Applicability of technology-enhanced visual glosses for explicit L2 vocabulary learning: The enhancement of metaphoric competence through the learning of English polysemous words

Takeshi Sato

Institute of Engineering of Tokyo University of Agriculture and Technology, Japan

Abstract

This study examines the efficacy of technology-enhanced visual glosses in explicit L2 vocabulary learning based on the concept of the image schema, which is a mental pattern of our bodily experiences. Although our previous studies could not confirm the advantage of the animated visual glosses for acquiring English prepositions, this study reexamines the learning effect of the animated glosses; we assume that the animation could enhance learners’ comprehension and production of the target prepositions (especially in their metaphorical sense). The findings suggest that the animated visual glosses were only effective in certain scenarios. More specifically, the animated image schema was more an effective gloss for students to produce metaphorical prepositions than for them to select the correct word in receptive tests. Thus, this study therefore shows that when examining technology for L2 learning, more analysis of the features of the target L2 knowledge should be made.

© 2016 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

One of the primary goals of second or foreign language (L2) acquisition is to improve learners’ comprehension of texts and discourse in the target language. In order to achieve this goal, L2 learners are taught various aspects of the target language, such as grammar and vocabulary. Since the 1990s, increasing attention has been paid to L2 vocabulary learning; this is because learners’ insufficient vocabulary has been consistently reported as a significant problem in their achievement of L2 proficiency [11,14,15,29,42]. Many studies, including computer-assisted language learning (CALL) research, have thus focused on how to most effectively learn L2 vocabulary (e.g., [6,28]).

Along with the Dual Coding Theory [38] in multimedia learning [37], which notes how representing information using verbal and visual codes can more effectively facilitate information recall of information than when using only one code, the efficacy of pictorial aids as visual glosses [35] in dictionaries and learning materials has been examined in CALL research. In fact, many studies conclude that the use of pictures as visual glosses can facilitate longer retention of target vocabulary (e.g., [1,7,19,27,35,48,49]). [5] noted the multimedia environment’s advantage for input comprehension by stating that “[i]n material in which visualization, sequence, motion and/or trajectory were essential, the animated presentation strategy showed significant advantage over text-only. (p.100)”, thus demonstrating how visual glosses can facilitate increased information retention [27].

However, this study asserts that the findings of previous studies are not sufficiently robust to present a clear conclusion on the
merits of technology-enhanced visual glosses for L2 vocabulary learning. First of all, the previous studies tend to give too much credit to technology in L2 learning and as a result, they ignore the linguistic differences of the target vocabulary. Although their conclusions seem to claim that technology-enhanced visual aids lead to successful learning of any vocabulary items, they do not mention the learning effect per their target vocabulary item. These studies’ random selection of the target vocabulary might indicate their ignorance of the fact that some words are easier or more difficult to learn. Often, some words are easier to learn because each form has only one meaning, while other words are more difficult because they have multiple meanings that require the users to select the appropriate meaning for the given context. [11] pointed out that the type of vocabulary item or its semantic structure may affect its level of learning difficulty. The studies mentioned above did not examine such differences and instead rather roughly conclude that pictorial aids are effective in any setting.

Furthermore, they examined the effectiveness of visual glosses in terms of incidental vocabulary learning (e.g. Refs. [1,5,7,17,28,48,49,52]), this emphasis is based on the claim that most words are learned incidentally [16]. The participants in these studies, therefore, should use visual glosses not to understand the target words, but to understand texts they read. [48]; for example, asked learners to read a narrative story on the Internet with verbal, visual, and both verbal and visual glosses and thereby illustrate which glosses could enhance retention of the target vocabulary learned incidentally. This research design indicates that the L2 learners did not use the visual glosses to understand the semantic features of the target words explicitly or intentionally. Therefore, the study’s finding as well as other related studies shown above do not support visual glosses’ effectiveness in learning certain vocabulary types explicitly.

Considering the limitations of previous studies focusing on computer-mediated visual glosses for incidental L2 vocabulary learning, this study focuses on the semantic features of the target vocabulary and therefore, the relevant theoretical perspectives from semantic research should be used in tandem with L2 vocabulary acquisition studies. In order to address the issue, this study takes up a theoretical framework from Cognitive Linguistics.

1. Application of cognitive linguistics in CALL research: image-schema

Cognitive Linguistics (CL) is the study of language as a cognitive function, that is, the study “to discover the ways in which linguistic structures are related to and motivated by human conceptual knowledge, bodily experience, and the communicative functions of discourse” [13]; p. 11). The most interesting claim of CL is that of the “embodiment of meaning” [53], [22] also illustrated that our concepts are meaningful because they are embodied or rooted in our bodily experiences. Metaphor, which is one of the most important concepts of CL, plays a pivotal role in mapping our bodily experiences to abstract ideas that underlie our ways of thinking [13]. To conduct CL-based language analysis called usage-based model [25], many studies address prepositions [8,9,22,44,46,50] and conclude that words develop prototypical semantic categories which are related to bodily experiences and the categories can systematically extend into abstract ones [12,22,25].

Furthermore, CL highlights how the patterns of our bodily experiences can serve as crucial mediators between language and concept, as well as facilitating metaphorical mappings from our bodily experiences to abstract domains [18]. Such patterns are called schema [18,21,22,25], [26] speculated that all the members of a category (such as meanings of a word) are conceptually motivated by a certain schema. This is because the schema is an instantiation of a prototypical member created by conceptualizing our bodily experiences, which could then motivate non-prototypical members (such as figurative senses of a word). Due to how all members of a category are systematically connected to the schema, the schema could help people understand both the literal and the figurative meanings. The process of polysemous network is illustrated in Fig. 1 [26]:

Considering the CL-based polysemous network that all meanings can be conceptually motivated by the schema and how figurative meanings are vital for L2 learners’ understanding [36], as cited in Ref. [33], I hypothesize that the schema can be utilized as a pedagogical device to aid L2 learners’ comprehension of the target vocabulary semantic network, as well as to understand and produce both the physical and figurative expressions of a word.

This study, which focuses on visual glosses, addresses the abstract representations of our bodily experiences [24] called the image schema [22], which is the core image of a word that connects it to a spatial concept [10], [18] also defined image schemata as “abstract patterns in experience and understanding that are not propositional” (p. 2).

Fig. 2 depicts a prototypical image schema called above-across schema [22]; a moving object (TR: trajector) is located at the vertically upper side of another object (LM: landmark), which is a typical spatial relationship expressed by the preposition over. The image schema is intended to cover all the meanings the word has, although each meaning appears to be different.

2. Polysemous words

In this sense, what word type should be learned using the schematic image? As shown in the previous CL studies analyzing the semantic structure of prepositions, an image schema can actively organize extended examples of polysemous words that involve several physical (or literal) and figurative (or metaphorical) senses, due to the special problems they pose for EFL learners. Common examples of polysemous words include not only prepositions (e.g., break, take, bring, get) [32] demonstrated the importance of polysemous words in L2 vocabulary learning. First, polysemous words are very frequently used in many types of contexts. Secondly, the words referring to positional relationships are important and finally, prepositions are one of the most typical polysemous word types and are used to construct and express many important concepts. However [31], argued that learning L2 spatial polysemous words involves some problems, which he presumes to stem mainly from learners’ perceptions. They are likely to think that spatial polysemous words are highly idiomatic and very often delexical, which means they are not the content words with any significant meanings. Therefore, a better way to learn how to use them naturally has to simply memorize as many set phrases as possible in their appropriate contexts. However, this study emphasizes that such rote learning results in learners’ failure to appropriately use spatial polysemous words. The complexity of their semantic networks makes them

![Fig. 1. The process of extension [26]; p. 271.](image-url)
difficult to comprehend, often leading to confusion in L2 learners despite how common these words are; this often results in learners’ difficulty in learning such words [4,14,31].

To address this difficulty, CL-based studies for L2 learning suggest that formal instruction can help students establish an organized semantic category for a word. Such a category can help learners recognize that all the information is in the same semantic category, as well as how metaphorical mapping occurs from the physical sense of a word to the figurative one. To do so, this study advocates the use of the image schema for visual glosses, as proposed by Refs. [2,30]. Although several studies apply the CL perspectives to L2 vocabulary learning. (e.g. Refs. [2,20,23,45,47], few studies have verified technology-enhanced visual glosses’ effectiveness, particularly in productive activities. This study, therefore, develops the visual glosses for L2 preposition with multimodal functions and examines their efficacy in explicit learning.

3. Technology-enhanced visual glosses

Of course, the advantage of developing image schemata in a multimedia environment is not only the spontaneous presentation of verbal and visual information as emphasized in previous studies. The multimedia environment could also produce visual glosses that could not be developed in paper-based materials: animation and three-dimensional images. [1] demonstrates that animation is a more efficient technique than still images in L2 vocabulary learning. [34] also claims that three-dimensional diagrams might be useful when they can be displayed dynamically. Therefore, it might be thought that the more technology-enhanced visual annotations learners are exposed to, the better learning outcomes they could acquire.

4. Our previous studies

Based on the supposition that image schemata can serve as effective visual glosses for explicit learning of spatial prepositions, I have conducted several verifications. [39] demonstrated the effectiveness of image schemata as visual glosses in explicit learning of target spatial prepositions (above, along, across, below, in, into, on, over) compared with verbal glosses such as “something is located on the upper surface of something else” as an example of above. Significantly different learning results are found between the visual and verbal glosses. [39] also examined the relative effects of two types of visual aids—still images and animated images. Each image is developed based on a central image schema [8], which conceptually depicts the prototypical meaning of each word. Fig. 3 shows the visual aid of the preposition over [43], which is a still picture that conceptually depicts the prototypical meaning based on Dewell’s central image schema (see Fig. 4).

Fig. 4 depicts a series of three-dimensional animations depicting the above-across schema like Fig. 2 or the central image schema like Figs. 3 and 4. After showing these glosses for a certain period, a vocabulary test was conducted to examine which type of visual aid best develops the understanding of the meanings of the eight target prepositions. As a result of our data analysis, no significant difference between the two visual aids was found. The other studies [39–41] conducted with different research designs also illustrate the same results: no significant difference was found between pictorial and animated aids in learning prepositions.

In sum, our previous studies demonstrated the effectiveness of image schemata as visual glosses displayed spontaneously together with verbal glosses on a computer screen. However, these studies failed to emphasize how beneficial the technology-enhanced image schemata (e.g. animations or 3D images) are in learning L2 spatial prepositions. It seems this issue is also rarely discussed in previous CALL studies.

5. Aim of this study

Therefore, the purpose of this study is to verify the efficacy of the visual glosses under the Cognitive Semantics framework [22,25]. To conduct this verification, my research design will be modified in the following respects. First of all, the target vocabulary is restricted: our previous studies [39–41] addressed eight spatial prepositions, but this research focuses on three: above, on, and over. This is because the translation of each prototypical meaning, which is a requisite for vocabulary learning, is the same: -no ue wo (ni) (located at the upper surface) in Japanese. For example, a prototypical meaning of over is translated into the sentence (1) below:

```
Hikouki ga yama (the mountain) no ue wo (over) tondeiru (is flying).
```

(1) An airplane is flying over the mountain.
Hikouki ga yama no ue wo tondeiru.

However, the Japanese translation of over ‘no ue wo’ is also used in the prototypical meaning of above and on. Overlapping the same translations with different prepositions should hinder learners’ understanding of how to appropriately select meanings based on the context. Thus, it will take longer to find the appropriate meaning according to the context that users encounter. It is also true that the proper use of the meaning would be difficult to determine even if all the translated definitions had been memorized. Therefore, focusing on the words might make the effectiveness of the visual glosses more explicit.

Secondly, the visual glosses were modified. As shown above, our previous studies used two types of visual gloss—still images and animated images, both of which are only schematic images. This study speculates that such images might not facilitate learners’ figurative understanding given the process of categorization [25] addressed this by stating that a schema is instantiated by a
prototypical member, then extended to non-prototypical members (e.g. the figurative members). That is, the schema would help learners to understand figurative meanings when the prototypical one is also represented.

Therefore, this research replaced the original animated images with still pictures [43]. The following figures, which have now been released on YouTube, are an example of the animated images used in this study. The animated images firstly display the picture of a prototypical sentence (where the balloon is flying over the mountain). The schematic image then depicts an object moving at the upper surface of the other object, namely the above-across schema [22] is foregrounded. As the animated images connect the prototypical meaning and the schematic image of over, the study hypothesizes that the L2 learners might easily understand the relationship between them, leading to a better understanding of prepositions (both literal and figurative meanings).

Finally, several tests were added to this research. Our previous studies conducted post-tests soon after the learners were given the images and had learned the target vocabulary using the images. The findings were obtained just after the intervention, so this study conducts not only a post-test but also a delayed test that was conducted one week after the intervention. In this study, another type of test is also conducted. Our previous studies focused on comprehensive tasks like a fill-in-the-blank or comprehension test, but this study also included a production test; here, the learners were asked to write English sentences using the target words. In addition, the method of analyzing the words has been modified. While the previous studies analyzed the data collectively, this study conducts analysis per word and meaning type to ascertain the difference between the words and between each word’s physical and figurative connotation.

By modifying the target words, the animated images, and the tests, this study verifies the efficacy of the image schemata as visual glosses in explicit L2 learning for spatial prepositions based on the following research questions.

6. Research questions

As our previous studies [39,41] examined the efficacy of the image schema as pictorial glosses for explicitly learning English prepositions, this study validates the effectiveness of animated glosses by comparing them with pictorial glosses. In concrete terms, the research focuses on whether static pictorial glosses or animated glosses is more effective for learning English spatial prepositions. The research questions below will thus be examined in the next section.

1. Can Japanese L2 learners utilizing animated glosses comprehend sentences with the target prepositions (above, on, over) better than those using pictorial glosses?
2. Can Japanese L2 learners utilizing animated glosses produce more sentences with the target prepositions than those using pictorial glosses?
3. Can visual glosses (either pictorial or animated) enhance both learners’ ability to understand and produce the target prepositions appropriately?

7. Research

7.1. Participants

Fifty-six undergraduate Japanese students (49 male and 7 female) from the author’s university participated in this research. All the participants were freshmen and sophomores in the Faculty of Technology (one group is from the Mechanical Engineering department and the other from the Physics department) and were students in my English classes, which is one of their compulsory subjects. As their English classes are separated according to their department rather than their proficiency, their English competencies should be randomized. One class from the Mechanical Engineering department (n = 30) was designated as a control group while the class from the Physics department (n = 26) was designated as the experimental group. Their scores in the G-TELP test (level 2 out of five levels, equivalent to 600 in TOEIC® test) confirmed the standardization of their English competence. A t-test
of their scores (grammar, listening, and reading sections) demonstrated no significant difference between the groups ($t (58) = -0.92, p > 0.05$).

7.2. Material

This research focuses on three spatial prepositions, above, on, and over. This is because all of these refer to the upper side of an object and are translated into “no ue ni” prototypically in Japanese. This means that it would be difficult to use the words efficiently for text comprehension and sentence production without instructing the learners on how each preposition differ from the others. The control group used pictorial images of the three prepositions that depict schematic images that conceptualize their prototypical meaning, such as the image schema. The control group used pictorial images extracted from an English-Japanese dictionary [43]. The experimental group, on the other hand, used the animated images uploaded on YouTube; these images were released on an English-learning website (COCONE) under the supervision of the chief editor of the English-Japanese dictionary (as seen in Fig. 6). As each image used for this research is provided by the same person under the same policy, both groups are shown the same images, although the configuration of the images differs.

7.3. Procedure

First of all, each participant in both groups was asked to answer 19 questions regarding the prepositions targeted by the research within the 7 min. The preposition test (pre-test) included four prepositions that consist of above (five questions: one physical and four figurative senses), on (five questions: two physical and three figurative senses), over (seven questions: three physical and four figurative senses), and across (one question: one physical and one figurative sense). These questions were developed with reference to the words’ meaning index from a bilingual dictionary [43]. As this dictionary is compiled from Cognitive Linguistics perspective, the physical and figurative meaning in the index are clearly distinguished. The words above, on, and over are, as shown above, the target vocabulary for this research while across serves as a distractor to prevent the unintentional choice of the correct word. Each question consists of an English sentence with its Japanese translation. With reference to the translation, the participants were asked to select the most appropriate preposition according to the context. After completing the pre-test, the students started the next activity; no feedback was given to the participants. After completing the pre-test, they were asked to look at three schematic images displayed on a screen located at the front of the classroom. They had to view the images with their peers because personal computers and displays for each participant could not be prepared for this research. For the control group, the static images were displayed for 1 min with the sound of a prototypical sentence of the preposition. For the experimental group, on the other hand, the animated images were displayed for 1 min with the sound of the prototypical sentence of the preposition. As each image was displayed for 1 min, the time it took for all the participants to view all the images amounted to 3 min.

Immediately following the display of all the visual images, another preposition test (post-test) was conducted. The questions on the post-test were the same as on the pre-test, consisting of 19 questions about the four prepositions, but the order of the questions was randomized. The participants were asked to answer the questions within the same time limit as for the pre-test. As with the pretest, no feedback was given to the participants after the post-test.

Next, a sheet of paper was distributed, and each participant was asked to write as many English sentences using the target prepositions as they could on the sheet. During the task, it was not permitted to use reference materials such as dictionaries. Collecting all the answer sheets from the participants, the experimenter (the author) told them that this task was the final activity of the tasks conducted at that day and did not inform them that further tasks would be held during the next class. This prevented participants from reviewing the prepositions they had learned or referring to the reference from which the images had been taken (nor did the author tell the participants the sources of the images).

One week later, during the next class, the final preposition test (delayed test) and production test were conducted. The author expected that no participant had prepared for the tests because he had not provided any information about the test in the previous class. The delayed test consisted of the same questions as the pretest and post-test, but the order of questions was again randomized. The method of answering the questions and the time limit were also the same as on the previous tests. After the students completed the delayed test, they were given a sheet of paper and then asked to write as many English sentences including the target prepositions as possible within the same time limit as for the previous production test. Just as in the previous test, it was not
permitted to use references during this task. After the production test, all the answer sheets for both the delayed preposition test and production test, correct answers for the test, and the references to the images were given. This concluded the research.

All the data collected from the participants were analyzed. The correct answers for each preposition test were statistically analyzed. In the reading test, one correct answer of each fill-in-the-blank test is graded as one point, so the number of the correct answers is equivalent to the participant’s total score. Regarding the production tests, two scorers (a Japanese EFL instructor and the author) examined the criteria for grading and then began marking individually. After completing the task, the scores given by each were combined to calculate the average scores for each participant. The average scores were used for the analysis of the research. The analysis for this research was conducted mainly to examine the differences between the control and experimental groups that were demonstrated in my research questions; for this purpose, a t-test and ANOVA were utilized.

### 7.4. Results

First of all, the results of the pre-test are demonstrated. The average score of the control group was 9.87 (SD = 0.30), whereas that of the experimental group was 11.15 (SD = 0.49). The results of the post-test, conducted after the intervention, are then demonstrated. The average score of the control group was 11.00 (SD = 2.18) while that of the experimental group was 11.27 (SD = 2.20). As a result of the t-test to examine the difference in the average of the groups, no significant difference was found (t(54) = 0.46, p > 0.05). As for the delayed test’s results, the average score of the control group was 9.93 (SD = 0.48) while that of the experimental group was 10.35 (SD = 0.47). The t-test to examine the difference between groups found no significant difference (t(54) = 0.61, p > 0.05). In summary, no significance was observed between the groups in any tests.

Next, the accuracy rate of the post and delayed tests’ questions are illustrated for each target preposition in Fig. 7 below. For the control group, the accuracy rate of the post-test for questions with above was 61%, on was 77%, and over was 33%, whereas the experimental group demonstrated an accuracy rate of 59% for above, 80% for on, and 37% for over. Comparing the accuracy rate per question between the groups (as shown in Table 1), the experimental groups obtained a higher accuracy rate in 3 out of the 5 questions on above, in 3 out of the 5 questions regarding on, and in 6 out of the 7 questions on over.

A comparison of the accuracy rate per question between groups was also conducted in the delayed test. Fig. 7 shows for the control group, the accuracy rate for the above questions was 55%, for the on questions was 66%, and for the over questions was 30%. Meanwhile, for the experimental group, the accuracy rate for the above questions was 61%, for on was 65%, and for over was 36%. As a result of comparison of the accuracy rate per question between the two groups shown in Table 2 below, the experiment group obtained the higher accuracy rate in 3 out of 5 above questions, 2 out of 5 on questions, and 5 over questions.

Next, the accuracy rate for questions with physical and figurative meanings of each preposition is compared in the post-test. In the control group, the accuracy rate for questions with physical senses was 59%, and for those with figurative sense was 47%. Meanwhile, in the experimental group, the accuracy rate for physical senses was 56% and for figurative senses was 51%. As in the previous analysis, the accuracy rate of the questions per question was calculated. The results shown in Table 1 demonstrate that the experimental group obtained higher accuracy rates in 4 out of 7 questions that referred to physical spatial relations, while they obtained higher accuracy rates in 9 out of 10 questions regarding the figurative spatial relationship. However, the t-tests per target preposition (total scores/scores of the figurative sentences) show no significant difference between the groups.

Next, the accuracy rate for physical and figurative meaning between groups was then compared in the delayed test. For the physical senses, the accuracy rate of the control group was 54% while that of the experimental group was 57%. Regarding figurative senses, on the other hand, the accuracy rate of the control group was 40% whereas that of the experimental group was 44%. Table 2 shows the result of the comparison per question between the groups, where the experimental group answered correctly 3 out of 7 questions with physical senses and 7 out of 10 questions with figurative senses. In the post-test, the experimental group also obtained a higher accuracy rate on questions including those on figurative meanings. Along with the results of the t-test in the post-test, however, no significant difference was found between the groups.

### 7.5. Production task

Just as the results of the fill-in-the-blank tests have been
discussed in the previous section, the results of the production tasks are here examined. In this task, the participants were asked to write sentences using the target prepositions within 5 min without using reference materials like dictionaries. The sentences were scored according to criteria discussed by the two graders in advance. More specifically, a grammatically correct sentence with a literal preposition use was given one point, whereas that with the figurative usage received an additional point. The average score of the graders was used for this research.

Fig. 8 shows the result of the production tasks conducted immediately after and one week after projecting the visual images. In the post-test, the average score of the control group was 4.17 (SD = 1.82) while that of the experimental group was 4.85 (SD = 1.74). As a result of the t-test, no significant difference was found between the groups (t(54) = 1.42, p > 0.05).

Among the scores of the production task, the scores of sentences with figurative senses were extracted and analyzed by t-test. The average score of the control group was 0.20 (SD = 0.66) whereas that of the experimental group was 0.77 (SD = 1.31). As a result of the t-test, a significant difference was found between the groups (t(54) = 2.01, p < 0.05). The statistical power (α = 0.05, two-sided) was 0.54.

Furthermore, the results of the delayed production task conducted one week after the intervention is examined as shown in Fig. 8 above. In the delayed test, the average score of the control group was 4.33 (SD = 1.90), and that of the experimental group was 4.46 (SD = 1.52). As a result of the t-test, a significant difference was found between the groups (t(54) = 1.42, p > 0.05).

Table 1
Accuracy rate for each question in the post-test.

<table>
<thead>
<tr>
<th>#</th>
<th>Correct answer</th>
<th>Type of meanings</th>
<th>Accuracy for control</th>
<th>Accuracy for experimental</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Above</td>
<td>Physical</td>
<td>0.73</td>
<td>0.56</td>
<td>-0.18</td>
</tr>
<tr>
<td>2</td>
<td>Above</td>
<td>Physical</td>
<td>0.63</td>
<td>0.59</td>
<td>-0.04</td>
</tr>
<tr>
<td>3</td>
<td>Above</td>
<td>Figurative</td>
<td>0.30</td>
<td>0.44</td>
<td>0.14</td>
</tr>
<tr>
<td>4</td>
<td>Above</td>
<td>Figurative</td>
<td>0.50</td>
<td>0.52</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>Above</td>
<td>Figurative</td>
<td>0.47</td>
<td>0.48</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>Across</td>
<td>Physical</td>
<td>0.33</td>
<td>0.30</td>
<td>-0.04</td>
</tr>
<tr>
<td>7</td>
<td>Across</td>
<td>Figurative</td>
<td>0.10</td>
<td>0.04</td>
<td>-0.06</td>
</tr>
<tr>
<td>8</td>
<td>On</td>
<td>Physical</td>
<td>0.90</td>
<td>0.78</td>
<td>-0.12</td>
</tr>
<tr>
<td>9</td>
<td>On</td>
<td>Physical</td>
<td>0.90</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>10</td>
<td>On</td>
<td>Figurative</td>
<td>0.80</td>
<td>0.96</td>
<td>0.16</td>
</tr>
<tr>
<td>11</td>
<td>On</td>
<td>Figurative</td>
<td>0.30</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>12</td>
<td>On</td>
<td>Figurative</td>
<td>0.73</td>
<td>0.85</td>
<td>0.12</td>
</tr>
<tr>
<td>13</td>
<td>Over</td>
<td>Physical</td>
<td>0.17</td>
<td>0.19</td>
<td>0.02</td>
</tr>
<tr>
<td>14</td>
<td>Over</td>
<td>Physical</td>
<td>0.17</td>
<td>0.19</td>
<td>0.02</td>
</tr>
<tr>
<td>15</td>
<td>Over</td>
<td>Physical</td>
<td>0.53</td>
<td>0.70</td>
<td>0.17</td>
</tr>
<tr>
<td>16</td>
<td>Over</td>
<td>Figurative</td>
<td>0.03</td>
<td>0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>17</td>
<td>Over</td>
<td>Figurative</td>
<td>0.13</td>
<td>0.19</td>
<td>0.05</td>
</tr>
<tr>
<td>18</td>
<td>Over</td>
<td>Figurative</td>
<td>0.77</td>
<td>0.52</td>
<td>-0.25</td>
</tr>
<tr>
<td>19</td>
<td>Over</td>
<td>Figurative</td>
<td>0.53</td>
<td>0.63</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 2
Accuracy rate for each question in the delayed test.

<table>
<thead>
<tr>
<th>#</th>
<th>Correct answer</th>
<th>Type of meanings</th>
<th>Accuracy for control</th>
<th>Accuracy for experimental</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Above</td>
<td>Physical</td>
<td>0.77</td>
<td>0.54</td>
<td>-0.23</td>
</tr>
<tr>
<td>2</td>
<td>Above</td>
<td>Physical</td>
<td>0.63</td>
<td>0.69</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>Above</td>
<td>Figurative</td>
<td>0.23</td>
<td>0.50</td>
<td>0.27</td>
</tr>
<tr>
<td>4</td>
<td>Above</td>
<td>Figurative</td>
<td>0.63</td>
<td>0.62</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>Above</td>
<td>Figurative</td>
<td>0.47</td>
<td>0.69</td>
<td>0.23</td>
</tr>
<tr>
<td>6</td>
<td>Across</td>
<td>Physical</td>
<td>0.47</td>
<td>0.54</td>
<td>0.07</td>
</tr>
<tr>
<td>7</td>
<td>Across</td>
<td>Figurative</td>
<td>0.17</td>
<td>0.08</td>
<td>-0.09</td>
</tr>
<tr>
<td>8</td>
<td>On</td>
<td>Physical</td>
<td>0.97</td>
<td>0.88</td>
<td>-0.08</td>
</tr>
<tr>
<td>9</td>
<td>On</td>
<td>Physical</td>
<td>0.97</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td>10</td>
<td>On</td>
<td>Figurative</td>
<td>0.77</td>
<td>0.88</td>
<td>0.12</td>
</tr>
<tr>
<td>11</td>
<td>On</td>
<td>Figurative</td>
<td>0.23</td>
<td>0.19</td>
<td>-0.04</td>
</tr>
<tr>
<td>12</td>
<td>On</td>
<td>Figurative</td>
<td>0.37</td>
<td>0.27</td>
<td>-0.10</td>
</tr>
<tr>
<td>13</td>
<td>Over</td>
<td>Physical</td>
<td>0.20</td>
<td>0.19</td>
<td>-0.01</td>
</tr>
<tr>
<td>14</td>
<td>Over</td>
<td>Physical</td>
<td>0.80</td>
<td>0.54</td>
<td>-0.26</td>
</tr>
<tr>
<td>15</td>
<td>Over</td>
<td>Physical</td>
<td>0.17</td>
<td>0.31</td>
<td>0.14</td>
</tr>
<tr>
<td>16</td>
<td>Over</td>
<td>Figurative</td>
<td>0.17</td>
<td>0.31</td>
<td>0.14</td>
</tr>
<tr>
<td>17</td>
<td>Over</td>
<td>Figurative</td>
<td>0.17</td>
<td>0.31</td>
<td>0.14</td>
</tr>
<tr>
<td>18</td>
<td>Over</td>
<td>Figurative</td>
<td>0.17</td>
<td>0.31</td>
<td>0.14</td>
</tr>
<tr>
<td>19</td>
<td>Over</td>
<td>Figurative</td>
<td>0.40</td>
<td>0.58</td>
<td>0.18</td>
</tr>
</tbody>
</table>
As in the first production task, the scores of the sentences with figurative senses were extracted and then analyzed by t-test. The average score of the control group was 0.93 (SD = 1.31), and that of the experimental group was 1.50 (SD = 1.65). According to the t-test, no significant difference was found between the groups (t(54) = 1.43, p > 0.05).

Finally, to examine the main effect and interaction of the treatments in the production task, which have not been conducted with our previous studies, a two-way ANOVA was conducted. The factorial design for this analysis is two factors (pictorial image and animated image) multiply (1) two factors (post-production test and delayed production test) and (2) two factors (post-production test and delayed production test focusing on figurative senses).

First of all, the result of (1) illustrates no significant difference between groups both in within-subject factors and between-subject factors (see Table 3). Also, no interaction effect was found. Furthermore, the result of (2) is illustrated, as seen in Table 4 below. A significant difference was found in within-subject factors. This means that cross-interaction was found between the post-production test and the delayed production test, which indicates that both images brought about some impact on the results of the production tests when the sentences with figurative senses were extracted. However, no interaction effect was observed.

8. Findings and conclusion

This study has discussed the impact of animated image schema as visual glosses in explicit L2 vocabulary learning by focusing on spatial prepositions for Japanese L2 learners, whose linguistic background is quite different from English. Unlike previous studies on L2 vocabulary learning using technology-enhanced visual glosses, our study relies on CL as a theoretical framework for the image schemata for understanding polysemous words’ semantic structures. Therefore, our study focused on three locative prepositions (above, on, and over), all of which refer to the upper side of an object and are translated into “no ue ni (wo)” in Japanese. This study also differs from our previous research on the following points. Firstly, the animated images were replaced with the ones displaying both the schematic and prototypical images. Furthermore, both receptive and production tests were conducted.

This study presupposed that the explicit learning using animated glosses can better facilitate learners’ textual comprehension and production which is based on the literature’s report on visual glosses’ positive impact on incidental vocabulary learning (e.g. [7,35,48,49]). Although this study finds no significant difference in some tests, this research has discovered new findings that were not seen in our previous studies.

I will now use the findings to address the three research questions.

1. Can Japanese L2 learners utilizing animated glosses comprehend sentences with the target prepositions (above, on, over) better than those using pictorial glosses?

| Table 3 |
| Two-way ANOVA (1) (groups and scores of the production tasks). |
| Source | SS | df | MS | F-ratio | p-value | p.eta 2 |
| A | 4.54 | 1 | 4.54 | 1.38 | 0.24 ns | 0.01 |
| B | 0.33 | 1 | 0.33 | 0.10 | 0.75 ns | 0.00 |
| A × B | 2.11 | 1 | 2.12 | 0.64 | 0.43 ns | 0.00 |
| Total | 363.5625 | 108 | 3.30 |
| Error | 356.68 | 108 | 3.30 |

*p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

This study could not answer this question since the advantage of the animated glosses was not be found in all receptive tests. However, higher accuracy rate could be seen when we focus on the questions about over, which can be one of the most polysemous words. In spite of this, no significant difference was also found even though other t-tests were conducted for each target preposition.

2. Can Japanese L2 learners utilizing animated glosses produce more sentences with the target prepositions than those using pictorial glosses?

Although no significant difference is found in the post and delayed production tests, a significant difference is found in the post-production test for questions on figurative meanings. Therefore, animated glosses might have some positive impact on the text production, particularly with words’ figurative connotations. This finding partly supports our claim based on [25]; where the representation of the image schema with the prototypical meaning could facilitate learners’ understanding of words’ non-prototypical meaning.

3. Can visual glosses (either pictorial or animated) enhance both learners’ ability to understand and produce the target prepositions appropriately?

As a result of two-way ANOVA, a significant difference is recognized in the within-subject factors in the production tasks. With this outcome, this study notes that the image schema can facilitate learners’ sentence production with target prepositions regardless of the image type. Furthermore, a marginal significance is found in the between-subject factors in the production tasks when only the sentences with figurative senses are graded. These findings could also answer the research question, but there are some limitations.

In conclusion, this study could discover a new finding that was not revealed in our previous studies in explicit L2 vocabulary learning with technology enhanced visual glosses: animated schematic images could enhance learners’ sentence production including the target prepositions, particularly with the figurative senses. In addition, animated images might be effective, particularly in comprehending and producing sentences with over, which has more meanings and is therefore harder to learn than other prepositions. However, no overall significant difference was found between the pictorial and animated images as our previous studies illustrated.

These findings show the pedagogical implications for L2 vocabulary acquisition with visual glosses. The result shows that there is no significant difference in the learning effect between pictorial glosses and animated glosses, which could indicate that students using technological aids may not always receive better learning outcomes than those using traditional materials. On the other hand, the more positive learning effect in the over questions when using animated glosses might imply that technology may affect learning
References


