Linguistic Research 29(3), 665-688

Measuring linguistic accuracy in an EFL writing class: An electronic communication channel*

Sun-Young Kim
(Mokpo National University)

Kim, Sun-Young. 2012. Measuring linguistic accuracy in an EFL writing class: An electronic communication channel. *Linguistic Research* 29(3), 665-688. Overemphasis on error correction may lead L2 students to perform writing tasks in a stressed condition, while an electronic communications channel (i.e., an online discussion board) tends to provide a social space to produce written communicative data through the interaction among peer students in a more natural setting. From a social-cultural perspective, this study examined the effects that corrective feedback could have on the improvement in writing accuracy in the use of prepositions and the subject-verb agreement, the most frequently permitted error categories by L2 writers, over a semester. Written communication data produced by 25 students participating in an online discussion board during the semester were used to examine the effects of feedback on writing accuracy. A two-factor ANOVA with repeated measures and descriptive statistics were performed to examine mean differences in accuracy scores over three treatment time and to analyze the improvement in writing accuracy observed on the discussion board. The results showed that the effect of both direct and indirect corrective feedback on accuracy levels in the use of two linguistic errors was found to be significant in the set of post tests conducted in class. However, such an improvement in writing accuracy was not immediate in written data associated with an electronic communication channel. Specifically, an analysis of the communication data L2 students produced through the interaction with their peers did not support the role of corrective feedback in students' writing accuracy. Unlike other studies emphasizing on the teaching practices of error correction at the local level, this study argues that the improvement in writing accuracy would be viewed as a natural progress of writing process. **(Mokpo National University)**

Keywords written corrective feedback, affective variables, written communication channel

* I'd like to express my appreciation to anonymous reviewers for their valuable criticism and suggestions. All remaining errors are of my own.

1. Introduction

An existing body of literature has examined the role of corrective feedback (hereafter CF)[1] in improving writing accuracy in L2 language classes, providing evidence supporting the effectiveness of grammar correction based on either indirect or direct corrections (Ashwell, 2000; Fathman & Whalley, 1990; Fazio, 2001; Ferris & Roberts, 2001; Polio, Fleck, & Leder, 1998). These studies emphasize students' values for error feedback from teachers and show evidence of studies demonstrating the efficacy of CF (Byrne, 1988; Edge, 1989; Hendrickson, 1978; Raimes, 1983; Ur, 1996). Many empirical studies now examine the relationship between teacher feedback and students' revision, employing analytical models; they provide sufficient amount of evidence of the short-term and long-term effectiveness of written CF (Bitchener, 2008, 2009; Bitchener & Knoch, 2010; Ferris, 2006; Sheen, 2010; S. Y. Kim, 2012). The majority of empirical studies supporting the efficacy of corrective feedback tended to favor indirect over direct feedback, on the ground that indirect feedback is better able to stimulate the students' meta-cognitive thinking process. As Haswell (1983) suggests, CF might reduce students' own written errors over time by letting them correct themselves immediately.

However, as Bitchener, Young, and Cameron (2005) address, there seems to be some doubt about a consistency of accuracy improvement over several semesters. A group of recent studies have argued against grammar correction in L2 writing classes (Gray, 2000; Krashen, 1992; Scarcella, 1996; Truscott, 1996, 1999, 2007, 2009), raising the issue of whether correction should be made in L2 writing classes to enhance accuracy. Truscott (1996, 2007) argues that CF should not be given on several grounds. He raises the issue of research design by showing that the evidence from controlled experiments does not justify the positive effect of grammatical error correction because the existing research does not make a clear distinction between correcting errors and providing no feedback at all. In such experimental studies, the teachers' feedback is not systematic but random, and the role of content-correction in writing accuracy is often ignored.

The traditional research on CF is often designed to measure students' writing

[1] In this study, the term "corrective feedback" is defined in the same vein as Lightbown and Spada (2006) as: Any indication to the learner that his or her use of the target language is incorrect. Corrective feedback can be explicit or implicit, and may or may not include metalinguistic information (p. 197).

accuracy under the encapsulated experimental conditions where instructional interventions with the set of performance tests are given to students in particular ways. Although this methodological approach is generally used to test isolated effect types of CF may have on the improvement in accuracy scores, it ignores the important role played by the affective variables, such as motivation, test anxiety, or students' fear of evaluation. Under experimental conditions where students are likely to be exposed to a higher level of affective variables, test results are subject to be skewed to levels of either over-performance or under-performance. For example, if the affective filter serves as a facilitator or barrier to learning in an L2 composition class, the result should be generalized across learning contexts. Affective factors can be raised or lowered as a result of classroom settings in which individual students interact with a teacher and their peers. However, the role of affective filter in CF has been rarely studied in prior research, though Truscott (1996, 2007) indirectly addressed the measurement issue through experimental design.

In the spirit of Truscott (1996), this study examines the effectiveness of CF, using students' data obtained from an electronic communication channel (i.e., an online discussion board) instead of relying only on data from the experimental classroom condition. An analysis of written communication data is different from that of in-class post tests in that students may produce this output in a less stressful place in order to communicate in writing. Such written data produced in a natural setting are considered to measure the improvement in the students' writing accuracy with a lower level of affective filter. Using the data obtained from both experimental and natural settings, this paper compares the improvement in accuracy scores in order to examine the effect of error feedback on writing accuracy. More specifically, this study measures the improvement in the writing accuracy of EFL college writers by comparing students' performance data collected from both the encapsulated experimental setting and the less stressful condition, the electronic communication channel. A descriptive statistic and a two-factor ANOVA with repeated measures are performed to examine the effectiveness of different types of error feedback on the students' writing over an extended period of time and across different writing contexts.

The students in the composition class, taught by the teacher/researcher, are divided into three groups: the direct CF group, the indirect CF group, and the self-correction group. The direct CF group is provided with summary end notes

about students' grammar problems but in-text correction while the indirect CF group is given the text with markings at the points of error. The self-correction group, used as a control group, engages in reading and writing practices (self-revision process) without any support from CF. Instead, they attend individual conferences with the teacher about writing content and organization for two semesters. Two research questions framed to investigate these aims are as follows:

1. Do types of corrective feedback help improve writing accuracy of L2 college writers in the use of two linguistic error categories?
2. Does the feedback effect, if any, persist when students' written communication occurring through an electronic channel (i.e., an online discussion board) is used as a means to measure their accuracy performance within a natural setting?

2. Literature review

2.1 Effectiveness of written corrective feedback

Since the 1970's, interactionist/cognitive theories have examined the facilitating effect of CF on L2 learners' acquisition. They focus on what happens inside the learner's head by emphasizing the role of attention and rehearsal that make up acquisition. Most of the recent studies on written CF (Ashewell, 2000; Bitchener & Knoch, 2010; Chandler, 2003; Fathman & Whalley, 1990; Fazio, 2001; Frantzen, 1995; Lalande, 1982; Robb, Ross, & Shortreed, 1986) have been based on this cognitive view in that they have examined which type of the written corrective feedback helps learners more attentive in noticing and correcting their errors.

On the one hand, Ashewell (2000) and Fathman and Whalley (1990) argued that indirect CF was more effective than direct CF because it caused learners to reflect and notice their errors, which led to self-correction and fostered long-term acquisition. On the other hand, others suggested that direct CF was more helpful to writers because it minimized students' confusion over teachers' feedback and facilitated immediate correction to a student's writing (Bitchener & Knoch, 2010). Chandler's (2003) study resulted in mixed findings on the effects of four CF types:

(a) correction (also called direct correction (Ellis, 2009)); (b) underlining with description; (c) description of type only; and (d) underlining. Chandler found that direct feedback had equal benefits with indirect metalinguistic feedback and that those two were more influential to students than the others.

The studies mentioned above compared learners' accuracy performance on pre-tests and (delayed) post-tests to measure the effects of CF. However, Truscott (2007) argued that this type of study failed in measuring change in students' ability in writing accurately, and thus gave no reliable result in learning. He argued that correcting students' errors did not show an improvement in writing accuracy in a new piece of writing even if the correction may eliminate the errors in a subsequent draft. Also, he raised the issue of research design by arguing that the evidence from controlled experiment did not justify the positive effect of grammatical error correction because the existing research does not make any clear distinction between correcting errors and providing no feedback at all.

Although the relation between types of feedback and outcomes is well established in the extant literature (Ferris, 2002, 2003, 2004; Kepner, 1991; Polio, Fleck, & Leder, 1998; Semke, 1984; Sheen, 2007, 2010), results driven empirically through quasi-experimental design can hardly be generalized and are specific to the learning setting from which they emerged. One factor needed to be considered as learner-specific is the affective filter. However, less is known about the relation of CF to affective filter. Studies with such design do not consider the role of the affective variables in writing improvement. According to Terrell (1977), affective rather than cognitive factors should be of primary concern in language classroom, and the correction of students' errors is "negative in terms of motivation, attitude, and embarrassment" (p. 330). Under the encapsulated experimental condition, students tend to be exposed to a higher level of affective variables due to testing anxiety. This may interfere with the measuring of their ability in writing accuracy, which is related to situation-specific anxiety. Such a specific type of event or situation can lead them to language anxiety due to their competitive nature. They are likely to become anxious when they recognize they are compared with learners from other groups and find themselves less proficient.

2.2 Corrective feedback in the computer-mediated communication

Chapelle (2001) mentioned the importance of written interaction through computer-mediated communication (CMC)[2] on learners' acquiring some linguistic forms such as English articles, third person singular *-s*, and the past tense *-ed* morpheme since it can increase the visual saliency of linguistic forms. In addition, Payne and Whitney (2002) pointed out that one of the advantages of CMC is that it can help to notice and produce target linguistic forms which need greater control due to increased processing and planning time. Thus, learners can benefit more in reviewing and reusing target language forms available in the input through CMC than a controlled experimental setting. By going through this process, they can show a natural progress of form-related writing development.

In spite of the potential advantages of CMC, there are a fairly limited number of outcome-based studies showing the effect of CF in CMC on improving grammatical competence through highlighting errors in certain syntactical features (Lowen & Erlam, 2006; Sachs & Suh, 2007; Sauro, 2009). Lowen and Erlam (2006) compared the effectiveness of different types of CF during small group text-chat interactions, but demonstrated no significant advantage for one CF type over the other and either feedback type over the control condition. They suggest that students' proficiency with the target form might not have been high enough for them to internalize the correct forms from the feedback during the short period of time. Similarly, Sachs and Suh (2007) demonstrated no significant difference in the target form accuracy between the groups that had different types of CF.

While the studies mentioned above do not show whether CF in the CMC context is more effective than that of the control condition (i.e., CF in experimental setting), many researchers argue that language outcomes would be significantly different if learners are exposed to learning contexts which allow them opportunities to focus on form (Doughty & Williams, 1998; Ellis, 1995; Norris & Ortega, 2000). Accordingly, the present study investigates the effectiveness of the different types of CF on the development of two grammatical features among advanced level English learners through a comparison of the classroom learning contexts using CMC. In this respect,

---

[2] Lee (2008) argues that CMC can help learners participate in affordable conditions by supporting both meaning-oriented communication and focus-on-form reflection needed to develop their language competence.

this study could be considered as an extension of CF research conducted under the CMC context but took an approach different from these studies. Specifically, it examined the impacts CF may have on the improvement in writing accuracy by considering both the control and CMC settings simultaneously. An existing body of research provides the direction for this study, or an investigation of the CF effects under two different research settings, as indicated by the research questions proposed.

3. Methods

3.1 Participants and context

The participants in this study were 25 L2 writers who were enrolled in the 'intermediate English composition class' in the department of English Education in a local university in Korea. This course was a part of the department writing program that was designed to help learners prepare for academic writing requirement and "Teaching Certificate Examination" in Korea. In this English composition class, learners learned about various aspects of writing an argumentative essay, or articulating thesis, developing their own argument, and refuting counterclaims. This process-oriented writing course could be understood as a reading-to-write class in that reading is connected to writing through various types of classroom activities (i.e., reading discussions, peer revisions, individual conferences, and discussion board). The course focused on how to construct a five-paragraph essay by offering some specific suggestions for writing the introduction, the body and the conclusion of the essay.

The teacher and researcher taught the course during the Fall semester in 2009. Most of the students were highly motivated to be good L2 writers in the areas of content and form to perform the writing tasks successfully. They considered writing proficiency as an essential part of preparing for "Teaching Certificate Examination in Korea." In this writing class, the discussion board as an alternative written communication channel available to the learners provided an opportunity to communicate with their peers without any restriction. Throughout this electronic channel, all of the participants were able to debate any issues discussed or missed in class. The amount of utterances the learners produced in such a low stressed

condition might serve as good student data to trace a natural progress of writing accuracy during the whole process of producing an essay.

The participant information on age, gender, and diagnostic test score at the beginning of the semester are reported in Table 1. The mean score for the TOEFL test was 70.1 out of 120 points, with the range of 57 to 92. The summary writing was given to the learners at the beginning of the semester to measure overall proficiency in writing. And the scores for both the TOEFL and summary writing were used to measure their L2 proficiency.

Table 1. Participant characteristics

Participants	Mean Age (standard deviation)	TOEFL test	
		Score (Range)	S. D.
Total Student(*N*=25)	23.1 (1.5)	70.1 (57 ~ 92)	10.5
Male (*n*=9)	24.5	68.1 (57 ~ 87)	10.7
Female (*n*=16)	21.8	72.2 (67 ~ 92)	10.6

Note: The score range for the TOEFL is 0 to 120.

3.2 Research design

To examine the effectiveness of CF on writing accuracy, the specific ways to form the group, to choose targeted linguistic errors, and to develop analytic procedures were illustrated in this section. For this purpose, the learners were divided into the three groups: the indirect CF group, the direct CF group, and the control group. Each CF group was provided with different types of CF; summary end notes about students' grammar problems but in-text correction, the text with markings at the point of error (see examples in Appendix), and individual conference with a teacher about content and organization. On the other hand, the control group got the conference session with the teacher after submitting the first draft of each given topic to satisfy ethical requirements.

3.2.1 Targeted linguistic errors

The selection procedure of Bitchener et al. (2005) was employed to choose two

targeted linguistic errors. This study identified the two linguistic errors occurring most frequently in the first writing task and in the written product from the discussion board. The range of error categories was identified in Table 2.

The most frequently permitted linguistic error by L2 writers was the use of prepositions (22.5% of the total errors), the personal pronouns (18.8% of the total errors), and subject-verb agreement (12.3% of the total errors). As indicated by many empirical studies (Bitchener et al., 2005; Ferris, Chaney, Komura, Roberts, & McKee, 2000; Ferris & Roberts, 2001; Sheen, 2007), L2 learners also had difficulty in the use of definite articles in their writing, with a corresponding number of 12.0%. Among the types of errors, the researcher decided to choose the two linguistic errors the students made most frequently: prepositions and subject-verb tense agreement.

Table 2. Number and percentage of error types

Types of Linguistic Errors	Number of Errors	Total Errors (%)
Prepositions	73	22.5
Personal pronouns	61	18.8
Subject-verb agreement	40	12.3
Definite articles	39	12.0
Indefinite articles	30	9.3
Nouns	18	5.6
Word choice	10	3.1
Adverbs	10	3.1
Demonstrative pronouns	9	2.8
Relative pronouns	8	2.5
Future	6	1.9
Modals	4	1.2
Capitalization	4	1.2
Comparatives	3	0.9
Subordinate conjunctions	3	0.9
Coordinate conjunctions	3	0.9
Passive	2	0.6
Others	2	0.6
Total	325	100.0

However, an error category of ‘personal pronouns’ was not included in the analysis on the ground that it was repeatedly treated in many prior researches

(Bitchener et al., 2005; Kepner, 1991; Polio, Fleck, & Leder, 1998: Sheen, 2007). The scripts were marked by the researcher and a native speaking instructor to identify and categorize error types, and an agreement rate of 94% indicated that the process of selecting linguistic errors was reliable.

3.2.2 Writing practices in class and out of class

In the intermediate English composition class, each participant was required to write 3 sets of an argumentative essay during the semester. This course could be considered as an integrated course of reading and writing in that the students developed their own ideas during the writing process using reading articles. Three main topics (i.e., religion, gender role, and mass media) as writing assignments were given to the students with three reading texts on each given topic. After engaging in reading and discussing each topic in class, the students were required to write an argumentative essay with five paragraphs during the semester. After submitting each draft, the students received CF from the teacher, and they were asked to revise it to resubmit it one week after they got the feedback.

Another writing task was related to the electronic discussion board, which was used as an extension of classroom discussion conducted in written form. Unlike in-class writing practices, the discussion board provided a social space to produce written data through an on-going interaction with other students. In this respect, such written products would be more appropriate to access a natural progress of writing accuracy in the use of the two linguistic errors by the students.

3.2.3 Analytic procedures

An analytic procedure is consistent with that of a tradition CF research in that this study employed an experimental set-up often used in the existing empirical studies in this field (Bitchener & Knoch, 2010; Chandler, 2003; Fathman & Whalley, 1990; Fazio, 2001; Frantzen, 1995). Specifically, in designing an experimental condition (i.e., ways of conveying instructional intervention, giving the time intervals between tests, and forming the control group), the researcher used an approach similar to these studies, as demonstrated below.

Two qualitative within-participant factors were analyzed: linguistic errors at three

levels (prepositions and subject-verb tense agreement) and times at three levels (week 4, week 8, and week 13 during the semester). In addition, the between-participants factor was analyzed based on CF at three levels (direct correction, indirect correction, self-practices). To form the groups at the beginning of the semester, all of the students in the composition class were classified into 3 proficiency groups according to the rank of English proficiency, which was measured by an average score of their Test of English as a Foreign Language (TOEFL). And the students in 3 proficiency categories were randomly chosen to establish the three experimental groups: two treatment groups (indirect CF group and direct CF group) and one control group.

As shown in Table 3, the experiment set-up consisted of two levels pertaining three subsections, or the traditional experimental setting and the natural setting. Specifically, to examine the effect of CF on the writing accuracy, this paper used both the post tests collected in class and written communicative data obtained from an electronic discussion board. Written products from an electronic discussion board is considered to be data produced in a natural setting, as compared with post tests produced in a controlled classroom setting.

Table 3. Experimental set-up

<table>
<tr><td rowspan="2">Initial Tasks
Pretest
(Week 1)</td><td colspan="4">Treatment
(Weeks 2-3)</td><td colspan="3">Three Subsequent Tasks
Posttest</td></tr>
<tr><td>Group</td><td colspan="3">Repeated Procedures</td><td>Week 4</td><td>Week 8</td><td>Week 13</td></tr>
<tr><td rowspan="3">Text Quality
(form and content)</td><td>Direct Group(N=8)</td><td>Feedback</td><td>Revision</td><td>No Practice</td><td colspan="3" rowspan="3">Text Quality
(content only)</td></tr>
<tr><td>Indirect Group(N=8)</td><td>Feedback</td><td>Revision</td><td>No Practice</td></tr>
<tr><td>Control Group(N=9)</td><td>No Feedback</td><td>No Revision</td><td>Additional Practice</td></tr>
<tr><td>Discussion Board</td><td colspan="4">Same procedures</td><td colspan="3">No test
(An analysis of cumulative data)</td></tr>
</table>

When it comes to the experimental setting, the students in each group took initial writing tasks at the beginning of the semester (week 1) to establish overall writing

proficiency. In a writing task, the students were required to summarize a given reading article with one or two paragraphs, using their own words. And in the following week, the students in both the indirect CF group and the direct CF group received feedback from their instructor and were required to use it in revising their own writing summaries. On the other hand, the students in the control group, defined as the self-practice group, practiced the self-correction of their writing summaries without any support from a teacher's feedback. The procedure associated with treatment was repeated for the first three weeks to provide the opportunities to learn ways of using types of feedback in their learning processes.

When it comes to the posttest procedure, the three post tests were performed to examine the effect of CF on writing accuracy over the course of the study. After the fourth week, all of the students would get content-related feedback only. The three types of post tests could be considered to be the immediate test conducted right after treatment (week 4), the intermediate test (week 8), and the delayed test (week 13). With regard to the discussion board, written communication data occurring at the natural setting were used as the writing tasks. All of the written data accumulated during the period of the three subsequent tasks (week 4, 8, and 13) were used as posttest measures, respectively. Generally, the first two tests could be relatively viewed as a short-term, while the final test was considered to be long-term. The number of form-related errors per sentence in the two linguistic error categories was used as a unit of analysis of the writing accuracy measures.

To examine mean differences in accuracy scores over the three treatment times, a two-factor ANOVA with repeated measures on one factor was performed. Specifically, all of the students were exposed to treatment time at three levels (weeks 4, 8, and 13) and between-participant factors at three levels (direct CF feedback, indirect CF feedback, and self-practice) with an accuracy measure as a dependent variable. Descriptive statistics for the three tests were used to describe the characteristics of each group. Since the mean differences in pre-test scores were not statistically significant (p = .348), a two-way repeated measured ANOVA was used to examine the effectiveness of CF on writing accuracy. As an aside, one-way ANOVAs with Turkey's post hoc pair-wise comparisons were chosen to analyze the existence of the group differences at a given point in time. For an analysis of chronical data, an independent measures t-test was used to analyze the mean differences in written communication data obtained from the discussion board over

three within-subject factors and across the groups. The results for group differences in writing accuracy and the improvement in writing accuracy over time are reported in the following section.

4. Results

4.1 Effectiveness of CF under the experimental setting

The descriptive statistics for the three groups were presented to illustrate the students' characteristics over the four different testing times and across the groups. The mean values with the corresponding standard deviations for each group were reported in Table 4. Since an accuracy score is defined as the number of errors per sentence unit, or an error score, the decrease in accuracy score between two subsequent tests denotes the improvement in writing accuracy.

Table 4. Mean accuracy scores by each group

Group	*N*	Pre-test (Week 1)		Immediate Test(W 4)		Intermediate Test(W 8)		Delayed Test (W 13)	
		M	SD	M	SD	M	SD	M	SD
Direct CF	8	2.39	0.34	2.01	0.32	1.70	0.33	1.63	0.31
Indirect CF	8	2.33	0.29	1.88	0.25	1.64	0.27	1.61	0.26
Control	9	2.29	0.32	2.13	0.29	1.99	0.29	2.01	0.33
Overall	25	2.34	0.31	2.01	0.35	1.78	0.31	1.75	0.32

The mean values for the pretest results were quite similar across the groups, with a mean error score of 2.39 for the direct CF group, 2.33 for the indirect CF group, and 2.29 for the control group, as shown in Figure 1. This indicates that the students in each group possessed the similar level of writing proficiency at the beginning of the semester. However, the mean differences between the two CF groups and the control group got wider over the four different testing times. Specifically, two treatment groups receiving either type of CF showed the substantial improvement in writing accuracy in the use of two linguistic error forms in the subsequent tests.

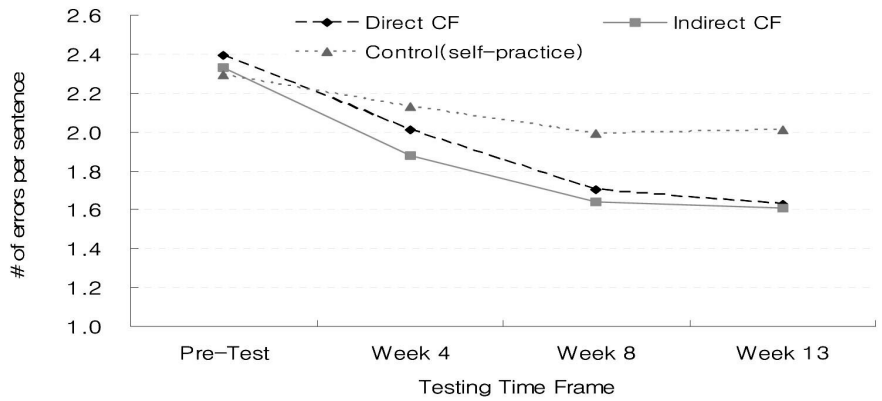

Figure 1. Effectiveness of corrective feedback over different testing time

On the contrary, the control group, practicing reading and writing activities (i.e., self-practices) without any support from error feedback, showed little improvement in accuracy in the subsequent tests, with the mean score of 2.29 for the pretest, 2.13 for the immediate test, 1.99 for the intermediate test, and 2.01 for the delayed test, respectively.

On the other hand, both the treatment groups were able to improve their mean writing accuracy in the subsequent tests (weeks 4, 8, and 13), though the accuracy gains were reduced over time. Unlike these treatment groups, the control group tended to show a slight improvement in the mean writing accuracy over the two subsequent tests, but such accuracy gains disappeared at the delayed post test. More importantly, both the direct and indirect CF groups showed the similar pattern of improving writing accuracy, widening the accuracy gap with the control group. This indicates that the effect of CF on writing accuracy was obvious though all three groups experienced the improvement in writing accuracy in the use of two linguistic error categories over time.

To compare the improvement of writing accuracy of the three groups, an analysis of ANOVA was performed. First, the one-way ANOVA was performed to check whether the group differences in writing accuracy existed at the beginning, and the result showed no statistically significant differences among the three groups at the time of the pre-test (F[3, 22] = 1.013, P = 395). Second, using a two-way repeated measured ANOVA with writing accuracy as the dependent variable, the treatment

(three levels) and CF types (three levels) as independent variables, accuracy differences between treatment groups were tested. The results showed that there was no significant interaction between *Time* and *CF* types. As reported in Table 5, two main effects were significant (F[3, 22] = 4.961, P = .001), pointing out the significant mean differences among the three groups and over three different testing horizon.

Table 5. Results from the two-way repeated measures ANOVA

Sources	*df*	*F*	*P*
Between Subject			
Corrective Feedback Types	2	4.916	.001
Within Subjects			
Time	3	2.901	.046
Time × Corrective Feedback Types	9	.0992	.201

Using post hoc comparison tests, the differences between pairs of groups were also examined. In particular, one-way ANOVA indicated that the group differences were not significant at the immediate post test (F[2, 22] = 1.345, P = .071). However, significant differences between the three groups were found in the intermediate post test (F[2, 22] = 3.993, P = .001) and in the delayed post test (F[2, 22] = 4.921, P = 0.01).

In short, the students in the two treatment groups showed a significant improvement in writing accuracy over the course of the study while accuracy gains for the control group were not significant. Specifically, accuracy differences between the two treatment groups and the control group were not immediate at the immediate post test, but such differences were reinforced at the subsequent post tests (the intermediate and delayed post tests).

4.2 Effectiveness of CF under the natural setting

In this section, the effectiveness of CF was tested using the written discussion data the learners produced through an on-line interaction with their peers during the semester. As compared with the posttest data produced in class, such written data were considered to be valuable in assessing a natural progress of writing accuracy over the course of the study. All of the students participated in the discussion board

as a channel to communicate and produced a total of 897 communication units (sentences) during the semester. The total communication units were chronically classified into 3 sets of data. That is, the data collected through weeks 1 to 4 were used for the immediate post test, written products associated with weeks 5 to 8 for the intermediate posttest, and ones obtained through weeks 8 to 13 for the delayed posttest.

Using an independent measures *t* test, the mean differences among three separate samples (the two treatment groups and the control group) were examined. For the immediate post test conducted right after treatment (week 4), the direct CF group (*Mean* = 2.29, *SD* = 0.33) showed a higher improvement in scores as compared with the indirect CF group (*Mean* = 2.41, *SD* = 0.38) and with the control group (*Mean* = 2.35 *SD* = 0.29). But, the difference in mean scores between the control and indirect CF groups was not statistically significant, $t(21) = 0.93$, $p > 0.05$, two tails. By the same token, the difference between the control and direct CF groups was also not statistically significant, $t(21) = 1.01$, $p > 0.05$, two tails.

Table 6. Differences in writing accuracy across the groups

	Pretest Scores	Posttest Scores (Weeks 1-4)	Posttest Scores (Weeks 5-8)	Posttest Scores (Weeks 9-13)
Direct CF Group (Improvement in Scores)	2.39	2.29 (0.10)	2.15 (0.14)	2.25 (-0.1)
Indirect CF Group (Improvement in Scores)	2.33	2.41 (-0.08)	2.27 (0.14)	2.19 (0.08)
Control Group (Improvement in Scores)	2.29	2.35 (-0.06)	2.26 (0.09)	2.29 (-0.03)
Total Mean Scores	2.34	2.35	2.23	2.24

Note: An accuracy score is defined as the number of error per sentence.

As shown in Table 6, for the intermediate post test (week 8), no significant differences in accuracy scores across the three groups were not found. Specifically, the direct CF group (*Mean* = 2.15, *SD* = 0.28) obtained a higher level of accuracy gains than did the indirect CF group (*Mean* = 2.27, *SD* = 0.37) and the control group (*Mean* = 2.26, *SD* = 0.32). For the delayed post test, the indirect CF group

(*Mean* = 2.19, *SD* = 0.39) showed a relatively high improvement in accuracy gain, while the accuracy scores for the two other groups were worsen. However, the differences in mean scores among the groups were not statistically significant at the 5% confidence level. Specifically, the mean score difference between the control and the indirect CF group was found to be insignificant, $t(21) = 0.88$, $p > 0.05$, two tails. Also, the difference between the control and the direct CF groups was not statistically significant, $t(21) = 0.69$, $p > 0.05$, two tails.

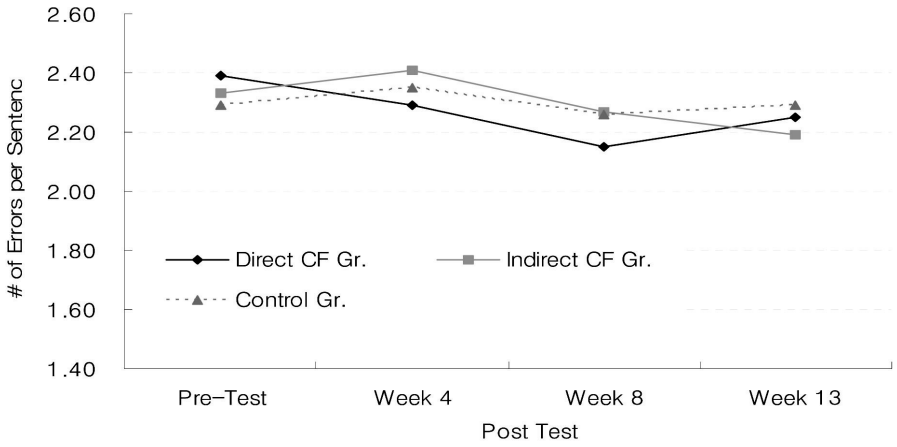

Figure 2. Mean accuracy scores across the groups and over time

In short, the effectiveness of CF was not documented from the analysis of the students' written communication. As opposed to the results from the experimental setting, all of the groups did not show the improvement in writing accuracy over an extended period of time. In particular, the group differences in accuracy scores were found not to be significant, indicating that the treatment effect was not reflected in the students' written communication via in an electronic channel.

5. Discussions and implications

In examining the effect of CF on writing accuracy of L2 writers, this paper took an approach which differs from the traditional method in this field. Specifically,

instead of post-test data collected in class, written communication data the learners produced in a natural setting (i.e., electronic discussion board) were used to examine the natural progress of writing accuracy in the two linguistic features. With regard to the first research question, using a traditional approach to the analytic procedure (i.e., experimental setting), the researcher investigated whether two types of CF helped to improve writing accuracy among L2 writers. The results showed that both the direct and indirect CFs have impact on form-related writing accuracy, but such an effect was not obvious for the learners in the control group. Under the traditional experimental setting, this study showed the effectiveness of CF, which supports the importance of error correction in L2 writing. As documented in other empirical studies (Ashwell, 2000; Ferris, 1997; Sache & Polio, 2007), the results emphasized on the role of CF in L2 writing classes.

With respect to the second research question, the effectiveness of CF was also tested using the written discussion data the learners produced through an electronic communication channel available to them outside of class. However, the results showed that the differences in writing accuracy across the groups were found not to be significant, thus providing evidence against CF. Such results are consistent with those discussed in other empirical studies (Chadler, 2003; Ferris et al., 2000; Frantzen, 1995). In many studies, the lack of effectiveness of CF is attributed to the way CF is delivered to L2 writers. Specifically, when instructional intervention is not appropriately practiced, it influences students' abilities to use instructional feedback in their writing practices (Truscott, 1996), thus casting serious doubt on the quality of CF.

Before discussing the effectiveness of CF, we as teachers need to take a close look at the two sets of data that produced contradicting results. That is, the post test data collected in class at a specific time were different from the students' written communication data in many aspects. First, an in-class post test is likely to expose learners to affective filters (i.e., test anxiety, motivation, or fear of negative evaluation), influencing test results either negatively or positively (Dornyei, 1996; Gardner, 1985; Lybeck, 2002). Second, such cross-sectional data might not be appropriate to evaluating the natural progress of form-related writing development. On the other hand, written data associated with an electronic discussion are likely to be obtained under a less stressful learning environment in that these data are produced in writing through interactions with peer learners. If the goal of CF is to

lead learners to the long-term growth in writing accuracy, longitudinal data they produced under a natural setting can be considered to be relevant to test an on-going process of the development of L2 writers.

With regard to the effectiveness of CF, the results were not sustainable in that the accuracy gains made in an experimental setting were not observed in the electronic communication channel. Specifically, an analysis of communication data did not show any pattern of improving writing accuracy across the groups over time. It illustrated the limited role of error corrections in L2 writing classes.

The findings provide important implications applicable to L2 writing classes. First, an attempt to provide an encapsulated instructional intervention in L2 writing class might not lead learners to the stable growth in writing development due to the overemphasis on error correction. From a socio-cultural perspective, the improvement in writing accuracy cannot be separated from content-related writing development in that form-related and content-related development in writing goes hand in hand. In this respect, we as teachers would view linguistic errors as part of the natural progress of language learning. Nevertheless, it is not still clear whether the interaction between form- and contend-related knowledge can serve as the sources of gains in accuracy, suggesting further research in this field. Thus, the issue of how to coordinate form- and content-related instructions in L2 English classes should be considered as an essential part of L2 teaching practices.

Second, from a pedagogical perspective, this study provides some implications for L2 composition classes. If an instructional approach to error correction is primarily targeted to achieving long-term accuracy, teachers need to provide content-related feedback coupled with additional reading and writing practices. These reading-writing practices can provide a social space in which connect form-related knowledge to content-related learning. When reading and writing practices are incorporated into teaching practices in an effective way, students are better able to correct linguistic errors from their written texts. In this case, the effect both form-and content related feedback may have on writing accuracy is likely to be reinforced each other. In this respect, it is interesting to examine how form- and content-related instructions influence and are influenced each other to improve writing accuracy in future research.

References

Ashwell, T. 2000. Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing* 9: 227-257.

Bitchener, J. 2008. Evidence in support of written corrective feedback. *Journal of Second Language Writing* 17(2): 102-118.

Bitchener, J. 2009. Measuring the effectiveness of written corrective feedback: A response to "Over-generalization from a narrow focus: A response to Bithener (2008)". *Journal of Second Language Writing* 18(4): 276-279.

Bitchener, J., & Knoch, U. 2010. Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing* 19(2): 207-217.

Bitchener, J., Young, S., & Cameron, D. 2005. The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing* 14(3): 191-205.

Byrne, D. 1988. *Teaching writing skills.* London: Longman.

Chandler, J. 2003. The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing* 12(3): 267-296.

Chapelle, C. 2001. *Computer applications in second language acquisition: Foundations for teaching, testing, and research.* Cambridge: Cambridge University Press.

Dornyei, Z. 1996. Moving language learning motivation to a larger platform. In R. L. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (pp. 38-51). Honolulu: University of Hawaii Press.

Doughty, C., & Williams, J. 1998. *Focus on Form in classroom second language acquisition.* Cambridge: Cambridge University Press.

Edge, J. 1989. *Mistakes and correction.* New York: Longman.

Ellis, R. 1995. Interpretation tasks for grammar teaching. *TESOL Quarterly* 29(1): 87-105.

Ellis, R. 2009. Corrective feedback and teacher development. *L2 Journal* 1: 3-18.

Fathman, A., & Whalley, E. 1990. Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.

Fazio, L. 2001. The effects of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing* 10: 235-249.

Ferris, D. R. 2002. *Treatment of error in second language student writing.* Ann Arbor: University of Michigan Press.

Ferris, D. R. 2003. *Response to student writing: Implications for second language students.*

Mahwah, NJ: Lawrence Erlbaum Associates.

Ferris, D. R. 2004. The "Grammar Correction" debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing* 13: 49-62.

Ferris, D. R. 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* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. R., & Roberts, B. 2001. Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing* 10: 161-184.

Frantzen, D. 1995. The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *The Modern Language Journal* 79: 329-344.

Gardner, R. C. 1985. *Social psychology and second language learning: The role of attitudes and motivation.* London, GB: Edward Arnold.

Haswell, R. H. 1983. Minimal markings. *College English* 45(6): 600-604.

Hendrickson, J. 1978. Error correction in foreign language teaching: Recent theory, research and practice. *The Modern Language Journal* 62: 387-398.

Kepner, C. G. 1991. An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal* 75: 305-313.

Kim, S. Y. 2012. The long-term effectiveness of written corrective feedback in an EFL writing class. *The Journal of Foreign Studies* 21: 9-35.

Krashen, S. D. 2002. *Explorations in language acquisition and use: The Taipei lectures.* Taipei, Taiwan: Crane.

Lalande J. F. 1982. Reducing composition errors: An experiment. *Modern Language Journal* 66: 140-149.

Lee, L. 2008. Focus-on-form through collaborative scaffolding in expert-to-novice online interaction. *Language Learning & Technology* 12(3): 53-72.

Lightbown, P. M., & Spada, N. 2006. *How language are learned* (3rd ed.). New York: Oxford University Press.

Loewen, S., & Erlam, R. 2006. Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning* 19(1): 1-14.

Lybeck, K. E. 2002. *The role of acculturation and social networks in the acquisition of second language pronunciation.* Unpublished doctoral dissertation. University of Minnesota.

Norris, J., & Ortega, L. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417-528.

Payne, J. S., & Whitney, P. J. 2002. Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal* 20(1): 7-32.

Polio, C., Fleck, N., & Leder, N. 1998. "If only I had more time": ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing* 7: 43-68.

Raimes, A. 1983. *Techniques in teaching writing.* New York: Oxford University Press.

Robb, T., Ross, S., & Shortreed, I. 1986. Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly* 20: 83-93.

Sachs, R., & Polio, C. 2007. Learners' use of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition* 29: 67-100.

Sachs, R., & Suh, B. 2007. Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 197-227). Oxford: Oxford University Press.

Scarcella, R. 1996. Secondary education in California and second language research: Instructing ESL students in the 1990s. *The CATESOL Journal,* 129-151.

Semke, H. 1984. The effects of the red pen. *Foreign Language Annuls* 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.

Sheen, Y. 2010. Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition* 32: 203-234.

Terrell, T. D. 1977. A natural approach to second language acquisition and learning. *The Modern Language Journal* 61(7): 325-337.

Truscott, J. 1996. The case against grammar correction in L2 writing classes. *Language Learning* 46: 327-369.

Truscott, J. 1999. The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing* 8: 111-122.

Truscott, J. 2007. The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing* 16: 255-272.

Ur, P. 1996. *A course in language teaching.* Cambridge: Cambridge University Press.

APPENDIX A

Group one: Direct feedback sample

First, reality shows give[s∅] viewers sincere pleasure. Because it focus[ES] on real situations, viewers cannot predict what will happen next. So, they can be interested and nervous and sad at times. For example, according to the article of[FROM]News Wave, hit program "One Night Two Days" recorded the highest ratings ever of variety shows. "One Night Two Days" has televised for five days and made viewers laugh and cry due to unexpected situations.

COMMENT:

Prepositions link nouns, pronouns, and phrases to other words in a sentence. The word or phrase that the preposition introduces is called the object of the preposition. The errors occurred on your essay as a result of using the wrong preposition (i.e., the article of News Wave --> the article **from** New Wave).

Regarding subject-verb agreement, the basic rule states that a singular subject takes a singular verb, while a plural subject takes a plural verb (i.e., reality shows give[S] ..., it focus[ES]...).

APPENDIX B

Group two: Indirect feedback sample

The second effect of these shows is that it is hind business. As I said it before, it is not acting it is getting a ringside view of other people's ostensibly private life, so audience is captivated with the program. Also, they tend to believe and make credit with that program which are profit to business part. People just follow what the people in the shows are using, wearing, speaking, and so on because they are at a seemingly real situation. This is called product placement.

Sun-Young Kim
Department of English Education
Mokpo National University
1666 Youngsan-ro, Cheonggye-myeon, Muan-gun
Cheonnam 534-729, South Korea
E-mail: sunyoung0412@mokpo.ac.kr

Received: 2012. 10. 30
Revised: 2012. 11. 26
Accepted: 2012. 11. 26