

## CHAPTER V

### SUMMARY, DISCUSSIONS, AND RECOMMENDATIONS

This chapter presents the summary of the study, the discussions of the results, the implications yielded by the study, and the recommendations for teachers and for further research.

#### 1. Summary of the study

##### 1.1 Purposes

The main purposes of this study were twofold. The first one was to investigate the effects of 'error treatments' and 'language abilities' and their interaction on the usage of English tenses, and second, to examine the retention of knowledge in students who received a different error treatment type. Error treatments comprised two levels— overt correction (OC) and self-correction (SC). Language abilities that served as a moderator variable had three levels— low achievers (LA), moderate achievers (MA), and high achievers (HA).

##### 1.2 Research design

The present study is a 2x3 factorial design. The diagram of the design is as follows:

Self-correction	$R$	$O_1$	$M_r$	$X_1$	$O_2$	$O_3$
Overt correction	$R$	$O_1$	$M_r$	$X_2$	$O_2$	$O_3$

The three observations include pretest, posttest, and delayed test.

##### 1.3 Samples

The samples are 210 first-year undergraduate students of Huachiew Chalermprakiet University, Thailand. The students were taking the course *GE 1063 English for Communication II* in semester 2, 2006.

The samples were reached by three steps. First, the *cluster random sampling* technique drew 6 groups from the 23 existing study groups. Second, all students were pretested and were then labeled as LA, MA, or HA by using the 30<sup>th</sup> and 70<sup>th</sup> percentile

ranks of the pretest scores. Last, the samples were matched and were then *randomly assigned* to each of the two study groups— OC and SC.

#### 1.4 Experimental materials

The experimental materials refer to the two CALL software programs that represent the OC and the SC treatments. The OC program represented the traditional ‘teacher correction’ in which the teacher would indicate that the answer is wrong, explain, and reveal the correct answer. The SC program, on the other hand, would not give the correct answer to students. Students were encouraged to find the correct answer by themselves. The program would provide ‘scaffolding’ feedback as the help to guide the students to think and to get to the correct answer.

The experimental materials, CALL, comprised 9 modules– 6 lessons and 3 tests. The contents cover seven most frequently used tenses in English. In each lesson, two related tenses were compared and contrasted. The three tests were basically similar to the six lessons in terms of the error treatment they provided. However, the coverage of contents of the tests was broader. Test 1 combined contents from lesson 1 and lesson 2, meaning that it covered three tenses. Test 2 covered five tenses and test 3 mixed up all seven tenses together. Another function that was added to the tests was “score” presentation in order to help students evaluate their progress in each particular practice area. On the last frame of the tests, there was a summary of the score presented both in raw data and in percentage.

#### 1.5 Instruments

The instruments used for collecting data in the present study include: Test of English tenses usage (TETU), the self-rating scales, and the opinionnaire.

**The Test of English Tenses Usage (TETU):** The test was developed by the researcher to serve in pre-testing and posttesting. Its contents cover seven English tenses that the subjects have practiced through the CALL programs. The TETU comprises 2 parts, starting from simple to more complicated tasks. The first part deals with *sentence completion tasks*. This part consists of 30 discrete-point items with 5 alternatives. The item was presented with a blank in which the students were asked to choose the appropriate verb form to fill. The second part is called *grammar in context* because the contexts are given together with the items. Normally, items in this part came in longer conversations. In addition, the conversations would cover more than one tense. There are 20 multiple-choice items in this part.

The scoring system for both parts was very objective. Students would receive one point if they chose the correct answer, or zero if they picked up the wrong choices. Total possible score for these parts was 50.

**Self-rating scales:** The self-rating scales were developed for eliciting the students' perception of their own learning. Each individual student was asked to rate his/her knowledge about the seven tenses from 4 (excellent) to 0 (none). The instruction was in Thai. The scales were employed twice, once before the treatment and again after the treatment.

**Opinionnaire:** The opinionnaire was designed to explore students' opinion towards the program that they were using. There were 13 items that covered the opinions towards the contents, design, self-motivation, and feedback. All items were printed in Thai. Students were asked to rate from 1 (strongly disagree) -5 (strongly agree) according to the extent to which they agree with the statement. The opinionnaire was distributed to the students at the end of the study.

### **1.6 Procedure**

At the very first week, all subjects were pre-tested. By using scores from the pretest, the subjects were labeled as high, moderate, or low achievers. After matching the scores, the subjects were randomly assigned to the two error treatment groups— overt correction or self-correction. In the second week, the researcher gave an orientation for the subjects, and then let them have the CD and the checklist. The subjects were also asked to answer the self-rating scales in this week. After they received the CD, the subjects practiced only with the program they received. They could come to the university lab which was reserved for this project and practiced or they could do it on-line at home if the lab time did not match their schedule. The researcher controlled their practice by checking results from the server that provided information about dates, time-in, time-out, number of trials, and sequence of the choices made by the students, and met the subjects after their regular class every week. The subjects practiced with the program every other day. Therefore, they could complete all nine modules within 3 weeks. If the subjects failed to follow the schedule on their checklist, the researcher would talk to them so that they could catch up with the others. The researcher checked the information from the database to make sure that they all completed the program. The posttest was conducted in week 6. Also, the opinionnaire and the self-rating scales were distributed to the students. After the posttest, the subjects stopped practicing with the program. They returned the CD and the checklist, so they did not get any more exposure

to the program. Still, they went to their regular English class. The delayed test was administered in week 12.

### 1.7 Results

The results and findings of the present study in response to the research questions can be summarized as follows:

1. Different methods of error treatment did not have a significant effect on the usage of English tenses of the students. Students in the OC and the SC groups made similar gains ( $F = .015, p > .05$ , one-tailed).

2. The students' language ability was found to be a significant source of the differences in the posttest gains ( $F = 146.525, p < .05$ , two-tailed) with a large effect size. Fifty-nine percent of the overall variance could be explained by this factor. It means that students in different ability groups (LA, MA, and HA) performed significantly different in the posttest. HA students performed better than MA students and MA students performed better than LA students.

Results from further investigation provide detailed information that tenses that all ability groups used differently in all of the three tests were simple present, simple past, and present continuous. However, the LA group used some tenses without significant difference from the MA group; for example, the use of past continuous tense in the pretest, the use of the use of future tenses in the delayed test, and the use of present perfect in all tests.

Regarding the easy or difficult tenses for different ability groups, it was found that from the pretest, the notable tenses used correctly by the LA group were present continuous and simple present. The MA group could perform best in simple present, present continuous, and future tenses. HA students could prominently use the following 5 tenses in order: simple present, present continuous, past continuous, simple past, and future tenses. From the posttest and the delayed test, results were not much different from the pretest, except that the past continuous became one of the top three tenses used correctly by all ability groups. Present perfect and past perfect were found the most difficult tenses for all groups, followed by simple past.

3. There was no significant interaction effect between types of error treatment and different language abilities on the usage of English tenses found in the posttest ( $F = .020, p > .05$ ). High achievers made greater gains than moderate achievers, and moderate achievers made greater gains than low achievers regardless of the error treatment type they received.

4. The decrease of scores of the SC group from posttest (test2) to delayed test (test 3) was small and insignificant while the decrease of scores of the OC group was greater and significant. This means that the self-correction treatment was better than the overt correction in terms of the retention. The SC students could maintain the knowledge in a higher percentage when tested by the delayed test.

5. The students' perceptions of their own learning were obtained by analyzing data from the self-rating scales. It was found that students in both SC and OC groups rated themselves at the post-treatment scales significantly higher than at the pre-treatment scales for every tense. It means that they perceived they had more knowledge about the 7 tenses after using the CALL than they had been before using the program. However, students who practiced with different error treatment type did not have significantly different means in the post-treatment scales. This means that the differences across OC and SC groups did not occur.

Results from the opinionnaire showed that, overall, students in both OC and SC groups had highly positive opinions towards the program that they used ( $\bar{x}_{oc} = 3.91$ ,  $\bar{x}_{sc} = 4.01$ ). When the means were tested by one-sample t-test, it was found that SC students had highly positive opinions towards the SC program in almost all items, except for item 8 that asked about their attention when practicing. On the other hand, the OC group had 3 items that they rated lower than the criteria of 3.5.

Results from Chi-square test by Fisher's Exact test revealed that in general the OC and the SC groups rated the opinionnaire insignificantly different. However, the SC group rated the materials significantly higher than the OC in terms of the 'feedback' (item 10).

## 2. Discussions

The findings will be discussed under seven topics: (2.1) the effects of types of error treatments on the usage of English tenses, (2.2) the effects of language abilities of the student on the usage of English tenses, (2.3) the interaction effect of error treatments and language abilities on the students' usage of tenses, (2.4) the retention, (2.5) the use of tenses by students with different abilities, (2.6) the students' perceptions of their own learning, and (2.7) the students' opinions.

## 2.1 The Effects of Types of Error Treatments on the Usage of English Tenses

Based on the finding from two-way ANOVA, the 'error treatment' factor did not have a significant effect on the posttest scores. This revealed that the two types of error treatment, namely OC and SC, did not have different effects on the usage of English tenses. Hypothesis 1 is then rejected. The question is: Why did the two types of error treatment have more or less the same effect on the students' usage of English tenses? This could be due to the following reasons:

### 1. The Characteristics of the materials

The first reason might be due to the characteristics of the materials (CALL). With regard to the six influences on 'noticing' discussed in Chapter II (see Figure 2.3), the two materials were similar in terms of *frequency* and *saliency*. It was obvious to the learners that the programs were designed to help them learn about tenses. The forms were prominently presented. Since the programs had parallel contents and the contents were presented in the same fashion, the third influence, *instruction*, was not relevant. The fourth influence discussed by Schmidt (1990; cited in Skehan, 1998: 48), *individual differences in processing ability*, has not been explored; it is then abandoned, assuming that both groups were not different. Regarding the *readiness to notice*, the fifth influence, Schmidt has applied cognitive processing principles concerning the acquisition order to claim that noticing also depends upon readiness. This implies that a prediction can be made about what the learner can profitably notice (Skehan, 1998: 50). The OC and SC groups contain the same number of students from each ability group. Therefore, we can probably say that both groups have similar average *current state of the interlanguage system*, overall. The last influence to be discussed is *task demands*. Among the six influences, this is the only one aspect in which the two programs are clearly dissimilar. The OC program has noticeably lower task demands than the SC program. Students who used the SC program had to engage more in the task to figure out the answers by themselves.

Considering the degree of explicitness, it was found that the SC program was more implicit while the OC was more explicit. Despite the dissimilarity in the degree of explicitness, both SC and OC programs were under the same type of feedback—that is *negative evidence* (see Figure 2.2). Negative evidence has been found effective in eradicating errors in many studies (Ellis, 2002; Ayoun, 2001; Carroll & Swain, 1993; Carroll, Swain, & Roberge, 1992; Mackey & Philp, 1998). So, by taking a broader view and looking at both OC and SC treatments as negative evidence, it is possible that the OC was as good as the SC in correcting errors in the immediate test. The findings in the

present study provide additional supports to the previous studies that are in favor of negative evidence.

An additional explanation is related to the nature of the materials and the tests. The structure of the experimental materials (CALL) was similar to that of the TETU. For example, both were in a multiple-choice format, the sequence of contents started with a simple and concrete item and ended with more complicated conversations or paragraphs. Even the item difficulty was about the same. The main difference between the CALL and the test was just the mode or channel in which they appeared. The former was given through computer while the latter used traditional paper-and-pencil mode. Both OC and SC groups could possibly perform well because they were familiar with the format and the structure of the test as they practiced with the program.

## 2. The advantages of CALL

As reviewed in Chapter II, CALL is a very useful language learning tool in today's world. It has numerous advantages over the traditional class. For example, it offers privacy and promotes individualization in students. Previous CALL studies report superiority of CALL over traditional classroom practice.

The present study used CALL as the experimental materials. Both OC and SC were delivered through computer. The use of CALL in the present study may have a role to explain why the OC and the SC groups performed more or less the same in the posttest.

Normally, English foundation courses at Huachiew Chalermprakiet University are taught through a normal lecture in a relatively large class. The CD was used just to provide the listening practice and it was mainly used by the teacher. When the students participated in this study, they had a chance to practice with the CALL by which they could better control their learning. Although the two treatment groups received different types of error treatment, a full explanation in Thai was provided for both programs. There was no time limit for each module. The students could spend as much time reading and making themselves understand the points. This might result in the similarity of the mean scores of the two groups.

At this point, it is also interesting to further investigate whether or not the CALL can help facilitate the usage of English tenses by comparing the means of the pretest to the means of the posttest. *Paired samples t-test* was applied. The results are presented in Table 5.1.

**Table 5.1: Comparison of the pretest and the posttest mean scores of the two error treatment groups**

Error Treatment		N	$\bar{X}$	S.D.	Mean Difference	df	<i>t</i>	Cohen's <i>d</i>
Overt	Pretest	105	20.39	8.392	-3.92	104	-6.558*	0.42
	Posttest	105	24.31	10.117				
Self	Pretest	105	20.48	8.563	-3.95	104	-6.663*	0.41
	Posttest	105	24.43	10.498				

\*  $p < .05$

It was found that, for both groups, the differences between the pretest and the posttest scores were significant. This means that both OC and SC groups have significantly higher scores in their posttest than in the pretest. The effect size calculated by Cohen's *d* reveals 'medium effect' for both groups, meaning that both CALL programs could reasonably help the students learn about the tenses. This finding appears to corroborate the results of other studies that reported the effectiveness of CALL as a language learning tool (Elkins, 1986; Avent, 1993; Maneekul, 1996; Nutta, 1996; Katekaew, 1997).

Additional support to this point can be drawn from the student's side. Results from the student self-rating scales showed that students in both treatment groups perceived that they knew more about the seven tenses than they had at the beginning of the study. So, this can be said that the current results provide supports that CALL (regardless of error treatment types) is beneficial for learners in language learning.

### 3. Time factor

Another factor that may explain why the students receiving OC and SC made the same gains in the posttest is the time factor. The student in the present study was exposed to the program only for 3 weeks of their practice period which is probably too short to see remarkable differences. In addition, the students were allowed to practice with each module only once (as an attempt to control equal exposure of the two groups). In fact, in real-life situations, students can use the program again and again to consolidate their knowledge, especially when scores from the test modules suggest that they do not achieve the pass criteria. Furthermore, the posttest was conducted in the fifth week which was the immediate week after the last practice module. Due to the short period of time,



students in both groups might still have fresh memory with them. It was not difficult, even for the OC group, to refer back to the questions they had just practiced with the programs when answering the immediate posttest. Therefore, there was a tendency that the OC and the SC groups performed similarly in the posttest due to the time restrictions.

It should be also noted from the researcher's observation, that the subjects took significantly less time to answer the posttest than they did in the pretest. Some students spent only 10-15 minutes to complete the test. As a result, the students might get lower scores than what they could when answering with full attention. This happening usually occurs in pretest-posttest research. In fact, it does not directly affect the current study because we are not focusing on the comparison of pretest to posttest scores. Rather, we compare the gains across groups. The decrease of attention occurs similarly in both SC and OC groups; then, it does not concern us.

The decrease of attention when doing the posttest may come from three reasons. First, the pretest was conducted at the very first lesson. The subjects thought it might be somewhat related to the course final grade, so they did it the best they could. Most of them spent the full one hour given. However, they later learned from the orientation and the research administration that this was just a research study that had nothing to do with the grade. Then, they spent less time when answering the posttest. The second reason is owing to the pretest-posttest effect. The present study uses the same instrument in pretesting and in posttesting. The students, then, had seen the questions before when answering the posttest, so they spent less time. The last reason comes from a positive view that the students had acquired the knowledge from the programs they used. The students knew more; thus, they used less time.

## **2.2 The Effects of Language Abilities of the Student on the Usage of English Tenses**

Based on the finding from two-way ANOVA, language abilities had a significant effect on the students' usage of English tenses ( $F = 146.525, p < .05$ ). Mean scores revealed that the high ability group outperformed the moderate group, and the moderate group outperformed the low group. The measures of effect size revealed a 'large' effect with the Eta squared of 0.59. As a result, hypothesis 2 was accepted. The question is: Why did different language abilities have a large and significant effect on the students' usage of English tenses? The possible reasons might be the following.

### 1. Characteristics of high and low ability students

The results presented in Table 3.2 and Table 4.1 reveal that the group that seemed to improve most is the HA group. This group had the largest mean difference at 4.35 points ( $\bar{x}_{\text{pretest}} = 30.36$ ,  $\bar{x}_{\text{posttest}} = 34.71$ ). Considering the mean differences, we might be able to say that the HA group benefited most from the program because the group had the greatest increase of mean scores.

As discussed in Chapter II, students with different abilities may differ in many ways. They have different attitudes towards autonomous learning (Soinam 1999; Supyan, 1994; cited in Sukamolson, 2000: 31). They apply different strategies when learning (Brandl, 1991). Also, they have different characteristics that might explain why the HA group outperformed the MA, and the MA outperformed the LA.

High ability students hold the following characteristics: can learn fast, have relatively higher concentration, are eager to learn, and have self confidence (Jan-aim, 1978 and Sukpanphotaram, 1984; cited in Leetaweekulsomboon, 1996:34). On the contrary, low ability students can learn slowly, need more time to digest what have been taught, lack self confidence, are easily influenced by the others, have low motivation and less attention.

Data from the server record sheets which recorded the students' performance during the practice period can help explain such characteristics. Average duration of time spent to complete each module as well as the number of items taken before they got to the correct answer are illustrated in Table 5.2.

**Table 5.2 Data from the server record sheets**

	Lessons	HA	MA	LA
1. Average duration of time in completing the module (minutes)	1	10.98	16.32	18.96
	2	12.52	14.67	20.03
	Test 1	12.69	15.21	20.56
	3	16.87	20.36	21.11
	4	15.55	19.53	22.55
	Test 2	16.54	22.37	22.57
	5	19.09	21.88	23.13
	6	18.66	21.01	23.34
	Test 3	18.85	26.64	25.31

	Lessons	HA	MA	LA
2. Number of choices taken to get to the correct answer	1	1.36	2.65	2.87
	2	2.01	2.95	3.06
	Test 1	1.93	2.88	2.96
	3	2.17	2.98	3.31
	4	2.03	2.56	3.14
	Test 2	2.05	2.43	2.99
	5	2.07	3.06	3.20
	6	2.64	3.12	3.07
	Test 3	2.15	2.79	3.01

Data from the record sheets confirmed the differences in the characteristics of high, moderate, and low ability students. It can be seen that high ability students took fewer choices to get to the correct answer in every tense they practiced. Also, they used less time to complete the module, especially when they practiced with the first tenses. This provides solid evidence that high ability students in the present study could learn things faster than lower ability ones.

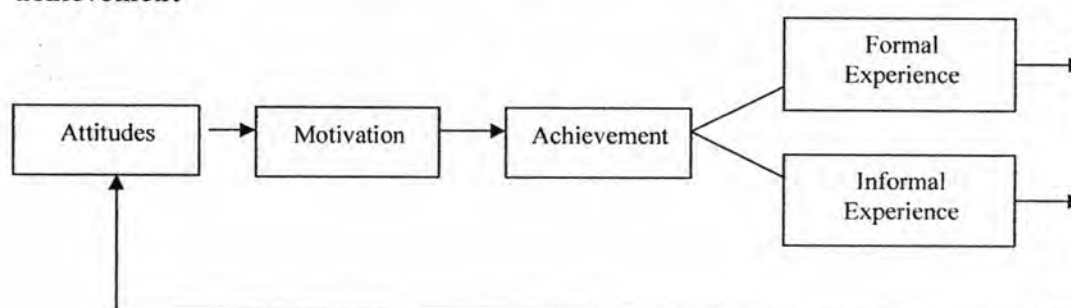
Also, from the researcher's own experience, high ability students tend to compete with one another. So, when they practiced with the program they were eager to learn whatever the program could help them with learning English better or overcoming their weaknesses. And because they generally had higher concentration and learn faster, they could improve most in terms of the mean differences. The results of this study are consistent with the previous research studies that reported effects of different language abilities on the achievement of students (Yindeetakul, 1986; Leela-ongart, 1987; Chollatarn, 1987; Sripathomsawad, 1988; Khansorn, 1988)

However, one might have a question when comparing the 'maximum scores' of the pretest and the posttest and find that the HA group got only 6 points up while the MA and the LA groups both improved 19 points. Yet, this does not mean that the programs were not beneficial for HA students. In general, a test comprises items of different difficulty levels. Some are easy and obvious and some are more complicated. To be fair to the HA group, the items that they could not do in the pretest are more difficult than the items that the MA and the LA could not answer. So, there was not much room for HA students to improve. Furthermore, the room was packed with tough items. The 6 points expansion of their ceiling was really impressive. Moreover, one of the HA students could even reach the perfect score of 50.

## 2. Attitudes of the students

It is often found that attitudes are causally related to attainment in language learning. Rokeach (cited in Smith, 1972: 16) defines attitudes as “a relatively enduring organization of beliefs around an object or a situation, predisposing one to respond in some preferential manner”. Brown (1987: 127) points out that everyone has both positive and negative attitudes and that learners seem to benefit from positive attitudes and that negative attitudes may lead to decreased motivation. And as suggested by Garner’s socio-educational model (1985), the decrease of motivation will influence language achievement (Figure 5.1).

**Figure 5.1: Gardner’s (1985) socio-educational model of affects and language achievement**



The relationship and causal effect of attitudes on language learning have been studied for a long time. Attitude is found to be a remarkable variable that has a significantly positive relationship with language achievement in different socio-cultural contexts. For example, Oller and his colleagues (Brown, 1987: 127) conducted several large scale studies of the relationship and found that positive attitudes towards self, the native language group, and the target language group enhanced proficiency. For the Thai context, attitude was reported as one of the significant predictors of students’ language achievement (Prapphal, 1981; Udomsamuthirun, 1998).

According to the literature review in Chapter II, students with differing abilities had significantly different attitudes towards autonomous learning (Soinam, 1999), and high ability students liked to use CALL more than low ability students (Supyan, 1994; cited in Sukamolson, 2000: 31). High ability students had more positive attitudes towards autonomous learning, so they had more motivation to learn with the CALL (Figure 5.1). As a result, they made greater achievement scores. According to the model, their achievement would, in turn, foster more positive attitudes which result in better

achievement. This is how attitudes attribute to the findings in the present study that high ability students made the greatest gain in this study.

Actually, the main effect of 'language abilities' is expected by the researcher as she posted it in the second hypothesis. However, the effect of language abilities solely is not the focus of this study because it is a moderator variable– the variable that we could not control or manipulate (Hatch & Farhady, 1982, 151) We, alternatively, would like to explore the interaction effect of different types of error treatment and different language abilities on the students' usage of tenses to see whether any particular error treatments work well with any ability groups.

### **2.3 The Interaction Effect of Error Treatments and Language Abilities on the Students' Usage of English Tenses**

There was no significant interaction effect between error treatments and language abilities on the usage of English tenses of the students in the posttest ( $F = 0.02, p > .05$ ). Hence, hypothesis 3 is rejected.

With reference to Figure 4.1, it was clearly demonstrated that the three lines were parallel. There was not a cross of the lines on the plot. No particular treatment was the preference for any ability groups. Low ability students in the OC group gained similar posttest means to the ones in the SC group. The situations were the same for the moderate and the high ability students. We can say that the OC and the SC students had similar gains regardless of their language abilities.

The reason might be due to the fact that the types of error treatment did not have a significant effect on the students' usage of English tenses due to the reasons discussed earlier. With the effect size of .001 conveying that the practical significance was almost equal to zero, there might be a tendency to find that language abilities did not yield a significant interaction effect on the same dependent variable either.

The finding in the present study is similar to the results reported by Maneekul (1996), Sripathomsawad (1988) and Khansorn (1988) discussed in Chapter II. Maneekul examined the effects of normal instruction supplemented with the CAI grammar game on the achievement of Thai vocational-technical students and reported that there was no significant interaction between the types of teaching methods (Lecture and Lecture supplemented with CAI) and levels of ability (high and low) on the achievement. Sripathomsawad, similarly, did not find any interaction between CAI picture options and levels of the students on their achievement. In his study, Khansorn also reported that

there is no significant interaction effect between the CAI feedback and student level on their achievement.

On the contrary, the finding in this study is not consistent with Yindeetakul (1986) who found that low ability students performed well when receiving the expository approach while high ability students gained higher scores in the group that received the discovery approach. This might be because Yindeetakul applied totally different approaches to her study while the two error treatments in the present study were not so different. OC and SC both use the same type of feedback that is negative evidence which has been found effective in many studies. So, this might result in different findings. Another point in which these two studies differ concerns the number of subjects who participated in the study. Yindeetakul had a total of 72 students which were divided into 4 groups; each contained 17-19 students while 210 students participated in the present study. Applying a parametric statistical analysis (two-way ANOVA) to a small sample size may result in a limitation in making a generalization.

The present study has provided empirical evidence that the language abilities had a significant effect on the usage of tenses; however, its interaction with the 'error treatment' did not. Students with differing ability may be different in many ways, for example, the way they choose to receive the feedback from the computer, the strategies they use to figure out the answer, or their attitudes towards autonomous learning (see details in Chapter II); however, they tend to have no preference in terms of the error treatment that works specially well for their ability.

## **2.4 The Retention**

### **2.4.1 The Retention in Different Error Treatment Groups**

The results revealed that the decrease of scores of the SC group from posttest (test 2) to delayed test (test 3) was small and insignificant while the decrease of scores of the OC group was greater and significant. This means that the SC students could maintain the knowledge in significantly higher percentage than the OC ones. Hence, hypothesis 4 is accepted. The question to be discussed here is: Why could the SC group retain the knowledge at a higher rate than the OC one in the delayed test? This might be due to the following reasons.

### 1. Degrees of Noticing

The main reason could be due to the fact that the program that the two study groups used provided different types of error treatment. When the student made an error, the OC program would indicate that the answer was wrong, reveal the correct answer, and provide a full explanation. The SC program, on the other hand, would not reveal the correct answer but would rather provide the scaffolding support to encourage the students to self-correct. According to the background theories discussed in Chapter II, the SC was expected to promote long-term learning.

In relation to the *Noticing Hypothesis*, the present study bears out the proof for the influence of task demands. The OC program had relatively low task demand compared to the SC program. The SC program required the students to participate more actively in their learning. Students had to try to identify the correct answer by themselves, with the help of the program in supplying scaffolding feedback / explanations. Such higher task demands could lead to the greater degree of *noticing* as suggested by Schmidt (1994; cited in Skehan, 1998). So, this could be the reason that the SC group was more successful than the OC.

### 2. Repetition of task as a rehearsal

Apart from the higher task demand of the SC program, another factor that fulfills the learning process is the 'rehearsal'. Skehan (1998) points out that rehearsal is the process that helps transfer information from the short-term to the long-term memory. According to the Noticing model (Figure 2.3), the noticed information can be processed in the working memory or go directly to the long-term memory if it is highly consciousness-raising. Either way, rehearsal is very important. If the information is stored in the working memory, rehearsal helps transfer the information to the long-term memory. If the noticed information goes directly to the long-term memory, rehearsal helps maintain the information.

The CALL modules were laid in a linear sequence of which the contents were 'topped up'. The students started with Lesson 1, then Lesson 2, and then Test 1. Of all the nine modules, there were 6 lessons and 3 tests. As the students finished every two lessons, they had a test which in fact was very similar to other lessons except that (1) it reported the score to the students while they were practicing and (2) it had a broader content that covered all contents from the previous modules. So, the contents of Test 1

covered the three tenses introduced in Lessons 1-2; the contents of Test 2 covered the five tenses from Lesson 1-4; and the contents of Test 3 covered all 7 tenses mixed in longer conversations/ passages. It can be seen that the students had chances to restructure their knowledge through the combination and repetition of the contents.

The noticing of the contents alone may, to some extent, lead to the success in language learning. However, when it is incorporated with rehearsal, it leads to a stronger tendency that the students would achieve the long-term goal. The SC students in the present study were not just trying to memorize the form but the program helped them notice the forms, and then restructured the new knowledge to the existing one by rehearsal. Therefore, SC students who had a greater degree of noticing incorporated with enough rehearsal could retain more of the knowledge.

### 3. The more active role in the learning process

Additional support to explain why the knowledge learned by SC lasted longer than one by OC can be drawn from the view of constructivists. Constructivists believe that all human beings construct their own version of reality (Brown, 2000: 11). In the context of language learning, learners are believed to be able to construct their own knowledge. Constructivists suggest that learners should be engaged more actively in their learning. Comparing the two error treatments applied in this study, it can be seen that the students who used SC program took a more active role than the OC group. When an error occurred, they had to try to figure out what was wrong by reading the feedback and/ or the explanations provided. If the first trial failed, they had to rethink again, and again, until they could get to the correct answer. On the contrary, OC students were not as much active as the SC in terms of the interaction with the program. When they chose the wrong answer, the program stated that the answer was wrong, revealed the correct answer, and provided a full explanation. Therefore, taking the view of the constructivists', the SC group which took a more active role in their learning could unsurprisingly perform better than the OC one.

### 4. The time factor

In terms of the administration of the study, the duration between the tests is another key point that should be discussed. The delayed test was conducted 6 weeks after the posttest. In general research practice, the duration is considered long enough to demonstrate the long-term effect. Additionally, the period of time covered the New Year



holidays on which the students were so happy and busy that they did not concentrate on the study for a while. The delayed test was conducted in the next two weeks after the holidays. The holidays made the proof about retention more valid because, in real life, students have to go through a lot of such situations that might distract them from their learning. Therefore, it can be said with more confidence that self-correction is an effective error treatment to promote higher retention.

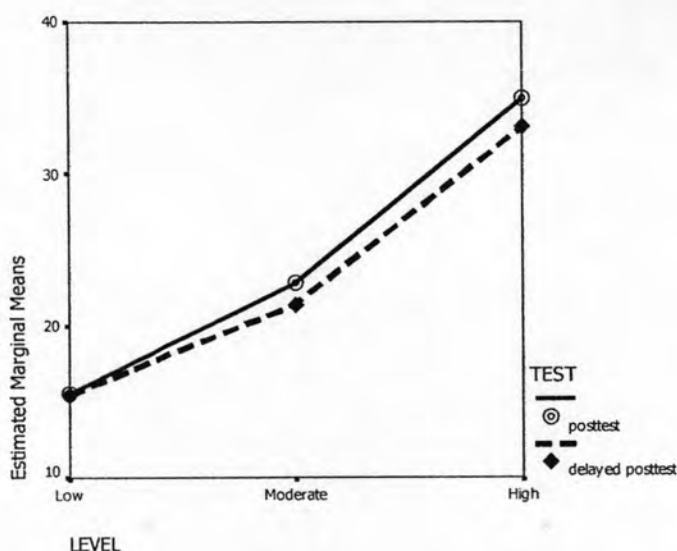
In short, it is obvious that students in the SC group performed better in the long run. The findings confirm that self-correction is better in promoting long-term learning. The students may demonstrate similar ability in the short-term assessment, but the delayed test proves that the knowledge gained by self-correction lasts longer. This finding provides additional supports to the previous studies reviewed in Chapter II that self-correction is more effective in promoting last-long learning; regardless of the channel where it is applied– in normal class or through CALL.

However, the findings in this study may look different from Ellis, Loewen, and Erlarm's study (2006) which reported that explicit teaching (metalinguistic explanations) of the past tense 'ed' was more effective than implicit learning (recasts) in the delayed imitation test. The main reason is that the two studies in fact focused or compared feedback at different layers. Recasts which were used in Ellis et al.'s study is implicit 'positive' evidence (Figure 2.2) because 'recasts' just provide grammatical models to the learners and do not give information about what is unacceptable in the learners' productions. So, what Ellis et al. actually studied was the comparison of negative and positive evidence. And the results pointed out the superiority of negative evidence. The present study, on the other hand, compared the two sub-types of negative evidence namely explicit (overt correction) and implicit (self-correction) feedback and found that implicit negative feedback could better facilitate long-term learning than explicit negative feedback.

### 2.4.2 The Retention in Different Language Ability Groups

Another dimension that is worth discussing concerns the retention of the knowledge in different ability groups. Although this is not one of the major research questions; it would, however, help us understand the students more clearly. Figure 5.2 depicts what happened with the retention of each ability group over time.

**Figure 5.2: Retention of each language ability group**



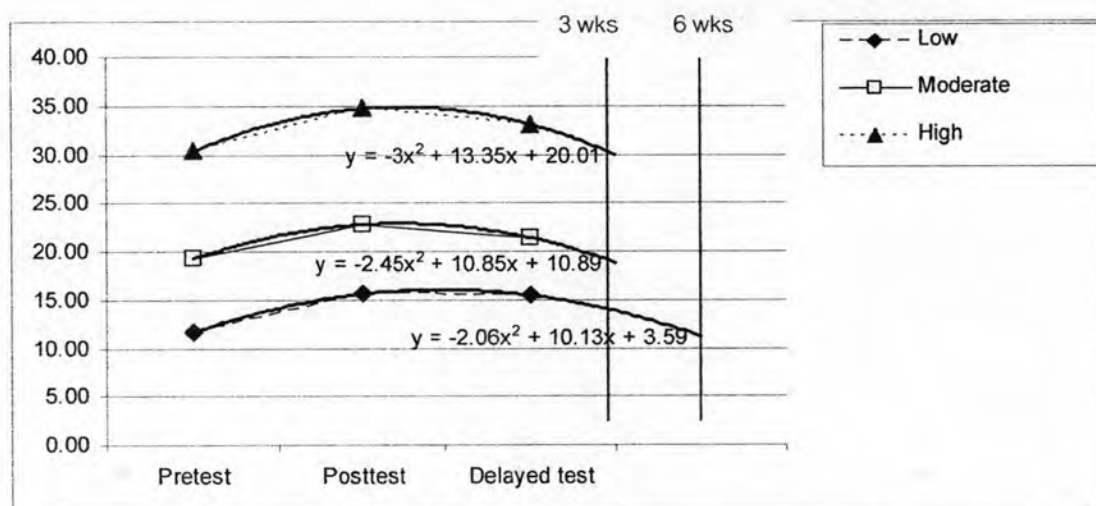
It is normal to find that the mean scores from the delayed test are lower than those of the immediate posttest. So, it is understandable to find that the HA and MA groups had some score drops. However, the LA group could surprisingly maintain their mean score ( $\bar{X}_{\text{posttest}} = 15.484$ ,  $\bar{X}_{\text{delayed test}} = 15.438$ ). This result is consistent with the findings of Bangert-Drown, Kulik, and Kulik's (1985; cited in Leetaweekulsomboon, 1996: 34) study which reported that CALL helped facilitate learner's achievement, especially in the group of low ability students.

The relatively high retention of the low group in the present study could be mainly due to the advantages of CALL in that it allows students to learn at their own pace. Low ability students who are normally left behind in normal classroom could spend as much time as they wanted reading and trying to understand the grammar rules. Besides, it allows privacy so that the low ability students felt at ease to use the language without worrying about being embarrassed. Furthermore, their newly learned knowledge

was not as complicated as the one learned by the MA and the HA groups; it could, thus, be maintained in a higher percentage.

Apart from trying to explain why students with different abilities retained the knowledge at different rates, it is interesting to further investigate their learning trend to see the tendency of their knowledge retention. Results in Figure 5.3 illustrate the means' plots as well as their trend lines as projected for the next 3 weeks.

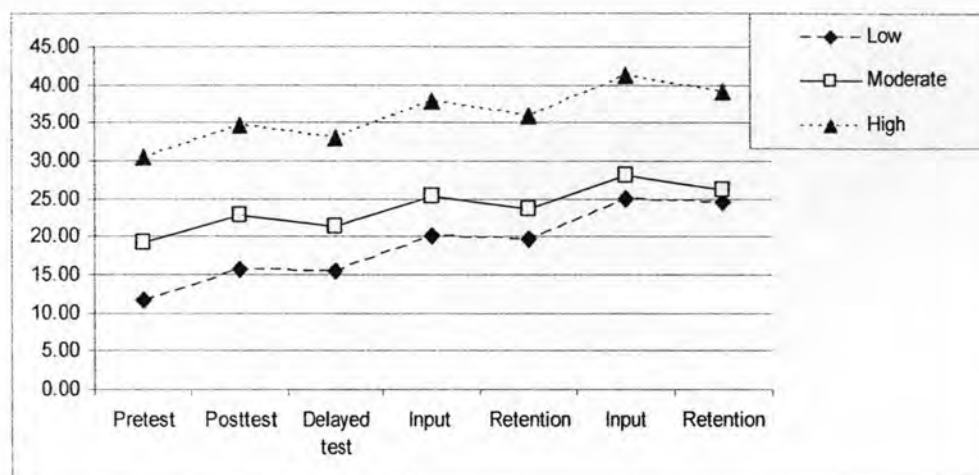
**Figure 5.3: Plots of means and trend lines of the three ability groups**



It can be seen that the three lines representing means of the three ability groups are almost parallel, with slightly different slopes. The figure reveals that mean scores from the delayed test were still higher than those of the pretest. This means that students in all ability groups could maintain quite a lot of knowledge when they were tested 6 weeks after the exposure to the last CALL module. If the students entirely stop being exposed to the program, the trend lines suggest that their knowledge would decrease and reach the level of their previous knowledge from the pretest within approximately 3 weeks (for the HA and MA groups) to 6 weeks (for the LA group).

However, based on the finding previously mentioned, if the students were exposed to the input (with the same focus of contents) again some time after the delayed test, it would be possible that they would be able to discontinue the decrease of scores and create rising trend lines again. Repeatedly and gradually, the students could become more proficient learners. The idea of providing continuous input is illustrated in Figure 5.4.

**Figure 5.4: An illustration of the retention in continuous learning**



### 2.5 The Use of Tenses by Students with Different Abilities

Results from the three observations analyzed by one-way ANOVA and Scheffe tests revealed that there were three tenses that students in all ability groups used differently in terms of the correctness of the answers in all the tests; they included simple present, present continuous, and simple past. However, the LA group used some tenses insignificantly different from the MA group; for example, the use of past continuous tense in the pretest, the use of the use of future tenses in the delayed test, and the use of present perfect in all tests.

Regarding the easy or difficult tenses for different ability groups, it was found that from the pretest, the notable tenses used correctly by the LA group were present continuous and simple present. The MA group could perform best in simple present, present continuous, and future tenses. HA students could prominently use the following 5 tenses in order: simple present, present continuous, past continuous, simple past, and future tenses. From the posttest and the delayed test, results were not much different from the pretest, except that the past continuous became one of the top three tenses used correctly by all ability groups. Present perfect and past perfect were found the most difficult tenses for all groups, followed by simple past.

The questions regarding the differences among the three ability groups in terms of their ability to use the tenses was already discussed earlier (see discussion 2.2). The

question to be discussed in this section concerns the difficulty and easiness of the tenses for Thai learners.

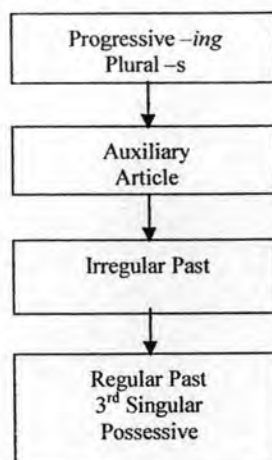
Results from the pretest revealed the differences in the use of tenses by students with different abilities. This can illustrate the picture of the differences in natural setting when the subjects had not been exposed to the treatments. The results (Figure 4.4) provide useful explanations about the mastery of tenses by individual ability group. The LA group comprised the weakest students. The two prominent tenses that they could do were simple present and the present continuous. The MA group, on the other hand, consisted of moderate ability students. In addition to simple present and the present continuous, they could notably answer the future tenses. The high ability group could do well in most tenses, except past perfect and present perfect. According to the findings, a potential 'instructional sequence' can be generated. Figure 5.5 demonstrates the order that is being discussed. Shadowy cells represent the tenses that students in each ability group could prominently achieve.

**Figure 5.5: The potential instructional order of tenses**

Language Ability	Simple Present	Present Cont.	Future Tenses	Simple Past	Past Cont.	Present Perfect	Past Perfect
Low							
Moderate							
High							

The next point to discuss is whether or not this pattern is consistent with that of the previous studies. Unfortunately, there is no particular study that directly examined that acquisition order of tenses. Early empirical evidence of the 'acquisition order' was provided by the so-called 'morpheme studies', which established the existence of a common acquisition order for a subset of English grammatical morphemes (Larsen-Freeman & Long, 1991: 88). Krashen (1977; cited in Larsen-Freeman & Long 1991) had reviewed over a dozen of morpheme studies available at the time and postulated a 'natural order'.

**Figure 5.6: Krashen's 'Natural Order' for ESL (Larsen-Freeman & Long, 1991: 90)**



Such morpheme study has been replicated with other groups of subjects. Other well-known developmental sequence studies include: *interrogatives in ESL* by researchers in the famous Harvard Project, ESL negation, German SL word order, and ESL relative clauses (for review, see Larsen-Freeman & Long, 1991).

Despite some limitations in some areas, Larsen-Freeman and Long (1991) propose that the morpheme studies provide strong evidence that Interlanguage exhibits common acquisition orders. Similarly, Ellis (1997) suggests that, on the basis of the findings, there must be a natural order of acquisition that all learners follow.

Comparing the findings from the present study to the 'natural order', it is confirmed that the progressive *-ing* (which is equivalent to *present continuous* in the present study) is acquired first. However, the *simple present* (which covers third person *-s*) is also found easy for this group of students. Another similarity is that tenses with 'past' time are acquired late. The most difficult tenses for the students in the present study, even for the advanced ones, are *present perfect* and *past perfect*.

Regarding the differences in the acquisition of 'third person *-s*', the results could firstly be explained by the fact that the two aspects are not completely parallel. The one in the present study, simple present tense, is much broader than the third person *-s*. It also covers the case for plural subject to which the *-s* inflection is not applied. So, students in the present study got more chance to be correct. Secondly, the different mode of testing may play a role. In morpheme studies, the data was generally collected through tape-recorded conversations while the present study collected data through the written test. This means that the students in this study had more time to think about their production, resulting in the greater percentage of correct answers they performed. Lastly,

the structure of the test itself may form part of the explanation. Since the test aims at measuring the usage of tenses, it does not pay much attention to the forms. The items simply state the sentence with a 'blank' in which the students are asked to choose the alternatives to fill. Each of the five choices represents one different tense in the correct form. So, when the student answers correctly, it only means that he can correctly identify the proper tense for that particular situation. However, if the test requires the student to write the answer by himself, it is possible that he would forget to put the third person -s.

Ellis (1997) point out that the order does vary somewhat according to the learner's first language. The late acquisition of 'past' tenses and the 'perfect' aspect might be a result of from the differences between the students' mother tongue, Thai, and the target language, English. The Thai verb has no inflected forms whereas the English verb does. In communications, Thai people have a different system to express 'time' and 'aspect'. Usually, situation and context preclude any ambiguity, but where there is a possibility of misunderstanding arising, structural words are added to clarify the time-reference (Smyth, 1987: 350). This might be because Thai speakers regard time as a separate unity by itself and, then, express it separately. Frequently, the time is indicated by adverbs of time (e.g. เมื่อวานนี้, ปีก่อน, เมื่อเช้า), by lexical items (e.g. ทานแล้ว) and by serial verb constructions (e.g. ไป+ซื้อของ+มา, อยากรไป+ดูหนัง). Due to the differences that we have discussed, the 'past' tenses which require 'inflected' form of verbs is the central problem for Thai students. The finding is consistent with Khaourai's (2002) study which found that the most frequent error, under the topic 'tense', was the use of the simple present tense instead of the simple past tense.

On the subject of 'aspects', the 'perfect' is found to be the most difficult aspect to acquire for all ability groups. This could also be explained by the influence of the students' first language. Wannadilok (1982: 207) did a contrastive analysis between Thai and English and presented examples of the 'perfect' sentences in both languages as follows:

- |                               |  |
|-------------------------------|--|
| Present perfect (habitual):   | (a.) <u>Has</u> he ever <u>eaten</u> a lot? เขาเคยกินมากี่   |
| Present perfect (completion): | (b) He <u>has</u> already <u>eaten</u> a lot. เขากินมากแล้ว  |
| Past perfect:                 | (c) He <u>had</u> already <u>eaten</u> a lot before he left his office yesterday.<br>เขากินมากแล้วก่อนเขาออกจากที่ทำงานเมื่อวานนี้ |

As discussed earlier, it can be seen from the examples that the verbs in Thai are not inflected. The verb ‘*มี*’ remains the same in every sentence.

It might be a bit surprising for language teachers to learn that most students (HA and MA groups) could perform better in the past perfect than in the present perfect (Figure 4.4). To many people, the latter should be taught after the former. Reviewing English language coursebooks, it is found that the past perfect is placed in the very last chapter, and of course, after the presence of the present perfect.

The incident that the students could do the past perfect better than the present perfect is probably due to their incomparable functions. The past perfect has two straightforward functions—to represent the past of the simple past and the past of the present perfect (Greenbaum & Quirk, 1990). The present perfect, in contrast, holds at least 3 functions; all are more complicated. The functions cover: (1) to refer to a state that began in the past and extends to the present, and will perhaps continue in the future; (2) to refer to one or more events that have occurred at some time within a period leading up to the present. This function can be divided into 2 subtypes: (2.1) the event or events are reported as news; usually they have occurred shortly before the present time, and (2.2) the event or events occurred at some more remote time in the past, but the implicit time period that frames the event or events leads up to the present; (3) to refer to past events that repeatedly occur up to and including the present (Greenbaum & Quirk, *Ibid.*). This is not all; there is a long note, containing another 5 remarks, attached to the functions of the present perfect. So, it can be seen that the functions of the past perfect are less complicated than those of the present perfect. The present perfect is very difficult for Thai speakers to understand and to use; it is also difficult for language teachers to explain and to help the students acquire it.

Turning to the posttest and the delayed test results, it is surprising to find that past continuous was one of the easiest tenses for all ability groups. This is due to the fact that the present study did not isolate formal instruction from the study of error treatment. And, because it was conducted during the semester, the results were somewhat influenced by the formal instruction. The students reported in the informal interview that they were all studying the *past continuous* in their regular class. Their teacher, apart from giving the lecture, gave them a lot of drills and homework to practice on this tense. The researcher has checked with all of the teachers and got a similar report. This means that the students got a lot more exposure to this tense, resulting in a huge improvement in the posttest scores. This finding disproves the statement of Lee and Van Patten (2003:129)



saying that explicit instruction and practice did very little, if anything, to alter acquisition orders. They claim that learners seemed to follow a particular path on their way to developing the second language system, regardless of the order in which grammatical features were taught. On the contrary, the findings in this study provide the evidence that practicing can really facilitate learning.

## **2.6 The Students' Perceptions of Their Own Learning**

It was found that students in both SC and OC groups rated themselves at the post-treatment scales significantly higher than at the pre-treatment scales for every tense. It means that they perceived they had more knowledge about the 7 tenses after using the CALL than they had before using the program. However, students who practiced with a different error treatment type did not have significantly different means in the post-treatment scales. This means that from the student's perceptions, both groups felt that they had more or less the same ability after practicing with the program.

The first point to be discussed is the finding that both groups rated their knowledge at the post-treatment scales higher than at the pre-treatment scales. This could be due to the fact that the students felt that they had really learned something. This view is supported by the results from the opinionnaire (item 7 and item 9) which reported that the students 'agreed' that the programs helped them understand and know more about tenses.

The second point for discussion is the finding that the OC and the SC groups rated their knowledge more or less the same. This might be due to two reasons. First, it might be due to the 'self-flattery' effect which normally occurs when using scales to collect data. Self-flattery is a phenomenon that the respondents tend to rate the scales similarly throughout the questionnaire, and generally at the middle point of the scales. Such happening can more or less affect the results which paid attention to the comparison of means. The second reason is due to the time that the scales were distributed. Since the post-treatment scales were collected on the same day as the posttest and the opinionnaire, the students felt tired after they had to answer the other two instruments, so they might not have fully concentrated when they answered the self-rating scales which was distributed the last. This factor indirectly attributes to the finding as it can stimulate the self-flattery effect which in turn results in the similarities of the response from the OC and the SC groups.

## 2.7 The Students' Opinions

Overall, the students in both OC and SC groups had positive opinion on the program they were using. The ranges of mean scores were close ( $\text{range}_{oc} = 3.45-4.31$  and  $\text{range}_{sc} = 3.52-4.44$ ). However, when examining into details, it is found that the means in every item of the SC group were greater than those of the OC group, except item six. Item six which states that 'the program is motivating' is the only item that the SC group did not have higher means than the OC group ( $\bar{x}_{sc} = 3.98$  and  $\bar{x}_{oc} = 3.99$ ). This might be due to the task demands that the students were required to participate more actively on the task. Most Thai students are not familiar with this type of learning that students take active role in their own learning. However, it was good enough that the SC group was not opposed to it, indicated by the almost equal means.

Although the ranges of mean between the two treatment groups looked similar, the results by testing the means with the criteria set at 3.5 showed that 12 out of 13 items (92%) were rated by the SC group higher than the criteria. However, only 10 out of 13 items (77 %) were rated higher than the criteria by the OC group. Particularly, one of the items that the OC group rated lower than the criteria was the one that asked about their attitudes towards the OC program. The item stated "I think I know more about tenses after practicing with the program". The SC group did not have a problem with this item. The only one item that they rated lower than the criteria was about their attention while practicing.

In relation to the results from Chi-square test of homogeneity, it is found that, in general, the proportions of the answers of the OC and those of the SC groups were not significantly different, except item 10. Item 10 contained the statement asking for their opinion on the feedback they were receiving. It is considered important to the present study which focuses the investigation on the effects of types of error treatment. In response to item 10 which states that 'the feedback is useful and helpful', about thirty-one percent of the SC students answered that they 'strongly agree' with the statement. Only about seventeen percent of the OC students 'strongly agree' with the statement. Most of them (68%) rated this item at 'agree'. This means that SC students rated the 'feedback' significantly higher than OC students.

The discussions above lead to the conclusion about students' opinion that from the in-depth analysis, the SC program is superior to the OC program. The two modes of the SC answers are in higher ranks than those of the OC. More importantly, the feedback

which is the key variable in this study got more positive rating from the SC group than the OC group.

### 3. Implications

Several implications can be drawn from the discussions as follows:

3.1 The major finding of this study points out that self-correction is more effective in eradicating long-term errors; therefore, it should be promoted. The implementation of self-correction not only has positive effects in the classroom, but also encourages the students to become autonomous learners which is the ultimate goal of all learning conditions.

3.2 Since self-correction is primarily based on 'knowledge of rules', students should be provided with the knowledge. The question about an uncertainty of teaching grammar in the classroom has been assured. Students need such information to successfully perform their self-correction. In the case of software, the program must provide enough and comprehensible information, so that the learners can study and apply the rules by themselves.

3.3 The discussion about the easiness and difficulty of tenses raises an implication for the instructional order that the tenses should be introduced to the students as follows:

**First group:** present continuous and simple present

**Second group:** future tenses

**Third group:** simple past and past continuous

**Fourth group:** past perfect and present perfect

However, the implication does not mean that the acquisition should be seen as analogous to building a wall, with one brick set in place before another is placed on top (Ellis, 1997: 22). Ellis points out that when learners acquire a grammatical structure, they do so gradually, moving through a series of stages en route to acquiring the native-speaker rule. To state it clearly, the researcher does not suggest that difficult tenses should not be instructed to beginners. In contrast, authentic language that normally consists of mixed tenses should be introduced, but the emphasis should only be put on the subjects that learners are developmentally ready to acquire. The sequence above provides a guideline for teachers about what to put the emphasis on at different learning stages.

3.4 The discussion about the advantages of practicing yields an implication that drill and practice could really facilitate learning. Drill and practice should not be merely viewed as the method resulting from behaviorism. Rather, meaningful and mindful practice could serve as a rehearsal that leads the new knowledge to the long-term memory. So, there is nothing wrong to make use of it in the classroom for such purpose.

3.5 The evidence from the opinionnaire concerning the motivation suggests that the self-correction feedback should be short, specific, and easy to understand because the demand of this kind of program from the learner is relatively high. Although self-correction is better than overt correction in promoting retention, the design of such CALL should be made with care. Learners have to get through the feedback, step by step, in order to search for the correct answer by themselves. Failure to identify the answer makes learners get stuck on the item, possibly leading to the decrease in motivation and enthusiasm.

#### **4. Recommendations:**

##### **4.1 Recommendations for the teachers**

4.1.1 It is recommended that teachers should use more self-correction technique with students. In small classes, the error treatment can be applied straight away by the teacher. A variety of self-correction techniques is reviewed and provided in Chapter II. However, it should be noted that some might be particularly good for specific type of errors. Careful selection of the techniques is recommended.

For large classes, it is almost impossible for the teacher to apply self-correction in the classroom because it would take a lot of the class time doing so. Furthermore, other students might get bored and feel less motivated to wait during the process. The teacher is recommended to use self-correction through CALL. Although the development of CALL may take a long time, it is worth doing, especially when we aim at the long-term goal.

4.1.2 It is also recommended that the teacher provides enough 'rehearsal' in the classroom. Self-correction alone may only reach to working memory, but together with the rehearsal, the knowledge would be stored in long-term memory.

4.1.3 When teaching, it is recommended that teachers should put emphasis on the point that the students are developmentally ready to learn. The instructional sequence should be organized accordingly.

completion  
4.1.4 In relation to grammar teaching, Swan (2002) suggests that if grammar is given too much priority, the course books become little more than grammar course. Students do not learn 'English'; they learn grammar. However, Swan argues that doing too little grammar is of course as damaging as doing too much. At this point, language teachers might be frustrated and wonder how to teach grammar without spending too much time on it. For this problem, it is recommended that the teacher uses CALL. With the help from the teacher in identifying their weaknesses or the areas that need improvement, students can work on their own to improve the grammatical problems. Teacher may provide some of the class time to discuss grammar points and assess the students understanding. Nevertheless, a lot more time in the classroom can be devoted to learn 'English'.

**4.2 Recommendations for further research:** For further research, the following is recommended.

4.2.1 Replications of this study should be conducted to see whether the results about the retention are confirmed; but with longer period of time and/ or more observations. The results will provide explanations of the retention rates after 6 weeks. The research will add up more information to illustrate a bigger picture. After many revisions, the findings will be more valid and more reliable.

4.2.2 This study should be replicated using students in other settings, e.g. students from public universities or students in other educational levels to find out if results are different from the results in the present study.

4.2.3 The characteristics of an effective self-correction program should be further investigated. The present study has applied the scaffolding technique in supplying the feedback. There are other techniques that have been used in eliciting self-correction, e.g. underlining, indicating types of error, reference to grammar rules, etc.(see Chapter II under the topic 'techniques used in error treatment' for more information). Some might be better than the others in promoting long-term learning. This should also be examined.

4.2.4 The isolation of formal instruction from the study is suggested. If possible, the study should be conducted during a school vacation or another proper time when the subjects do not have the influence from formal instruction.