Application of online annotations to develop a web-based Error Correction Practice System for English writing instruction

Shiou-Wen Yeh [a,*], Jia-Jiunn Lo [b], Ho-Ming Chu [b]

[a] Graduate Institute of Teaching English to Speakers of Other Languages, National Chiao Tung University, Taiwan, ROC
[b] Department of Information Management, Chung-Hua University, Taiwan, ROC

A R T I C L E   I N F O

A B S T R A C T

Error correction and peer feedback have been recognized as vital in second language writing development. This study developed a web-based error correction practice mechanism which was attached to an online annotation system for EFL writing instruction. In this system, students input new essays with the Essay Editor and the teacher marked students' errors with the Annotation Editor. Students could read the corrected essay and the results of error analysis through the Viewer. Based on the results of error analysis through the Error Correction Practice Recommender, the system recommended an essay to the student to practice error correction and to implement peer feedback. Following the peer feedback exercise, the student could compare his/her corrections with the teacher's corrections on the same essay through the Viewer. A pretest–posttest study was also conducted to evaluate the effects of using the system on 35 EFL students' writing performance and peer feedback performance. The experiment consisted of four rounds of writing and peer feedback practice. The results of students' writing error ratios and error correction ratios showed that the system was effective in improving students' written accuracy and error correction performance in the peer feedback process. A range of recommendations for future research are discussed.

1. Introduction

Corrective feedback, also known as "error/grammar correction" (Bitchener & Ferris, 2012, p. viii), refers to any indication to the learners that their use of the target language is incorrect (Lightbown & Spada, 1999). For second language (L2) writing teachers, providing feedback to students is an important but challenging task, which involves complex factors (Hyland & Hyland, 2006). For instance, in order for corrective feedback to be successful, it needs to be "processed and acted upon" (Wigglesworth & Storch, 2012, p. 368). The processes require more constructive opportunities (Paulus, 1999; Yeh & Lo, 2009), effective techniques (Ferris, 1995, 2006; Hyland & Hyland, 2006; Lee, 1997), and careful training and modeling (Berg, 1999; Min, 2005; Stanley, 1992; Zhu, 1995). In the past, a number of studies have been conducted to explore how to enhance corrective feedback activities in L2 writing (e.g., Bitchener, 2008; Ferris, 2006; Hyland & Hyland, 2006; Kubota, 2001; Lee, 1997; Liu & Hansen, 2002; Lundstrom & Baker, 2009; Paulus, 1999; Rinehart & Chen, 2012).

---

* Corresponding author.
  E-mail addresses: shiouwen@mail.nctu.edu.tw (S.-W. Yeh), jlo@chu.edu.tw (J.-J. Lo), flairming@gmail.com (H.-M. Chu).

As an alternative to paper-based corrective feedback, computer-mediated corrective feedback has been used to enhance L2 writing (e.g., Ho & Savignon, 2007; Lowry, 2002; Tuzi, 2004; Ware & O'Dowd, 2008; Yang, 2010). The special features of online technology provide an interactive approach to support corrective feedback. For instance, with networked computers, learners can do peer reviews online anywhere at any time (Ho & Savignon, 2007). Specifically, learners exchange writings and feedback through the Internet (Tuzi, 2004) and are able to observe each other's writing processes and corrective feedback recorded online (Yang, 2010). Different from face-to-face corrective feedback, computer-mediated feedback not only offers opportunities for learners to compare one's own texts with those revised by others, it can also "reduce psychological pressure on learners who do not like to give feedback in face-to-face situations" (Ho & Savignon, 2007, p. 273).

With the advancement of computer technology, it is important to design interactive learning environments to support corrective feedback and peer review in writing. In the past, some researchers have attempted to apply online annotation technology in error feedback and error analysis (e.g., Guardado & Shi, 2007; Yeh & Lo, 2009). Annotations are the notes or glosses a reader makes to himself or herself, which are a natural way to record comments in specific contexts within a document (Wolfe, 2002). Annotation tools can scaffold different note taking styles and information strategies, which can help students learn to move from reading to writing (Bargeron, Gupta, Grudin, & Sanocki, 1999; Wolfe, 2002). Online annotation systems also allow a group of readers to make notes on the same copy of a text and provide readers opportunities for interaction with and learning from others in the context of a common text. Such features transform the web into an inter-active medium in which students are no longer limited to viewing content passively but are actively giving and sharing commentary (Bargeron et al., 1999; Lo, Yeh, & Sung, 2013; Wolfe, 2002). To expand the research in computer-mediated corrective feedback, the current study applied online annotations to support error correction practice in an EFL (English as a Foreign Language) context.

The purpose of this study was two-fold: (1) we wished to design and develop a web-based error correction practice system to implement error correction and peer feedback activities by applying online annotation techniques and (2) evaluate the effects of the system on students' written accuracy and peer feedback performances. The system not only allowed the users to give and receive feedback but also provided a convenient interface for students to implement error correction practice on essays with similar error distributions. In addition, the system provided an environment for teachers to do strategy training for error correction practice and peer feedback activities.

2. Literature review

2.1. Corrective feedback in second language writing

Corrective feedback includes responses consisting of an indication that an error has been committed, the provision of the correct language form, or an offer of metalinguistic information about the error (Ellis, 2007). Many researchers are concerned with whether corrective feedback has any effect on written accuracy (e.g., Bitchener & Knoch, 2010; Ferris, 2006; Ferris & Roberts, 2001; Sheen, 2007) and writing development (e.g., Bitchener, 2008; Ferris, 2006; Hyland & Hyland, 2006; Kubota, 2001). For instance, Ferris and Roberts (2001) conducted an experimental study with 72 ESL (English as a Second Language) students who were randomly assigned to three groups: a "codes" group, a "no codes" group, and a "control group" that received no error marking. Their findings showed that the correction ratio of the two groups who received error feedback was significantly higher than the correction ratio of the control group. The effects of corrective feedback in reducing the number of errors were also confirmed in Ferris' (2006) study with 92 ESL students, which found a significant reduction in the number of errors from the first draft to the last draft.

Several studies have also investigated students' attitudes toward corrective feedback and suggest that L2 students need and expect different types of feedback on their written errors (e.g., Ferris, 1995; Ferris & Roberts, 2001; Hyland, 2003; Lee, 1997; Rinehart & Chen, 2012). For instance, in Ferris and Roberts' (2001) study, students preferred feedback with error la-bels attached to errors rather than feedback that was simply marked but not labeled. Hyland's (2003) study revealed that students believe repeated feedback will eventually help them and that without the feedback they will fail to note the errors and improve. Accordingly, as Rinehart and Chen (2012) suggested, L2 learners' preferences for different types of feedback at the revision stage should be carefully considered.

Processing feedback in pairs has also been shown to help learners to engage more deeply with the feedback (Wigglesworth & Storch, 2012). To date, a number of studies have looked at the benefits of peer feedback on L2 writing (e.g., Berg, 1999; Freeman, 1995; Lee, 1997; Liu & Hansen, 2002; Lundstrom & Baker, 2009; Min, 2005; Paulus, 1999; Zhao, 2010). For instance, Freeman (1995) found that peer feedback could improve students' awareness of their own work and enhance a deeper understanding of the language. In addition, it is easier for students to detect errors in peer-written essays than to detect errors in self-written essays (Lee, 1997). Peer feedback can especially benefit students' developmental processes in writing classes (Lundstrom & Baker, 2009). Since reviewing peers' work requires students to implement revising tasks that include detecting errors in texts from readers' perspectives and recommending solutions to correct errors, students can become more independent and active (Rinehart & Chen, 2012) and ultimately improve their writing skills (Cho & MacArthur, 2011).

Although peer feedback provides a number of advantages, some researchers argue that it has limitations in the L2 classroom. For instance, the quality of feedback might be affected by limitations in students' knowledge, experience, and language abilities (Paulus, 1999; Rinehart & Chen, 2012; Saito & Fujita, 2004). Paulus' (1999) study revealed that the effects of

peer feedback on improvements in student revisions are much smaller than that of teacher feedback (peer revision influenced 13.9% of all revisions in the study, while teacher feedback influenced 34.3%). In addition, L2 writers tend to focus heavily on surface errors (Paulus, 1999; Tuzi, 2004). Ware and O'Dowd (2008) found that the feedback provided by peers is often limited in accuracy. In Liou and Peng's (2009) study, some students gave overly critical comments and others gave complementary feedback "since they were reluctant to criticize their peers" (p. 515).

It is also important for teachers to consider learners' cultural backgrounds when adopting the use of peer feedback activities in class because students from different cultural backgrounds often have different expectations for peer review tasks. For instance, Chinese EFL learners might avoid direct criticism of their peers' writing to maintain group harmony and mutual face-saving (Nelson, 1997; Nelson & Carson, 1995). As Nelson (1997) explained, Chinese students are more concerned with "the group's social dimensions than with providing their peers with suggestions to improve their essays" (Nelson, 1997, p. 80).

Another factor that affects the effects of peer feedback is training. A number of studies have stressed the importance of training students to provide more feedback (Berg, 1999; Min, 2005, 2006; Stanley, 1992) and appropriate comments (Min, 2005), to engage in more interaction and negotiation (Zhu, 1995), to produce improved revision types and writing outcomes (Berg, 1999), and to help less-advanced reviewers gain confidence in the peer review processes (Min, 2005). For instance, Stanley (1992) trained students to become effective peer evaluators and found that when students were coached in effective peer response strategies, the number of revisions made increased substantially. In Zhu's (1995) study, students receiving training not only generated more feedback but also engaged in more active interaction and negotiation. Ware and O'Dowd (2008) further claimed that teachers have to go further to dedicate sufficient class time to modeling and scaffolding effective feedback strategies.

The above studies highlighted some advantages and challenges of corrective feedback activities in L2 classroom and reported positive results of trained peer response on students' language development, attitudes, and communication about writing. Appropriate training can also encourage students to write more and gain confidence. Following the discussions of peer feedback training and scaffolding peer feedback strategies, computer-mediated corrective feedback seems to be a potentially powerful tool to enhance L2 writing.

2.2. Computer-mediated corrective feedback

To date, many studies have been conducted in the area of computer-mediated corrective feedback. For example, Liu and Sadler (2003) compared the computer mode and the traditional mode for peer review. As pointed out by the researchers, the students felt it not as intimidating to give correct feedback in the setting of computer-mediated peer feedback as compared to in the setting of face-to-face interactions. Computer-mediated peer feedback has also been proposed as a solution to provide anonymity (Li & Steckelberg, 2006). In Li and Steckelberg's (2006) study, the researchers employed a web-mediated system to facilitate peer assessment. Research results showed that the anonymity provided by this web-mediated peer assessment system gave students a more comfortable environment and less pressure from peers.

Similar results were found in Ho and Savignon's (2007) study. In an Asian EFL academic writing context, the researchers conducted a study examining the use of face-to-face peer review and computer-mediated peer review. This study investigated the attitudes of 37 EFL students toward the use of face-to-face peer review, in which students were able to talk with peers during the review session, and computer-mediated peer review, in which students used email and some features in common word processing programs. Learners reported two major advantages of computer-mediated peer review. First, it offered more flexibility than face-to-face peer review. Many indicated that since they and their partner did not need to be logged on to the computer at the same time, they could comment when convenient and at their own pace.

Liou and Peng's (2009) study suggested that EFL students can be taught rhetorical strategies as scaffolds for successful computer-mediated peer review. They conducted a case study to examine the training effects of peer reviews on 13 EFL students' comments, the quality of revisions to those comments, and their perceptions when composing in weblogs. Based on Tsui and Ng's (2000) writing cycle, four writing assignments were designed for the students. The results showed that, although the training did not make the students more willing to adopt peers' comments (with adoption rate decreasing from 48.9% to 47.7%), the revision success actually increased from 67.8% to 91.8%. The authors therefore suggested that the quality of training is essential to make computer-mediated peer review effective.

Tuzi (2004) developed a database-driven website, which allowed students to add new essays to the system, provide comments to the author, and revise the original essays. Tuzi's (2004) study showed that students who received e-feedback training "developed better quality responses, which contained more specific suggestions for improving an essay" (p. 222). Another concern of Tuzi's (2004) study was the impact of the interface on the e-feedback given by the students. In response to the research need regarding the interface of computer-mediated corrective feedback, Yeh and Lo (2009) developed an annotation system for online corrective feedback and error analysis. The system consisted of five facilities: Document Maker, Annotation Editor, Composer, Error Analyzer, and Viewer. An experiment was also conducted to evaluate the effectiveness of the system. Their empirical results revealed that applying the online annotation system can improve students' error recognition abilities and, therefore, can enhance their error correction and corrective feedback in EFL writing. The researchers further suggested that online annotation tools for manipulating, rearranging, searching, displaying, and sharing annotations could be used to support EFL corrective feedback, especially the collaboration between teachers and students outside the classroom.

In a more recent study, Chen (2012) conducted a study to investigate the effectiveness of blog-based peer review activities in EFL writing courses. This study involved 67 first-year English majors with similar English proficiency levels. As the findings indicated, 74% of the students stated that the peer preview activities on the blog supported improvement in their writing skills. They also perceived the blog-based peer review approach as a useful tool and as a positive experience for achieving their writing goals. This study demonstrated a promising direction for research on the potential for web-based peer review to support EFL students' academic writing abilities. The author also confirmed the necessity of peer feedback training to avoid incorrect feedback and confusion among peers.

2.3. Summary of literature review

While the above literature shows that online technology provides special features to support error correction and peer review, currently, there is a shortage of such learning environments for L2 learners. To expand the research in the area of computer-mediated corrective feedback training and online annotation technology, the current study developed a corrective feedback practice system and examined its benefits for EFL writing development and error feedback performance. The system features were drawn from the reviewed literature and are summarized as follows: (1) providing users with convenient annotation tools to detect and correct errors (Bitchener, 2012; Kubota, 2001; Wigglesworth & Storch, 2012; Yeh & Lo, 2009); (2) analyzing teachers' error feedback and recommending error feedback practice with the most similar error distribution pattern (Paulus, 1999; Yang, 2010); (3) providing error feedback management so that teachers can trace the results of error analysis and students' learning progress (Tuzi, 2004; Yeh & Lo, 2009); (4) structuring peer feedback modeling with teacher's feedback so that students can compare the teacher's and the peer's feedback (Ware & O'Dowd, 2008); and (5) applying anonymity to the system to minimize the impact of peer pressure (Li & Steckelberg, 2006).

This study further examined the effects of using the system on EFL students' writing outcomes and error feedback performance. Two specific research questions were addressed in this study: (1) does the Error Correction Practice System help EFL students to improve their written accuracy?, and (2) does the Error Correction Practice System help EFL students to improve their error correction accuracy when giving feedback to peers?

3. Materials and method

3.1. Design and development of the system

The Error Correction Practice System was based on client-server architecture (Fig. 1). The client side included the Essay Editor, the Annotation Editor, and the Viewer. The server side included the Database, the Error Analyzer, and the Error Correction Practice Recommender. After logging into the system, the student created his/her new essay with the Essay Editor and the written essay was stored in the Database. The teacher then retrieved the essay from the Database to correct and mark students' writing errors with the Annotation Editor. The marked errors were then stored in the Database for error analysis through the Error Analyzer. Students and teachers could read the corrected essay and the results of error analysis through the Viewer. Based on the results of error analysis from a group of students through the Error Correction Practice Recommender, the system could recommend an essay stored in the Database to the student for error correction practice on the Annotation Editor to implement peer review activities. The results of the error correction practice were then stored in the Database. Following the error correction practice, the student could compare his/her corrections with the teacher's corrections on the same essay through the Viewer.

3.1.1. Essay Editor
In this study, FCKeditor (http://ckeditor.com/) was employed as the Essay Editor. Based on JavaScript, FCKeditor is an online WYSIWYG (What You See Is What You Get) editor where students input their essays (Fig. 2). As the essay is edited, the
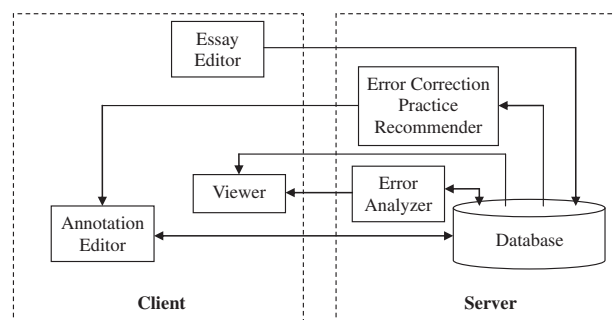

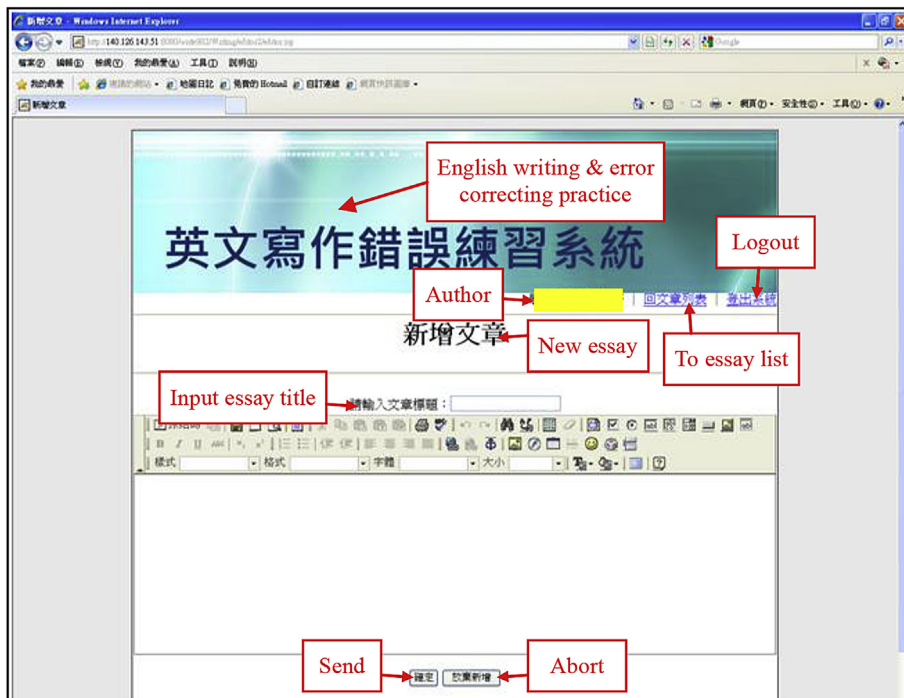**Fig. 1.** The Error Correction Practice System architecture.

Fig. 2. The screenshot of Essay Editor.

system will convert it into HTML format and save it in the Database so that it can be displayed with web page browsers for error correction marking.

3.1.2. Annotation Editor

The Annotation Editor allowed users to make error corrections on web-based documents with online annotations in the same way as the traditional paper-based correction approach. Users such as teachers and students made correction markings through the Annotation Editor (Fig. 3). Since one of the key aims of this system was to allow users to be able to compare the original works and the corrective feedback, the Annotation Editor was designed to be implemented under "read-only" status in that the content of the original essay should not be changed while making correction marks.

To make a correction mark, the user first highlighted the text to be annotated, which was named "annotation keywords." Then, the user assigned an error code by using two pull-down menus to indicate the major error category and error type. Under each major error category, there were different numbers of error types. In this system, five major error categories were applied: (1) writing style, (2) composition structure, (3) sentences, (4) words and phrases, and (5) agreement, tense, and voice (Yeh & Lo, 2009). After assigning the error type, the user then selected one of the annotation tools to place the error correction mark into the annotation keywords. In this study, the annotation tools included "Delete," "Replace," "Move," "Insert-Before," "Insert-After," and "Highlight." "Replace" was applied when wrong words (annotation keywords) were used and should be replaced by other words. If there were missing words before or after the annotation keywords, the tools "Insert-Before" or "Insert-After" could be applied. If the annotation keywords were misplaced, the tool "Move" could be applied. Unlike other annotation tools, the tool "Highlight" did not change the texts. It was used when the user wanted to express their suggestions without changing the text. The system stored annotation information in the Database. As the user moved the cursor over the annotation mark, the annotation information was shown and the user could delete the correction mark by clicking the "Delete this annotation" button (Fig. 3).

There were two display modes in the annotation editor: "annotation mode" and "review mode." The user could freely switch between the annotation mode (Fig. 3) and the review mode (Fig. 4). In the annotation mode, annotation tools were provided to the user for making annotation marks. On the other hand, in the review mode, the annotation tools and annotation marks were hidden. The purpose of the review mode was to provide an environment for users to review the corrected essays easily without showing the correction marks to reduce the cognitive load.

3.1.3. Error Analyzer

The Error Analyzer accessed the Database and analyzes students' errors to display the statistical results of error distributions as requested by the user. The error ratio of error type i of student j in writing practice p is computed as $E_{pij}/\sum_{i=1}^{e}E_{pij}$, where E_{pij} is the number of errors of error type i, identified by the teacher, that student j has made in writing practice p. In the
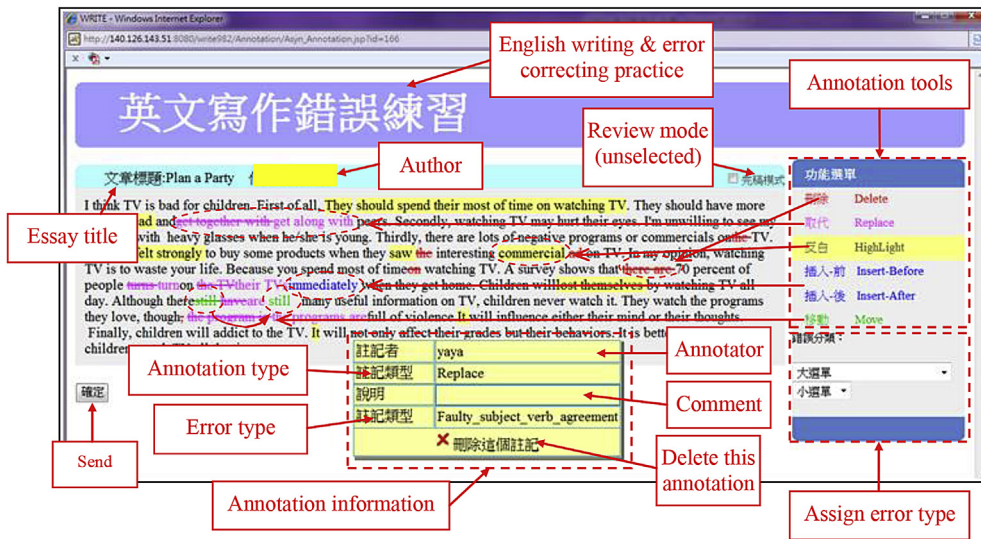
Fig. 3. Screenshot of Annotation Editor (annotation mode).

equation, *e* is the total number of error types. The error ratio of error type *i* of a group of students in writing practice *p* is computed as $\sum_{j=1}^{n_p} E_{pij}/\sum_{i=1}^{e}\sum_{j=1}^{n_p} E_{pij}$, where n_p is the number of students that participated in writing practice *p* (Yeh & Lo, 2009).

3.1.4. Viewer

Through the Viewer, a student could compare his/her original text of an essay (the upper window of Fig. 5) with corrective feedback from the teacher (the lower window of Fig. 5). It helped students recognize which parts of their essays were corrected. When the user moved the cursor over each annotation mark, he/she would get detailed error feedback in a pop-up window (see Fig. 3). Furthermore, in this system, users could compare a student's error correction practice on a peer's essay (the upper window of Fig. 6) with the corrective feedback from the teacher on the same essay (the lower window of Fig. 6). It helped students evaluate whether they correctly identified writing errors in a peer's essay. Like the Annotation Editor, users could freely switch between annotation mode to view correction marks and review mode to read the corrected essay easily without seeing the correction marks on the Viewer.

3.1.5. Error Correction Practice Recommender

In this system, the Error Analyzer analyzed the teacher's corrections on every student's essay. This data was recommended to students for error correction practice. From the Database, the Error Correction Practice Recommender retrieved a peer's essay with the most similar error distribution pattern for error correction practice purposes. The error ratio vector of student *j* in essay *p* is represented as $(E_{p1j}/E_{pj}, E_{p2j}/E_{pj}, \cdots, E_{pej}/E_{pj})$. If the error ratio vector of another student *k* in essay *q* is $(E_{q1k}/E_{qk}, E_{q2k}/E_{qk}, \cdots, E_{qek}/E_{qk})$, the error ratio similarity between student *j* in essay *p* and student *k* in essay *q*, $S(pj, qk)$ is
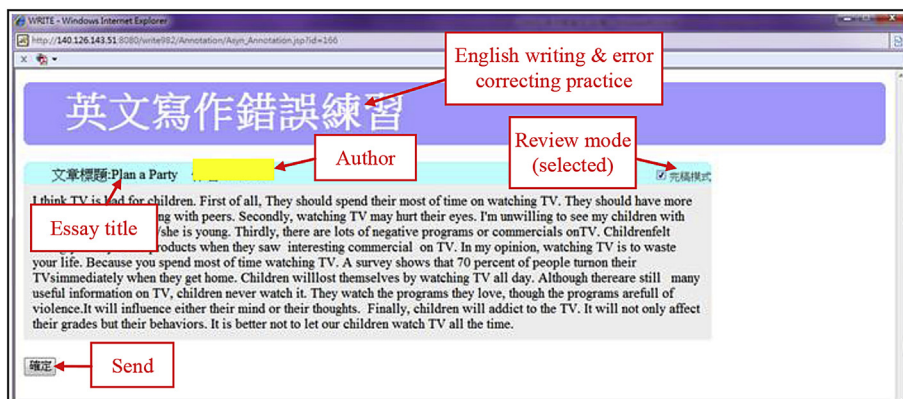

Fig. 4. Screenshot of Annotation Editor (review mode).

Fig. 5. Screenshot of Viewer-Comparisons between the student's original essay (upper) and teacher's correction (lower).
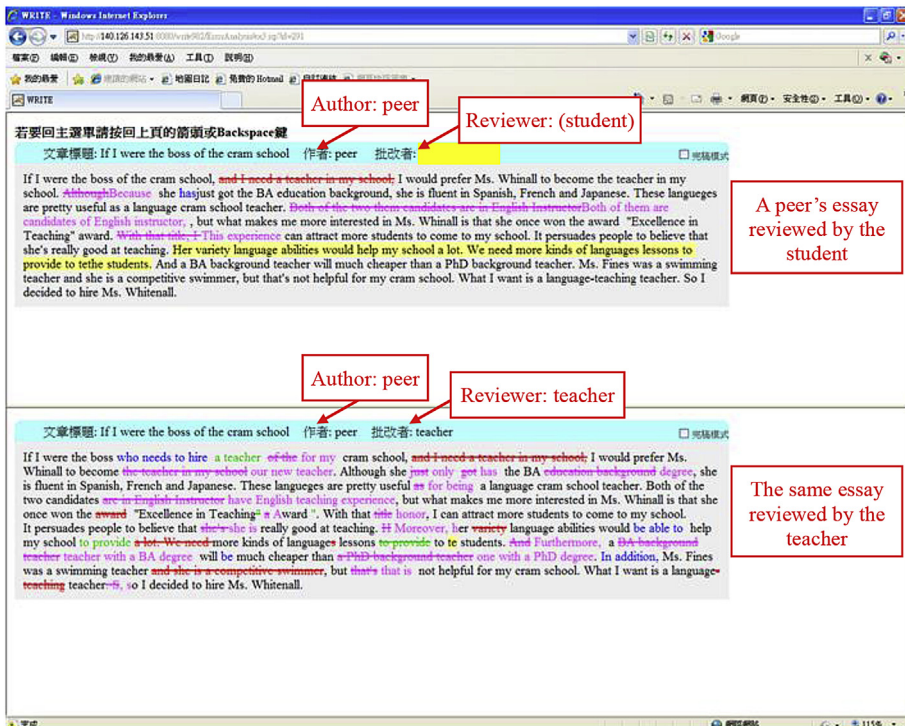

Fig. 6. Screenshot of Viewer-Comparisons between the student's correction (upper) and teacher's correction (lower).

computed by Equation (1). The larger $S(pj, qk)$ is more similar between these two essays in error distribution patterns. The system then recommended the essay with the largest similarity value to student j for error correction practice through the Annotation Editor. While implementing error correction practice, the author's name was represented as a "peer" to ensure anonymity. After practice correcting a peer's essay, the student could check the correctness of this correction practice by comparing the correction marks made by himself/herself and the correction marks made by the teacher on the viewer.

$$S(pj, qk) = \frac{\sum_{i=1}^{e} \frac{E_{pij}}{E_{pj}} \cdot \frac{E_{qik}}{E_{qk}}}{\sqrt{\sum_{i=1}^{e} \left(\frac{E_{pij}}{E_{pj}}\right)^2 \sum_{i=1}^{e} \left(\frac{E_{qik}}{E_{qk}}\right)^2}}, \text{ for all } q's \text{ and } k's \tag{1}$$

3.2. The evaluation

3.2.1. Setting

A pretest/posttest study was conducted to evaluate the effectiveness of the system on 35 EFL students enrolled in a freshman English writing course at a university located in northern Taiwan. This course met 2 h per week for 18 weeks in the semester, with the objective of helping students to be familiar with the process and strategies involved in organizing, producing, and evaluating their writing. Types of essays practiced included giving instructions, writing descriptions, expressing opinions, comparing and contrasting, writing about causes and effects, writing personal/business letters, and writing summaries.

The participants had an average of 7–10 years of previous English instruction. Most of these students had no experience with similar systems. To ensure that the instructor and all the participants had familiarity with the error correction practice system, an instructional manual was designed in advance, which introduced the functions and usages of the system. An in-class training in how to use the system also took place during the first week of the course. Participation in the training session involved approximately 50 min of each subject's time.

3.2.2. Data collection procedure

The experiment began during the second week of the semester and the data collection phase, consisting of four rounds of writing and error feedback practice, lasted for eight weeks. Each writing feedback cycle lasted for two weeks: one week for writing and the other for error feedback. In each week, the course included two class hours. In the first hour, the instructor gave lectures on writing or error feedback; in the second hour, the students practiced writing or error feedback. In the eight-week data collection phase, four topics were assigned: "plan for a party" (essay type: giving instructions), "my relax place" (essay type: writing descriptions), "is TV good for children?" (essay type: expressing opinions), and "if I were a …" (essay type: writing personal/business documents). The content of the lectures and the four writing topics was based on the textbook used in this course, which was *Ready to Write*: *A First Composition Text* (Blanchard & Root, 2003). This textbook was written for beginning to intermediate level EFL students. Each of the four rounds included the following three steps (see Fig. 7).

Step I. Writing practice
The writing practice in the first week involved developing an essay on a topic assigned by the instructor. In the first hour, the instructor gave lectures on writing. At the beginning of the second hour, the instructor announced and explained the essay topic for writing practice, and then the students spent 30 min writing with the system's Essay Editor (Fig. 2). After the students' essays were edited, the system automatically converted them into HTML format and saved them in the database for error correction marking.

Step II. Teacher feedback
After the classes, the instructor provided error feedback for each writing using the Annotation Editor (Fig. 3). Explanations for or comments on each error could be added to each markup. After the instructor completed the feedback, the Error Analyzer then analyzed the teacher's corrections on every student's essay. The data was then recommended to students for error correction practice.

Step III. Error correction practice
In the following week, students participated in the 2-h peer review classes to give each other feedback on their writing. In the first hour, the students first viewed the teacher feedback in the system's Viewer. Then, the instructor discussed and modeled peer feedback strategies and techniques for providing appropriate feedback on each other's work. In the second hour, the students practiced peer feedback with the system. From the database, the Error Correction Practice Recommender retrieved a peer's essay with the most similar error distribution pattern for error correction practice purposes. In each round of the writing and peer feedback cycle, the instructor reviewed 35 articles and each student reviewed one article recommended by the system.

3.2.3. Data analysis

Two constructs were employed to examine the effectiveness of the system on students' writing performance and error feedback performance: (1) the *writing error ratios* and (2) the *correct error-correction ratios*. The *writing error ratios* were used
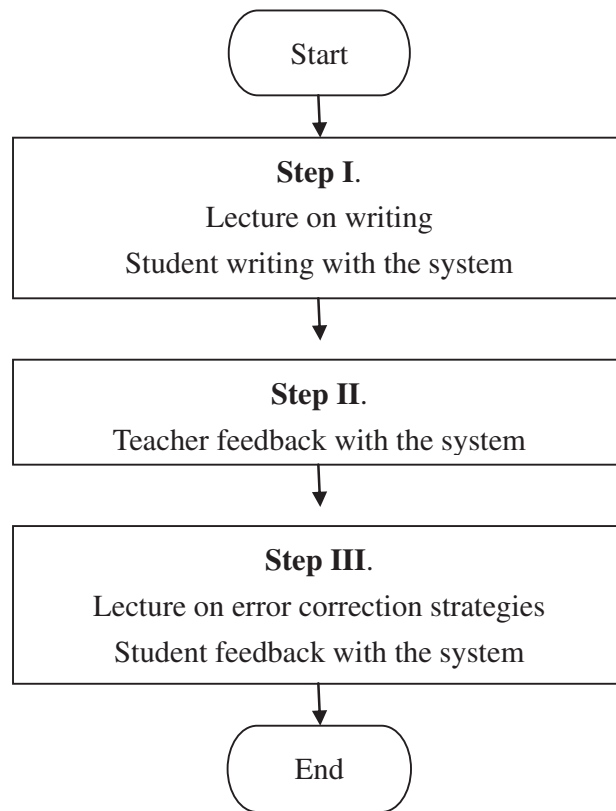
Fig. 7. The three-step writing-feedback cycle.

to evaluate students' writing performance. For each essay, students' *writing error ratios* were computed as the total number of writing errors corrected by the instructor divided by the total number of words in the essay. In this study, students' *writing error ratios* for the first writing practice were used as the pretest performance while their *writing error ratios* of the last (fourth) writing practice were used as the post-test performance.

The *correct error-correction ratios* represent the correctness of students' error corrections when giving feedback to peers. If a piece of text was determined as incorrect by both the student and the teacher, it was defined as a correct error correction. The *correct error correction ratios* were calculated as the total number of correct error corrections of students divided by the total number of incorrect texts determined by the teacher. In this study, the *correct error correction ratios* were used to examine the level to which the students identified the errors correctly.

4. Results

Based on students' pretest performances (first writing practice), students were evenly divided into two groups for further analysis: a high performance group (Mean = 8.6%) with 18 students and a low performance group (Mean = 20.27%) with 17 students (Table 1). Table 1 also lists the descriptive statistics of students' *writing error ratios* in the second, third, and fourth (the post-test) writing practices.

The *t*-test results of students' *writing error ratios* between the high and low performance groups are listed in Table 2 and the paired *t*-test results of students' *writing error ratios* between the pretests and post-tests for both high and low performance groups are listed in Table 3.

Table 2 shows that although there were significant differences in students' pretest performances between the high and low level groups, after using the system, there were no significant differences in students' *writing error ratios* between these two groups. Moreover, Table 3 shows that there were significant differences between the pretest and post-test for the low performance group and no significant differences for the high performance group.

Table 4 lists the descriptive statistics of students' *correct error correction ratios*. The *t*-test results of students' *correct error correction ratios* between the high and low performance groups are listed in Table 5 and the paired *t*-test results of students' *correct error correction ratios* between pretests (the first error correction practice) and post-tests (the fourth error correction practice) for both high and low performance groups are listed in Table 6.

As shown in Table 5, after using the system, there were no significant differences in students' *correct error correction ratios* between these two groups. There were significant differences between the pretest and post-test in *correct error correction*

Table 1
Students' writing error ratios.

Writing practice	Group	Number	Mean (%)	S.D. (%)
1st (pretest)	High performance	18	8.6	2.50
	Low performance	17	20.27	10.50
2nd	High performance	18	9.63	3.37
	Low performance	17	11.66	4.55
3rd	High performance	18	10.97	3.29
	Low performance	17	9.13	3.84
4th (post-test)	High performance	18	10.51	5.09
	Low performance	17	11.51	4.83

Table 2
t-Tests of students' writing error ratios between two groups.

Writing practice	Equality of variance	F-test for variance		t-test for mean		
		F	P	t	d.f.	P
1st	Equal	11.070	0.002	−4.586	33	0.000***
	Unequal			−4.467	17.706	0.000***
2nd	Equal	2.250	0.143	−1.505	33	0.142
	Unequal			−1.492	29.410	0.146
3rd	Equal	0.980	0.329	1.524	33	0.137
	Unequal			1.517	31.593	0.139
4th	Equal	0.092	0.764	−0.596	33	0.555
	Unequal			−0.597	32.999	0.554

***$P < 0.001$.

Table 3
Paired *t*-tests of students' writing error ratios (post-test−pretest).

Group	Difference (%)	S.D. (%)	t	d.f.	P
High performance	1.91	5.41	1.495	17	0.153
Low performance	−8.76	12.31	−2.934	16	0.010*

*$P < 0.05$.

ratios for the low performance group and no significant differences in *correct error correction ratios* for the high performance group as illustrated in Table 6.

5. Discussion

5.1. Effects of the system on writing performance

Table 2 shows that although there were significant differences in students' pretest performance between the high-level and low-level groups, after using the system, there were no significant differences in students' *writing error ratios* between these two groups. This implies that the system eliminated the gap in students' writing outcomes between these two groups. This finding confirms previous studies which hold that appropriate corrective feedback has effects on learners'

Table 4
Students' correct error correction ratios.

Error correcting practice	Group	Number	Mean (%)	S.D. (%)
1st (pretest)	High performance	18	21.20	12.37
	Low performance	17	10.34	8.86
2nd	High performance	18	20.54	12.83
	Low performance	17	11.44	8.57
3rd	High performance	18	28.19	14.71
	Low performance	17	21.12	13.98
4th (post-test)	High performance	18	26.64	13.11
	Low performance	17	30.29	15.30

Table 5
t-Tests of students' correct error correction ratios between two groups.

Writing practice	Equality of variance	F-test for variance		t-test for mean		
		F	P	t	d.f.	P
1st	Equal	2.440	0.128	2.972	33	0.005**
	Unequal			3.000	30.825	0.005**
2nd	Equal	2.718	0.109	2.452	33	0.020*
	Unequal			2.479	29.792	0.019*
3rd	Equal	0.020	0.888	1.455	33	0.155
	Unequal			1.457	32.998	0.154
4th	Equal	1.617	0.212	−0.760	33	0.452
	Unequal			−0.757	31.592	0.455

$^*P < 0.05$; $^{**}P < 0.01$.

written accuracy (Bitchener & Ferris, 2012) by improving the correction ratio (Ferris & Roberts, 2001) and by reducing the number of errors from the first draft to the final draft (Ferris, 2006). It is also consistent with Berg's (1999) study, which reported the influence of trained peer response on students' writing outcomes.

As discussed previously, for corrective feedback to be effective in terms of language development it needs to be processed and acted upon (Wigglesworth & Storch, 2012). Compared with paper-based written corrective feedback, the interface of the current system provided interactive tools for learners to "notice" and "process" the feedback received, as suggested by Wigglesworth and Storch (2012, p. 363). In the system, with the Annotation Editor, users could create and manipulate error corrections on digitized documents with online annotations. When using the Annotation Editor, after the teacher or student highlighted the text to be annotated, he/she then assigned an error code by using the two pull-down menus with error types to indicate the major error category and error type. After assigning the error type, the user selected one of the annotation tools to place the error correction mark into the annotation keywords. When using the system, the students were no longer limited to viewing content passively on the web but actively gave and shared feedback and commentary. In a sense, the system created a dynamic and interactive environment with metalinguistic information and learning supports that was helpful for students to visualize their thinking and linguistic behaviors in writing.

In the context of L2 learning, metalinguistic knowledge about the target language is essential for successful language learning because beginning language learners are often overwhelmed by too much unfamiliar vocabulary, confusing rules, different writing systems, and social customs (Blake, 2000). As stated by Blake (2000), "L2 learners must develop their own metalinguistic awareness in order to stimulate a change in their interlanguage" (p. 120). In this system, corrective feedback responses created by the Annotation Editor included not only an indication that an error has been committed, but also the provision of metalinguistic information about an error. The current system indeed provided a mechanism for users to "give and receive" metalinguistic information, varying according to error type. With the Viewer function, when a student viewed his/her corrected articles in the system, he/she could freely switch between the annotation mode and the review mode to view the annotations with metalinguistic information. In this regard, the system appeared to be a means of supporting the development of critical thinking by allowing students to compare their corrections with the teacher's corrections on the same essays.

The error correction tools of the Annotation Editor can also be compared with the "New Comment" tool in Microsoft Word. When using the "New Comment" tool to proofread a paper, the user first highlights the area where he/she would like to create a note, and Word adds a balloon linking to that text where the user can add his/her notes. Then the user can use the "Track Changes" tool to keep track of the changes made to a document. If there is stored information about changes made to the document, the user can choose to display those changes or to hide them. In fact, the "Track Changes" tool in Microsoft Word can be compared with the Viewer in the current system. However, the current system provided students with more learning support for corrective feedback activities than the tools in Word. For instance, the Error Analyzer of the system analyzed students' errors to display the statistical results of error distributions. The error distributions of each writing practice for an individual student could help the student realize the most severe barrier he/she faces in writing. The error distributions for a group of students could also help the teacher realize the unclear concepts most students have in writing so he/she can adjust the instructional strategies. The system indeed offers new ways of gathering corrective feedback and peer feedback information from students.

Table 6
Paired *t*-tests of students' correct error correction ratios (post-test−pretest).

Group	Difference (%)	S.D. (%)	t	d.f.	P
High performance	5.44	18.33	1.258	17	0.225
Low performance	19.96	12.32	6.677	16	0.000***

$^{***}P < 0.001$.

5.2. Effects of the system on error feedback performance

Table 5 shows that after using the system, there were no significant differences in students' *correct error correction ratios* between the two groups. Similar to the results regarding students' writing improvements, the system benefited students' error correction performance for low performance students and eliminated the gap in error correction accuracy between these two groups. Findings of this study not only lend support to the consensus that training is important for successful peer response (Berg, 1999; Bitchener, 2012; Min, 2005, 2006; Stanley, 1992), they also contribute to general peer feedback research that computer-mediated peer feedback training with online annotations improves EFL students' error feedback accuracy and "helps them become effective responders" (Tuzi, 2004, p. 232). Such issues were not addressed by previous research.

In addition to the error correction tools as discussed in Section 5.1, the system also presented a unique feature, an Error Correction Practice Recommender, which was an innovative design of this research. In this system, the Error Correction Practice Recommender retrieved a peer's essay with the most similar error distribution pattern for error correction practice purposes. This feature is supported by Coit's (2006) view that by reading and writing within a small difference level of proximal development, students are more likely to pick up vocabulary, grammatical structures, and structural devices used by the students. In addition, by reviewing or giving feedback on peers' writing, students can gain insight into their own learning (Cho & MacArthur, 2011; Lundstrom & Baker, 2009; Topping, 1998). In the study, after practicing correcting a peer's essay, each student could check the correctness of this correction practice by comparing the correction marks made by himself/herself and the correction marks made by the teacher on the viewer. Such design also supports Paulus' (1999) view that "peer feedback and teacher feedback can complement each other" (Paulus, 1999, p. 267).

Based on Li and Steckelberg's (2006) suggestion, the system displayed a supportive environment with anonymity. In paper-based peer assessment, it is extremely hard to maintain an anonymous environment (Li & Steckelberg, 2006). In the present study, while conducting error correction practices, the author's name was represented as a "peer" to ensure anonymity. The feature of anonymity allowed students to respond anonymously without having to face the writers. With this feature, the reviewers could "feel more comfortable stating their true thoughts" (Tuzi, 2004, p. 220) and this might minimize the impact of peer pressure.

As Blake (2000) claimed, the conditions for second language acquisition can be enhanced by having learners negotiate meaning or solve problems with other learners. In this study, when students started to review peers' works, they actually encountered linguistic and learning problems, be they grammatical, pragmatic, or lexical in nature. As discussed earlier, if students are untrained, their feedback tends to focus on surface errors (Paulus, 1999; Tuzi, 2004) and be limited in accuracy (Ware & O'Dowd, 2008). The system indeed provided awareness-raising activities with instructor's examples of when and how to provide feedback. With this function, students learned how to work with their peers' writing "in a sensitive and efficient way" (Ware & O'Dowd, 2008, p. 56).

5.3. Limitations and recommendations

The results in Table 3 show that there were significant differences between the pretest and post-test in *writing error ratios* for the low performance group and no significant differences for the high performance group. Similar findings are revealed in Table 6. There were significant differences between the pretest and post-test in *correct error correction ratios* for the low performance group and no significant differences in right error correction ratios for the high performance group. The results suggested that the system significantly improved the writing performance and peer feedback accuracy of the low performance students. However, the effect of the system was not obvious in the writing performance and peer feedback accuracy of high performance students.

The above results are similar to some studies regarding the relationship between the use of error feedback and the learner's proficiency level. For instance, Semke (1984) and Mantello (1997) found that coded feedback is more effective for weak students. Lundstrom and Baker (2009) conducted a study to examine the benefits to students of giving feedback and found students at the lower proficiency level gained more improvements than those at the higher proficiency level. As Ferris (1995) suggests, novice writers usually tend to have difficulties in diagnosing writing errors in their own texts. In addition, Ferris (2006) has observed that "students at lower levels of L2 proficiency may not have sufficient linguistic knowledge to self-correct errors even when they are pointed out" (p. 83). In the current study, the reason might also be that students in the high performance group already possessed a set of learning strategies or range of metalinguistic information. However, more research is needed to confirm this assumption.

It is important to note that the findings in Tables 3 and 6 should not be interpreted to mean that the system was not helpful for high-level students. It is suggested that future studies may use a qualitative approach to address how students of high proficiency levels can benefit from the system. Studies may also consider presenting a comprehensive picture of students' learning processes while conducting peer review and comparing teacher correction in the future. In addition, due to time constraints, this study administered four rounds of writing and peer feedback cycles in eight weeks. In each round, each student reviewed one article recommended by the system. It is suggested that future research could extend the training period and examine the long-term effects of the system on students' writing and peer feedback performance.

6. Conclusions

As the online composition classroom has become more common on university campuses, it is urgent to look for innovative ways to meet the needs of students and teachers. This study was an attempt to develop an online annotation system for error correction practice for EFL writing instruction. Researchers and instructional system developers could continue efforts to exploit the interactive potential of online annotation technology and web-based learning for corrective feedback. For those who are interested in creating similar systems in the future, we suggest that computer supported collaborative learning mechanisms could be integrated with online annotations to allow a group of peers to review the same essay collaboratively and synchronously. Specifically, online discussions could be integrated into corrective feedback processes. Since, as stated by Ede and Lunsford (1992), writing is "social engagement in intellectual pursuits" (p. 15), educators and researchers could provide an environment to help students understand the processes and collaborative features of corrective feedback.

Acknowledgments

References

Bargeron, D., Gupta, A., Grudin, J., & Sanocki, E. (1999). Annotations for streaming video on the web: system design and usage studies. *Computer Networks, 31*(1999), 1139–1153.
Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing, 8*(3), 215–241.
Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118.
Bitchener, J. (2012). Written corrective feedback for L2 development: current knowledge and future research. *TESOL Quarterly, 46*(4), 855–860.
Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. New York: Taylor & Francis.
Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207–217.
Blanchard, K., & Root, C. (2003). *Ready to write: A first composition text* (3rd ed.). New York: Pearson Education Inc.
Blake, R. (2000). Computer in mediated communication: a window on L2 Spanish interlanguage. *Language Learning & Technology, 4*(1), 120–136.
Chen, K. T. (2012). Blog-based peer reviewing in EFL writing classrooms for chinese speakers. *Computers and Composition, 29*(4), 280–291.
Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology, 103*(1), 73–84.
Coit, C. (2006). A student-centered online writing course. In P. Zaphiris, & G. Zacharia (Eds.), *User-centered computer aided language learning*. Hersley, PA: Information Science Publishing.
Ede, L., & Lunsford, A. (1992). *Singular text/plural authors: Perspectives on collaborative writing*. Carbondale & Edwardsville: Southern Illinois University Press.
Ellis, R. (May 17–20, 2007). Corrective feedback in theory, research and practice (abstract). In *Presented at the 5th international conference on ELT in China & the 1st congress of Chinese applied linguistics*. Beijing, China: Beijing Foreign Language Studies University. Retrieved March 8, 2014 from the World Wide Web http://www.celea.org.cn/2007/edefault.asp.
Ferris, D. (1995). Teaching ESL composition students to become independent self-editors. *TESOL Journal, 4*(4), 18–22.
Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues*. New York: Cambridge University Press.
Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: how explicit does it need to be? *Journal of Second Language Writing, 10*(3), 161–184.
Freeman, M. (1995). Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education, 20*(3), 289–300.
Guardado, M., & Shi, L. (2007). ESL student' experience of online peer feedback. *Computers and Composition, 24*(4), 443–461.
Ho, M. C., & Savignon, S. J. (2007). Face-to-face and computer-mediated peer review in EFL writing. *CALICO Journal, 24*(2), 269–290.
Hyland, F. (2003). Focusing on form: student engagement with teacher feedback. *System, 31*(2), 217–230.
Hyland, K., & Hyland, F. (2006). Interpersonal aspects of response: constructing and interpreting teacher written feedback. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues*. New York: Cambridge University Press.
Kubota, M. (2001). Error correction strategies used by learners of Japanese when revising a writing task. *System, 29*(4), 467–480.
Lee, I. (1997). ESL learners' performance in error correction in writing: some implications for teaching. *System, 25*(4), 465–477.
Li, L., & Steckelberg, A. L. (2006). Perceptions of web-mediated peer assessment. *Academic Exchange Quarterly, 10*(2), 265–269.
Lightbown, P. M., & Spada, N. (1999). *How languages are learned*. Oxford, UK: Oxford University Press.
Liou, H. C., & Peng, Z. Y. (2009). Training effects on computer-mediated peer review. *System, 37*(3), 514–525.
Liu, J., & Hansen, J. G. (2002). *Peer response in second language classrooms*. Ann Arbor, MI: University of Michigan Press.
Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes, 2*, 193–227.
Lo, J. J., Yeh, S. W., & Sung, C. S. (2013). Learning paragraph structure with online annotations: an interactive approach to enhancing EFL reading comprehension. *System, 41*(2), 413–427.
Lowry, P. B. (2002). *Improving distributed collaborative writing over the internet using enhanced processes, proximity choices, and a java-based collaborative writing tool* (unpublished Ph.D. dissertation). The University of Arizona, USA.
Lundstrom, K., & Baker, W. (2009). To give is better than to receive: the benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing, 18*(1), 30–43.
Mantello, M. (1997). Error correction in the L2 classroom. *Canadian Modern Language Review, 54*, 127–131.
Min, H. T. (2005). Training students to become successful peer reviewers. *System, 33*(2), 293–308.
Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118–141.
Nelson, G. L. (1997). How cultural differences affect written and oral communication: the case of peer response groups. *New Directions for Teaching and Learing, 70*(summer), 77–84.
Nelson, G. L., & Carson, J. G. (1995). Social dimensions of second language writing instruction: peer response groups as cultural context. In D. Rubin (Ed.), *Composing social identity in written language*. Hillsdale, N.J.: Erlbaum.
Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing, 8*(3), 265–289.
Rinehart, D., & Chen, S. J. (2012). *The benefits of a cycle of corrective feedback on L2 writing*. Saarbrücken, Germany: Lambert Academic Publishing.
Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research, 8*(1), 31–54.
Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annuals, 17*, 195–202.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*, 255—283.

Stanley, J. (1992). Coaching student writers to become effective peer evaluators. *Journal of Second Language Writing, 1*(3), 217—233.

Topping, K. (1998). Peer assessment between students in Colleges and Universities. *Review of Educational Research, 68*(3), 249—276.

Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing, 9*(2), 147—170.

Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition, 21*(2), 217—235.

Ware, P., & O'Dowd, R. (2008). Peer feedback on language form in telecollaboration. *Language Learning & Technology, 12*(1), 43—63.

Wigglesworth, G., & Storch, N. (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing, 21*(4), 364—374.

Wolfe, J. (2002). Annotation technologies: a software and research review. *Computers and Composition, 19*(4), 471—497.

Yang, Y. F. (2010). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology, 42*(4), 687—700.

Yeh, S. W., & Lo, J. J. (2009). Using online annotations to support error correction and corrective feedback. *Computers & Education, 52*(4), 882—892.

Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: a comparative study in a Chinese English writing classroom. *Assessing Writing, 15*(1), 3—17.

Zhu, W. (1995). Effects of training for peer response on students comments and interaction. *Written Communication, 12*, 492—528.