Glosses, comprehension, and strategy use

Myong Hee Ko
Seoul National University

Abstract

This study, using both qualitative and quantitative measures, investigates how different types of gloss conditions affect Korean college students' reading comprehension. One hundred and six undergraduates at a university in Korea participated in the study. Twelve were assigned to think aloud, and the rest (ninety-four) took part in the main study only. They read the material under one of three conditions: no gloss, Korean gloss (L1 glossing), and English gloss (L2 glossing). After reading, they were asked to take a multiple-choice reading comprehension test and to answer a questionnaire. The results of the quantitative analyses indicated that only the second language (L2) gloss condition significantly affected students' reading comprehension. However, the think-aloud protocols revealed that both types of glossing made their reading comprehension smoother and faster than was possible for those who read without glosses. In other words, even first language (L1) glosses enabled them to comprehend more easily while reading, although statistics did not indicate a significant difference between the no gloss and L1 gloss conditions. When surveyed, learners showed their preference for glosses in the margin: more than 62% of the learners favored L2 glosses for their reading material.

Keywords: gloss, strategy, reading comprehension, think-aloud, Korean glosses, English glosses

Introduction

Marginal glossing is one way to help a learner comprehend reading materials. By offering additional notes or information beyond the text in the margin on the same page or on another page, glosses guide the learner and assist as a mediator between the text and the learner. Glosses have various functions in helping to decode the text by providing additional knowledge in specific content, skills, strategies, and definitions of difficult words. In the case of second language (L2) learning, gloss generally means information on important words via definitions or synonyms (Nation, 2001; Richgels and Mateja, 1984; Stewrat and Cross, 1991; Stewrat and Cross, 1993).
The two most important reasons to use glosses are to assist reading comprehension and aid vocabulary learning. In general, four advantages result from glossing. First, glosses can help readers understand new words more accurately by preventing incorrect guessing. Deriving meaning from context can be difficult and risky because of readers’ lack of language or reading strategies (e.g., Bensoussan and Laufer, 1984; Hulstijn, 1992; Kruse, 1979; Nation, 2001; Stein, 1993). Second, glossing can minimize interruption while reading is in process. Since glossing provides definitions for low frequency words, L2 readers do not have to constantly look them up (Nation, 1990; Nation, 2001). Third, glosses may help readers build a bridge between prior knowledge or experience and new information in the text. In other words, interactions among gloss, reader, and text may promote comprehension and retention of the content of the text. Besides these points, glosses in key words can help readers recall their background knowledge and connect it to the text (Stewrat and Cross, 1993). Fourth, glosses can make students less dependent on their teachers, allowing for greater autonomy. Since not all students have problems with the same words, they can look up just the words they do not know (Jacobs, 1994; Nation, 1990). Some studies have shown that students prefer to have glosses in their L2 language reading materials (e.g., Jacobs, Dufon and Fong, 1994). At any rate, the use of vocabulary glosses in L2 reading materials is a common practice (Holley and King, 1971; Jacobs, et al., 1994).

There have been some studies done on the effect of glossing in enhancing L2 language reading comprehension. Holley and King (1971), Johnson (1982), Jacobs et al. (1994) showed no significant effect for glossing in L2 reading comprehension, whereas Davis (1989) and Jacobs (1994) demonstrated that glosses can enhance it. Bell and LeBlanc (2000) compared use of first language (L1) glosses to use of L2 glosses and reported no significant difference between them in their effect on L2 reading comprehension.

Holley and King (1971) compared different types of glosses. U.S. students learning German (N=110) were randomly assigned to read the same passage but with different types of glosses: side-of-page; bottom-of-page; or glosses on an attached sheet. The subjects took multiple-choice vocabulary and reading comprehension tests. The results indicated that the three different forms of glossing, when placed at different positions in the text, did not make a significant difference on either vocabulary or reading comprehension tests.

Johnson (1982) was primarily interested in the role of cultural schema – knowledge about Halloween in this case – on L2 reading comprehension. She also tested the effect of the treatment of vocabulary under four different conditions: no help with vocabulary; studying the definitions of words before reading; reading a passage with the target words glossed; and studying the target words before reading and reading the text with glosses. A cloze test was administered to measure retention of the target words. After reading the passage, seventy two advanced ESL university students were asked to recall the story in their L2 and to recognize exact sentences from the passage. The dependent measures – cloze test, recognition task, and written recall protocol – indicated that knowledge of Halloween facilitated reading comprehension, but exposure to the different types of vocabulary aids did not significantly affect the comprehension.
Jacobs et al. (1994) investigated the effects of L1, L2, and no gloss on foreign language reading comprehension along with foreign language vocabulary learning. Native speakers of English enrolled in the fourth semester of Spanish (N=85) were assigned to one of three different gloss conditions: no gloss, English (L1) gloss, and Spanish (L2) gloss. After reading a passage written in Spanish with different gloss conditions, they were asked to recall the story. The findings again showed that there was no significant difference among the different conditions, and participants expressed their preference for L2 glosses.

In contrast, Davis (1989) and Jacobs (1994) indicated that glossing resulted in a positive effect on L2 reading comprehension. Davis (1989) tested whether marginal glosses would improve comprehension of a text. U.S. students in a French class (N=71) were divided into three text conditions and read a short story. The first condition was a read-write-reread group, in which participants read the passage for fifteen minutes, wrote what they could remember for ten minutes, and then reviewed the passage for five minutes. The second condition was to offer a vocabulary guide before reading, in which they were guided by questions and comments on the story and given definitions of some vocabulary. They studied the guide for ten minutes, read for fifteen minutes, and then wrote. The third condition was a vocabulary guide during reading (gloss): they were given the same questions and definitions of the same vocabulary in glossed form. They read the text for twenty-five minutes, and then wrote. Subjects who received vocabulary help either before or during reading did significantly better than those who received no help. In other words, it showed that those who read a text with glossing performed significantly better than those without glossing.

Jacobs’ (1994) study showed a positive effect between glossing and foreign language reading comprehension. U.S. students learning Spanish (N=166) were randomly assigned to read two types of Spanish reading texts, that is, with no gloss and with English (L1) glossing. Subjects were asked to write in their L1 as much as they could recall of the content of the story. The results showed that the glossed group performed significantly better than the other group.

Bell and LeBlanc (2000) investigated the type of glossing frequently used for computer-based reading. U.S. undergraduates (N=40) enrolled in third semester Spanish were divided into two groups: L1 gloss and L2 gloss. They were asked to read a short story and then to take a multiple-choice reading comprehension test. The findings showed there was not a significant difference between the L1 and L2 group, although the participants preferred L1 glosses over L2 glosses.

The previous studies mentioned above on glossing in L2 reading comprehension are diverse, and the outcomes are inconclusive on the effect of glossing for L2 reading comprehension. Moreover, there have been only two studies which compared the effect of L1 and L2 glosses in L2 reading comprehension. Jacobs et al. (1994), and Bell and LeBlanc (2000) reported that there was no statistical difference between L1 and L2 gloss types measured by either recall or multiple-choice reading comprehension tests, respectively. However, Jacobs et al. (1994) mentioned that their participants preferred L2 glosses, whereas Bell and LeBlanc (2000) indicated that their participants favored L1 glosses.

Additional studies are required to provide further information on the effect of glossing on L2 reading comprehension. Also, more empirical evidence on the effect of different types of glosses

http://nflrc.hawaii.edu/rfl
(L1 versus L2) is needed along with reader preference. Accordingly, this study compares a non-gloss condition with a gloss condition, as well as L1 gloss versus L2 gloss conditions. It also reports L2 reader preference on two different types of glosses. Moreover, this study investigates the kind of strategies foreign language readers use while reading. Three different gloss conditions are examined using a think-aloud procedure in order to gain information on their use of strategies. This approach enables examination of participant behavior, both qualitatively and quantitatively. The following research questions were asked to address these issues:

1. Is reading comprehension affected by gloss type?
2. Are reading strategies affected by gloss type?
3. Which type of glossing do learners prefer?

**Method**

**Participants**

In total, 106 undergraduates at a university in Korea participated. Their reading abilities appeared to belong to either an intermediate or a high-intermediate level according to the results of cloze testing. They were enrolled in the second semester of a freshman English class, a reading course that met twice a week for 50 minutes. The course was a requirement for graduation. The students were mostly freshmen, majoring in either liberal arts or social science. They were a homogeneous group in terms of age and English education background. In general, their ages ranged from 19 to 21 and they had all received English instruction for six years before entering the university. None of them had lived in English speaking countries prior to entering the university.

**Materials**

An article, "The Cellist of Sarajevo", from the advanced level of The NorthStar series was used (Miller and Cohen, 1998). A piece of non-fiction originally published in *Reader's Digest*, it is a moving story about a cellist in Sarajevo who played his cello on the street while bombs and bullets flew during the war in 1992. The length of text was 931 words and its readability was 10.4 on the Flesch-Kincaid Grade Level readability scale. In other words, the appropriate level was for native speakers who are in the fourth month of the tenth grade. Twenty-two words in the text were glossed since more glossed words in a two page text was deemed too much for students to learn. The words were selected based on two considerations. The first was degree of difficulty. A pilot test was conducted where the identical text was given to students with the same background as those in this study. They were asked to mark the words they found difficult or did not understand. Words marked most frequently were selected for glossing. Nineteen words were selected using this method of frequency of marks. Naturally, the students in this pilot test were not included in the main study as they were exposed to the text before the experiment. The second consideration was degree of importance, i.e., keywords deemed critical to understanding the text based on my professional experience in teaching reading to Korean university students. Three words were selected in this way. All the target words appeared once except for three words, *unaccompanied*, *shell* and *massacre* as shown in Table 1.

http://nflrc.hawaii.edu/rfl
Any type of production test was avoided because the reading time was limited to 15 minutes and judge the plausibility of the distracters. Then each item and distracter was revised and refined had experience in writing multiple-choice questions were consulted to check each item and to choices. Questions were matched to all parts of the passage so that the test could check for the reading (see Appendix B). They were expected to choose a correct answer among four A multiple-choice reading test consisting of 25 items in English was given to the students after the reading (see Appendix B). They were expected to choose a correct answer among four choices. Questions were matched to all parts of the passage so that the test could check for overall understanding of the story. Two native speakers who were teaching at the university and had experience in writing multiple-choice questions were consulted to check each item and to judge the plausibility of the distracters. Then each item and distracter was revised and refined twice after piloting the reading comprehension test to students who had similar backgrounds. Any type of production test was avoided because the reading time was limited to 15 minutes and the story was quite long. It was thought students would have difficulty if they were asked to write a summary of the story.

### Table 1: Target Words and Korean and English Glosses

<table>
<thead>
<tr>
<th>Target Words</th>
<th>Korean Glosses</th>
<th>English Glosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unanimous</td>
<td>모의감사</td>
<td>Modest, not attracting attention</td>
</tr>
<tr>
<td>Unanimousness</td>
<td>모의감사</td>
<td>instrumentally Played alone</td>
</tr>
<tr>
<td>Superannuated</td>
<td>수리상감사</td>
<td>Retired</td>
</tr>
<tr>
<td>Mortar</td>
<td>머트어</td>
<td>Artillery, gun</td>
</tr>
<tr>
<td>Shell</td>
<td>-shell</td>
<td>Armor, bullet</td>
</tr>
<tr>
<td>Carnage</td>
<td>칼리카지</td>
<td>血海, killing many people</td>
</tr>
<tr>
<td>Massacre</td>
<td>массаж</td>
<td>Killing many people</td>
</tr>
<tr>
<td>Current attitude</td>
<td>현재의 조건</td>
<td>formal suit (clothes) worn by man</td>
</tr>
<tr>
<td>Century</td>
<td>세기</td>
<td>A round hole made by an explosion</td>
</tr>
<tr>
<td>Merciful</td>
<td>미친, 순수</td>
<td>Very kind</td>
</tr>
<tr>
<td>Cellar</td>
<td>지하실</td>
<td>A room under a house, or building</td>
</tr>
<tr>
<td>Composition</td>
<td>조성</td>
<td>A piece of music</td>
</tr>
<tr>
<td>Payback</td>
<td>보상</td>
<td>Quit</td>
</tr>
<tr>
<td>Power</td>
<td>힘</td>
<td>Great anger</td>
</tr>
<tr>
<td>Mocking</td>
<td>모개</td>
<td>Signaling someone to come toward you</td>
</tr>
<tr>
<td>Hung</td>
<td>굶는다</td>
<td>Throw with great power</td>
</tr>
<tr>
<td>Tainted</td>
<td>오염</td>
<td>Old and torn</td>
</tr>
<tr>
<td>Profound</td>
<td>깊은</td>
<td>Deep</td>
</tr>
<tr>
<td>Cruelling</td>
<td>구울링</td>
<td>Apple to apple in English</td>
</tr>
<tr>
<td>Shrewish</td>
<td>흉기</td>
<td>Make someone feel less wanted or angry</td>
</tr>
<tr>
<td>Serene</td>
<td>평온한</td>
<td>A French military leader who fought against the English</td>
</tr>
</tbody>
</table>

The reading text was adapted into three different forms: a text with no gloss; a text with L1 glosses with Korean definitions or synonyms; and a text with L2 glosses with English definitions or synonyms at the bottom margin of the page (see Appendix A). Two native Koreans who majored in English and two English native speakers checked the L1 and L2 definitions in order to obtain the most appropriate definitions of the target words for the context. The reading text was two pages long with eleven target words provided in the margin of each page. Each glossed item was boldfaced in the text.
A questionnaire was designed to gather information about students’ attitudes toward the presence of glosses and toward their preferences between two types of glosses. Samples of the two types of glosses were included for those who read the text with no gloss. In addition, a question regarding how much they looked at the glosses was included (see Appendix C).

Procedures

All 106 took a cloze test. The cloze test Watanabe (1997) used in his study was employed since my students had a similar background in terms of their English education and age. The passage contained 277 words in total, with 25 blanks.

Afterward, they were split into two groups: the quantitative study and the qualitative study. Ninety-four were selected for the main study (quantitative) and twelve for the qualitative part. The ninety-four students were randomly assigned to one of three groups: no gloss, L1 gloss, and L2 gloss. Although the participants were randomly assigned, a one-way ANOVA was carried out to determine whether there were any significant differences among the three groups at the beginning of the study. The absence of any significant differences among the three groups confirmed that their level of proficiency was approximately equivalent. They were requested to read the text for 15 minutes, after which the reading materials were collected by the researcher. Then, they were given 15 minutes to take the comprehension test. Once the test materials were collected they were asked to complete the questionnaire.

The twelve think-aloud participants were selected based on the result of a cloze test – strictly 19 out of 25 – and their achievements in the reading class. Those with similar English proficiency were chosen because reading strategies can vary depending on L2 proficiency. The twelve were randomly assigned to one of the three conditions (four participants per condition). They were chosen because the level of the text seemed to be challenging for their level. Low level students were not chosen because they would not use many strategies if the level of text were too difficult for them. In other words, if they were overwhelmed with the difficulty of the text, they would probably simply rely on guessing strategies. High level students were avoided because they would probably not use many strategies if the level of text were too easy for them. In this case, their reading process might be so automatic that even the readers would not notice they were using any strategies.

Before doing the actual think-aloud task, the procedure was clearly explained to the students and they all had a chance to practice with a different text. Their think-aloud protocols were recorded and transcribed verbatim. Their reading comprehension test scores were not included with the scores of the main test group because they had more than 15 minutes to complete the reading assignment. The additional time might have enabled them to encode the overall story more deeply into their memories. Furthermore, the reading scores in the think-aloud group were not analyzed separately because everybody took a different amount of time in thinking aloud and, thus, it was thought to be invalid to compare their levels of reading comprehension. The transcripts were evaluated qualitatively by the investigator and a colleague of another school in Seoul. The interrater reliability is discussed in detail in the results section.
Results

Quantitative analysis of the main study was conducted using the SPSS PC+ Version 10.0 software package (Norusis, 1990); qualitative analysis was carried out on the 12 think-aloud data. The alpha level for the whole study was set at 0.05. The total possible score on the reading comprehension test was 25 points. The reliability of the reading comprehension test was 0.61 measured by K-R 20. Table 2 shows descriptive statistics for the reading comprehension test for the three groups.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>No gloss</td>
<td>31</td>
<td>19.58</td>
<td>3.52</td>
</tr>
<tr>
<td>L1 gloss</td>
<td>30</td>
<td>20.90</td>
<td>2.14</td>
</tr>
<tr>
<td>L2 gloss</td>
<td>33</td>
<td>21.61</td>
<td>2.28</td>
</tr>
<tr>
<td>Total</td>
<td>94</td>
<td>20.71</td>
<td>2.81</td>
</tr>
</tbody>
</table>

The main effect for glossing was calculated. A one-way ANOVA was performed to investigate whether there was any significant effect of glossed conditions on reading comprehension. The result shows that the probability was significant (p <.05) between the groups. Sheffe's post hoc analysis was performed in order to locate specific significant differences among the three different gloss conditions. The Scheffe's test revealed that there was a significant difference between the no gloss condition and L2 gloss condition.

Next, qualitative analysis was performed with respect to the L2 think-aloud data. It involved reviewing the transcribed protocols and coding them for specific reading strategies. The investigator and a second rater went over the protocols and checked the participants' reading strategies and their understanding of the story. The taxonomy of reading strategies was based on Anderson (1991), Block (1986), and Grabe and Stoller (2002). Additional strategies noted among the Korean students, such as confirming prediction, deriving word meaning from context, elaborating content, and understanding main ideas, were also added to the reading strategy list. Table 3 shows the 19 strategies used for classifying the protocols. The two raters reached a consensus in case of conflicts in coding for reading strategies. The interrater reliability between the raters was 0.87.
Table 3: Reading Strategies

<table>
<thead>
<tr>
<th>Strategies</th>
<th>No gloss</th>
<th>L1 gloss</th>
<th>L2 gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirming prediction</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Commenting on one's own idea</td>
<td>4</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Deriving word meaning from context</td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Elaborating on content</td>
<td>11</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Wild guessing</td>
<td>15</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Understanding main idea</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Making inferences</td>
<td>2</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Monitoring comprehension</td>
<td>20</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Paraphrasing</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Predicting content</td>
<td>7</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Predicting subject of the text</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Questioning content</td>
<td>18</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Rereading</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Skipping</td>
<td>36</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Summarizing</td>
<td>9</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Using background knowledge</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Using discourse markers to see relationships</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Using syntactic clues</td>
<td>10</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Using morphological clues</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total Number of Strategy Used</td>
<td>154</td>
<td>68</td>
<td>60</td>
</tr>
</tbody>
</table>

When the no gloss and glossed conditions were compared, those who read the text under the no gloss condition used far more strategies in terms of number than those who read glossed texts, as shown in Table 3. The strategies employed were mostly related to situations when readers confront unknown words and unknown content. Hence, they made a lot of guesses, questioned the content, skipped difficult words, derived word meanings from context, and used syntactic clues. As Table 3 indicates, they frequently monitored their comprehension. The qualitative protocols revealed that they monitored constantly due to uncertain comprehension. In general, low level strategies, such as guessing, skipping, and questioning content, were used by those in the no gloss condition. Participants under the gloss conditions made relatively more inferences, belonging to high level reading strategies than those under the no gloss condition. It seems they were able to derive more and deeper meanings based on the facts since they understood the content fairly well.

Participants under the two glossed conditions were very similar in terms of the number of strategies they used. They used far fewer strategies than the participants under the no gloss condition, and they generally did not use strategies used by no gloss condition participants. One interesting difference was that those in the L1 gloss condition often employed the strategy of commenting on their own idea when compared to those in the L2 condition.

The survey finding showed that in general readers preferred to have glossed reading materials. As shown in Table 4 below, 94% preferred to have glosses: about 62% wanted L2 glosses, and
about 32% wanted L1 glosses. The most prominent reasons provided by the survey were that glosses facilitated understanding of the content, prevented long interruptions resulting from looking up words in the dictionary, provided the most appropriate definitions for the context, promoted learning new words as a by-product of reading, and prevented wrong guesses arising from insufficient contextual clues.

<table>
<thead>
<tr>
<th>Glosses</th>
<th>No gloss</th>
<th>L1 gloss</th>
<th>L2 gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>6%</td>
<td>32%</td>
<td>62%</td>
</tr>
</tbody>
</table>

**Discussion**

Overall, only the L2 gloss condition indicated significant effects on the reading comprehension test. The inconsistent results between previous research and this current study may be attributed to three factors: the importance of glosses in the context; the level of learner proficiency; and different types of assessment in reading comprehension.

First, the degree of importance of the target glosses in the context may be a reason for different results. If the target words are key words and very crucial in understanding the story, they may play an important role in understanding the content of the text. When selecting the target vocabulary, researchers need to consider this matter along with the frequency of marks of unknown words. In other words, solely determining the target words by the frequency of marks of unknown words is not enough because it may not always represent the most important words in a given context.

Second, the learners' level of proficiency may be another variable for different outcomes. If their level of proficiency is high enough to understand definitions in the target language, L2 glosses can be more effective than L1 glosses. In this study, more than 62% of the participants preferred L2 glosses according to the questionnaire. They opined that they could connect meanings more effectively in the context with L2 glosses than with L1 glosses. In other words, English could be explained more effectively and clearly through English than through their L1. In fact, participants who preferred L2 glosses had higher mean scores (M=21.4) than those who favored L1 glosses (M=19.5) on the reading comprehension test. On the other hand, if students' level of proficiency is relatively low and they cannot decode the definitions in the target language, then L2 glosses may not be very effective in promoting reading comprehension. In that case, L1 glosses may be more effective.

It is difficult to judge the proficiency level of participants from previous studies and the present study. Participants in Jacobs et al. (1994), and Bell and LeBlanc (2000) were in the third and fourth semester, respectively, in Spanish at universities in the U.S. The participants in the present study were in their second semester in English, but had had six years of English instruction before entering university. Thus, further study is necessary to investigate the influence of proficiency level on the effect of different types of glosses in reading comprehension.
Third, the nature of the reading comprehension test can affect reading comprehension results. It seems that currently available testing devices have some drawbacks. Davis (1989) argues that either cloze tests or target language protocols are less efficient measures compared to native language recall protocols (e.g., Johnson, 1982) because those procedures might have hindered the participants from expressing their understanding or revealing their thoughts precisely. However, L1 written protocols also seem to have limitations. There can be cases where students understand the text perfectly but are not good at expressing their thoughts accurately on paper. In other words, if they lack writing skills or are not expressive, they may not be able to express their thoughts overtly, just as there are students who cannot express their thoughts orally in their mother tongue. As with L2 cloze tests, L1 cloze tests also will not be able to judge reader understanding of the text accurately since the context and other written information may provide hints concerning the answers, and the number of blanks and other test format decisions can influence performance.

Moreover, the multiple-choice test also seems to have flaws. The possibility that the L1 condition did not significantly affect the reading comprehension test results in this study could be due to test method. The reliability of the multiple-choice reading comprehension test was more or less low (r=0.61). The low reliability suggests the possibility that the MC test might not have measured the participants' reading comprehension accurately and consistently. This may be a possible reason for the L1 gloss condition not showing any significant difference, whereas the L2 gloss condition indicated a significant effect on reading comprehension.

Bell and LeBlanc (2000) also pointed out that their ten-question multiple-choice test could have been an inadequate instrument to measure their participants' reading comprehension. Regarding one of the ten MC questions, only 17.7% of the L1 gloss group answered correctly, whereas 47.6% of the L2 gloss group was able to select the right choice. They pointed out that although the instrument was not the primary concern of their study, the result gave them reason for speculation. Thus, they also discussed the possibility that a choice of measurement can influence participant reading comprehension results. Hence, a consistent assessment needs to be considered in order to come up with more reliable results.

With respect to strategy use under the three different conditions, the think-aloud protocols revealed that there were some similarities between glossed and unglossed conditions. In general, the most widely used strategy for the three groups was, of course, verbatim translation. An interesting phenomenon was that participants tended to stick to one definition without paying attention to a given context. This might be due to their habitual use, before entering college, of a word list which treats English and Korean definitions as having a one-to-one relationship. It seems they are not aware of the fact that words can be translated in different ways depending on the context. In addition, when encountering unknown words, they tended to skip them without making a sufficient effort to use the contextual clues.

The think-aloud protocols also demonstrated some differences between glossed and unglossed conditions. Those who read the text without glosses used far more strategies than those who read the glossed texts. However, they were mostly low level strategies. They did a lot of guessing and skipping difficult words. They also employed a lot of questioning about the content when facing the difficult text, and monitored their comprehension constantly in order to
decode sentences. Those who read the glossed texts did not use as many strategies in terms of number and variety. They used high level strategy such as making inferences. Overall, they used far fewer guesses when encountering difficult words, instead basically translating the content sentence by sentence smoothly with help of the glosses in the margin. According to the survey, participants mostly did use the glosses. Of course, the protocols revealed that they sometimes encountered unglossed words that they did not know. In these situations they either guessed, tried to derive meanings from contextual clues, or simply skipped. However, there was evidence that they sometimes misinterpreted the sentence due to their lack of syntactic knowledge or general reading ability, even if the glosses were provided. This indicates that there are other variables in addition to difficult vocabulary that can obstruct reading comprehension. 

On the whole, think-aloud data revealed that glosses facilitated learners’ reading comprehension. Glosses evidently reduced the time needed to understand the content and showed smoother translation from English to Korean. Although there was no statistical difference under the L1 gloss condition, the results were still positive compared to the unglossed condition in terms of better mean scores. The think-aloud protocols also revealed that students' reading processes were smoother and faster. Most importantly, the L2 gloss significantly improved reading comprehension.

Conclusion

This study compared how three different gloss conditions affected Korean college students’ reading comprehension, using quantitative and qualitative methods. Multiple-choice reading comprehension test results showed that those who read the L2 glossed texts performed significantly better than those who read the text with either L1 glosses or without glosses. However, the think-aloud protocols of the representatives of each group revealed that although the statistics did not indicate a significant difference between the non-gloss and L1 gloss conditions, they definitely produced a different phenomenon. The transcribed think-aloud protocols demonstrated that those who used the glossed texts read more smoothly and faster with the help of the glosses and did not use as many low level strategies, such as guessing and skipping. Instead, they used more high level strategies such as making inferences. In short, glosses enabled them to comprehend more easily while reading. Thus, those who are in this field should not neglect the positive outcomes of glosses in reading materials.

However, there is a limitation in the qualitative analysis due to the small number of participants. More studies with a sufficient number of students in this area are required in order to draw better conclusions. At the same time, there should be more studies comparing the relationship between the gloss types and reading comprehension in order to discover the effects of the three variables mentioned in the discussion above in order to reach a more precise conclusion. Further research is needed to determine how L2 language learners' reading comprehension is influenced by the degree of importance of target words in context, the level of learner proficiency, and the nature of the reading comprehension test. Further investigations can build on the findings of this study to discover practical and valuable applications in teaching and learning L2 reading.
Acknowledgements

I would like to express my sincere gratitude to two anonymous reviewers for their insightful comments and suggestions on the earlier drafts of this article. I would also like to thank Yoon-ah Shin at Seoul National University for her gracious assistance. Any errors are mine.

References


**Appendix A**

As a pianist, I was invited to perform with cellist Eugene Friesen at the International Cello Festival in Manchester, England. Every two years a group of the world's greatest cellists and others devoted to that unassuming instrument – bow makers, collectors, historians – gather for a week of workshops, master classes, seminars, recitals and parties. Each evening, the 600 or so participants assemble for a concert.

The opening-night performance at the Royal Northern College of Music consisted of works for unaccompanied cello. There on the stage in the magnificent concert hall was a solitary chair. No piano, no music stand, no conductor's podium. This was to be music in its purest, most intense form. The atmosphere was supercharged with anticipation and concentration. The
world-famous cellist Yo-Yo Ma was one of the performers that April night in 1994, and there was a moving story behind the musical composition he would play.

On May 27, 1992, in Sarajevo, one of the few bakeries that still had a supply of flour was making and distributing bread to the starving, war-shattered people. At 4 p.m. a long line stretched into the street. Suddenly, a mortar shell fell directly into the middle of the line, killing 22 people and splattering flesh, blood, bone and rubble.

Not far away lived a thirty-five-year-old musician named Vedran Smailovic. Before the war he had been a cellist with the Sarajevo Opera, a distinguished career to which he patiently longed to return. But when he saw the carnage from the massacre outside his window, he was pushed past his capacity to absorb and endure any more. Anguished, he resolved to do the thing he did best: make music. Public music, daring music, music on a battlefield.

For each of the next 22 days, at 4 p.m., Smailovic put on his full, formal concert attire, took up his cello and walked out of his apartment into the midst of the battle raging around him. Placing a plastic chair beside the crater that the shell had made, he played in memory of the dead Albinonis Adagio in G minor, one of the most mournful and haunting pieces in the classical repertoire. He played to the abandoned streets, smashed trucks and burning buildings, and to the terrified people who hid in the cellars while the bombs dropped and bullets flew. With masonry exploding around him, he made his unimaginably courageous stand for human dignity, for those lost to war, for civilization, for compassion and for peace. Though the shellings went on, he was never hurt.

After newspapers picked up the story of this extraordinary man, an English composer, David Wilde, was so moved that he, too, decided to make music. He wrote a composition for unaccompanied cello, "The Cellist of Sarajevo," into which he poured his feelings of outrage, love and brotherhood with Vedran Smailovic. It was "The Cellist of Sarajevo" that Yo-Yo Ma was to play that evening.

Ma came out on stage, bowed to the audience and sat down quietly on the chair. The music began, stealing out into the hushed hall and creating a shadowy, empty universe, ominous and haunting. Slowly it grew into agonized, screaming, slashing furor, gripping us all before subsiding at last into a hollow death rattle and, finally, back to silence.

When he had finished, Ma remained bent over his cello, his bow resting on the strings. No one in the hall moved or made a sound for a long time. It was as though we had just witnessed that horrifying massacre ourselves. Finally, Ma looked out across the audience and stretched out his
hand, **beckoning** someone to come to the stage. An indescribable electric shock swept over us as we realized who it was: Vedran Smailovic, the cellist of Sarajevo!

Smailovic rose from his seat and walked down the aisle as Ma left the stage to meet him. They **flung** their arms around each other in an exuberant embrace. Everyone in the hall erupted in a chaotic, emotional frenzy – clapping, shouting and cheering. And in the center of it all stood these two men, hugging and crying unashamedly: Yo-Yo Ma, a suave, elegant prince of classical music, flawless in appearance and performance; and Vedran Smailovic, dressed in a stained and **tattered** leather motorcycle suit. His wild, long hair and huge mustache framed a face that looked old beyond his years, soaked with tears and creased with pain. We were all stripped down to our starkest, deepest humanity at encountering this man who shook his cello in the face of bombs, death and ruin, defying them all. It was the sword of **Joan of Arc** – the mightiest weapon of all.

Back in Maine a week later, I sat one evening playing the piano for the residents of a local nursing home. I couldn't help contrasting this concert with the splendors I had witnessed at the festival. Then I was struck by the **profound** similarities. With his music the cellist of Sarajevo had defied death and despair, and celebrated love and life. And here we were, a chorus of **croaking** voices accompanied by a **shopworn** piano, doing the same thing. There were no bombs and bullets, but there was real pain – dimming sight, crushing loneliness, all the scars we accumulate in our lives – and only cherished memories for comfort. Yet still we sang and clapped.

It was then I realized that music is a gift we all share equally. Whether we create it or simply listen, its a gift that can **soothe**, inspire and unite us, often when we need it most – and expect it least.

**composition**: a piece of music  
**hushed**: quiet  
**furor**: great anger  
**beckoning**: signaling someone to come toward you  
**flung**: threw with great force  
**soothe**: make someone feel less worried or angry  
**profound**: deep  
**croaking**: a harsh low sound  
**shopworn**: old and dirty  
**tattered**: old and torn  
**Joan of Arc**: a French military leader who fought against the English

---

**Appendix B**

Reading Comprehension Test

*Choose the correct answers.*

1. What is the main idea of the story?
   a. How terrible Sarajevo is.
   b. How music affects people.
   c. How great Yo-Yo Ma's performance is.
   d. How unusual Vedran Smailovic's life is.
2. Why was the author invited to the festival?
   a. To play the violin
   b. To help Yo-Yo Ma
   c. To help organize the festival
   d. To perform with cellist Eugene Friesen

3. Who went to the concert?
   a. A group of the world's greatest cellists
   b. World famous musicians and music fans
   c. People who are devoted to popular music
   d. World famous cellists and people devoted to this modest instrument

4. What is true about the opening-night performance?
   a. Only the cello was played.
   b. The conductor was excellent.
   c. There was everything except a piano.
   d. All kinds of musical instruments were played.

5. What was the atmosphere like at the festival on that night?
   a. The audience was a little gloomy.
   b. The audience's expectation was high.
   c. The audience was calm and peaceful.
   d. The audience was noisy and unfocused.

6. Why was the music Yo-Yo Ma played at the festival special?
   a. Because it was calm and beautiful.
   b. Because it was composed by Yo-Yo Ma.
   c. Because it was based on his life experience
   d. Because it was connected with a heart touching story

7. What is the best description of the street of Sarajevo in May 1992?
   a. People killed each other to get food.
   b. Bombs were dropping among people standing in lines.
   c. Starving people attacked most of the bakeries on the street.
   d. People could hear the sound of bombing but they were safe walking on the street.

8. What did Vedram Smailovic do for a living before the war?
   a. He was a leader of the local orchestra.
   b. He was a cellist with the Sarajevo Opera.
   c. He played cello with the famous local music band.
   d. He was a music professor at the University of Sarajevo.
9. What did Vedran Smailovic decide to do when he saw the carnage?
   a. He decided to make music.
   b. He decided to call Yo-Yo Ma.
   c. He decided to give up his music.
   d. He decided to perform at a concert.

10. How did Vedran Smailovic dress when he walked out of his apartment with his cello?
    a. He put on worn-out clothes.
    b. He put on a fancy wedding tuxedo.
    c. He put on a heavy gray coat and black pants.
    d. He put on a formal dark suit worn by musicians.

11. According to the text, what is the tone of Albinoni's *Adagio in G minor*?
    a. Very sad
    b. A little gloomy
    c. Somewhat light
    d. Very slow but fresh

12. Where did the terrified people hide while the bombs dropped and bullets flew?
    a. In the concert hall
    b. In a neighbor's attic
    c. In rooms under houses
    d. In the round hole made by an explosion

13. What did Vedran Smailovic do on the street for 22 days in May 1992?
    a. He played the cello alone.
    b. He played the cello with Yo-Yo Ma.
    c. He played the cello with other musicians.
    d. He helped poor people by giving away bread and playing the cello.

14. After newspapers published the story of Vedran Smailovic, what did David Wilde, an
    English composer, decide to do?
    a. He decided to meet him.
    b. He decided to write to him.
    c. He decided to invite him to his concert.
    d. He decided to compose a piece of music.

15. What was the title of the music that Yo-Yo Ma played on that evening at the concert?
    a. The Hero of Sarajevo.
    b. The Anguished Cellist.
    c. The Cellist of Sarajevo.
    d. The Greatest Cellist in Sarajevo.
16. How was Yo-Yo Ma's music at the festival?
   a. It was calm and peaceful.
   b. It was noisy and loud all the way.
   c. It gradually intensified and then returned to silence.
   d. It was cheerful in the beginning and very sad in the end.

17. Right after Yo-Yo Ma finished playing the music, what did he do?
   a. He knelt down on the stage.
   b. He remained bent over his cello.
   c. He slowly walked away from the stage.
   d. He smiled at the audience, waving his hands.

18. After he finished playing the music, what did Yo-Yo Ma do looking out across the audience?
   a. He called someone's name.
   b. He signaled someone to come to the stage.
   c. He shouted at someone to come to the stage.
   d. He shook his cello and called someone's name.

19. What did the two men, Yo-Yo Ma and Vedran Smailovic, do after finishing the music?
   a. They shook hands.
   b. They hugged each other.
   c. They bowed to each other.
   d. They kissed each other's cheeks.

20. How was Vedran Smailovic's appearance when he met Yo-Yo Ma?
   a. He wore concert attire.
   b. He was badly wounded.
   c. He looked neat and elegant.
   d. His clothes were dirty and torn.

21. What was the reaction of the audience looking at Vedran Smailovic meeting Yo-Yo Ma?
   a. They were calm and peaceful.
   b. They were moved and overwhelmed.
   c. They were not interested in their meeting.
   d. They were very sad to see the two men's behavior.

22. What was Vedran Smailovic's cello in May 1992 compared to?
   a. Hitler
   b. Albinoni
   c. Alexander the Great
   d. The Sword of Joan of Arc
23. When the author played the piano for the nursing home, what did he realize?
   a. He found great similarities.
   b. He found slight similarities.
   c. He found only great contrasts not similarities.
   d. He found that nothing could compare with the splendors of Yo-Yo Ma's concert.

24. Which of the following is not found in the situation of the nursing home?
   a. Bombs and bullets
   b. Cherished memories
   c. A chorus of croaking voices
   d. The sound of a shopworn piano

25. What did the author come to realize about music in the end?
   a. It makes people feel calm and united.
   b. It is a gift which can separate individuals
   c. It is a gift which can help lonely old people.
   d. It is a gift which can inspire great musicians

Appendix C
Questionnaire

1. Do you like to have glosses in the reading material? Why or why not?

2. If you said yes in #1, which type of gloss do you prefer? (see example below)
   a. Korean gloss
   b. English gloss

   e.g., No amount of kicking or screaming could dissuade my mother. She wanted us to learn the language of our heritage.

   Korean gloss dissuade: 끔찍하게 피스
   English gloss dissuade: change someone's decision

3. If you were in either group #2 or #3, how much did you look at the glosses?
   Almost all words
   About half
   A few
   None

About the Author

Myong Hee Ko is a candidate in the Advanced Graduate Certificate Program at the University of Hawai‘i at Manoa. Her research interests include reading, vocabulary acquisition, and program development/evaluation. E-mail: myongheeko@yahoo.co.kr