

# Examining the Role of the Computer in EFL Instruction

**Li-Ling Chen**

Chung Hwa College of Medical Technology, Tainan, Taiwan

## **Abstract**

*The purpose of this study was to explore whether the addition of CAI to traditional instruction would have an effect on reducing Taiwanese beginning EFL students' written error rates. A quasi-experimental research design was conducted at a private college located in southern Taiwan. A post-writing assessment was administered for both the control group and the experimental group after the treatment. Data were analyzed through a one-way ANOVA on overall error rates as well as a series of one-way ANOVA on error rates within individual error categories. The major finding on overall error rates demonstrates that there was no statistical difference between the control group and the experimental group. However, findings on individual error categories suggest that CAI as an instructional aid may have produced a significant effect on reducing students' error rates in the error category of nouns and prepositions. Although the supplemental CAI program in this study did not produce statistically significant effects on reducing beginning EFL learners' overall written error rates, no evidence was found suggesting that CAI had a negative effect on learning. The research concluded that CAI as a supplement to traditional instruction leaning could be at least as good as traditional instruction.*

## **Introduction**

The Education Department of Taiwan has made great efforts to motivate students to learn English so as to enable Taiwan to be more competitive in the global economy. Many researchers (Cheng, 2003; Gonzalez-Bueno & Perez, 2000; Jan, 2000, 2002; Lin, 2003; Liou & Yeh, 2000; Shih & Cifuentes, 2003; Sotillo, 2000; Sun, 2000; Wei, 2002) have investigated the effects of technology on language learning. The results tend to favor the incorporation of technology into language instruction. Some researchers argue that multimedia technology can provide authentic cultural contexts and can help promote students' learning interests (Chen, Kao, Scanlon, & Lee, 1998; Kramersch and Andersen, 1999). Some studies examined the use of computers in honing L2 writing skills, such as grammar and grammatical accuracy, error feedback, and the writing process (Gonzalez-Bueno & Perez, 2000; Ogata, Feng, Hada, & Yano, 2000; Thorson, 2000). The results

demonstrate that technology has a positive effect on language learning. This study was conducted to examine whether the use of computer-assisted instruction in EFL (English as a Foreign Language) grammar instruction could improve student learning outcomes.

### **Review of Literature**

#### *Computer Technology in EFL Education*

Technology has found a welcome home in foreign language education. Language instruction that integrates technology has become popular and has had a tremendous impact on language education. Numerous EFL research studies (Blake, 2000; Cheng, 2003; Cheng & Liou, 2000; Egbert, 2002; Higgins, 1993; Kramersch & Andersen, 1999; Skinner & Austin, 1999; Strambi & Bouvet, 2003; Willetts, 1992; Williams & Williams, 2000) suggest that integration of technology can improve academic performance, enhance motivation, and promote learning. To examine how technology supports teaching and learning, Chatel (2002) conducted interviews and observations with eight classroom teachers and four ESL teachers and asserted that the selection of appropriate software and websites help ESL students learn and apply English. Lasagabaster and Sierra (2003) conducted a similar research study to examine the attitude of 59 undergraduate students toward CAI software programs. The findings revealed that the students had a positive attitude toward learning another language with computers.

Also, results from some research studies suggest the value of incorporating technology into EFL instruction (Carey & Gregory, 2002; Cheng, 2003; Godwin-Jones, 2002; Gonzalez-Bueno & Perez, 2000; LeLoup & Ponterio, 2003). Liu, Moore, Graham, and Lee (2002) investigated the literature relating to how computer-based technology had been used in language instruction during the past decade (1990 – 2002) and found a shift

in the research focus. Current research, unlike that conducted in the early 1990s when the value of technology was still questioned, is now centered on how to integrate technology into language instruction to make teaching and learning more effective.

Through a meta-analytic methodology, Bolk, Oostdam, Otter, and Overmaat (2002) reviewed 42 studies about CAI programs employed in teaching beginning readers. The findings suggest that CAI programs generally tend to be effective and have a positive effect on beginning learners. In addition, Warden and Chen (1998) applied CAI in evaluating Taiwanese students' writing error types. The researchers scrutinized writing samples from 381 college students. All the subjects utilized a custom-designed computer program. While one group received feedback generated by the computer software, the other group did not. The results suggest that computer generated feedback helped the learners reduce error rates in their writings.

While a great number of studies confirm the positive effects that technology may have on language learning, not all research studies favor the use of technology in language instruction. Liou, Wang, and Hung-Yeh (1992), for example, implemented a multimedia project in grammar instruction. The result showed that instruction plus computers did not positively impact achievement more than traditional instruction. Such a non-significant phenomenon was also supported in some studies, in which the effectiveness of CAI was not found significantly better than traditional instruction (Chirstmann, Badgett, & Lucking, 1997; Fletcher-Flinn & Gravatt, 1995; Lowe, 2001; Shute & Gawlick-Grendel, 1996; Spotts, 1992; Yaakub, 1998). In spite of these and others' results, Joy and Garcia (2000) contend that many non-significant difference

findings may result from uncontrolled variables, as perfect controls over all the factors that may impact a study's results are sometimes difficult in educational settings.

## **Methodology**

### *Research Question*

The following question was developed to guide the research: Is there a significant difference in students' written English error rates, as measured by a post-writing assessment, between those Taiwanese beginning EFL students who receive CAI instruction to support traditional instruction and those who receive traditional instruction alone? The following null hypothesis was established to examine this question:

*There is no statistically significant difference in error rates between students who receive CAI instruction as an instructional support to traditional instruction and those who receive traditional instruction alone.*

$$H_0: \mu_{x1} \geq \mu_{x2}$$

$$H_a: \mu_{x1} < \mu_{x2}$$

Significance level:  $\alpha = .05$

x1: error rate of the experimental group

x2: error rate of the control group

### *Research Design*

This study, a quasi-experimental posttest only research design, involved a control group and an experimental group, with 50 beginning Taiwanese EFL learners in each group who had in-house TOEFL scores less than or equal to 430. The criterion of a TOEFL score of 430 is not a standard to distinguish beginning English learners but rather a way used in this study to identify beginning EFL students. Since all subjects in the two groups shared similar English proficiency levels, as defined as beginning English learners with TOEFL scores less than or equal to 430, the two groups were considered homogeneous.

The independent variables in this study were traditional English grammar instruction alone and the use of CAI as an instructional support to traditional instruction. The dependent variable was the written grammar error rates as measured by a post-writing assessment. Both groups were administered a post-writing assessment after the treatment. The post-writing assessment involved a writing assignment of narrative essays. An error analysis of the written essays was on the basis of a predetermined set of error categories.

### *Subjects*

The target school for this study is a private medical technology college located in southern Taiwan. All of the students are Taiwanese EFL learners with Mandarin as their first language. Although the participants in this study were not in the same department, all the subjects have the same background of culture and language, with Mandarin as their first language. All of them had received formal English education at school for about six years. Both groups shared similar English proficiency levels and were considered as homogeneous.

### *Instrumentation*

Narrative essays were the data collected in the post-writing assessment. Errors in these narrative essays were identified and classified into different error categories by two raters who are experts on English grammar. Each is a native speaker of English. The researcher used their completed score sheets to calculate error rates in SPSS. The error rates in this study were based on absolute frequencies of errors relative to the total words a certain group of participants have written (Huang, 2002).

A total of 15 error categories were used in this study (see Appendix A), among

which nine were based on Horney's (1998) study, and six were from the researcher's pilot study. To augment inter-rater reliability, an expert validated guide for error analysis as well as a scoring sheet were created by the researcher to provide detailed error types and English grammatical rules for the raters to follow.

### *Error Analysis Training*

Two raters were used to establish inter-rater reliability for the study. Each rater analyzed the collected essays from the post-writing assessment. Both raters are experts in English grammar and are native speakers of English. Both received training from the researcher, following an expert validated guide for error analysis, and corrected 30 writing samples that served as a pilot study for the purpose of examining inter-rater reliability. The 30 writing samples were from a writing assignment of another English class at the same institution.

### *Delivery of Instruction*

Instruction methods in this study involved traditional class instruction, referring to lectures without computer aids, and instruction with an aid of a computer tutorial. The multimedia project used in this study was created with Knowledge Adventure's *HyperStudio*<sup>TM</sup> 4.0 and was developed by the five steps of the ADDIE model: analysis, design, development, implementation, and evaluation. The duration of treatment in this study was four weeks, four hours a week, with a total of 16 hours of instruction for both groups. The students in the control group were taught in a traditional classroom, while those in the experimental group were instructed in a computer lab, with one computer per student. The treatment was administered via a computer tutorial module. The content covered parts of speech: nouns, articles, pronouns, verbs, adjectives, adverbs,

prepositions, coordination, and subordination. In addition to grammar explanation, exercises were provided as well. Exercises in the control group were done by papers and pencils, whereas exercises in the experimental group were practiced via the computer software with immediate feedback on students' answers. While the way of instructional delivery was different between the control group and the experimental group, the duration of instruction and material used for both groups were the same. During instruction, students in both groups were requested to develop short paragraphs for the grammar sections that had been taught.

### *Data Collection*

In order to ensure that all writing samples collected from the subjects were non-revised first drafts, the subjects were required to write in class during one 100-minute period. All the writing samples were encoded with numbers instead of student names to maintain confidentiality of the participants. To reduce writing anxiety as a factor in the study, the following writing prompt was presented to both groups:

*Write a short essay with the topic "The Most Memorable Thing in My Life." Minimum page requirement is one page. Think about an event that is most memorable or unforgettable for you. Write a story about what happened. Be sure to narrate an event and include specific details in your response.*

*A narrative essay may include three main parts: introduction, body, and conclusion. Provide background information that helps your readers understand the thesis; develop more examples and details in the body paragraph; end with personal opinions or the influence the event has had on you.*

## Results

### *Inter-Rater Reliability*

The Pearson product-moment correlation coefficient was used to establish the inter-rater reliability through computing the correlation between Rater 1 and Rater 2. The computed Pearson correlation coefficient  $r$  for overall errors was .877. Correlation is significant at the 0.01 level, two-tailed. The result  $r = .877 (n = 100), p = .000$  indicated that there was a significantly positive and strong relationship between errors identified by Rater 1 and Rater 2. The Pearson correlation coefficients for individual error categories are listed in Table 1.

**Table 1. Inter-Rater Reliability Coefficients for Individual Error Categories (Correlation is Significant at the 0.01 Level, Two-Tailed)**

Error Categories	Pearson Correlation Coefficient ( $r$ )	Sig. ( $p$ value)
Verbs	.806	.000
Subject Omission	.805	.000
Syntax	.733	.000
Punctuation	.685	.000
Pronouns	.631	.000
Articles	.586	.000
Capitalization	.568	.000
Adjectives	.538	.000
Adverbs	.488	.000
Lexicon	.486	.000
Spelling	.467	.000
Sentence Fragment	.438	.000
Nouns	.428	.000
Conjunctions	.421	.000
Prepositions	.394	.000

The obtained Pearson correlation coefficients  $r$  for all the above 15 error categories were larger than zero with  $p$  values that equaled .000, which suggests a significantly positive relationship between Rater 1 and Rater 2, although the extent varied within different error categories.



Although the values of coefficient  $r$  varied somewhat in individual error categories, the coefficient for overall errors was convincing. Since all of the inter-rater reliability coefficients obtained for overall errors and errors within individual error categories indicated significantly positive relationships between Rater 1 and Rater 2, it was assumed that the scoring method was reliable for further statistical tests to the research question.

#### *Descriptive Statistics for Post-Writing Assessment*

Data collected totaled 14,302 words, of which 6,600 words were from the control group and 7,702 words came from the experimental group. Total errors found were 3,345, and 3,332 by Raters, 1, and 2, respectively. Errors identified by both raters were considered significantly and positively related. The researcher did not use the average of error counts identified by the two raters, but rather selected one of the raters instead. Since total errors found by the two raters were similar and because their results were considered significantly and positively related, choosing either rater was deemed acceptable. Data analysis in this study was based on the errors identified by Rater 2.

While 1,518 errors were found in the control group, 1,814 errors were detected in the experimental group. Of these errors, errors in the use of verbs (5.05%), errors in punctuation (3.56%), errors in lexicon (3.10%), errors in syntax (2.36%), errors in capitalization (2.32%), errors in subject omission (1.80%), errors in prepositions (1.03%), and errors in articles (0.99%) were the eight most frequent error types (see Table 2).

**Table 2. Total Errors Found in This Study**

Error Category	Number of Reported Errors	Error Rate (%)
Verbs	722	5.05
Punctuation	509	3.56
Lexicon	444	3.10
Syntax	338	2.36
Capitalization	332	2.32
Subject Omission	257	1.80
Prepositions	147	1.03
Articles	141	0.99
Spelling	103	0.72
Pronouns	92	0.64
Nouns	70	0.49
Conjunctions	65	0.45
Sentence Fragments	59	0.41
Adjectives	40	0.28
Adverbs	13	0.09

*Total words written by both groups equaled 14,302. Total errors equaled 3,332.*

Table 3 lists error frequencies found in the control group. It is noted that the five error categories where the greatest number of errors occurred were slightly different from those showed in Table 2. While the first three most frequent errors in Table 3 were the same as those in Table 2, the control group had errors in capitalization as the fourth highest errors and syntax as the fifth highest errors.

Table 4 presents error occurrences for the experimental group. Like the control group, the experimental group had the same five most frequent error types as showed in Table 2 but with a slight difference in the error order. While the experimental had the second highest errors in lexicon and third highest errors in punctuation, Table 2 illustrates that the second highest errors were punctuation and the third highest errors were lexicon for the control group.

**Table 3. Error Frequencies and Error Rates for the Control Group**

Error Category	Error Frequency	Error Rate (%)
Verbs	303	4.59
Punctuation	246	3.73
Lexicon	166	2.52
Capitalization	150	2.27
Syntax	135	2.05
Prepositions	100	1.52
Subject Omission	83	1.26
Articles	78	1.18
Spelling	57	0.86
Nouns	50	0.76
Pronouns	49	0.74
Conjunctions	36	0.55
Sentence Fragments	32	0.48
Adjectives	27	0.41
Adverbs	6	0.09

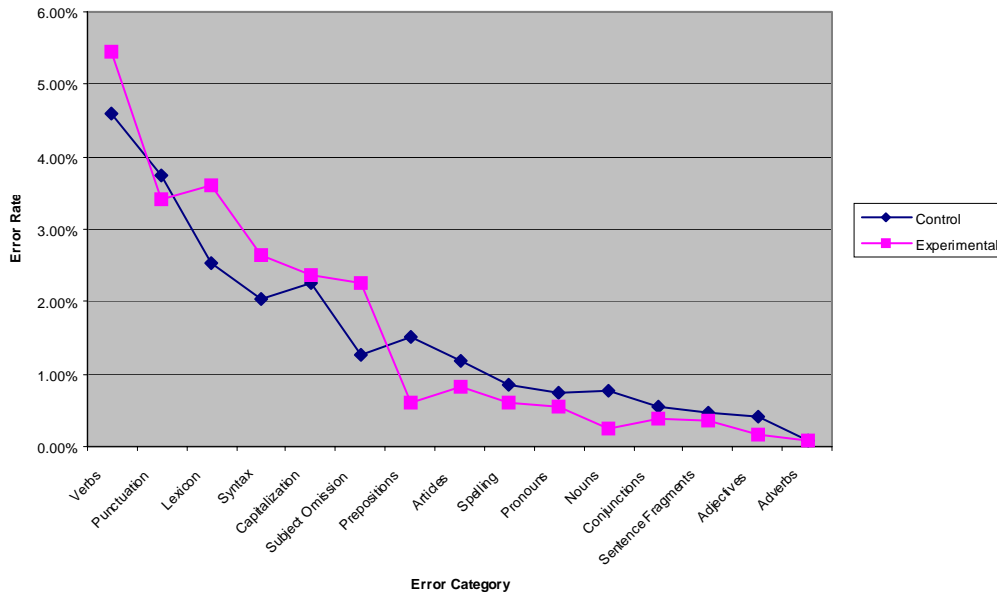
*Total words written by the control group equaled 6,600. Total errors equaled 1,518.*

**Table 4. Error Frequencies and Error Rates for the Experimental Group**

Error Category	Error Frequency	Error Rate (%)
Verbs	419	5.44
Lexicon	278	3.61
Punctuation	263	3.41
Syntax	203	2.64
Capitalization	182	2.36
Subject Omission	174	2.26
Articles	63	0.82
Prepositions	47	0.61
Spelling	46	0.60
Pronouns	43	0.56
Conjunctions	29	0.38
Sentence Fragments	27	0.35
Nouns	20	0.26
Adjectives	13	0.17
Adverbs	7	0.09

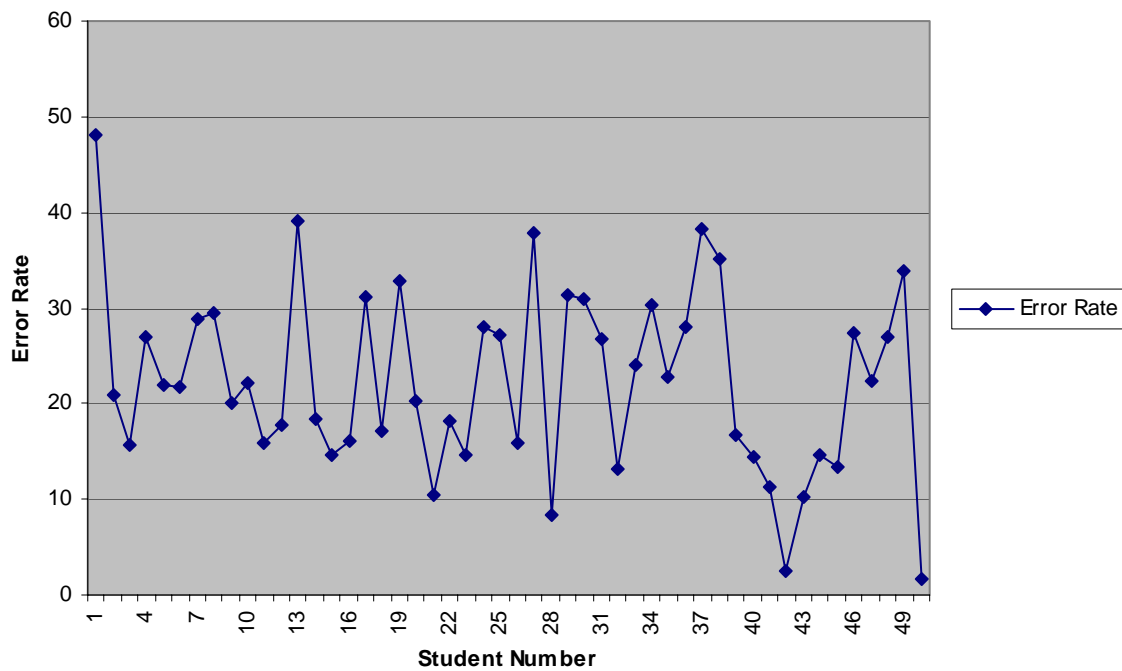
*Total words written by the control group equaled 7,702. Total errors equaled 1,814.*

Figure 1 shows the comparison of the participants' performance between the two groups within different error categories. While the control group demonstrated more errors within the categories of nouns, articles, prepositions, adjectives, pronouns, conjunctions, sentence fragments, punctuation, and spelling than the experimental group, the figure reveals that the experimental group had higher error percentages than the control group within the categories of verbs, syntax, lexicon, capitalization, and subject omission.



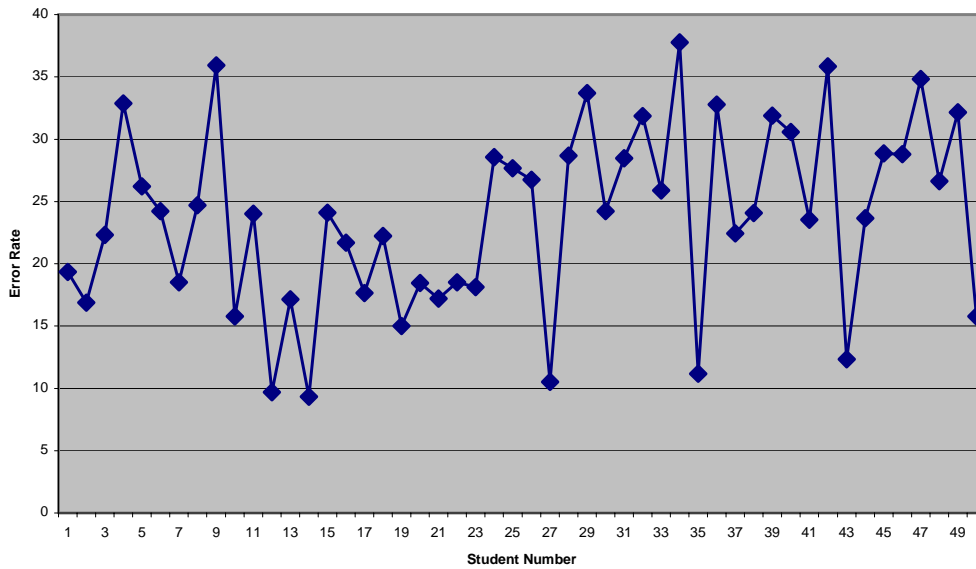
**Figure 1. Error Occurrence for Both Control and Experimental Group**

Figure 2 indicates the distribution of errors for the control group. Forty-six students in the control group had error rates between 10% and 40%. While the majority of the students had error rates between 10% and 40%, two students in the control had error rates less than 2%.



**Figure 2. Error distribution for the control group**

Figure 3 presents the distribution of errors for the experimental group. The distribution shows smaller variance in error rates among the individual students in the experimental group. Forty-nine students in the experimental group had error rates between 10% and 40%. There was only one student whose error rate was outside this range. The error rate of 9.32% for this student was just beyond the range.



**Figure 3. Error distribution for the experimental group**

The minimum error rates and maximum error rates differ between the two groups. Although the means of the two groups were similar, the control group had a much larger range (46.35%) than that of the experimental group (28.44%). The standard deviations also account for the greater variance in error rates for the control group than the experimental group (see Table 5).

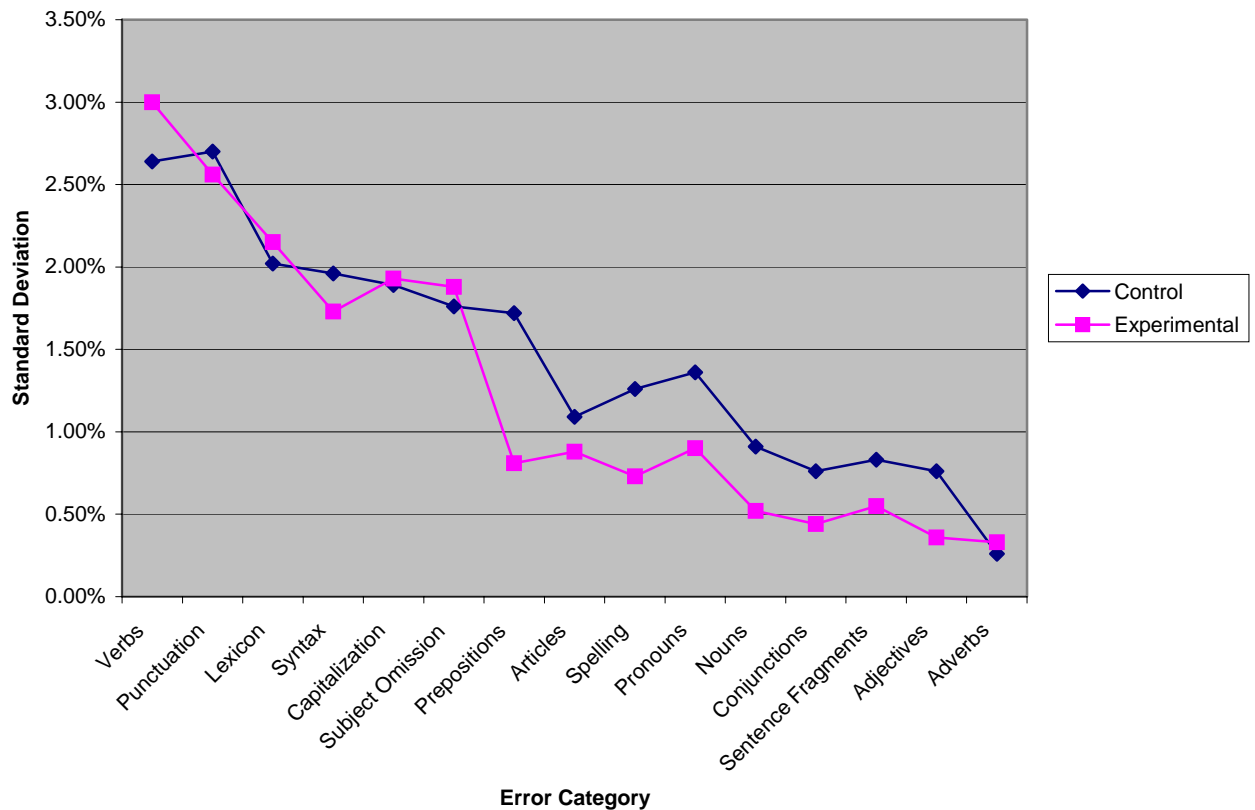
**Table 5. Error Rates for the Post-Writing Assessment**

Group	N	Mean (%)	Min. Error Rate (%)	Max. Error Rate (%)	Range (%)	SD (%)
Control	50	22.39	1.75	48.1	46.35	9.69
Experimental	50	23.76	9.32	37.76	28.44	7.42
All Subjects	100	23.07	1.75	48.1	46.35	8.62

*Means and Standard Deviations for Error Categories*

Tables 6 through 20 indicate means and standard deviations within different error categories for the two groups. The control group had higher standard deviations within the following ten error categories: (1) punctuation, (2) syntax, (3) prepositions, (4) articles, (5) spelling, (6) pronouns, (7) nouns, (8) conjunctions, (9) sentence fragments,

and (10) adjectives, while the experimental group had higher standard deviations within the following five error categories: (1) verbs, (2) lexicon, (3) capitalization, (4) subject omission, and (5) adverbs (see Figure 4).



**Figure 4. Standard deviation of error rates for the control group and the experimental group within different error categories**

A comparison of standard deviation within all 15 error categories revealed that the control group had larger variance within ten error categories, while the experimental group had larger variance within five error categories. This explains why the control group had a larger variance in error rates than the experimental group.

#### *Error Rates for Individual Categories*

Errors in the use of verbs were very common for all the subjects as both the control group and the experimental group had the highest incidence of errors in the verb

category (5.05%), 303 errors for the control group, and 419 errors for the experimental group, as indicated in Tables 2 - 4. The experimental group had a higher mean error rate and a larger standard deviation, as displayed in Table 6. The higher standard deviation helps explain the greater variation in the error rates of verbs for the experimental group. The experimental group did not perform better than the control group within this category.

**Table 6. Error Rates for Errors in the Use of Verbs**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	4.52	0	11.39	11.39	2.64
Experimental	50	5.60	0.88	12.62	11.74	3.00
All Subjects	100	5.06	0	12.62	12.62	2.85

The majority of the subjects had difficulty in the correct use of English punctuation, which was evidenced by the high error occurrences found in the error category of punctuation for both groups. The two groups had similar mean error rates, as indicated in Table 7, whereas the control group had smaller variation in error rates within this error category.

**Table 7. Error Rates for the Errors in Punctuation**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	3.46	0	7.04	7.04	2.70
Experimental	50	3.45	0	9.68	9.68	2.56
All Subjects	100	3.46	0	9.68	9.68	2.61

Lexicon is another error category with high error incidences for both groups. While 166 errors in lexicon were detected in the control group, 278 errors in lexicon were observed in the experimental group. The control group had better performance and smaller variation in error rates than the experimental group within this error category (see Table 8).



**Table 8. Error Rates for the Errors in Lexicon**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	2.34	0	7.83	7.83	2.02
Experimental	50	3.51	0	7.65	6.65	2.15
All Subjects	100	2.92	0	7.83	7.83	2.15

“Errors in syntax” is one of the eight most frequent error categories for both the experimental group and the control group. A total of 338 errors in syntax were observed, of which 135 errors were for the control group and 203 errors were for the experimental group. As displayed in Table 9, the control group outperformed the experimental group, whereas the experimental group had smaller variation in error rates within this category.

**Table 9. Error Rates for Errors in Syntax**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	1.93	0	7.14	7.14	1.96
Experimental	50	2.57	0	6.43	6.43	1.73
All Subjects	100	2.25	0	7.14	7.14	1.87

A total of 332 errors in capitalization were noticed, of which 150 errors were for the control group and 182 errors were for the experimental group (see Tables 2 - 4). Aside from the lower mean error rate, the control group, as indicated in Table 10, had smaller variation in error rates within this category.

**Table 10. Error Rates for the Errors in Capitalization**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	2.08	0	9.32	9.32	1.89
Experimental	50	2.40	0	9.18	9.13	1.93
All Subjects	100	2.24	0	9.32	9.32	1.91

The performance for both the control group and the experimental group appeared to be different in the errors of subject omission. While only 83 errors in subject omission

were noticed in the control group, 174 errors in subject omission were observed in the experimental group. Table 11 demonstrates a lower mean error rate for the control group. Additionally, the lower standard deviation of the control group, as showed in Table 11, explains the smaller variance in error rates for the control group within this error category.

**Table 11. Error Rates for the Errors in Subject Omission**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	1.24	0	8.86	8.86	1.76
Experimental	50	2.17	0	7.55	7.55	1.88
All Subjects	100	1.71	0	8.86	8.86	1.87

“The use of prepositions” is another error category where the experimental outperformed the control group. While 100 errors were reported in the control group, only 47 errors were observed in the experimental group. In addition to the lower mean error rate, Table 12 also indicates that the experimental group had a smaller standard deviation, which suggests that the experimental group had smaller variation in error rates within this category.

**Table 12. Error Rates for Errors in the Use of Prepositions**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	1.61	0	8.86	8.86	1.72
Experimental	50	0.62	0	3.53	3.53	0.81
All Subjects	100	1.11	0	8.86	8.86	1.43

Seventy-eight errors in the use of articles were noted in the control group, whereas 63 errors in the use of articles were detected in the experimental group (see Tables 2 - 4). While the control group demonstrated a mean error rate up to 1.23%, the mean error rate for the experimental group was 0.87%, as showed in Table 13. The experimental group performed better than the control group in the category of articles.

The standard deviations displayed in Table 13 also demonstrate that the experimental group had smaller variation within this error category.

**Table 13. Error Rates for Errors in the Use of Articles**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	1.23	0	3.77	3.77	1.09
Experimental	50	0.87	0	3.40	3.40	0.88
All Subjects	100	1.05	0	3.77	3.77	1.00

Table 14 illustrates that the experimental group outperformed the control group within the error category of spelling, with a lower mean .68%. The smaller standard deviation of the experimental group, as displayed in Table 14, also demonstrates that the experimental group had smaller variation in error rates within this error category.

**Table 14. Error Rates for the Errors in Spelling**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	0.91	0	4.92	4.92	1.26
Experimental	50	0.68	0	2.91	2.91	0.73
All Subjects	100	0.80	0	4.92	4.92	1.03

As displayed in Table 15, the experimental group performed better than the control group within the error category of pronouns, as a lower mean error rate was obtained for the experimental group. The obtained standard deviation also explains why the experimental group had smaller variance in error rates within this error category.

**Table 15. Error Rates for Errors in the Use of Pronouns**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	0.82	0	7.83	7.83	1.36
Experimental	50	0.61	0	3.91	3.91	0.90
All Subjects	100	0.71	0	7.83	7.83	1.15

A total of 70 errors in the use of nouns were found, among which 50 errors were for the control group and 20 errors were for the experimental group (see Tables 2 - 4). As indicated in Table 16, the experimental group had a lower mean error rate that accounts for better performance than the control group. In addition, the smaller standard deviation for the experimental group also explains that the experimental group had smaller variation in error rates within this error category.

**Table 16. Error Rates for Errors in the Use of Nouns**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	0.75	0	3.08	3.08	0.91
Experimental	50	0.28	0	2.86	2.86	0.52
All Subjects	100	0.52	0	3.08	3.08	0.78

Thirty-six errors in conjunctions were identified in the control group, while 29 errors in conjunctions were reported in the experimental group. Table 17 indicates a lower mean error rate for the experimental group, which explains that the experimental group outperformed the control group within this error category. The smaller standard deviation for the experimental group accounts for smaller variance in error rates within this error category.

**Table 17. Error Rates for Errors in the Use of Conjunctions**

<b>Group</b>	<b>N</b>	<b>Mean (%)</b>	<b>Min. Error Rate (%)</b>	<b>Max. Error Rate (%)</b>	<b>Range (%)</b>	<b>SD (%)</b>
Control	50	0.53	0	3.79	3.79	0.76
Experimental	50	0.38	0	1.56	1.56	0.44
All Subjects	100	0.46	0	3.79	3.79	0.62

The performance of the experimental group was better than that of the control group within the error category of sentence fragments. A total of 59 errors were detected in errors in sentence fragment, of which 32 errors were reported for the control group, 27

errors were reported for the experimental group. Table 18 indicates that the experimental group had a smaller standard deviation that suggests smaller variance in error rates within this category.

**Table 18. Error Rates for Errors Sentence Fragments**

Group	N	Mean (%)	Min. Error Rate (%)	Max. Error Rate (%)	Range (%)	SD (%)
Control	50	0.49	0	3.90	3.90	0.83
Experimental	50	0.33	0	1.94	1.94	0.55
All Subjects	100	0.41	0	3.90	3.90	0.71

A total of 40 errors in the use of adjectives were found in the subjects' English writing samples, of which 27 errors were reported for the control group and 13 errors were reported for the experimental group (see Tables 2 – 4). Table 19 reveals that the experimental group outperformed the control group with a lower mean error rate.

**Table 19. Error Rates for Errors in the Use of Adjectives**

Group	N	Mean (%)	Min. Error Rate (%)	Max. Error Rate (%)	Range (%)	SD (%)
Control	50	0.38	0	3.23	3.23	0.76
Experimental	50	0.18	0	1.33	1.33	0.36
All Subjects	100	0.28	0	3.23	3.23	0.60

Both the control group and the experimental group had the fewest error occurrences in the use of adverbs. A total of 13 errors in adverbs were observed, of which 6 errors were for the control group and 7 errors were for the experimental group. Both groups had similar error mean rates and standard deviations (see Table 20).

**Table 20. Error Rates for Errors in the Use of Adverbs**

Group	N	Mean (%)	Min. Error Rate (%)	Max. Error Rate (%)	Range (%)	SD (%)
Control	50	0.09	0	0.99	0.99	0.26
Experimental	50	0.10	0	1.94	1.94	0.33
All Subjects	100	0.10	0	1.94	1.94	0.30

### *Data Analysis*

To address the research question, a one-way analysis of variance (ANOVA) was conducted on the subjects' overall error rates to determine whether the use of CAI instruction as an instructional support to traditional teaching has a significant impact on reducing students' written errors. The results in Table 21 display a  $p$  value of 0.393 which is higher than 0.05. Therefore, the research is unable to reject the null hypothesis. In other words, the results in Table 21 ( $F(1, 98) = 0.736, p = .393$ ) indicate that the use of CAI instruction as an instructional support to traditional teaching in this study did not produce a significant difference in the overall error rates between the control group and the experimental group. That is, the students who received CAI instruction to support traditional instruction did not have significantly lower error rates than those who received traditional instruction alone.

**Table 21. One-Way ANOVA for Overall Error Rates**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.005	1	.005	.736	.393
Within Groups	.722	98	.007		
Total	.727	99			

Although the above one-way ANOVA suggests that CAI instruction did not produce a significant difference in the students' overall error rates between the two groups, from previous descriptive statistics, student performance for the two groups varied within different error categories. Therefore, the researcher decided to further examining each error category to see whether significant differences existed within any error categories.

The results of a series of ANOVAs on individual error categories reveal that significant differences were found within four categories: Lexicon, Subject Omission, Prepositions, and Nouns (see Table 22). The control group had lower error rates within categories of Lexicon and Subject Omission, whereas the experimental group had lower error rates within categories of Prepositions and Nouns.

**Table 22. Results of One-Way ANOVAs for Individual Error Categories**

Error Category	F	Sig.
Verbs	3.712	.057
Punctuation	.001	.980
Lexicon	7.791	.006*
Syntax	2.996	.087
Capitalization	.694	.407
Subject Omission	6.585	.012*
Prepositions	13.674	.000*
Articles	3.147	.079
Spelling	1.254	.265
Pronouns	.755	.387
Nouns	10.236	.002*
Conjunctions	1.506	.223
Sentence Fragments	1.302	.257
Adjectives	2.960	.089
Adverbs	.039	.843

\*Significance level:  $\alpha = .05$

### Conclusions

The major finding on overall error rates demonstrates that there was no statistical difference between the control group and the experimental group. However, other findings on individual error categories indicate that the experimental group performed significantly better than the control group in the error categories of Nouns and Prepositions. Although the experimental group significantly outperformed the control group within these two categories, the researcher could not find literature to support an

assertion that CAI helps reduce errors in particular parts of speech. More research is needed to examine this effect.

On the other hand, it was found that the control group performed significantly better than the experimental group within the error categories of Lexicon and Subject Omission. Lexicon in this study refers to word choice, which is based on a better understanding of the target language and culture. Grammar instruction for the two groups covered the parts of speech, such as nouns, articles, pronouns, verbs, adjectives, adverbs, prepositions, and conjunctions. Students' vocabulary ability and understanding of the target culture may primarily relate to their pre-existing English proficiencies. Likewise, although English is a subject-prominent language different from Mandarin that tends to be topic-comment, the grammar instruction in this study did not cover it as a section for more practice. The significantly better performance of the control group might have been caused by the potential diversity between the two groups with little relation to the treatment. More research is needed in this area.

The result  $F(1, 98) = 0.736, p = .393$  derived from the one-way ANOVA (see Table 21) indicates that there was no significant difference in overall error rates between students who received CAI instruction as a supplement to traditional instruction and those who received traditional instruction, sans computer. This result of non-significant difference in student overall performance was consistent with Liou, Wang, and Hung-Yeh's (1992) study in which the authors implemented a multimedia project to motivate students and correct recurrent grammatical weaknesses to enhance writing instruction. The results in the study of Liou et al. showed that instruction plus computers did not positively impact achievement more than traditional instruction. Such a non-significant



phenomenon was also supported in some studies, in which the effectiveness of CAI was not found significantly better than traditional instruction (Chirstmann, Badgett, & Lucking, 1997; Fletcher-Flinn & Gravatt, 1995; Lowe, 2001; Shute & Gawlick-Grendel, 1996; Spotts, 1992; Yaakub, 1998).

While the finding of this study was in harmony with some studies that indicated the effects of computer-assisted instruction are overestimated (Iheanacho, 1997; Liou, et al., 1992; Patron, Miller, Chisamore, & Lee, 1999; Wood, 2001), such a result contradicted other research that supports computer-assisted instruction (Cheng, 2003; Gonzalez-Bueno & Perez, 2000; Jan, 2000, 2002; S. F. Lin, 2003; Liou & Yeh, 2000; Liu, et al. 2002; Murray, 2000; Shih & Cifuentes, 2003; Sotillo, 2000; Sun, 2000; Wei, 2002). Joy and Garcia (2000) contended that many non-significant difference findings resulted from uncontrolled variables as perfect controls over all the factors that may impact a study's results were sometimes difficult in educational settings. Uncontrolled variables, such as the size of the control group and the experimental group and variance between the two groups including prior knowledge, learning styles, and learner familiarity with technology (Joy & Garcia, 2000), could also have happened during this study. The following paragraphs discuss some possible factors affecting the study's results, including (1) variance, (2) group size, (3) self-learning pace and path vs. teacher-controlled learning pace and path, (4) duration of treatment, (5) length of essays and (6) design and development of CAI

### *Variance between the Two Groups*

Although students in the two groups obtained similar mean error rates for the overall performance, variance existed between the two groups. While the experimental group had a slightly higher mean error rate ( $M = 23.76\%$ ) than the control group ( $M = 22.39\%$ ), the control group had a greater standard deviation in error rates ( $SD = 9.69$ ) than the experimental group ( $SD = 7.42$ ). The variance that existed between the two groups could be a factor that impacted the study's result.

### *Group Size*

In Liao's (2004) meta-analysis of CAI and students' achievement, the author noticed that the effect size of CAI varied with group sizes and suggested that CAI could be more effective if used in individual or small group settings. As large group instruction was involved in this study, this may have diminished the effects of CAI on learning.

### *Self-Learning Pace and Path vs. Teacher-Controlled Learning Pace and Path*

Ehsani and Knodt (1998) declared that CAI projects promote learning as CAI provides a level of flexibility that allows students to determine their own learning pace and path. Nevertheless, since CAI was used in this study as a supplement to traditional instruction, the instructors dominated the learning schedule and learning pace. The added value of flexibility that CAI brings to learning may have been removed, or at least greatly reduced, by the use of teacher-controlled instruction in this study.

### *Duration of Treatment*

Aside from the above factors that could possibly affect the study's results, time likely was a factor, as well. Liao (2004) pointed out that the duration of treatment was a critical variable in many of the studies examined in the meta-analysis. Rachal (1993)

indicated that the design flaw of too short treatment periods could diminish the effects of CAI. The 16-hour instruction designed in this study covering eight main parts of speech as well as related exercises may not have been long enough. More time might be needed when the languages (Mandarin and English) are so dissimilar. It is possible that the treatment effects could have been increased if fewer parts of speech were covered. Further research is recommended.

#### *Length of Essays*

Since beginning EFL learners do not easily master English grammar, it is likely that more errors could be detected in longer essays (Kwok, 1998). The experimental group had more words than the control group (7,702 words vs. 6,600 words) and more errors were found in the experimental group as well (1,814 errors vs. 1,518 errors or an error rate of 23.76% vs. 22.39%). This finding was consistent with Kwok's (1988) study, in which Kwok analyzed EFL college students' English essays and found that students who wrote longer essays made more errors. It could be possible that subjects in the control group were self-monitoring their words for some reason.

#### *Design and Development of CAI*

Although materials used in the *HyperStudio*<sup>TM</sup> tutorial were vetted by subject matter experts, the design and development of the CAI tutorial could have been a factor impacting the study's results as well. The researcher was the creator of the CAI tutorial. However, the researcher's limited computer programming knowledge might have reduced the quality of the learning experience for some students.

## **Recommendations for Future Research**

As stated previously, possible reasons for the non-significant result could be the variance among students as well as defects of the research design. It is suggested that potential diversity among students be diminished as much as possible to reduce the variance as a factor that impacts a study's results. In addition, group size and the duration of treatment are also critical variables that should be controlled in future research to avoid possible impacts on results. Other suggestions for future research, such as (1) providing remedial instruction, (2) narrowing the scope of error analysis, and (3) supplementing the essays with objective tests are discussed in the following paragraphs.

### *Remedial Instruction*

Error analysis is a good device to detect learning difficulties. However, the supplemental CAI program in this study was grammar instruction before error analysis, which did not cover all aspects of participants' errors. As remedial instruction is designed particularly for errors made by learners after error analysis, it may have better effects on reducing students' errors than instruction before error analysis.

### *Scope of Error Analysis*

This study examined the subjects' English essays with 15 error categories. For better and more persuading results, further researchers may limit errors to fewer error types or may even focus on only one or two error types and provide instruction accordingly. This may provide a better understanding of how computer-assisted instruction affects student performance on specific aspects rather than overall grammar instruction that probably leads to a too broad range and decreases the effects that CAI could have on student performance.

### *Supplement with Objective Tests*

The current study explored the participants' errors in their English writings through error analysis. However, reliable inter-rater reliability was challenging for such a subjective scoring method, especially when the writing samples were created by beginning EFL learners who may have been deeply interfered by their first language in the process of writing in English.

Another problem for error analysis is that students may have an incorrect understanding of English that does not appear in their papers, as they may avoid using structures that they find problematic for them. Therefore, the analysis might be inadequate to account for a broader range of errors. It is recommended that objective tests that are carefully designed to cover major learning content should be used as well for future research to evaluate student performance. While error analysis may help researchers understand what learning difficulties students have, objective tests or standardized tests may help establish more convincing inter-rater reliability and cover more learning content to offset the defects of error analysis.

### **Contributor**

Li-Ling Chen, Ed.D., is a Chinese Literature and English language instructor at Chung Hwa College of Medical Technology, Tainan, Taiwan. Her recent research focuses on errors made by Taiwanese beginning EFL learners as well as the effect of CAI on language learning.

## References

- Carey, J. O., & Gregory, V. L. (2002). Students' perceptions of academic motivation, interactive participation, and selected pedagogical and structural factors in Web-based distance learning. *Journal of Education for Library and Information Science*, 43, 6-15.
- Chatel, R. G. (2002). New technology, new literacy: Creating a bridge for English language learners. *The New England Reading Association Journal*, 38(3), 45-49.
- Cheng, Y. (2003). The effects of web-based instruction on Chinese EFL students' learning outcomes. *Dissertation Abstracts International*, 64(2), 382. (UMI No. 3081482)
- Cheng, Y. C., & Liou, H. C. (2000, November). *Using electronic bulletin board as a virtual community to aid college English*. Paper presented in Learning Societies in the New Millennium: Creativity, Caring & Commitments. International Conference on Computers in Education/International Conference on Computer-Assisted Instruction, Taipei, Taiwan. (ERIC Document Reproduction Service No. ED454819)
- Chirstmann, E. P., & Badgett, J. L. & (2000). The comparative effectiveness of CAI on collegiate academic performance. *Journal of Computing in Higher Education*, 11(2), 91-103.
- Egbert, J. (May 2002). A project for everyone: English language learners and technology in content-area classrooms. *Learning and Leading with Technology*, 29(8), 36-41.
- Ehsani, F., & Knodt, E. (1998, July). Speech technology in computer-aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning and Technology*, 2(1), 45-60. Retrieved June 7, 2004, from <http://llt.msu.edu/vol2num1/article3/index.html>
- Fletcher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer-assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research*, 12(3), 219-242.
- Godwin-Jones, B. (2002, September). Emerging Technologies: Technology for prospective language teachers. *Language Learning & Technology*, 6(3), 10-14. Retrieved June 7, 2004, from <http://llt.msu.edu/vol6num3/pdf/emerging.pdf>
- Iheanacho, C. C. (1997). *Effects of two multimedia computer-assisted language learning programs on vocabulary acquisition of intermediate level ESL students*. Retrieved Feb. 23, 2005, from <http://scholar.lib.vt.edu/theses/available/etd-11397-193839/unrestricted/Clems.pdf>

- Jan, L. C. (2000). An experimental computer and Internet based English writing course in CCIT. *Journal of C.C.I.T.*, 19, 229-261.
- Jan, L. C. (2002). Interactive online discussion in an Internet assisted English writing course. *Journal of C.C.I.T.*, 31(1), 39-58.
- Joy, E. H., & Garcia, F. E. (2000, June). Measuring learning effectiveness: A new look at no-significant-difference findings. *JALN*, 4(1), 33-39. Retrieved Feb. 23, 2005, from [http://www.sloan-c.org/publications/jaln/v4n1/pdf/v4n1\\_joygarcia.pdf](http://www.sloan-c.org/publications/jaln/v4n1/pdf/v4n1_joygarcia.pdf)
- Kramsch, C., & Andersen, R. W. (1999, January). Teaching text and context through multimedia. *Language Learning & Technology*, 2(2), 31-42. Retrieved May , 7, 2005, from <http://ilt.msu.edu/vol2num2/article1/>
- Kwok, H. L. (1988). *Why and when do we correct learner errors? An error correction project for an English composition class*. Retrieved Apr. 16, 2005, from <http://sunzi1.lib.hku.hk/hkjo/view/45/4500101.pdf>
- Lasagabaster, D., & Sierra, J. M (2003, September). Student' evaluation of CLL software programs. *Educational Media International*, 40(3/4), 293-304.
- Liao, Y-K, C. (2004). *CAI/CAL and students' achievement in Taiwan: A meta-analysis*. Paper presented at National Educational Computing Conference, New Orleans, LA. Retrieved, Feb. 25, 2005, from [http://center.uoregon.edu/ISTE/NECC2004/handout\\_files\\_live/KEY\\_238668/CAImetaapperNECC.pdf](http://center.uoregon.edu/ISTE/NECC2004/handout_files_live/KEY_238668/CAImetaapperNECC.pdf)
- Lin, S. F. (2003). Modeling a supplemental course web site for EFL vocabulary acquisition. *Dissertation Abstracts International*, 64(01), 118. (UMI No. 3077889).
- Liou, H. C., & Yeh, H. Y. (2000, November). *A study of using web articles to support college English students' ideas in writing*. Paper presented in Learning Societies in the New Millennium: Creativity, Caring & Commitments. International Conference on Computers in Education/International Conference on Computer-Assisted Instruction, Taipei, Taiwan. (ERIC Document Reproduction Service No. ED454819)
- Liou, H-S, Wang, S. H., & Hung-Yeh, Y. (1992). Can grammatical CALL help EFL writing instruction? *CALICO Journal*, 10(1), 23-43. Retrieved Feb. 25, 2005, from <http://calico.org/journalarticles/Volume10/vol10-1/Liou,etal.pdf>
- Liu. M., Moore, Z., Graham, L., & Lee. S. (2002, Spring). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990-2000. *Journal of Research on Technology in Education* 34(3), 250-273.

- Ogata, H., Feng, C., Hada, Y., & Yano, Y. (2000). Online markup based language learning environment. *Computers and Education*, 34(1), 51-66.
- Patron, L. G. Miller, K. Chisamore, C. Lee, E. (1999). *Effects of computer-enhanced instruction on vocabulary acquisition among adult second language learners*. Retrieved Feb. 25, 2005, from <http://education.concordia.ca/~krista.miller/FinalCALL.pdf>
- Rachal, J. D. (1993). Computer assisted instruction in adult basic and secondary education: A review of the experimental literature 1984-1992. *Adult Education Quarterly*, 43(3), 165-172.
- Shih, Y. D., Cifuentes, L. (2003). Taiwanese intercultural phenomena and issues in a United States – Taiwan telecommunications partnership. *Educational Technology Research and Development*, 51(3), 82-90.
- Shute, V. J., & Gawlick-Grendel, L. A. (1996). *Experimental approach to teaching and learning probability: Stat lady*. Lack Land Air Force Base, TX: Galaxy Scientific Corporation (NTIS No. ADA 316 969/5/XAB).
- Skinner, B., & Austin, R. (1999). Computer conferencing: Does it motivate EFL students? *ELT Journal*, 53(4), 270-277.
- Sotillo, S. M. (May 2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology*, 4(1), 82-119. Retrieved May 7, 2005, from <http://llt.msu.edu/vol4num1/sotillo/default.html>
- Spotts, D. S. (1992). *Comparison of computer-based training and conventional classroom training for technical instruction*. (Report No. AFIT/CI/CIA-94-010). Wright-Patterson Air Force Base OH: Air Force Institute of Technology (NTIS No. ADA 281 825).
- Strambi, A., & Bouvet, E. (2003, September). Flexibility and interaction at a distance: A mixed-mode environment for language learning. *Language Learning & Technology*, 7(3), 81-102. Retrieved May 7, 2005, from <http://llt.msu.edu/vol7num3/strambi/default.html>
- Sun, Y. C. (2000, March). *Using on-line corpus to facilitate language learning*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, Vancouver, British Columbia, Canada. (ERIC Document Production Service No. ED442275)
- Thorson, H. (2000). Using the computer to compare foreign and native language writing processes: A statistical and case study approach. *Modern Language Journal*, 84(2), 155-170.



- Warden, C. A., & Chen, J. F. (1998). The application and impact of PC-based software in evaluating Taiwan students' writing error types. *Journal of ChaoYang University of Technology*, 3, 215-225.
- Wei, C. L. (2002). DFLL learners' perceptions of Internet assisted English writing instruction: A case study. *Journal of Humanities, College of Liberal Arts, NCHU*, 32, 483-503.
- Willetts, K. (1992). *Technology and second language learning*. Washington DC: ERIC Clearinghouse on Language and Linguistics. (ERIC Document Production Service No. ED350883)
- Williams, H. S., & Williams, P. N. (2000). Integrating reading and computer: An approach to improve ESL students reading skills. *Reading improvement*, 37(3), 98-101.
- Wood, J. (2001). Can software support children's vocabulary development? *Language Learning & Technology*, 5(1), 166-201. Retrieved Feb. 25, 2005, from <http://llt.msu.edu/vol5num1/wood/default.pdf>

## **APPENDIX A**

### **Error Categories Used in This Study**

- Errors in the use of nouns
  - Singular/Plural
- Errors in the use of articles
- Errors in the use of pronouns
  - Incorrect case forms
  - Missing possessives
- Errors in the use of verbs
  - Tense
  - Subject-verb agreement
  - Auxiliary
  - Verb omitted
- Errors in the use of prepositions
  - Prepositions omitted
  - Wrong prepositions
  - Unnecessary prepositions
- Errors in the use of adjectives
  - Wrong form (confusion of adjectives and adverbs)
  - Comparative/Superlative forms
- Errors in the use of adverbs
  - Wrong form (confusion of adjectives and adverbs)
  - Comparative/Superlative forms
- Errors in the use of conjunction
  - Coordination
  - Subordination (adverbial clauses, relative clauses, and nominal clauses)
  - Missing conjunctions
- Errors in sentence fragments
  - Incomplete sentences
- Errors in syntax
  - Word order (incorrect sentence structures)
- Errors in lexicon
  - Word choice
- Errors in punctuation
- Errors in spelling
  - Misspelling
- Errors in capitalization
- Errors in subject omission