

Type of multimedia gloss and L2 proficiency: A computer-based study.

Abstract

The present study will be conducted within an attentional framework [Schmidt's "noticing hypothesis" (1990, and elsewhere)] to investigate what type of multimedia gloss is more beneficial for computerized text comprehension and vocabulary learning. Following Bowles (2004), online measures will be also employed to examine the amount of words *noticed* and their relationship to the independent variables: type of gloss and level of proficiency in the L2.

Adult L2 students of Spanish enrolled in two different college levels will be randomly assigned to one of four conditions (no glosses, textual gloss, visual gloss, textual+visual gloss). A pre-, post-, delayed posttest design with recognition, production, and comprehension tasks as the dependent variables will be followed. Participants will also fill in two questionnaires that will tap into their impressions on the glosses and to control for potential outside exposure to the target vocabulary words in the study. Both quantitative and qualitative analyses will be carried out to investigate the effect of gloss type and language proficiency on measures of the dependent variables as well as their relationship with the *noticing* of the target words.

Keywords

Multimedia gloss, type of gloss, Noticing hypothesis, L2 proficiency, online measures, off-line recall protocols, incidental vocabulary learning

1. Introduction.

Ever since the advent of computer-assisted language learning (CALL), teachers and researchers alike have tried to devise ways in which computer technology can be of use in foreign language classes. Perhaps one of the most investigated phenomena has been glosses, either in reading passages or in listening comprehension activities. These would be substitutes for the traditional dictionary with which the student had to switch focus from the text to the dictionary with the consequent waste of time and effort. On the contrary, multimedia glosses come integrated into the text or listening passage. In the case of reading computerized passages for comprehension, glosses are most commonly accessed when the students click the hyperlinked words.

During the past twenty years, a considerable number of studies have investigated the effectiveness of marginal glosses under different premises (Bowles, 2004; Chun & Plass, 1996; Hulstijn, Hollander, & Greidanus, 1996; Jones, 2004; Jones & Plass, 2002; Lomicka, 1998; Nagata, 1999) as well as the theoretical underpinnings supporting their use as a valid tool for SLA (Hulstijn & Laufer, 2001; Mayer, 2002; Schmidt, 1990 and elsewhere). Glosses, usually placed at the margin of the text, can be textual, visual, both textual and visual, and auditory. The present study addresses the first three types under the same theoretical framework utilized in Bowles (2004), so that its results can be easily interpreted within this strand of research.

Researchers have mainly investigated two issues regarding multimedia annotations: on the one hand, whether they really aid the student's overall comprehension of the text, and which exact type does more so; on the other hand, studies have tried to show if these types of annotations promote incidental vocabulary learning. The most common measures in these studies are "off-line" recall protocols and post reading tests (e.g., Al-Seghayer, 2001). However, given that "off-line" protocols may be dependent on memory and, therefore, may not fully reflect learners' comprehension processes, "online measures" currently employed in SLA research conducted within an attentional framework (e.g., Leow, 1997, 2000, Rosa & leow, 2004; Rosa & O'Neill, 1999) will be employed in this study to investigate what type of available gloss (textual, pictorial or both) better aids different proficiency level L2 students to comprehend a written passage and acquire more vocabulary words when the goal is reading comprehension.

2- Review of the literature.

Multimedia glosses as an aid for overall comprehension of the text

Glosses act as substitutes for the dictionary. However, they do not interrupt the reading process as much, since the definition is easily available in the text. Traditionally, they have been used to promote comprehension of the text and incidental vocabulary learning (Bowles, 2004). In general, when comparing marginal glosses, whether multimedia or traditional, with the absence of any type of gloss, glosses have been shown to be of help to the student in the comprehension of a written text (Bowles, 2004; Davis, 1989; Lomicka, 1998). Davis (1989) investigated whether marginal glosses improved comprehension of a literary text read in a foreign language by intermediate-level college students. In this study, 71 students in 201 French classes at the university level were randomly assigned to three conditions: control, vocabulary guide before reading the text, and gloss condition. The first group read the passage, then wrote everything they could remember, and then reread the passage looking back to the text as much as they wanted. The second group studied the vocabulary guide for ten minutes and then read the passage as much as they wanted for 15 minutes. The last group had access to the gloss and could read the text as many times as they wanted in 25 minutes.

The number of pausal units recalled was counted and then averaged for each group and a One-Way ANOVA was run for statistical significance. Results show that both presenting the students with a vocabulary guide before reading the text and providing glosses when reading the text made the students recall significantly more than those with no help. The use of glosses, Davis claimed, "ensures a more fluent reading of the selection" (p. 45).

In this study, therefore, glosses are not shown to be a significantly better aid than a vocabulary guide in recalling a text. However, its results should be considered with caution for at least the following reasons: first, students in each cell carried out very different tasks in a different amount of time. As a consequence, possible confounding variables were not controlled for. Second, recall protocols can be argued not to be comprehension measure. They may just be assessing memory retention of the text read, as argued in the literature (e.g., see Myers, 1990).

In a study with higher internal validity, Lomicka (1998) investigated whether glossing aided L2 comprehension of a written text, and whether glossing hindered fluency in the L2. She was also interested in the relationship between number of glosses consulted and level of comprehension achieved and inferences created. She stated that the most common measure in the literature to assess the influence of glosses had been recall protocols, and argued that this type of measure was not a very accurate way of depicting underlying processes or comprehension. She drew on the use of online protocols that L1 literature had made to assess the creation of causal inferences to argue that this kind of measurement was a much more valid one. Metalinguistic think-alouds allowed her to check whether the students launched causal inferences from the glosses and created a situation model. Twelve participants enrolled in a second semester French course at an American university took part in the pilot study and were randomly assigned to one of the three conditions: a control group, a group with access to traditional glosses, and a group with access to multimedia extended glosses. Participants in all conditions thought aloud to check for inferences and comments made about the structure of the text or the glosses. Statistical analyses indicated no significant differences in the amount of inferences created by each group. She concluded, however, that the data gathered from the think-aloud protocol indicated that multimedia annotations seemed to have a positive effect on comprehension of a written text and the construction of a situation model. The author argued that her results were not significant due to the small number of participants and that future research was needed to further investigate these issues.

Bowles (2004), in her study to investigate whether computerized glosses versus traditional glosses had an effect on vocabulary acquisition and text comprehension, also employed think-aloud protocols. However, her motivation to do so came from a very different theoretical background, namely, Schmidt's (1990) noticing hypothesis that postulates that learners have to "notice" the form in question before it can be processed further¹. To measure noticing, Bowles (2004) used online protocols to establish what type of gloss induced learners to "notice" more words, and to investigate whether this potential increase in noticing had an impact on text comprehension and vocabulary acquisition. A final sample of 50 participants enrolled in college-level first year Spanish courses was randomly assigned to one of three groups: a control group, a traditional gloss group, and a multimedia gloss group.

¹ Schmidt's noticing hypothesis (1995) proposes that learners must "notice" a form in the input before that form can be processed any further. Noticing, according to Schmidt, is paying attention with a low level of awareness. A number of studies have provided empirical support for the role of noticing and awareness in L2 behavior and development (see for example: Leow, 1997; Rosa & O'Neill, 1999).

Participants' performances on a reading comprehension task and immediate and delayed recognition and written production tasks were submitted to quantitative and qualitative analyses. Her results indicate that both experimental groups, multimedia and traditional gloss, had an advantage over the control in amount of noticed words, text comprehension, and acquisition of target vocabulary. These results therefore support the use of glosses, either multimedia or traditional.

At the other end of the spectrum, Jacobs, Dufon & Hong (1994) found that glossing did not significantly affect recall in general. These researchers explored the effect of vocabulary glossing on recall and vocabulary learning as well as the attitudes and preferences for every kind of gloss. Eighty five native speakers of English studying second-year Spanish at the university level participated in this study. Participants from 8 sections of fourth semester Spanish were randomly assigned to one of three conditions: control, Spanish gloss, and English gloss. Right after reading the passage in the three different conditions, participants were asked to write everything they could recall, and then, asked to translate the list of 32 glossed items from Spanish to English. At this time they also filled in a questionnaire that tapped on their preferences on glossing. Four weeks later, as a delayed post test, the translation task was repeated.

The main findings in this study were that glossing did not significantly affect recall in general, even if there was a trend that favored students who had glosses. However, posthoc analysis of the scores on the recall measure showed that those students with higher proficiency recalled more if they had had access to a glossed word. In the translation task, those who had glosses outperformed those with no access to glosses. Regarding vocabulary learning, superior scores for those students who were presented with glosses disappeared after four weeks.

Although these results would seem to contradict what was reported by the studies reviewed above, it should be noted that the post-reading assessment tasks were not similar: participants in this study were asked to write everything they could remember about the text and write translations of the target words. It can be argued that these types of measures do not measure the degree of comprehension or acquisition, since they rely heavily on short-term memory. Another possible explanation for the contradictory results could be reached looking at the differences in proficiency level across studies. It could, therefore, be argued that annotations have a different impact on learners with different proficiency in the L2.

In summary, the literature in this strand of research shows that the presence of glosses, either multimedia or traditional, appears to aid in the comprehension of an L2 written text. However, given the different research designs, participants' proficiency levels, and some of the limitations already mentioned, there is need to be cautious when interpreting the results.

The influence of multimedia annotations on vocabulary learning when the goal is reading comprehension

As far as incidental vocabulary learning in reading comprehension is concerned, the literature also seems to support the use of glosses. Incidental vocabulary learning has been defined as the picking up of new words during listening or reading activities when the general goal of the activity is comprehension rather than learning those new words (Hulstijn, Hollander & Greidanus, 1996).

Hulstijn et al. (1996) argued that research had shown that the rate at which intermediate and advanced students of a foreign language acquire or “pick up” new words is very low. They proposed the use of glosses as a possible means to improve these rates. These authors set out to investigate whether glossing and frequency of occurrence can be used to counter the low occurrence of incidental vocabulary learning. 78 advanced French students from three different universities participated in this experiment. They were assigned to one of three groups: a group that had access to textual glosses, a group that only had access to the traditional dictionary, and a control group. There were 16 target, high frequency words, and 16 low frequency words in the reading text. Three posttests were administered: a recognition and recall test, a test assessing their prior knowledge of the target words, and, finally, a third test that measured again their knowledge of those words. The researchers concluded that frequency of occurrence had a greater beneficial effect on incidental vocabulary learning both when the gloss has been accessed and when the word has been looked up in the dictionary than when no meaning of the word was available. In addition, marginal glosses better promoted incidental vocabulary learning, given that in some cases students did not use the dictionary; when it was indeed used, the results were as good as with glosses. However, half of the words that had been glossed, looked up, or appeared once or three times on the text were forgotten. In summary, this study supports the positive role of multimedia annotations on incidental vocabulary learning.

Drawing from the L2 input enhancement literature and L1 reading literature, Watanabe (1997) explored the effects of text modifications and task on incidental L2 vocabulary learning when reading. Two hundred and thirty one Japanese speaking students took part in this experiment in which they were randomly assigned to one of ten conditions. These 10 conditions were made up by three modification types (appositives, gloss, and multiple-choice gloss), two control conditions, and two variations in the task (with and without translation). Pre-, post-, and delayed posttests were administered and all the students in every group spent the same amount of time on the treatment. This consisted of a reading passage with 16 glossed words on the margin, either on the single gloss format, or in the multiple-choice format.

Results confirmed that students who had access to either glossing format outperformed those participants who had not had access to glosses in three unexpected vocabulary posttests. In addition, no significant results were found for the scores of those participants who could choose the type of gloss or those who were provided with the translation.

Finally, Rott and Williams (2003) utilized think-aloud protocols when qualitatively exploring the effect of multiple-choice glosses and periodic second language text reconstruction on lexical acquisition. These authors investigated the students' processing strategies and how access to a number of glosses influenced inferencing strategies and form-meaning mapping. They also investigated the effects of number of occurrences of target words on noticing and form-meaning mappings as well as the impact on an output activity on these variables. Twelve fifth-semester L1 English learners of German at college-level formed the final sample of participants in this study. These participants were randomly assigned to one of two conditions: [+ gloss],[- gloss]. The reading passage included four target words, for which they were pre-tested, and seven distractors. The participants had to retell in German each one of the four sections into which the text was divided immediately after reading it. To measure immediate vocabulary gain, two immediate vocabulary post tests were administered. In addition, to assess L2 readers' processing behavior, a concurrent think-aloud procedure was used.

In general, the results of this study revealed a difference in reading strategies and vocabulary test scores for both groups. In particular, those students who had access to glosses looked for concrete meanings and strong form-meaning mapping; students with no access to glosses aimed at overall comprehension of the text; if students re-encountered the target words their form-meaning connections were strengthened as much as in the output phase.

This study, however, has some limitations that should be kept in mind when interpreting its results. On the one hand, only twelve students took part in this experiment, seven and five participants in each group; on the other hand, it lacks an operationalization of "noticing" and depth of processing. The theoretical framework on which they developed their study was Hulstijn and Laufer's (2001) Involvement Load Hypothesis. Their basic contention was that "the retention of unfamiliar words is, generally, conditional upon the degree of involvement in processing these words" (p. 545). "On line" think aloud protocols are one tool that researchers have made use of to gain insight into cognitive processes and "noticing". In this study, however it was implied that it was the gloss and the kind of activity that promoted "noticing" and certain depth of processing but lacked the operationalization of those variables. Consequently, there is no objective way to address the degree of processing or "noticing".

In summary, as seen in the research reviewed, the claims made by these studies concerning vocabulary learning have to be interpreted within their very own context, since most of the vocabulary measures employed could be argued to depend too heavily on memory. These studies do not attempt to discover and measure any possible internal process which may indicate that learning is taking place.

What type of gloss is more effective for reading comprehension and vocabulary learning?

Several types of glossing when reading for comprehension have been addressed in the literature and could be subsumed under textual and visual glossing and a mixture of the two. Texts, pictures, or dynamic videos placed at the margin of texts when the students click on the hyperlinked word help the reader understand the meaning of difficult words in the text. The first two types and their combination are going to be addressed in the present study.

In a widely cited study, Chun and Plass (1996) investigated, on the one hand, incidental vocabulary learning when the goal was reading comprehension, and on the other, the effect of visual information for vocabulary learning and students' look-up behavior. They used CyberBuch, a program the authors themselves developed, that allowed the participants to access text, pictures, and video glosses. The authors reported the results for three different experiments in which 160 second year German learners took part. After being exposed to a video preview of the story of around two minutes, participants read the text with access to every gloss (text, text + picture, and text + video), took a vocabulary post test, and underwent a written "off line" protocol in which they wrote everything they remembered about the story. Their results were threefold: firstly, incidental vocabulary acquisition rates of 25% in production tasks and 77% in recognition tasks were attested; secondly, significantly higher scores were found for words annotated with text and picture than those glossed with text only or text + video; thirdly, a significant weak to moderate positive correlation between using certain annotation type and using this same type for the retrieval of words was found.

Al-Seghayer (2001) investigated what type of multimedia gloss, dynamic video or still picture, was more effective in aiding vocabulary acquisition. Thirty participants studying English as a second language in a U.S. university took part in this within-subject design study. They were placed at the intermediate level according to their performance on the reading section of the placement test at the university. All the participants were exposed to the same treatment condition (printed gloss alone, printed gloss along with still picture, and printed text along with dynamic video) to be thereafter tested on the dependent variable measured through two types of vocabulary tests: recognition and production. They were also administered a questionnaire to tap on their impressions as to what type of gloss helped them more. Finally, participants were interviewed to find out what mode they thought conveyed the meaning of the lexical items more accurately.

Results of the analyses carried out indicated that learners presented with text + video had scores that were significantly better than the text + picture condition. In addition, results also showed that the text only condition was significantly more effective than the text + video condition.

However, this study has some limitations that could be factored in when explaining the contradicting results reported in this study and the one by Chun and Plass (1996),

reviewed above. The measures used in this study included a post vocabulary test, taken right after reading the passage, a questionnaire tapping on their impressions as to what type of gloss helped them better learn vocabulary words, and an interview in which they were asked what particular mode better conveyed the meaning of a particular word. On the one hand, the research design in this study did not include a pre-test to control for prior knowledge; on the other hand, while the questionnaire and the interview are valid tools to gather students' impressions about the glosses and their usefulness, they do not provide insight into the internal processes needed for any learning to take place. Furthermore, the production and recognition post vocabulary tests taken immediately after reading the passage might not have been only assessing learning, but also memory retention.

Kost, Foss, and Lenzini (1999) also measured how different types of glosses influenced incidental vocabulary growth in a non-multimedia environment. Fifty six second-semester students of German at an American university participated in this study. They were assigned to one of three marginal gloss conditions: textual, pictorial, and a combination of the two. Vocabulary learning was measured by one production task and two recognition tasks administered right after reading the passage, measuring short term memory, and two weeks later, to measure vocabulary retention. Results showed that those participants who had access to both textual and pictorial glosses outperformed those under the other two conditions in the recognition of target words on both short term memory and retention tests. The authors argued that these results were due to the different degrees of cognitive effort needed to process: "the mapping of pictures onto the mental model provides a stronger bond than the mapping of words due to the different representations of their information" (p. 94).

It is relevant to mention that no significant results were found for the production tasks and it is generally agreed that in these types of tasks, further processing and more exposure to the target words are needed. However, this study's research design lacks in any type of measure that allows researchers to gain a deeper insight into the cognitive processes involved in the learning of those new words

Chun, Mayer, and Leutner (1998) investigated the effects of different types of glosses according to the students' preferred mode on text comprehension and learning of new words. One hundred and three second year students of German at the university level participated in this study. On the first day, participants filled in a personal questionnaire and read the computerized text at their own pace, consulting the glosses as much as they needed. On the second day, they took the vocabulary and comprehension post-tests. Finally, participants were grouped according to their preference for glosses using log-file data from the computer program used. These files contained a record of the type of annotations the participants used. They were divided into verbalizers, visualizers, and those having no strong preference. The results of this study provided justification for the generative theory of multimedia learning (Mayer, 1997). Participants performed better on the posttests when both visual and textual information were selected, moderate when only one mode was selected, and worse

when neither was selected. In addition, participants comprehended the text better when they could choose the gloss in their preferred mode.

The theoretical justification for this study is provided by the generativist multimedia learning theory proposed by Mayer (1997). Based on work carried in the psychological field, Mayer (2002; 1997) proposes that it is through two channels that human beings represent and manipulate knowledge: a visual-pictorial and an auditory-verbal channel. Therefore, the use of textual and pictorial glosses would enter the cognitive system through those two channels. He argues that “meaningful learning occurs when learners engage in active processing within the channels, including selecting relevant words and pictures, organizing them into coherent pictorial and verbal models, and integrating them with each other and appropriate prior knowledge” (Mayer, 2002: 60).

However, these authors’ claims about vocabulary learning have to be carefully considered mainly due to the following research design limitation: no pre- or delayed posttest were administered, therefore, the post vocabulary test might be just showing previous knowledge or good memory skills, without the new words having been processed with the help of the glosses.

The influence of glosses has also been investigated in listening comprehension activities with results that add to the literature on positive effects of a combination of textual and visual glosses (Jones, 2003; 2004; Jones & Plass, 2002). In these studies, both the listening passage and the glosses were presented through the computer. First, participants listened to the digitalized passage in its entirety and, then, accessed the different types of glosses thorough different screens on the computer.

To sum up, the literature on different types of multimedia glosses appears to indicate that the combination of textual and visual glosses has a more beneficial effect on comprehension and vocabulary learning than either type in isolation. As mentioned above, however, any generalization about vocabulary learning should be closely considered, since the measures utilized in these studies involve memory skills to certain degree. In addition, the different theoretical backgrounds from which these above studies originate make difficult any general conclusion about what types of glosses are more beneficial for the L2 learner in reading a computerized text.

Differences in L2 proficiency

There is a wealth of research that has explored the effects of L2 proficiency on the acquisition process of both an L2 (Leeser, 2004; Leow, 1995; Ortega, 2003; Shook, 1994; VanPatten, 1989, 1990) and L3 (Cenoz, 2000, 2003; Lasagabaster, 2000; Muñoz, 2000).

One prediction for the role of L2 proficiency in L2 acquisition can be found in VanPatten’s model of input processing (VanPatten, 1996, 2004). This model claims that the student will process for meaning first and, then, if there are still cognitive available resources, will focus on form. VanPatten (2004) posits three basic claims: 1)

students focus primarily on the extraction of meaning from the input. 2) learners should “notice” the input for acquisition to happen. 3) the processes of noticing are constrained by a limited working memory capacity. The consequences of this for more proficient language learners seem clear: the more proficient, the less working memory capacity will be used to focus on meaning and the more resources there will be for focusing on a certain form.

In an earlier study, VanPatten (1990) tested whether students of various proficiency levels could consciously attend to meaning and form when processing language input. A total of 202 students of Spanish at the university level participated in this study. First semester, fourth semester and third year conversation classes at that university were the proficiency levels selected. Whole classes were randomly assigned to complete one of four listening tasks: control where participants had to listen for content only, listening for content and simultaneously noting a lexical item, listening for content and simultaneously noting a definite article, and, finally, listening for content and simultaneously noting a verb morpheme. Each item occurred 11 or 12 times in the passage, and participants were told to write a mark anywhere on the paper each time they heard the item, so that processing would not be interrupted. As for comprehension, it was measured by free written recalls in English. In general, his results confirm that less proficient students have greater difficulty in attending to form and meaning at the same time than more proficient students.

Leow (1993) also included second/foreign language experience as an independent variable in his study on the effects of simplification, type of linguistic item on learners' intake of linguistic items in written and aural input. First and second-year learners of Spanish at the university level were randomly assigned to one of four conditions: a simplified or unsimplified passage including the present perfect, or a simplified or unsimplified passage including the present subjunctive. Participants were tested in a pre and post-exposure multiple choice recognition task. Results of the analyses included a significant effect for language experience, that is, higher levels of language experience performed significantly better than lower levels on the postexposure task.

Shook (1994) also investigated the differences between two proficiency levels of learners of Spanish in the way they processed grammatical information presented via written input. Specifically, this author investigated how drawing attention to the linguistic item, the linguistic item itself, and language proficiency affected FL/L2 learner-readers' intake. One hundred and twenty five first- and second-year university students of Spanish participated in this study. Two different reading passages in three different attention conditions were chosen to present the two different linguistic items: the present perfect and a relative pronoun. Production and recognition tasks for both linguistic items served to assess participants' abilities in producing and recognizing the grammatical forms after being exposed to the text. No significant main effect for language experience was found in any of the analyses carried out. More specifically, no significant difference was found in the amount of grammatical information taken in between the two proficiency levels. Second-year learners, however, significantly outperformed first-year learners in the production of the present perfect but not on the

relative pronoun tasks. The author concluded that “the amount of experience with certain FL/L2 structures may bring about differences in the ability of the subjects to produce the grammatical items” (p. 85).

Ortega (2003) meta-analyzed 25 studies and their use of syntactic complexity measures as indices of college-level L2 writers’ overall proficiency in the L2. In particular, this research synthesis focused on “studies that investigated the differences between quantitative differences in the syntactic complexity of L2 written texts and proficiency differences” (p. 496) in college level L2 learners in second or foreign language contexts. In this vein, 21 cross-sectional studies were chosen for analysis. In these, at least two levels of proficiency in the L2 were compared in their examination of L2 written syntactic complexity. In addition, four longitudinal studies were chosen. These investigated the learners’ linguistic development as given by changes in the learners’ syntactic complexity over time. The researcher identified the six more widely used measures of syntactic complexity across studies and used the results of each individual study on those measures for her analysis. Results of the analysis carried out made the author conclude that “studies that established group differences on holistic ratings tended to yield narrower changes” (p.512) in syntactic complexity across different proficiency groups. In addition, it seems that in the studies included in the analysis ESL learners produced more complex writing than FL learners.

These results should be borne in mind, while acknowledging that there are some intrinsic limitations to this meta-analysis and meta-analyses in general; namely, the writing tasks, timing, elicitation procedures, and sample size, varied widely across studies.

Finally, Leiser (2004) investigated how grouping participants by proficiency level (high-high, high-low, low-low) affected the production (amount, type, and outcome) of language-related episodes (LREs) during collaborative tasks. Four sections of a fourth-semester content based Spanish American Geography course on the social and physical geography of Latin America participated in this study. Grammar and oral skills were not emphasized at any point during the course. Forty-two students participated in this study; they were grouped according to the oral proficiency in Spanish by the instructor of the course to make up 21 dyads. First, participants just listened to a passage in Spanish. Then, they listened to the passage a second time and took down some notes, but not complete sentences. Finally, participants working in pairs put together their notes and tried to reconstruct their own written version of the passage. It is in this phase that they produced LREs as their search for the desired words and sentences. The experiment took place in two sessions: a practice session, and a week later, the real session where the interactions between the members of the dyads were recorded. The results of the analyses carried out showed that the level of proficiency of the dyads had an effect on how much participants focused on certain types of form as well as on how well they solved the problems that the reconstruction of the passage created in the pair.

In summary, the literature on the effects of varying levels of L2 proficiency on acquisition of different aspects of an L2 supports that differing L2 proficiency levels have an impact on the acquisition process; therefore, differences in the way novice and expert language learners learn the L2 can be expected.

Summary and limitations of previous research

The above review of the literature has shown that glosses, either multimedia or traditional, appear to be beneficial for the comprehension of a passage (Bowles, 2004; Lomicka, 1998). Regarding the claims made by studies such as Hulstijn et al. (1996) and Watanabe (1997) in their investigation of how multimedia glosses aid in the acquisition of vocabulary learning, it has to be kept in mind that there was no real measurement of any internal process that may indicate acquisition; instead, these studies utilized measures that assessed retention of the target words. In some cases, word retention might be the result of the word having been learnt; in other cases, it might just be a consequence of good memory skills. In addition, those studies that have compared textual and pictorial glosses in reading comprehension and vocabulary learning have, for the most part, shown that a combination of textual and pictorial glosses is more beneficial to the students (Al-Seghayer, 2001; Kost et al., 1999; Plass et al., 1998). Finally, differences in the acquisition process due to different proficiency levels can be expected from the above review (Leow, 1993, 1995; VanPatten, 1989, 1996, 2004).

The most serious limitation of most of the reviewed studies is that they used written recall protocols and vocabulary tests as measures of comprehension and vocabulary learning respectively, which seem better fit to measure short-term memory than comprehension and learning. As argued above, good results on these types of measures might not always be the result of the acquisition process having taken place. Memory-related factors have too strong of a role in these measurements and cannot be discarded. To address this important issue, in addition to comprehension, production, and recognition tasks, on-line think-aloud protocols will be employed during the treatment so that a cognitive component can be added to the equation. This type of introspective technique will make possible to reveal the learner's cognitive processes involved in the comprehension of a computerized text and the learning of new words with the help of multimedia glosses. Schmidt's (1990, and elsewhere) attentional framework will serve as the theoretical framework on which to base the analysis of these underlying cognitive processes. This author proposes two levels of awareness: awareness at the level of noticing and awareness at the level of understanding. The former is considered to be a superficial level and the minimum requirement for input processing, whereas the latter is considered to be a deeper level of awareness leading to restructuring and system learning. Support for Schmidt's "noticing hypothesis" and the different levels of awareness has been found in a number of studies (Leow, 1997; Rosa & Leow, 2004; Rosa & O'Neill, 1999).

In this manner, in addition to the results achieved by the comprehension, recognition and production tasks, a further level of analysis will be available through the use of

introspective protocols. These will allow us to tap on the participants' cognitive processes and possible different levels of awareness so that a clearer picture of the comprehension and acquisition processes can be achieved and more robust claims can be made.

The present study will then add to the literature of this strand of research in at least the following ways: on the one hand, by using the same theoretical framework that Bowles (2004) utilized [Schmidt's "noticing hypothesis" (1990, and elsewhere)], we will add the cognitive component missing in most of the studies investigating the impact of different types of multimedia glosses in reading comprehension and vocabulary learning. In this way, not only will we add to the literature in ease of interpretation of different studies, but also we will be able to investigate cognitive processes that have not been subject of study in this particular area of research. On the other hand, this study addresses an issue which has not been investigated so far in this strand of research: the impact that different L2 proficiency levels may have on the effectiveness of the different types of glosses in comprehending a computerized written passage and learning the targeted words.

The present study investigates the effects that different types of glosses, namely textual, pictorial, and textual + pictorial, and different levels of L2 proficiency have on text comprehension and vocabulary learning when the goal is just comprehension. Like Bowles (2004), it is based on the theoretical framework of attention, which maintains that the role of attention is critical in the acquisition process of an L2 (Robinson, 1995; Schmidt, 1995, 2001; Tomlin & Villa, 1994).

Based on this theoretical foundation and on the literature review above, the present study operationalized and measured attention through think-aloud protocols following previous studies (e.g., Leow, 1997, 2000; Rosa & Leow, 2004; Rosa & O'Neill, 1999) Participants with two different proficiency levels in the L2 read a text in four conditions (no gloss, text gloss, picture gloss, and text + picture gloss) while asked to think-aloud. As in Bowles (2004), these different conditions were used to assess whether any of them promoted noticing and whether this noticing led to a better comprehension of the text and vocabulary learning of the target words. Specifically, answers to the following research questions are sought:

- 1) Does type of gloss have an effect on adult L2 learners' recognition and production of target words, as measured by recognition and production tasks? If so, does the effect last over time?
- 2) Does level of L2 proficiency have an effect on adult L2 learners' recognition and production of target words, as measured by recognition and production tasks? If so, does the effect last over time?
- 3) Does type of gloss have an effect on adult L2 learners' comprehension of a computerized reading passage, as measured by a comprehension task?

4) Is there a relationship between type of gloss and adult L2 learners reported noticing of target words in the input?

5) Is there a relationship between adult L2 learners' level of proficiency and their reporting of noticed words in the input?

3. Participants

The original pool of participants will consist of 6 sections of fourth-, and sixth-semester college-level Spanish classes. To control for prior knowledge only those participants who score 6 out of 20 on the pre-tests (chance score) and who indicate no outside exposure on the debriefing -questionnaire administered at the end of the study- will be included in the final sample. Participants will be also eliminated for failing to produce intelligible think-alouds, attend all sessions, or back-tracking to the previous items on the post-exposure tasks.

4. Materials

An internet-based passage from an online newspaper will be utilized.

Piloting of the experimental text: a pool of 20 randomly-selected fourth-, and sixth-semester college-level Spanish learners will underline all the unknown words in the text. Only words that more than half of the participants underline will be selected to be glossed.

Glossing: as a result of the pilot test 20 words will be glossed. In the experimental conditions, the words will be hyperlinked. When the participants click on them a box will appear on the upper right hand side corner of the screen with a definition (textual gloss group), a picture (pictorial gloss group), or a combination (textual + pictorial group). Participant in the control group will be exposed to the text with no glosses.

5. Assessment tasks

Pre-test / Post-test recognition tasks. The pre-test, immediate post-test, and delayed post-test will consist of the same 20 item multiple-choice recognition task. The target words in Spanish will be followed by four possible definitions in English, or by four different pictures for the students to choose from, or by a combination of text and picture depending on the gloss group to which they were assigned. The only difference between the tests will be that the items will be in different order with different distractors.

Pre-test / Post-test production tasks. The pre-test, immediate post-test, and delayed post-test will consist of the same 40 item production task. Participants will be given a list of words in English, or pictures, or a combination of text and pictures (20 target words and 20 distractors) depending on the gloss group they were in, and asked to write their equivalent in Spanish. Participants will perform the production task first so that the multiple choice recognition task will not provide additional exposure.

Reading comprehension task: The comprehension task will be administered only in the immediate post-test. It will consist of 12 multiple-choice comprehension questions

in English. Participant will be allowed to refer back to an unglossed version of the text to help them answer the comprehension questions if they need to.

6. Procedure

Two weeks prior to exposure, participants will take the production and recognition pre-tests to assess their knowledge of the target words.

On the day of the exposure, students in the three groups will be given the instructions to complete the experimental treatment, including instructions on how to think aloud. Time spent on task will be recorded.

Three weeks after the exposure, a delayed post-test with scrambled items will be administered. To control for outside exposure, a debriefing questionnaire will be administered after the delayed post-test.

Tracking: Participants in every experimental group will be tracked to find out how many times they accessed a gloss. Participants in the control group will be asked to underline every unknown word. In addition, the total time every group will spend on task will be recorded.

Coding think-aloud data: At least two raters will code the data according to the following criteria (Bowles, 2004): any comments about the targeted word, read the gloss out loud or clicks on the hyperlinked word.

7. Analysis

In order to explore the possible effects that type of gloss and L2 proficiency may have on recognition, production and comprehension tasks as measured in this study, One-Way repeated measures ANOVA (4x2x3) analysis will be used.

To assess the relationship between type of gloss/L2 proficiency and reported noticing of target words two separate correlational analyses will be carried out. To compare the correlations of both matrixes a Fisher's z Transformation will be utilized so that the two Independent rs can be compared.

8. References

- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232.
- Bowles, M. A. (2004). L2 glossing: To CALL or not CALL. *Hispania*, 87(3), 541-552.
- Cenoz, J. (2000). Research on Multilingual Acquisition. In J. Cenoz & U. Jessner (Eds.), *English in Europe : the acquisition of a third language*. Clevedon ; Buffalo: Multilingual Matters.
- Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review, *The International Journal of Bilingualism* (Vol. 7, pp. 71-87).
- Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183-198.
- Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *Modern Language Journal*, 73(1), 41-48.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary

- use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539-558.
- Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17(1), 19-28.
- Jones, L. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language, Learning & Technology*, 8(3), 122-143.
- Jones, L., & Plass, J. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, 86(4), 546-561.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, 32(1), 89-113.
- Lasagabaster, D. (2000). Three Languages and Three Linguistic Models in the Basque Educational System. In J. Cenoz & U. Jessner (Eds.), *English in Europe : the acquisition of a third language*. Clevedon; Buffalo: Multilingual Matters.
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8(1), 55-81.
- Leow, R. P. (1993). To simplify or not to simplify: A look at intake. *Studies in Second Language Acquisition*, 15, 333-355.
- Leow, R. P. (1995). Modality and intake in second language acquisition. *Studies in second language acquisition*, 17, 79-89.
- Leow, R. P. (1997). Attention, awareness and foreign language behavior. *Language Learning*, 47, 467-505.
- Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: aware vs. unaware learners. *Studies in second language acquisition*, 22, 557-584.
- Lomicka, L. L. (1998). "To Gloss or not to gloss": An investigation of reading comprehension online. *Language Learning & Technology*, 1(2), 41-50.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32(1), 1-19.
- Mayer, R. E. (2002). Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. In D. F. Halpern & M. D. Hakerl (Eds.), *Applying the science of learning to university teaching and beyond* (pp. 55-72). San Francisco: Jossey-Bass.
- Muñoz, C. (2000). Bilingualism and trilingualism in school students in Catalonia. In J. Cenoz & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 157-178). Clevedon: Multilingual Matters.
- Myers, J. L. (1990). Causal relatedness and text comprehension. In D. A. Balota, G. F. d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 361-375). Hillsdale, NJ: Erlbaum.
- Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals*, 32(4), 469-479.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518.

- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*, 25-36.
- Robinson, P. (1995). Attention, memory and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.
- Rosa, E., & Leow, R. P. (2004). Awareness, different language conditions, and second language development. *Applied Psycholinguistics, 25*, 269-292.
- Rosa, E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. Another piece to the puzzle. *Studies in second language acquisition, 21*(4), 511-556.
- Rott, S., & Williams, J. (2003). Making form-meaning connections while reading: A qualitative analysis of word processing. *Reading in a Foreign Language, 15*(1), <http://nflrc.hawaii.edu/rfl>.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.
- Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-63). University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge applied linguistics series. Cambridge University Press.
- Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning, 5*(2), 57.
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in second language acquisition, 16*, 183-203.
- VanPatten, B. (1989). Can learners attend to form and content while processing input? *Hispania, 72*, 409-417.
- VanPatten, B. (1990). Attending to form and content in the input: an experiment in consciousness. *Studies in second language acquisition, 12*, 287-301.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood: Ablex Publishing Corporation.
- VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research and Commentary*. Mahwah, NJ: Lawrence Earlbaum.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition, 19*(3), 287-307.