CAN GRAMMATICAL CALL HELP EFL WRITING INSTRUCTION?

Hsien-Chin Liou, Samuel H. Wang, and Yuli Hung-Yeh
National Tsing Hua University at Hsinchu, Taiwan, R.O.C.

ABSTRACT

This is a research report of a one-year study which addresses whether and in what way grammatical CALL can help English writing instruction in an EFL setting, namely, in Taiwan, R.O.C. The study started with the implementation of CALL courseware which contains ten lessons of grammar exercises. The courseware featured (a) empirically based contents, results from error analysis of Chinese EFL students’ common mistakes; (b) a sentence-level drill and practice with various formats of guided Chinese-English translation, error correction, and sentence combining; (c) adaptive remedial instruction for individual students while practicing on-line; (d) integration of error anticipation and keyword matching for answer judging as well as item-specific feedback messages; and (e) verbatim recording of student input. To evaluate the effectiveness of this CALL strategy, we designed an experimental study in which two groups of college freshman EFL majors did the pre-test and post-test tasks with ten-week CALL intervention for the experiment group while the control group worked on the same contents in a paper-and-pen homework setting). Results suggested that the combined effect of classroom instruction and grammatical CALL is helpful for writing instruction, or at least not detrimental.

KEYWORDS

Keyword matching, error anticipation, error analysis, courseware development, item-specific response, verbatim recording of student input and history, quasi-experimental design.
INTRODUCTION

This project was motivated by pragmatic concerns and our interest in the development of computer-assisted language learning (CALL). In language teaching, writing teachers are often bothered by the time invested in grading students' compositions. Recurrent grammatical errors seem never to disappear (also demonstrated in other foreign language instruction: see a German case in Lalande, 1982 for example). English-as-a-foreign-language (EFL) writing teachers are more burdened with the task because students tend to be negatively influenced by their first language (LI) ("interference" as it is called in the early literature of language learning) partly due to a lack of native speakers' input and opportunities or needs to practice English in the learning environment. Likewise, students in our department often make grammatical mistakes in writing, in spite of the fact that they have received six years of obligatory EFL instruction. It is, furthermore, observed that students demonstrate idiosyncratic weaknesses even though they are generally exposed to identical instructional materials. To alleviate the writing teachers' burden and accommodate students' individual needs, we turned to the computer for help. The specific goal of this project was to provide remedial grammar exercises for those students who show deficiency in language structures, in order to see whether the exercises help improve the quality of student compositions.

CALL has been developed in Taiwan, R. 0. C. since the mid 1970's. Good and useful software for writing instruction at the college level, however, is seldom seen in the market or at academic institutions. Early efforts were invested in adopting lesson materials from the PLATO system. Unfortunately, effectiveness comparison research proved "no difference" between the CALL user-group and the control group (Yao, 1984). The direct borrowing strategy was banned later due to the criticism that materials developed abroad were not suitable for our learners. More recently in the 1980's, cooperation between government and universities flourished when it came to junior high school and vocational high school instruction (see for example, Chen, 1991). Needs at the college/university levels themselves have seldom been addressed, however. As far as electronic courseware was concerned, the tendency was simply to transpose the materials from the standard textbook (junior high schools in Taiwan au use the same textbook designed by the government). Undoubtedly, courseware development requires very much effort and ingenuity. One of the problems is that commercial courseware developers do not know much about EFL learning theories or practices (Chen, 1988; Nash, et al., 1989 hold similar views while addressing other issues about
writing instruction). For example, vocabulary software packages abounded but simply reflected a naive flash-card design. There are several other shortcomings in current available software packages. First, many software packages require users to type in simply a letter or a number. This tends to reinforce uncontextual language learning, which has been thoroughly discredited by recent EFL theories and practices. Moreover, when given a chance to type in a phrase or, more rarely, a sentence, answer judging in the software packages is usually unsatisfactory. Almost all of them simply use naive string matching techniques. Keeping online records of learning history may have to be limited in a personal computer environment. Nonetheless online records of learning history can help teachers/tutors locate students’ weak areas, in a way which human observation cannot capture. Commercial software is weak in these respects. Many software packages for vocabulary and reading comprehension have become available in the last several years, but those for grammar or writing appear to be much rarer.

Thus, we decided to launch a project to develop a grammar program to meet our needs while trying to overcome some of the disadvantages in the available software. In Section II, special features which characterize our English grammar courseware are discussed, incorporating pedagogical considerations. In Section III, a research study is presented for the purpose of evaluating the effectiveness of the courseware.

THE COURSEWARE

As mentioned above, the motivation for this project was to use CALL to enhance writing instruction in the areas of language use, specifically, to correct recurrent grammatical weaknesses for Chinese EFL learners. The context for the CALL courseware use was remedial instruction. It was designed in a drill and practice format. The courseware was implemented on an IBM personal computer platform. The program was written in the Turbo C programming language in the color graphic mode. To elaborate the features of the courseware, some design considerations are explicated in the following.

The courseware contains ten lessons, each of which has ten question items. It features (a) empirically based contents, (b) sentence-level exercises, (c) adaptive remedial instruction, (d) improved answer judging, and (e) comprehensive recording of student learning history. By empirically-based contents, we mean that the contents of the
courseware are based on error analysis of common mistakes in Chinese EFL students' compositions. The corpus came from 51 essays written by freshman EFL majors. Errors in the essays were analyzed according to a coding system which was mainly adopted from Chiang (1981) with some modification (also referring to Chen's work, 1979). The errors were tallied and the ratio for each category was calculated. 428 errors were identified and categorized into 35 sub-categories. The total items in the courseware were distributed mainly in accord with the error ratio distribution; variation was adjusted based on our pedagogical experiences and other materials (Canning & Canning 1986; Kei, 1982; King, 1989). In addition, the exercise items were arranged in a neat format and saved as data files so that update, expansion, or modification could be done easily, an advantage which makes the package more generic to accommodate contents of various proficiency levels of learners or teachers who know very little about computer programming.

Second, phrase or single-letter input tends to reinforce uncontextual learning of discrete language points; thus, we made our courseware accommodate sentence-level input. There are three exercise formats: guided Chinese-to-English translation, English sentence correction, and English sentence combining. The translation exercises are implemented in lessons one, three, five, and seven; the correction exercises are in lessons two, four, six, eight, and nine; and the sentence combining exercises are in lesson ten. The translation exercises are guided by giving some key words so that the input does not display chaotic possibilities. Figure 1, 2 and 3 demonstrate examples of the translation, the correction, and the sentence combining formats. These formats provide more contextualized practice compared with units of single words or phrases. Similar cognitive operations would seem to be involved in original composition: (a) translation, at least for the lower level learners, and (b) sentence combining or sentence correction in the revision/editing process on all levels.

Third, to make good use of CALL's adaptivity in individualized instruction, we designed adaptive remedial practice for each item in each lesson. In the English grammar courseware, 100 main items were devised for the ten lessons. For each main item, there were three practice items addressing the same grammar/usage point as provision for students who need more drills; this amounted to 400 items in this courseware. Three attempts allowed for each main item. Main items test one aspect of English grammar or word usage, such as run-on/fragment sentences, usage of verb root form as a subject, and expressions which sound Chinese due to L1 interference. The
Figure 1. An example item of the translation exercises.

Figure 2. An example item of the sentence correction exercises.
LESSON 10

PROBLEM 1:  Sentence combining/reconstruction

Celestine Sibley is a newspaper woman. She works for the Atlanta Constitution.

Celestine Sibley is a newspaper woman who works for the Atlanta Constitution.

Press any key to proceed.

Figure 3. An example item of the sentence combining exercises.

LESSON 1

PROBLEM 1:  Translate to English using the supplied keywords.

我家有六個人

my family, six persons
There are six persons in my family.

F10: quit  F3: check

Figure 4. A translation item and its major pattern.
focus point tested is called a major pattern for each item. For instance, in Figure 4, the problem is to ask students to translate the sentence "There are six persons in my family." from Chinese to English, given two key phrases, 'my family' and 'six persons'.

The major pattern in this example is 'there is'. In the program, the answer judging routine has, before proceeding, to check if the major pattern appears in student input. This will be explained in the fourth point shortly.

For a flow chart within each main item, see Figure 5. If the student makes mistakes in the area of major patterns and fails after three attempts, s/he is routed to the remedial practice unit which contains three related question items. In the remedial practice unit, there is a provision for recycled items. Each practice item may be attempted twice. If the student is correct in the first attempt for the first practice item, the program skips the last two items but saves and presents them after the student finishes the ten main questions. Otherwise, s/he has to practice the second and third items. In the recycled pool, the second practice item is presented to the student first. If s/he is correct in the first trial, the third one is skipped; otherwise, s/he has to practice items 2 and 3.

Fourth, the answer judging routine features keyword matching and error anticipation, an improvement over very inflexible string precise match (for a detailed analysis of various answer-handling techniques, refer to Pusack, 1983). As mentioned above, each item in our grammar courseware tests one major aspect (i.e., major pattern) of English grammar or word / expression usage. For answer judging, we made a list of possible good, acceptable, and wrong answers. The program tries to search the input for matches of any of the anticipated answers / errors. If the input matches, corresponding messages such as "Good", "Acceptable" and 'I think you can do a better job" are displayed in the screen. For expected wrong answers, a corresponding hint message is also displayed at this point. If the input does not match any of the expected correct answers and thus is judged as incorrect, then a string search is done. The major pattern designated in each of the good and acceptable answers is searched and an attempt is made to match the student input string. Finally a complete answer judgment is made. If a match is made, the program knows that the student is weak in this area and provides an on-line hint, which directs the student a step closer to the correct answer. If it is not matched in either of the first two attempts, a general wrong message, "I think you can do a better job", is given. If the pattern is matched, after three attempts, the student is routed to the practice unit because his/her weakness is targeted.
Figure 5. Flowchart within each item.
Meanwhile, the on-line message is somewhat different from the one for the first two trials: the student is given "Sorry! It's still not correct." and "More practice will be good for you" and an elaborate explanation about the focus point in this item (see Figure 6 for an example). If none of the major patterns in the correct and acceptable answers is matched after three attempts, this means the student made a mistake which was unanticipated by the system. S/he is then given a message indicating that the answer is incorrect, but is not routed to the practice unit because s/he is not weak in the target area but in some other respect. In this way, the student is directed to the next main item. Hint and explanation messages are mainly in students' L1, Chinese.

Lastly, the program records student logon data (identification number, date), total time on each lesson, score, and verbatim records of each item responded in each attempt. Table 1 shows a segment of the record kept for each lesson in a particular subject's file.

![Figure 6. An example of Chinese explanation message.](image)
If anyone comes, please inform me.

It looked fragile, so I handled it like china.
It looks like fragile, so I handle it as a china.
He started crying and looked like a hungry baby.
He started to cry like a hungry baby.

I prefer eating fresh things like tomato.
I prefer eating fresh thing like tomato.

A good teacher like him should be respected.
A good teacher like him should be respect.

I do not remember meeting him before.
I can not remember him when I meet him.
I stopped playing the piano at the age of eight.
I am not old enough to drive car.
You are not old enough to drive.
You are not old enough to drive a car.
They have not enough time to talk to her.
They do not have enough time to talk to her.

Table 1. A sample of the record for student CALL learning history.
THE RESEARCH STUDY

To evaluate the effectiveness of this CALL strategy, a research study with an experimental design was set up. Data collected in this study were used to answer the following research questions:

1. How do groups differ as a function of CALL use? Is CALL in this study better for a particular group of subjects?

2. Which aspect of subjects' writing performance is better enhanced by CALL compared with those who do not use CALL, sentence4evel or essay-type tasks?

3. How do subjects like to use CALL to enhance their writing proficiency? What exercise format is regarded the most useful by the subjects?

Answers to these questions were used to uncover a general research inquiry: can grammatical CALL help EFL writing instruction?

Research Design

In the design, students' writing performance before and after they took the grammar lessons was assessed. Subjects were 42 college freshman EFL majors: 4 males and 38 females with an average age of 19 years. They were assigned into two groups while taking the "Grammar and Writing" course (G & W), which is a course required in the Department of Foreign Languages. Because pair work was an important part in the course, the subject groups were not equal. The control group (who did not take CALL lessons) had 20 subjects, whereas the experiment group (who took CALL lessons) had 22 subjects.

Six types of data were collected using the following six research instruments in an attempt to answer the research questions we posed in the beginning.

a. Questionnaire A (about background and attitudes toward EFL learning in general; in Chinese): This surveys subjects' (a) background information, including grade, senior high school attended, native language (or dialect), experience in English-speaking countries, length of time for EFL learning, and when they started EFL learning; (b) attitudes toward EFL learning, including anxiety level in general EFL learning, preference as a school subject, sense of achievement in EFL learning,
anxiety level in EFL classes, confidence in EFL writing, perception about importance of grammar for EFL writing, and perception about importance of communicative effectiveness for EFL writing; and lastly (c) previous computer experiences, including computer experience, when they started using computers, software they have used, and CALL experience.

b. A 20-question test (objective measure of student grammar performance; items randomly drawn from the 100 items in the courseware). This was a test which contains 20 items: ten of them were in a Chinese-English translation format; another ten, in an English-sentence correction format.

c. An essay (subjective measure of student writing performance): This was an essay-writing task where subjects were asked to write an essay in 30 minutes, given a topic.

d. Questionnaire B (responses toward each of the ten grammar lessons; in English): This was a simple one-page sheet with questions surveying feelings/impression about each item and overall impression about the lesson, including causes for unhappiness and joy, and items too difficult or too easy.

e. Questionnaire C (perception about CALL and the courseware; in Chinese): Questions were divided into two major sections. Section I surveys students’ attitudes toward CALL in general, including comparison with textbooks, immediate feedback, increased interest in English due to CALL, time saving for answer checks using CALL, justification of wider use of CALL in college, confidence in writing due to CALL use, grammar error corrected due to CALL practice, and perception about usefulness of each of the three exercise formats. Section II surveys specific questions concerning the grammar courseware developed in this project: contents, screen design, feedback and on-line messages, on-line operations, perception of effectiveness in various aspects of the courseware, and drawbacks in the courseware.

f. The homework sheets: Items in each of the grammar lessons were chosen (main items are obligatory whereas the practice items were chosen at random) and printed out on paper. Corresponding answer sheets were also prepared.
During the study, regular G & W classes were given by two instructors. Much control was exercised to ensure that the control group and the CALL group received the same teaching materials and class activities without attempting to claim so. Class activities consisted of a lecture about the grammar points, exercise with the grammar part, vocabulary presentation by pairs of students, peer editing, group evaluation of an essay written by one of the students, and other exercises in the textbooks (Interactions 1: A Writing Process Book and the series on a communicative grammar). The two instructors co-taught the two groups. A syllabus for each unit was designed by either of the two instructors and agreed upon beforehand. The lectures of the grammar points were given by either of them to the whole class, 42 students. Then, the other exercise parts were done separately within each group.

The research procedures using the instruments above were as follows. In the pre-test, Questionnaire A, the 20-question test, and an essay (titled: "My Expectations about College Life") were given in subjects' first period of the G & W class. The pre-test lasted about one hour. After the pre-test was given, regular G & W proceeded, which provided a two-hour class weekly. Two weeks after the pre-test was completed, the CALL lessons were given. Each of the ten lessons was provided in a computer laboratory each week; the experiment group (the CALL group) had one lesson every week. Questionnaire B was given when the CALL group was working on each lesson. While they were working on-line, an observer was present in the lab to provide help or resolve confusion, if any. At the same time, the control group was given the homework sheets. The homework sessions were simulations of students working on their assignments at home. After they finished, they came to pick up the answer sheets to be checked by themselves. Both groups spent as much time as they needed. The English grammar CALL sessions lasted for ten weeks. When the semester ended, the subjects completed the CALL lessons, and the homework sessions, and the G & W class. One week later, the post-test was given, which included the 20-question test (with the same items in a different order), another essay (titled "My College Life"), and Questionnaire C.

DATA ANALYSIS AND RESULTS

To answer the research questions posed in this study, variables were identified, methods of data analysis were used to assign values to the variables, and results were reported in an attempt to answer the questions. The two independent variables were performances in the 20-question test and the essay-writing tasks. The affective
variables were responses elicited in the three questionnaires.

**Methods of data analysis**

Regarding the two important independent variables, scoring of the 20-question test was straightforward; each item scored 5 points and came to a total of 100. For translation items, the raters, or the two G & W instructors, obtained consensus on the major grammatical points tested in the questions and provided the results. A master graduate student was responsible for double checking the discrepancies, if any. AR the essays were read by two native speakers (associate professors in linguistics), who were asked to grade the essays holistically. They were given the ESL Composition Profile (Jacobs, et al., 1981), Form C and free choice on what proportions should be allocated to the five components in the Profile: content, organization, vocabulary, language use, and mechanics. To ensure the interrater reliability, they both agreed on changes made on this form and graded every essay written in the pre- and post-tests for all the subjects without knowing which was written in which situation.

Methods of analyzing data collected from the three questionnaires were based on descriptive statistics. Questionnaire A, surveying the background information, was for understanding the students' background and learning experiences. Questionnaire B, the survey for each CALL lesson, was mainly for eliciting feedback about each lesson for future improvement of the courseware. Questionnaire C, surveying perception and attitudes about the courseware, served to show students' affective reaction to CALL use in this context.

All the data for the independent variables were transformed and coded in numeric formats and typed in a data file. Procedures in a statistical package (SPSS for PC) were used to analyze the data. The t-test (with independent samples) procedures were used to compare the two achievement variables in pre-test and post-test for group differences between the CALL group and the control group. The paired t-test procedures (with dependent samples) were used to determine if there is a difference within group across the period of CALL or homework use.

**Preliminary results of data analysis**

The results we obtained concerned student performance before and after ten-week CALL use and their attitudes toward CALL use and the courseware.
Student performance

T-test results regarding grammar and composition performance in both pre-test and post-test sessions show that there is no significant difference in the groups (see Tables 2 and 3). That is, results in Table 2 indicated that the two groups performed equally before the experiment was conducted, given the grammar test and the essay-writing tasks. Results in Table 3 indicated that the CALL group did not perform better than the control group after ten weeks of CALL use. The possible causes for the lack of difference may be attributed to the small number of subjects or the fact that CALL is no better than the traditional homework counterpart.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Groups</th>
<th># of Cases</th>
<th>Mean</th>
<th>SD</th>
<th>T Value</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>59/62</td>
<td>17/14</td>
<td>-.69</td>
<td>NS</td>
</tr>
<tr>
<td>Comp 1</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>69/71</td>
<td>6/10</td>
<td>-.79</td>
<td>NS</td>
</tr>
<tr>
<td>Comp 2</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>77/80</td>
<td>8/9</td>
<td>-1.25</td>
<td>NS</td>
</tr>
</tbody>
</table>

comp1/comp2: composition score by the first/second rater

Table 2. Pre-test performance between CALL group and control group.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Groups</th>
<th># of Cases</th>
<th>Mean</th>
<th>SD</th>
<th>T Value</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>77/72</td>
<td>13/12</td>
<td>1.39</td>
<td>NS</td>
</tr>
<tr>
<td>Comp 1</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>77/74</td>
<td>5/6</td>
<td>1.50</td>
<td>NS</td>
</tr>
<tr>
<td>Comp 2</td>
<td>CALL/Ctrl</td>
<td>22/20</td>
<td>80/82</td>
<td>4.8/5.1</td>
<td>-.82</td>
<td>NS</td>
</tr>
</tbody>
</table>

Table 3. Post-test performance between CALL group and control group.
Not surprisingly, paired t-test results (with dependent samples) show that the two groups as a whole improved after one semester G & W and CALU homework instruction/exposure as indicated in the significant differences in Table 4.

<table>
<thead>
<tr>
<th>Measure</th>
<th>DIF Mean</th>
<th>SD</th>
<th>T Value</th>
<th>SIG Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar</td>
<td>13.93</td>
<td>13.29</td>
<td>6.79</td>
<td>&lt;.000**</td>
</tr>
<tr>
<td>Comp1</td>
<td>5.88</td>
<td>8.20</td>
<td>4.65</td>
<td>&lt;.000**</td>
</tr>
<tr>
<td>Comp2</td>
<td>2.88</td>
<td>9.10</td>
<td>2.05</td>
<td>&lt; 47*</td>
</tr>
</tbody>
</table>

As a whole (# of cases = 42)
CALL (n = 22) vs. Control group (n = 20)

Table 4. Comparison of performance of all subjects after and before the treatment

Respectively, the CALL group seems to have improved more than the control group as indicated in Table 5. Comp2, scores graded by the second rater, is not shown to be significantly different.

<table>
<thead>
<tr>
<th>CALL</th>
<th>Grammar</th>
<th>18.18</th>
<th>15.48</th>
<th>5.51</th>
<th>&lt;.000**</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAII</td>
<td>Comp1</td>
<td>8.09</td>
<td>8.28</td>
<td>4.58</td>
<td>&lt;.000**</td>
</tr>
<tr>
<td>CALL</td>
<td>Comp2</td>
<td>3.91</td>
<td>9.15</td>
<td>2.00</td>
<td>&lt;.058*</td>
</tr>
<tr>
<td>Ctrl</td>
<td>Grammar</td>
<td>9.25</td>
<td>8.50</td>
<td>4.87</td>
<td>&lt;.000**</td>
</tr>
<tr>
<td>Ctrl</td>
<td>Comp1</td>
<td>3.45</td>
<td>7.58</td>
<td>2.04</td>
<td>&lt;.056*</td>
</tr>
<tr>
<td>Ctrl</td>
<td>Comp2</td>
<td>1.75</td>
<td>9.15</td>
<td>.86</td>
<td>&lt;.403 NS</td>
</tr>
</tbody>
</table>

Table 5. Comparison of performance before and after the treatment for individual groups.
In summary, paired T (with dependent samples) procedures indicated that not surprisingly, all subjects improved in their grammar tests (prob. <.000) and composition writing (prob. <.000 and .47 for the two raters). In looking at both groups’ improvement respectively, the CALL group improved regarding the grammar test and two composition scores from the raters. However the control group improved in only the grammar test but did not show progress regarding one of the two composition scores.

Lumping the two composition scores together, we found the result indicated that the CALL group (p<.001*) improved but the control group did not (p>.1 NS) as shown in Table 6.

<table>
<thead>
<tr>
<th>Groups</th>
<th># of Cases</th>
<th>Mean</th>
<th>SD</th>
<th>T Value</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALL</td>
<td>22</td>
<td>6.0</td>
<td>7.28</td>
<td>3.86</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Ctrl</td>
<td>20</td>
<td>2.6</td>
<td>7.38</td>
<td>1.57</td>
<td>&gt;.1 NS</td>
</tr>
</tbody>
</table>

*Table 6. Comparison of two group’s performance after combination of two scores.*

In order to know more about group differences, we used the proportion of increase score (post-pre / pre), instead of the raw scores, as the independent variable and made another comparison (see Table 7). Comparing the proportion indicated that the CALL group significantly improved in their grammar test at the.05 level (prob. =.036) but no significant group difference was found regarding the composition scores (prob. = .263 NS). In this case, the comp measure was the average of the two scores from different raters.

<table>
<thead>
<tr>
<th>Measure</th>
<th># of Cases</th>
<th>Mean</th>
<th>SD</th>
<th>T Value</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar</td>
<td>CALL 22</td>
<td>.41</td>
<td>.468</td>
<td>2.13</td>
<td>&lt;.036*</td>
</tr>
<tr>
<td></td>
<td>Ctrl 20</td>
<td>.17</td>
<td>.165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>CALL 22</td>
<td>.091</td>
<td>.116</td>
<td>1.14</td>
<td>.263 NS</td>
</tr>
<tr>
<td></td>
<td>Ctrl 20</td>
<td>.048</td>
<td>.131</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 7. Comparison of group difference/progress proportion.*
Student attitudes

From the background questionnaire, it is shown that their experiences with language learning were, as expected, quite homogeneous: their LI is Chinese or its dialects; most of them have not visited English-speaking countries; most of them have six years of EFL education and started learning English in junior high school. Their experiences with learning English, however, varied in terms of anxiety levels, sense of achievement, confidence, etc. About half of them regard grammar important in learning to write English; the other half regard communicative effectiveness important. In terms of computer literacy, the results indicated that they had very limited experience with computer use, let alone CALL use.

Regarding exercise formats which subjects regarded helpful, subjects rated sentence correction the highest, 86% out of a 4-point agreement scale (i.e., 86% of them agree that this format is helpful), translation the second (59%), and sentence combining, the last (50%). Still over half of them believe that these three exercise formats are helpful for EFL writing instruction.

In a section about general attitudes toward CALL use, nine question items were designed to elicit the subjects’ responses. As summarized in Table 8, except item 7, over half of the subjects agreed that CALL is helpful in one way or another.

<table>
<thead>
<tr>
<th>Attitudes Toward CALL.</th>
<th>% Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CALL better than text</td>
<td>59%</td>
</tr>
<tr>
<td>2. immediate feedback helpful</td>
<td>82%</td>
</tr>
<tr>
<td>3. promotion of interest in EFL learning</td>
<td>73%</td>
</tr>
<tr>
<td>4. less pressure in CALL than in classroom</td>
<td>73%</td>
</tr>
<tr>
<td>5. time-saving for answer checking in CALL</td>
<td>73%</td>
</tr>
<tr>
<td>6. promotion of CALL use in college</td>
<td>86%</td>
</tr>
<tr>
<td>7. more communicative writing after CALL use</td>
<td>36%</td>
</tr>
<tr>
<td>8. more errors self-corrected in writing after CALL use</td>
<td>77%</td>
</tr>
<tr>
<td>9. more consciousness-raising of errors in writing</td>
<td>68%</td>
</tr>
<tr>
<td>after CALL use</td>
<td></td>
</tr>
</tbody>
</table>

*Table 8. Subject's attitudes toward CALL.*
In summary, it is suggested that the combined effects of classroom instruction and either of CALL or traditional homework exposure may contribute to the subjects’ improvement in both test and essay-writing tasks. But we did not show that instruction plus CALL impacted more of the achievement than instruction plus homework. Interestingly, the CALL group improved more on both tasks but the control group improved only on the grammar test task. As far as the questionnaire results are concerned, most of the responses were positive toward the courseware. At least we are sure that none of the subjects think this CALL strategy is detrimental to their learning. Specifically they regarded the sentence-correction exercises as the most useful format, with translation the second and sentence-combining the third. There were many more variables of language learning processes we did not examine in this study. One of the most ambitious future projects will be to analyze the student record files to extract their interlanguage systems as initiated by Garrett (her dissertation work was discussed in 1987 but completed in 1982). In the following semester (Spring of 1992), we are asking the control group to work on the grammar courseware. We will investigate the learning processes more closely by using interview techniques in addition to the pre/post-test measures and questionnaires. Hopefully we can get a dearer picture.

CONCLUSION

This project aims to address a general research question: can grammatical CALL help EFL writing instruction? The steps toward the answers are developing a piece of CALL grammar courseware and conducting a quasi-experimental study. Though the study did not show CALL to be superior to paper/pencil homework, we did not find any evidence that indicates CALL is detrimental. The combined effect of instruction and CALL seemed to contribute to more learning gain than that of instruction and homework/exposure alone. The student attitudes showed that they liked CALL and thought CALL useful for language learning. The tentative answer we provide at this stage of enquiry is a conditional yes: when combined with classroom writing instruction, grammatical CALL is helpful for writing performance. Clearly, more research studies should be conducted to uncover the complex learning behavior underlying the process.

In terms of courseware refinement, we plan to make the courseware a template-like package so that a language teacher can easily devise his/her own material to fit in the program. The courseware can thus operate with any desired contents at various levels.
In addition within each acceptable and good answer for each item, some kind of markup (e.g., Hart, 1989) can be designed to make answer-judging more diagnosis-specific so that users can learn to approximate the best answers.

REFERENCES


**AUTHORS' BIODATA**

Hsien-Chin Liou is currently associate professor in Foreign Languages, National Tsing Hua University at Hsinchu, Taiwan, R. O. C. She teaches CALL, EFL Writing, and introductory Computational Linguistics. Her research interests include CALL for writing instruction, multimedia courseware development, and CALL research.

Samuel H. Wang is currently an associate professor in Foreign Languages and Linguistics at National Tsing Hua University, Taiwan, R. 0. C. His research areas include experimental phonology, phonetics, psycholinguistics, machine translation, and CALL.
Yuli Hung-Yeh is an instructor in the Department of Foreign Languages, National Tsing Hua University. She is also a candidate in the Ed. D. Program of the Institute of Instructional Systems Technology of Indiana University at Bloomington.

AUTHOR’S ADDRESS

Hsien-Chin Liou
Foreign Languages
National Tsing Hua University
101 Sec. 2 Kuang-Fu Road
Hsinchu, Taiwan, R. O. C. 30043

PH: 886-35-715131 ext. 4405
FAX: 886-35-718657
E-mail: hcliu@fl.nthu.edu or twnthut048@twnmoe10.bitnet