focused on the process of constructing coherent situation models. Those studies did not consider what readers need to accomplish the process (Rapp, 2008). In addition, they failed to replicate actual reading conditions because materials were intentionally controlled.

1.2 The Event-Indexing Model and the Verb-Clustering Test

Since the event-indexing model was one of the theories explaining situation models suggested in an L1 reading study (Zwaan, Langston, & Graesser, 1995), event-indexing model can move research past this impasse. This model provides abundant pedagogical implications in two ways.

First, the event-indexing model refers to specific elements included in situation models. According to the event-indexing model, situation models consist of at least five dimensions: protagonist, time, space, causality, and intentionality — all of them necessary for successful reading comprehension (Zwaan et al., 1995; Zwaan & Radvansky, 1998). It is assumed that readers construct dimensions for each clause and maintain coherence of each dimension. More constructed dimensions are needed to succeed in constructing accurate situation models.

Second, the verb-clustering test used in the event-indexing model paradigm is useful because the test allows participants to read texts in a naturalistic condition.
(Zwaan, 1999). By adopting the test, situation models were investigated without any intervention in the reading process. The test consisted of some texts and the clustering test. The clustering test shows a list of the ten target verbs and a column of seven boxes and requires participants to write in the same box the verbs that are thought to be related to the narrative. The verbs categorized in the same group are associated with the reader’s mental representations (Iseki & Kawasaki, 2006). Similarly, the information implied by the verbs is connected in the reader’s representation. These verbs are presented in root form and in random order to avoid giving participants redundant clues. Iseki and Kawasaki (2006) used the test and examined the difference between L1 readers’ situation models of narrative text and expository text. Their results supported the validity of the test.

Therefore, this study investigated EFL learners’ mental representations based on the event-indexing model paradigm. The purpose of this study is to examine the utility of the verb-clustering test and verify the potential implications of the studies on situation models for classroom reading instruction.

1.3 Studies in the Event-Indexing Model Paradigm

Few of the event-indexing model studies using the verb-clustering test covered L2, except for Zwaan and Brown (1996). They compared situation models of L1 and L2 readers, and found that L2 readers could include only three dimensions in situation models, whereas L1 readers could include four. This result indicated that the difference in languages influences perception of dimensions in situation models. In other words, L1 readers could make more dimensions in situation models than L2 readers. However, it was unclear what factors caused this difference. Therefore, this study investigates the influence of language proficiency on situation models. On the basis of Zwaan and Brown (1996), it is anticipated that proficient readers can include more dimensions in situation models than less advance readers.

This study also examines the effects of pre-reading instructions that facilitate the construction of situation models. As indicated in Zwaan and Brown (1996), L1 learners could construct more elaborate situation models than L2 learners, and more dimensions were necessary for more successful reading comprehension. Therefore, better reading instruction was defined as instruction that enhances reader perception of dimensions in situation models.

As some L1 studies investigated, situation models do not always include the spatial dimension in the naturalistic reading process. In other words, even L1 readers have limited ability to build the spatial dimension in situation models. For example, Hakala (1999) suggested that readers needed to recognize the importance of spatial
information for text comprehension so that the situation models contain the spatial dimension. Then, reading instruction that enables readers to construct the spatial dimensions is necessary for elaborate situation models and profound reading comprehension. Therefore, this study also investigates whether pre-reading instructions that make readers focus on spatial information enhance the spatial dimension in situation models, which in turn contribute to more elaborate situation models.

1.4 Research Questions (RQs)
RQ1: Does EFL learners’ reading proficiency affect their construction of situation models?
RQ2: Does reading instruction that makes learners focus on spatial information contribute to their construction of situation models?

2. Method
2.1 Participants
A total of 122 Japanese high school students participated in this study. The data of one participant was excluded because of nonperformance on the test. The participants were divided into two groups according to their performance on a cloze test (see 2.2) used as a reading proficiency test (Aitken, 1977): an upper group \( n = 62 \) and a lower group \( n = 59 \).

2.2 Materials
Two materials were used in this study: the cloze test as a reading proficiency test and verb-clustering tests (see Appendix A). STEP Eiken’s pre-second grade test was used as text for the cloze test. Thirty-two blanks were made: one after every seven words except for the first and last sentences.

Each verb-clustering test consists of a short narrative, the clustering test, and ten comprehension questions. This study used four original narrative texts that were used in Zwaan et al. (1995) for the verb-clustering tests. Two of the four verb-clustering tests were given to participants as a counterbalance. In the second test, the instruction (see Appendix B) was added that participants should read the narrative with focus on spatial information. After the verb-clustering test, ten true-or-false questions on the narrative were presented. These questions were included to see whether participants took these tests seriously.

Six difficult words in the narratives were replaced with familiar words in case those words prevented participants from reading based on a survey about difficult words for university students and the JACET 8000 word list (2003). In the survey, ten students
read the original narratives and reported unknown words. In addition, ten unique verbs were selected as target verbs in each narrative. Table 1 shows the properties of the narratives.

Table 1 Readability of the Narratives Used in the Verb-Clustering Tests

<table>
<thead>
<tr>
<th>Title</th>
<th>Token</th>
<th>Flesh-Kincaid Grade level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czar and his daughter</td>
<td>88</td>
<td>3.1</td>
</tr>
<tr>
<td>The boy and his dog</td>
<td>120</td>
<td>2.3</td>
</tr>
<tr>
<td>John at Leone's</td>
<td>79</td>
<td>1.9</td>
</tr>
<tr>
<td>The ant and the pigeon</td>
<td>103</td>
<td>3.3</td>
</tr>
</tbody>
</table>

2.3 Procedure

Participants were informed about the directions beforehand and then were tested all together. They were not allowed to use any dictionaries. First, they took the cloze test within 10 minutes. Second, they were given the verb-clustering tests and instructed to attempt them following written directions. After they read the narrative at their own pace, they immediately attempted the clustering test and comprehension questions. After beginning the clustering test, they were not allowed to go back and reread. The time for these tests was unlimited to make sure that participants could construct situation models from the narratives.

2.4 Analysis

Following previous studies on the event-indexing model (e.g., Zwaan et al., 1995; Zwaan & Brown, 1996), multiple regression analysis was conducted. In the verb-clustering test paradigm, the dependent and independent variables were computed for each verb pair. The verb pairs were used as items in this analysis. Since each narrative had ten target verbs to be clustered, 45 verb pairs could be made in each narrative.

Figure 1. The Model for the Regression Analysis
Consequently, 180 items were analyzed. Figure 1 indicates the model for regression analysis of this study. The dependent variables were the clustering scores. They showed the proportion of participants who placed a verb pair in the same box. They could range from 0 (no participants judged that the verbs were related) to 1 (all participants judged that the verbs were related). As this study examined the effects of language proficiency (upper and lower proficiency groups) and the effect of instruction (without and with instruction), four dependent variables were obtained.

The independent variables were continuity of information implied by verb pairs. They consisted of situation model level factors, textual (surface structure and text-base) level factors, and the lexical factor. There were five situation model level factors: (a) protagonist identity, (b) temporal relatedness, (c) spatial relatedness, (d) causal relatedness, and (e) intentional relatedness. There were three textual level factors: (f) surface connections, (g) surface distance and (h) argument overlap. The lexical factor was (i) lexical relatedness. Protagonist identity indicated whether the same character caused the events indicated by two verbs. Temporal relatedness indicated whether the events occurred in the same time paradigm. Spatial relatedness reflected whether the events occurred in the same place. Causal relatedness indicated whether the events had a cause and effect relationship. Intentional relationship indicated whether the events shared the same goal. The scores of protagonist, time, space, causality, and intentionality were rated as 0 (those information elements implied by a verb pair were not consistent) or 1 (those information elements implied by a verb pair was consistent) by two judges who followed those definitions explained in Zwaan et al. (1995). All the discrepancies were resolved by discussion, and the inter-rater reliability was 88%.

Following Iseki and Kawasaki (2004), the other four factors were decided by one judge. Surface connections showed whether two verbs were in the same sentence. Argument overlap indicated whether two verbs shared the same argument. Surface connections and argument overlap were each rated as 0 (a verb pair did not fulfill the condition) or 1 (a verb pair fulfilled each condition). Surface distance indicated the number of content words existing between the two verbs. In contrast to the previous five variables that were rated by two judges, these three variables were rated by one judge in order to prevent scoring differences from accruing, per Zwaan and Brown (1996). Finally, lexical relatedness indicated the intrinsic relatedness of verb pairs in the lexicon. Following Zwaan et al. (1995), estimates of lexical relatedness were obtained by having another group of participants perform the verb-clustering tests without reading the narratives. Therefore, twenty undergraduates performed the tests without reading the narratives. Lexical relatedness could range from 0 (no participants judged that the verbs were related) to 1 (all participants judged that the verbs were related) in a similar
3. Results and Discussion

3.1 Reading Proficiency Test

Table 2 Descriptive Statistics of the Reading Proficiency Test

<table>
<thead>
<tr>
<th>Proficiency</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>62</td>
<td>11.15</td>
<td>2.51</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>Lower</td>
<td>59</td>
<td>4.83</td>
<td>1.84</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Full marks = 32

Table 2 shows the descriptive statistics for the cloze test. There was significant difference between participants’ proficiency in the upper and lower groups: \( t(119) = 15.7; p = .000 \); Reliabilities of Cronbach’s \( \alpha = .75 \).

3.2 The Verb-Clustering Test

Table 3 Intercorrelations Between the Predictor Variables

<table>
<thead>
<tr>
<th>Verb pairs (n = 180)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.Protagonist</td>
<td>- .13( \uparrow )</td>
<td>.25**</td>
<td>.10</td>
<td>.09</td>
<td>.20**</td>
<td>-.15*</td>
<td>.64**</td>
<td>-.07</td>
<td></td>
</tr>
<tr>
<td>2.Time</td>
<td>-</td>
<td>.71**</td>
<td>.48**</td>
<td>.32**</td>
<td>.46**</td>
<td>-.54**</td>
<td>.19*</td>
<td>-.14( \uparrow )</td>
<td></td>
</tr>
<tr>
<td>3.Space</td>
<td>-</td>
<td>.46**</td>
<td>.16*</td>
<td>.43**</td>
<td>-.54**</td>
<td>.25**</td>
<td>-.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.Causality</td>
<td>-</td>
<td>.35**</td>
<td>.33**</td>
<td>-.48**</td>
<td>.14( \uparrow )</td>
<td>-.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.Intentionality</td>
<td>-</td>
<td>.27**</td>
<td>-.26**</td>
<td>.08</td>
<td>-.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.Surface connections</td>
<td>-</td>
<td>-.30**</td>
<td>.08</td>
<td>-.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.Surface distance</td>
<td>-</td>
<td>-.20**</td>
<td>.13( \uparrow )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.Argument overlap</td>
<td>-</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.Lexicon</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. \( n = 180 \). \( \uparrow p < .10 \), \( * p < .05 \), \( ** p < .01 \).

Before conducting multiple regression analysis, the data points that indicated instruction flaws were excluded. Moreover, the data points whose comprehension scores were outside 3SD from the mean were eliminated because they showed that participants did not comprehend the narratives. As a result, 41 data points (18 in the first test and 23 in the second test) were removed. Table 3 shows bivariate correlations.
among the predictor variables. The correlation between temporal relatedness and spatial relatedness was .71, which indicated that multicollinearity problems could appear. If multicollinearity problems occur, the results of multiple regression analysis would be irrelevant. Therefore, two indications — the variance inflation factor (VIF) and the minimum variance of predictor variables — were used to determine whether such problems are present. The results showed that multicollinearity problems did not appear. (The highest absolute VIF was 2.45 and the lowest minimum variance was .254.)

Table 4 indicates the beta weights from the multiple regression analyses. In the first test, the results of the upper and lower groups were the same. In both groups, the effects of protagonist identity, spatial relatedness, and lexical relatedness were significant ($p < .01$). In the second test, however, the upper and lower group results were different. In the upper group, the effects of protagonist identity, temporal relatedness, spatial relatedness, causal relatedness, surface connections, and lexical relatedness were significant ($p < .05$), whereas in the lower group, spatial relatedness, surface connections, and lexical relatedness were significant ($p < .05$).

### 3.3 RQ1: Does EFL learners’ reading proficiency affect the construction of situation models?

As to the naturalistic reading condition, situation models constructed by the upper group were similar to those by the lower, in that protagonist identity, spatial relatedness, and lexical relatedness influenced their situation models. Although Zwaan and Brown (1996) concluded that situation models constructed by L1 and L2 readers were different, the result of this study indicates the possibility that reading proficiency does not influence situation models.

Consequently, the result shows that the difference in situation models between L1 and L2 readers is independent of reading proficiency. However, one cannot conclude that reading proficiency does not affect reading process. That would allow one to assume that the upper groups were not proficient enough to construct more elaborate situation models. Zwaan and Brown (1996) suggested that the construction of situation models was influenced by linguistic threshold. They pointed out that L2 readers consumed more resources in lower level processing and fewer resources for the construction of situation models. Therefore, the upper group participants in this study were obliged to construct similar situation models to those of the lower group.
### Table 4 Summary of Regression Analysis for Variables Predicting Clustering Tests  
(N = 180)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Upper without Instruction</th>
<th>Upper with Instruction</th>
<th>Lower without Instruction</th>
<th>Lower with Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protagonist</td>
<td>.23</td>
<td>.11</td>
<td>.12</td>
<td>.15</td>
</tr>
<tr>
<td>Time</td>
<td>.02</td>
<td>.14</td>
<td>.05</td>
<td>.06</td>
</tr>
<tr>
<td>Space</td>
<td>.14</td>
<td>.25</td>
<td>.03</td>
<td>.05</td>
</tr>
<tr>
<td>Causality</td>
<td>0</td>
<td>.08</td>
<td>0</td>
<td>-04</td>
</tr>
<tr>
<td>Intentionality</td>
<td>.06</td>
<td>.05</td>
<td>.07</td>
<td>.05</td>
</tr>
<tr>
<td>Surface Connections</td>
<td>.07</td>
<td>.13</td>
<td>.08</td>
<td>.06</td>
</tr>
<tr>
<td>Surface Distance</td>
<td>0</td>
<td>0</td>
<td>.06</td>
<td>0</td>
</tr>
<tr>
<td>Argument Overlap</td>
<td>.05</td>
<td>-.04</td>
<td>.00</td>
<td>.03</td>
</tr>
<tr>
<td>Lexical Relatedness</td>
<td>.24</td>
<td>.15</td>
<td>.03</td>
<td>.13</td>
</tr>
</tbody>
</table>

Note. $R^2 = .42$ for the Upper Group Without Instruction; $R^2 = .62$ for the Upper Group With Instruction; $R^2 = .34$ for the Lower Group Without Instruction; $R^2 = .31$ for the Lower Group With Instruction. † $p < .10$, *$p < .05$, **$p < .01$

### 3.4 RQ2: Does the reading instruction that makes learners focus on spatial information contribute to the construction of situation models?

Under the naturalistic reading condition, situation models of the upper and lower groups were similar: they contained protagonist identity, spatial relatedness, and lexical relatedness. However, the reading instruction had distinct effects on situation models relating to language proficiency. In the upper group, the instruction encouraged the learners to include information of temporal relatedness, causal relatedness, and surface connections in their situation models.

Then the relationships between spatial relatedness and the three encouraged variables (temporal relatedness, causal relatedness, and surface connections) were analyzed. As presented by Table 2, those variables were more highly correlated to spatial relatedness than any other variables were. Therefore, Hotteling's t-test was used to determine whether differences between correlations were statistically significant. As
the correlation coefficient between spatial relatedness and argument overlap was the highest, other than those between spatial relatedness and the three encouraged variables, the former was compared to latter. The results showed that the coefficients between spatial relatedness and (a) temporal relatedness, (b) causal relatedness, and (c) surface connections were significantly higher than that between (d) spatial relatedness and argument overlap: (a) and (d), $t(179) = 6.773, p < .01$; (b) and (d), $t(179) = 2.674, p < .01$; (c) and (d) $t(179) = 1.981, p < .05$.

These results indicate that the reading instruction stimulated the upper group participants to include in situation models temporal and causal dimensions that were strongly related to spatial information. Therefore, the upper group learners constructed more elaborate situation models upon receiving the reading instruction. Based on Rinck and Weber (2003) that investigated the relationships between dimensions, it is valid to conclude that the reading instruction facilitated situation model inclusion of not only spatial dimension but also of temporal and causal dimensions.

In the lower group, the instruction discouraged learners from including the protagonist dimension in situation models. In brief, the lower group learners did not construct elaborate situation models upon receiving the reading instruction. The protagonist dimension was recognized as one of the most important types of information (Zwaan & Radvansky, 1998). Furthermore, Therriault, Rinck, & Zwaan (2006) showed that L1 readers focusing on spatial dimension were not constrained from comprehending the protagonist dimension; that is, spatial dimensional focus had no influence on protagonist dimension focus.

Therefore, this result showed that the lower group learners could not utilize the reading instruction for the construction of situation models. According to Carrell (1989), good readers could make use of a general reading strategy, but poor readers tended to use a local reading strategy. Poor readers were apt to have trouble in dealing with text descriptions, such as an individual sentence, and were not able to achieve global comprehension. The reading instruction seemed to force learners into using a general reading strategy. The lower group learners could not use the instruction.

Moreover, Kikuchi (1997) gave text structure information to their L1, L2 proficient, and moderate readers before reading. From the results, she argued that any reading instruction conducted before reading did not facilitate learners' comprehension. In addition, she suggested that the effects of such instructions on learners were not clear.

In the upper group learners, the reading instruction increased the complexity of their situation models because the surface connection information was strongly related to spatial information in the narratives. However, the lower group learners were not assisted similarly because they seemed not to be able to use the instruction. In this study,
the participants who did not read narratives to make models were excluded based on the comprehension questions. In other words, the analyzed data points were collected from participants who succeeded making models of narratives. Thus, this result shows that the lower group learners receiving the instruction could not make elaborate situation models and strengthened surface and text-base level representations.

As Zwaan and Brown (1996) indicated, L2 readers reinforced the text-base representations when they could not construct elaborate situation models. These results suggest that the reading instruction for learners to focus on spatial information had different effects on learners depending on their reading proficiency. More proficient learners can make use of the instructions and construct elaborate situation models. On the other hand, less proficient learners cannot construct simple situation models; instead, they reinforce surface and text-base representations.

4. Conclusion

This study investigated (a) the effects of language proficiency on the construction of situation models and (b) the effects of pre-reading instructions encouraging readers to focus on spatial information with the verb-clustering test. The investigation suggested that language proficiency does not affect situation models, but this result requires further verification. This contention is supported by the fact that reading instructions affected groups of readers differently based on their language proficiency. More proficient readers could make use of the reading instructions, which enabled them to include other dimensions related to spatial relationships in their situation models. On the other hand, less proficient readers could not use the reading instructions, with the result that their situation models were superficial and they instead strengthened their text-base models. This suggests that reading instruction that makes readers focus on spatial information does not always enhance their construction of situation models and the effects depend on the readers’ language proficiency.

All the results taken together have some pedagogical implications. Teachers should motivate readers to include more information in their mental models, especially information about the protagonist, time, space, cause, and intention. Moreover, teachers must take readers’ language proficiency into account when they give reading instructions. These instructions can lead readers to better comprehension if their proficiency is high enough to utilize the instructions well.

Finally, a limitation of this study should be noted. One cannot specify which dimensions in situation models are constructed by EFL learners solely from the findings of this study. Reader factors as well as text factors affect situation models. Since the texts used in this study were limited to narratives and the number of texts was only four,
text factors should be taken into account in future research.

Acknowledgement
I would like to thank Yuji Ushiro, Haruka Shimizu, Akari Kai for their comments.

References
JACET. (2003). JACET list of 8000 basic words. Tokyo: JACET.


**Appendix A: One of the narratives and the verb-clustering test**

**Czar and his daughter**

Once there was a Czar who had three lovely daughters. One day the three daughters went walking in the woods. They were enjoying themselves so much that they forgot the time and stayed too long. A dragon kidnapped the three daughters. As they were being dragged off they cried for help. Three heroes heard their cries and set off to rescue the girls. The heroes fought the dragon. Then the heroes returned the daughters to their palace. When the Czar heard of the rescue, he rewarded the heroes.

Below are the original forms of the verbs that appeared in the text. Group them in the boxes below, based on how you recall the story (originally written in Japanese):

verbs: cry, fight, kidnap, enjoy, reward, return, stay, forget, walk, drag

Appendix B: The instruction that make readers focus on spatial information

“Read the narrative carefully with paying attention to changes of places where events occurred in the story.” (originally written in Japanese)