

Evidence in support of written corrective feedback

John Bitchener*

School of Languages, Auckland University of Technology, Private Bag 92006, Auckland 1020, New Zealand

Abstract

The extent to which ESL learners benefit from written corrective feedback has been debated at length since Truscott (1996) mounted a case for its abolition. Ten years later, the debate continues, not only because little attention has been given to testing its efficacy over time but also because studies that have investigated the issue have not always been well designed and have produced conflicting results (Ferris, 2004, 2006). This article presents the results of a 2-month study of the efficacy of written corrective feedback to 75 low intermediate international ESL students in Auckland, New Zealand. Assigned to 4 groups (direct corrective feedback, written and oral meta-linguistic explanation; direct corrective feedback and written meta-linguistic explanation; direct corrective feedback only; the control group received no corrective feedback), the students produced three pieces of writing (pre-test, immediate post-test, and delayed post-test) that described what was happening in a given picture. Two functional uses of the English article system (referential indefinite “a” and referential definite “the”) were targeted in the feedback. The study found that the accuracy of students who received written corrective feedback in the immediate post-test outperformed those in the control group and that this level of performance was retained 2 months later.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Written corrective feedback; Error correction; Accuracy improvement; Treating linguistic errors

Introduction

Prior to 1996 when Truscott claimed that written corrective feedback (error correction on L2 student writing) is ineffective and harmful, the assumption that corrective feedback helps L2 writers improve the accuracy of their writing had not been challenged. In fact, as Truscott (1996, 1999) and Ferris (1999) explained, research evidence was limited in terms of the range of studies that had attempted to address the question of efficacy and in terms of the quality of the research design. Although a decade has now passed and considerable debate has been presented in journal articles and conference papers, limited research has been undertaken on this key issue. In considering that

* Tel.: +64 9 921 9999x7830; fax: +64 9 921 9978.

E-mail address: john.bitchener@aut.ac.nz.

which has been published, it is clear that a conclusive answer to the question will not be possible unless researchers make a concerted effort to conduct well designed studies that examine over time the effectiveness of different corrective feedback options on new pieces of writing and by comparing them with the texts of students who do not receive corrective feedback. An attempt at addressing these needs lay behind the design and execution of the study reported in this article.

The aim of the study was twofold: (1) to investigate whether targeted corrective feedback on ESL student writing results in improved accuracy in new pieces of writing over a 2-month period and (2) to investigate whether there is a differential effect on accuracy for different corrective feedback options. Each component of these aims was carefully selected to address the design limitations of earlier studies and thereby provide a more robust platform from which answers to both questions might be sought. Thus, the study focuses on one targeted error category (two functional uses of the English article system) rather than a myriad of error categories. Secondly, it examines longitudinally, by means of a pre-test/post-test design, the effectiveness of corrective feedback on new pieces of writing within the same genre rather than single or multiple text revisions across different genres. Thirdly, it incorporates a control group (one that does not receive corrective feedback) so that its error ratios can be compared with those of the treatment groups. Fourthly, it investigates the extent to which three under-researched direct feedback options, typically offered in L2 classrooms, determine accuracy performance: direct corrective feedback; direct corrective feedback plus written meta-linguistic explanation; and direct corrective feedback plus written and oral meta-linguistic explanation. A justification of this focus is presented in the following review of the published literature.

Background literature

The first part of this section presents a literature review of empirical studies that argue for and against the practice of corrective feedback and the relative merits of different types of feedback. The second part reviews the research design issues that need to be addressed so that future research is able to provide clearer answers to the key questions concerning the value of written corrective feedback.

Part one: empirical studies

A number of studies have claimed that corrective feedback on ESL student writing is effective. However, the design of very few has included a control group so it has not been possible to make a comparison between those who do and those who do not receive corrective feedback. Studies that fail to compare the effects of corrective feedback and no corrective feedback do not provide evidence of the effectiveness of corrective feedback (Ferris, 1999, 2004; Truscott, 1996, 2004). Of those that have made the comparison (Ashwell, 2000; Fathman & Whalley, 1990; Ferris & Roberts, 2001; Kepner, 1991; Polio, Fleck, & Leder, 1998), three claim that corrective feedback had a positive effect on accuracy but, in each case, there were design issues that need to be taken into account when considering the value of their claims. These issues and others concerning the design of these studies are presented in Table 1.

Although the findings of Fathman and Whalley (1990) offer positive evidence of the value of corrective feedback, it should be noted that their post-test only required a revision of the pre-test rather than the writing of a new text. Like their study, the Ferris and Roberts (2001) investigation found a positive effect for both types of written corrective feedback, but it also involved only text revisions. In Ashwell's (2000) study, gains in accuracy between drafts one and three were evident

Table 1
Control group studies claiming WCF improves accuracy

Study	Participants	WCF type	Duration	Effective	Limitations
Fathman and Whalley (1990)	72 ESL learners (intermediate) USA college	(1) Indirect underlining (2) Content comment (3) Content comment and indirect underlining (4) Control	A few days	Yes Groups 1 and 3 outperformed groups 2 and 4	(1) New texts not measured; text revision only (2) Not longitudinal (3) Focus on all errors
Kepner (1991)	60 Spanish learners (intermediate) USA college	(1) Direct error correction (2) Control	1 semester	No	(1) No pre-test measurement (2) No control over journal entry length (3) No control over texts written out-of-class (4) Analytical flaws
Polio et al. (1998)	65 ESL learners USA university	(1) Error correction; editing instruction; text revision (2) Control	7 weeks	No	(1) Different instruments in post-test (journal entry v in-class essay)
Ashwell (2000)	50 EFL learners Japan university	(1) Content comment then indirect underlining and coding (2) Indirect underlining and coding then content comment (3) Mix of (1) and (2) (4) Control	1 semester	Yes Accuracy gains for groups 1–3 in draft 3	(1) New texts not measured; text revision only (2) Effect of intervening variables possible
Ferris and Roberts (2001)	72 ESL learners USA college	(1) Indirect underlining and coding (2) Indirect underlining (3) Control	1 semester	Yes Groups 1 and 2 outperformed group 3	(1) New texts not measured; text revision only

in the text revisions of all three groups that received form-focused feedback but not in those redrafted by the control group. Like the previous two studies (Fathman & Whalley, 1990; Ferris & Roberts, 2001), Ashwell's study did not require students to write new texts. While the findings of these three studies signal a positive effect for written corrective feedback on error reduction, two further studies (Kepner, 1991; Polio et al., 1998) report the opposite. Polio et al. (1998) claim that differences in post-test scores for the treatment and control groups were not significant, but it needs to be realised that different instruments were used in the two tests and that this, therefore, raises the possibility that instrument variability might have had an effect on the findings. Similarly, a number of design and execution shortcomings in the study by Kepner (1991) mean that its claims also need to be read with this in mind.

Other studies (Chandler, 2000; Ferris, 1995, 1997; Ferris, Chaney, Komura, Roberts, & McKee, 2000; Lalande, 1982) that have not included a control group have nevertheless claimed that corrective feedback is beneficial to students who receive it. These are outlined in Table 2.

While the findings from this research are interesting and may well be indicative of the effectiveness of corrective feedback, they were not compared with the accuracy scores of a control group so cannot be read as evidence. Thus, they can really only be regarded as offering insights into the relative effectiveness of different types of feedback.

Studies comparing different types of corrective feedback

The major focus of studies that have investigated the effectiveness of different types of corrective feedback has been the extent to which direct or indirect feedback facilitates improved accuracy. Although these terms have not always been used consistently in the literature, *direct* corrective feedback may be defined as the provision of the correct linguistic form or structure above or near the linguistic error (Bitchener, Young, & Cameron, 2005; Ferris, 2003). It may include the crossing out of an unnecessary word/phrase/morpheme, the insertion of a missing word/phrase/morpheme, or the provision of the correct form or structure. Additional forms of direct feedback may include written meta-linguistic explanation (the provision of grammar rules and examples at the end of a student's script with a reference back to places in the text where the error has occurred) and/or oral meta-linguistic explanation (a mini-lesson where the rules and examples are presented, practiced, and discussed; one-on-one individual conferences between teacher and student or conferences between teacher and small groups of students).

On the other hand, *indirect* corrective feedback indicates that in some way an error has been made. This may be provided in one of four ways: underlining or circling the error; recording in the margin the number of errors in a given line; or using a code to show where the error has occurred and what type of error it is (Ferris & Roberts, 2001; Robb, Ross, & Shortreed, 1986). Rather than the teacher providing an explicit correction, students are left to resolve and correct the problem that has been drawn to their attention.

In earlier years, a stronger case had tended to be made for the special value of providing students with indirect feedback rather than direct feedback. Lalande (1982) and James (1998) explained that indirect feedback requires learners to engage in guided learning and problem solving and, therefore, promotes the type of reflection that is more likely to foster long-term acquisition. But as SLA researchers of oral L2 production have found, learners must first "notice" (Schmidt, 1990) that an error has been made. Once the error has been noted, indirect feedback has the potential to push learners to engage in hypothesis testing—a process which Ferris (2002) and others (see Doughty & Williams, 1998) suggest may induce deeper internal processing and promote the internalization of correct forms and structures.

Table 2
Studies without control group predicting WCF improves accuracy

Study	Participants	WCF type	Duration	Effectiveness
Lalande (1982)	60 German FL learners (intermediate) USA university	(1) Direct error correction (2) Guided learning and problem solving	10 weeks	Improvement Group 1 outperformed group 2 in post-test
Ferris (1995)	30 ESL learners USA university	Selective indirect underlining	1 semester	Improvement but inconsistent in some error categories and essays
Ferris (1997)	47 ESL learners USA university	Teacher commentary and selective indirect underlining	1 semester	Improvement
Ferris et al. (2000)	92 ESL learners USA university	Mix of direct, indirect (coded and uncoded); notes (marginal and end-of-text); text revision	1 semester	Improvement 81% accurate revision by end of semester
Chandler (2000)	30 ESL learners USA college	(1) Indirect underlining and revision (2) Indirect underlining only	1 semester	Improvement Group 2 reduced errors by one third in essay 5

While not dismissing the value of indirect feedback, those more in favour of a direct approach have explained that teachers and students prefer direct feedback (Ferris et al., 2000; Ferris & Roberts, 2001; Komura, 1999). In addition, they suggest that direct feedback reduces the kind of confusion that can result when students fail to understand or remember the meaning of error codes used by teachers. Ferris and Roberts (2001) explain how this can easily occur with lower proficiency learners. Leki (1991) and Roberts (1999) have also pointed out that students sometimes feel that indirect feedback does not provide them with sufficient information to resolve more complex errors such as idiosyncratic and syntactic errors. More recently, Chandler (2003) explained that the greater cognitive effort expended when students are required to use indirect feedback to make their own corrections is offset by the additional delay in knowing whether their own hypothesized correction is in fact correct. Weighing up the relative merits of the various claims is not possible, however, unless the findings of well-designed empirical studies are considered.

A limited number of studies have investigated whether direct or indirect corrective feedback is more facilitative of improved accuracy in L2 writing. Five studies (see Table 3) have compared the two approaches while a smaller body of research has compared the effectiveness of different types of indirect feedback and different types of direct feedback.

Some studies (Semke, 1984; Sheppard, 1992) have also compared direct and/or indirect treatments with content/comments feedback. Comparing first the relative effects of direct and indirect feedback, it can be seen in Table 3 that two studies (Ferris & Helt, 2000; Lalande, 1982) report an advantage for indirect feedback. On the other hand, Robb et al. (1986) and Semke (1984) found no significant differences across feedback types. The first of these is particularly compelling because it was carefully designed and considered multiple drafts of completed compositions. However, the findings of Semke's study are problematic because of the type of writing the students were doing (weekly journal entries) and because the post-test written accuracy measure consisted of only a 10-minute free-write. Another reason for being tentative in making firm conclusions from this conflicting and limited body of evidence is the positive findings for direct feedback reported by Chandler (2003).

Table 3
Studies comparing the effectiveness of direct and indirect WCF

Study	Participants	WCF types	Duration	Most effect
Lalande (1982)	60 German FL learners (intermediate) USA university	(1) Direct error correction (2) Indirect coding and error logs kept	10 weeks	Indirect
Semke (1984)	141 German learners USA university	(1) Comments (2) Direct corrections (3) Direct corrections and comments (4) Indirect (coded) corrections	10 weeks	No difference
Robb et al. (1986)	134 EFL learners Japan college	(1) Direct corrections (2) Indirect coded feedback (3) Indirect highlighted feedback (no codes) (4) Indirect marginal feedback	1 year (34.5 contact hours)	No difference
Ferris and Helt (2000)	92 ESL learners USA university	Mix of direct, indirect (coded and uncoded); notes (marginal and end-of-text); text revision	1 semester	Indirect
Chandler (2003)	31 ESL learners Hong Kong	(1) Direct and indirect underlining (2) Error description and indirect underlining	1 semester	Direct

Given the conflicting nature of these findings, it is worth noting the findings of oral corrective feedback in SLA research—a body of research that has tended to be more carefully designed and more focused on specific linguistic categories. Studies by Carroll and Swain (1993), Ellis (1998), and Ellis, Loewen, and Erlam (2006) have reported a significant advantage in L2 production tasks for direct feedback over indirect feedback. This said, it should also be acknowledged that there may be salient differences between SLA work in oral feedback and written feedback in second language writing studies.

Another group of studies has investigated the effectiveness of different types of indirect feedback (coded and uncoded). None of these studies (Ferris & Roberts, 2001; Ferris et al., 2000; Robb et al., 1986) found any difference between coded and uncoded options. However, only the study by Robb et al. (1986) examined the effect of corrective feedback on new pieces of writing over time. The other two studies only measured the effect of corrective feedback on text revisions.

Apart from one recent study (Bitchener et al., 2005), scarcely any attention has been given to an investigation of the effect of different types of direct feedback on accuracy improvement. Bitchener et al. (2005) compared the effect of direct correction only with direct correction and meta-linguistic explanation (both written and oral) and found that learners who received the three forms of direct feedback significantly outperformed those who only received direct feedback, indicating that the addition of meta-linguistic explanation makes a difference to the reduction of error ratios. From this rather limited research base and its conflicting findings on the relative merits of direct and indirect feedback options, firm conclusions will only become available if further research, incorporating both types within the design of a single study, is carried out.

Part two: design issues

This review of the published literature not only reveals a need for an on-going examination of the relative merits of (1) providing or not providing ESL student writers with corrective feedback

and (2) providing direct or indirect feedback, but also reveals the need for research that manages to eliminate the design and analytical flaws highlighted in Section 2.1 of this review. Following those that have already been identified by Ferris (2004), four in particular are critical: (1) a control group that does not receive corrective feedback must be compared with treatment groups; (2) a longitudinal measurement of accuracy improvement in new pieces of writing by means of pre-test and post-test comparisons; (3) instruments that are valid measures of progress; and (4) an intensive targeting of one or only a few error categories at a time. Each of these issues is discussed below and each is addressed in the design of the present study (see Section 3.1 following).

Researchers are in complete agreement with the need for studies that include a true control group if the efficacy of corrective feedback is to be effectively addressed (Ferris, 2004, 2006; Truscott, 2004). As Ferris (2004, 2006) points out, it has been the ethical concern with not providing some students with corrective feedback while others receive corrective feedback that has been one of the reasons this crucial design issue has been ignored in much of the earlier research. She offers two suggestions about how this issue might be addressed. The first is to provide one group of students (the “control” group) with “summary end notes” on their errors while another group (the “experimental” group) receives in-text corrections. The second suggestion, involving a case study approach where the progress of student volunteers receive different treatments, would to some extent eliminate this problem, but the difficulty of convincing sufficient participants to take part in a study over time and in their own personal time might create another difficulty—an insufficient sample size. In settings where students are studying English over a number of semesters, another approach might be to focus on the targeted error category with one class during one semester and with another class in another semester after the data collection has been completed.

The second and third issues concern the measurement of accuracy in new pieces of writing. In order to measure the effectiveness of corrective feedback, an immediate post-test needs to be administered so that the effect of other variables between the treatment and a new piece of writing can be eliminated. If a post-test is to be a valid measurement of progress, a comparable pre-test needs to be included in the research design. Avoiding the use of different genres (for example, journal entries in one test and argumentative essays in another) will enable valid text comparisons to be made. To measure retention over time, delayed post-tests need to be incorporated into the design. In doing so, it is not possible to eliminate the effect of intervening variables between an immediate post-test and a delayed post-test (for example, students may access further instruction or undergo further practice outside of class time), but a delayed post-test can be used to measure the level of retention that was observed in an immediate post-test.

The fourth issue to be addressed is the number of error categories that are targeted. Virtually all of the published research (except Bitchener et al., 2005; Sheen, 2006) has provided corrective feedback on 15 or more error categories. Although studies by Ferris and colleagues (referred to above) have reduced theirs to five broad categories, the effectiveness of this approach is nevertheless questionable. If categories are too broad, it is not possible to determine exactly where an error lies. Furthermore, it has been suggested (Schwartz, 1993; Truscott, 1999) that different domains of linguistic knowledge (and therefore different linguistic error categories) are acquired in different ways. Assuming this is the case, it is even more important that error categories not be too broadly constituted. An intensive targeting of one or only a few error categories makes further sense when one considers the difficulty that ESL learners experience in

trying to cope with information overload. Instead of comparing outcomes across a range of grammatical forms and structures, oral corrective feedback studies in SLA research (Doughty & Varela, 1998; Han, 2002; Iwashita, 2003; Long, Inagaki, & Ortega, 1998; Lyster, 2004; Mackey & Philp, 1998) have reported reliable results from studies that have intensively targeted a single linguistic feature. On-going research into written corrective feedback would do well to emulate these examples. Such was the motivation for the following study.

The study

Design

This study was designed to answer three research questions:

- Does accuracy in the use of two functions of the English article system vary over time?
- Does accuracy in the use of these features vary according to the type of corrective feedback provided?
- Does accuracy in the use of these features vary as a result of the interaction of feedback type and time?

Accuracy in the use of two functional uses of the English article system was measured over a semester by means of a pre-test, a post-test immediately after the administering of the treatment (corrective feedback), and a delayed post-test after 2 months. Four groups of low intermediate ESL students took part in the study: group one received direct corrective feedback above each targeted error as well as written and oral meta-linguistic explanation; group two received direct corrective feedback above each targeted error and written meta-linguistic explanation; group three received direct corrective feedback above each targeted error; and group four was the control group. Each 30-minute piece of writing required the participants to describe what was happening in a different picture.

Participants

The study was conducted in two private language schools in Auckland, New Zealand. Four intact low intermediate classes from each school ($n = 75$) took part. Both schools followed the same process when assigning students to classes: a standardized grammar test, a writing test, and a one-on-one interview for students new to the school or an achievement test for students who had moved from one proficiency level to another. The schools describe their approach to the teaching of English as communicative, with an equal focus on reading, writing, speaking, and listening. All students were international visa students studying English in New Zealand for a specified time. All students had spent less than 6 months in New Zealand. Depending on whether they enroll full-time or part-time, students in both schools receive 3–5 hours of English language instruction 5 days a week. Of the 33 male and 42 female students, most were of East Asian origin: 27% from Korea, 26% from Japan, 10% from PR China, and others (3–7%) from countries such as Taiwan, Thailand, Vietnam, Russia, Switzerland, Saudi Arabia, Chile, and Brazil. The average age of the students was 22.7 years. The majority of the students (88%) had received formal classroom instruction in the English language for an average of 8 years. The classes were randomly assigned to one of the four groups (group one = 17 students; group two = 18 students; group three = 20 students; group four = 20 students).

Target structures

Two functional uses of the English article system were chosen as the target structures: the referential indefinite article “a” when referring to something for the first time (first mention) and the referential definite article “the” when referring to something that has been mentioned before (subsequent mentions). Other functional uses of the definite and indefinite articles were not targeted in the study.

These functions were targeted because, as a growing literature reports, learners across English language proficiency levels experience difficulty in the use of the English article system (Bitchener et al., 2005; Butler, 2002; Ferris, 2002, 2006; Master, 1995). Difficulties occur when deciding whether the definite or indefinite article should be used and whether, in fact, either article is required. Accuracy in the use of these functions in the pre-test revealed a mean score of 57.73%, thereby indicating a partial mastery of the functions.

Compared with earlier studies on the value of written corrective feedback for improving the accuracy of ESL student writing (see Ferris, 2002, 2003, 2006), where sometimes as many as 15 or more linguistic forms and structures were investigated, this study has heeded the example of effective SLA studies on oral corrective feedback and investigated the effect of targeting only two potentially “treatable” (Ferris, 2002, 2003; Truscott, 1996) error categories.

Treatment

The treatment that was provided was corrective feedback (error correction) on errors that had been made in two functional uses of the English article system. The type of corrective feedback provided varied according to the group/class in which the students were placed. Groups one, two, and three received corrective feedback on their errors, but group four did not receive this feedback because it was the control group. For ethical reasons, they were provided with a single, brief, general comment such as “clear outline of what is happening in the picture” or “well organized piece of writing”.

Group one received direct error correction above each targeted error category, as well as written and oral meta-linguistic explanation. The written meta-linguistic explanation included a simple explanation of the two targeted functional uses of the definite and indefinite articles together with an example of their use. Attached to their pre-test pieces of writing, the students received the following explanation and illustration:

1. Use “a” when referring to something for the first time.
2. Use “the” when referring to something that has already been mentioned.

Example

A man and **a** woman were sitting opposite me. **The** man was British but I think **the** woman was Australian.

Oral meta-linguistic explanation took the form of a 30-minute mini-lesson. During this lesson, the researcher explained the rules and example that the students had received on their returned texts. Additional examples were illustrated on the whiteboard and discussed with the class. The students were then given a short “controlled practice” exercise (see Appendix A) and asked to fill

the gap in each sentence with “a” or “the”. The students were given 5 minutes to complete the exercise. The lesson concluded with a plenary discussion of the answers.

Group two received direct error correction above each targeted error category and written meta-linguistic explanation. Group three only received direct error correction above each targeted error category. Group four did not receive feedback on the targeted features.

The written and oral meta-linguistic feedback was only provided in the treatment session that took place 2 weeks after the pre-test and on the same day as the immediate post-test. Direct error correction above the targeted error categories was provided on all pieces of writing (pre-test, immediate post-test and delayed post-test) for students in groups one, two and three. Group four never received corrective feedback on the targeted error categories.

Instruments

Each of the three tests required students to describe what was happening in a given picture. Each picture was of a setting where a wide range of people were doing various activities. Picture one for the pre-test was about a beach scene, picture two for the immediate post-test was about a park picnic, and picture three for the delayed post-test was about a camping site. Picture descriptions were chosen because the range of people, objects, and activities illustrated would predispose the students to using the English article system. However, it was acknowledged that students would be able to avoid such uses if they were uncertain about which use was appropriate and choose other determiners such as “one”, “two”, “this”, and “that”. Across the three writing tasks, no student made fewer than six uses of “a” and “the”.

Because the students were at a low intermediate level of proficiency, some of the key vocabulary items (concrete nouns) were provided around the margins of each task with arrows pointing to the relevant person, object, or activity. It was decided that this would lower the anxiety level for the students if unknown words were provided. Additionally, they were allowed to ask the researcher for a particular word if necessary. The students were given 30 minutes to complete their writing on each occasion.

Procedures

One week prior to the pre-test, and in accordance with the requirements of the university’s ethics committee, the students and the teachers (in separate sessions) in the four classes at both schools were provided with information sheets about the study and were given the opportunity to ask questions before signing a consent form. Because of the amount of marking required for each class, the data collection at the first school took place 1 month before that at the second school. The data were collected by the researcher (an experienced teacher of students at this proficiency level) to ensure consistency of procedure and input across the two schools. Before the data were collected, the teachers were briefed on the procedures and input to be provided and agreed that they were in keeping with the approach taken by them when students do writing tasks. Thus, it was agreed that the approach taken in the research would be familiar to the students.

On day 1, the pre-test was administered. Two weeks later, the treatment (corrective feedback) was provided. This involved the researcher visiting each of the classes and, for groups one, two, and three, returning the students writing and asking them to look at the corrections for approximately 5 minutes. Group one was then given a 30-minute lesson (oral meta-linguistic explanation). Immediately after the lesson, the students were asked to do a second piece of

writing (immediate post-test). For group two, after looking over the corrections that had been made on their texts and reading the attached written meta-linguistic explanation, the students were asked to do their second piece of writing. For group three, the second piece of writing was done 5 minutes after the students had looked over the corrections that had been made to their errors on the pre-test. For group four, the second writing task took place as soon as the pre-test writing had been returned.

The second piece of writing for all groups was returned 1 week after it had been written. Direct error corrections were again written above the errors made by students in groups one, two, and three but no written or oral meta-linguistic explanation was given to groups two and three. It was decided that feedback should be provided as soon as possible after the students had completed this piece of writing while the task was still relatively fresh in their memories so that they might be more motivated to take note of the feedback provided.

The delayed post-test was administered 2 months later. The students were not told when the researcher would be returning to conduct the delayed post-test. The reason for this was to eliminate the possibility of any student studying the feedback that they had been given on the pre-test and immediate post-test. The researcher did not want the students to be primed in any way beforehand. The third piece of writing was returned to the students 1 week later. Again groups one, two, and three received only direct corrections above their errors.

Analysis

Obligatory uses of the targeted features were identified and corrected for each text written by each student in the four groups on each of the three testing occasions. For the texts of students in group four, the control group, this was done on a photocopy of each script so that students in this group did not receive the targeted feedback. Error identification and correction was carried out by the researcher. A trained teacher/researcher in the primary researcher's institution did an interrater reliability check on this analysis. The initial rate of agreement was 91%. Collaborative analysis of the other 9% of instances resulted in 100% agreement.

Accuracy on each occasion was calculated as a percentage of correct usage for each script given the range of obligatory occasions arising in each script. For example, in any one script, three correct uses of the targeted features from 10 obligatory occasions meant a 30% accuracy rate. Group means and standard deviations were then calculated for each feedback group on each of the three testing occasions. Tests of statistical significance were carried out by means of mixed design ANOVAs.

Results and discussion

The descriptive statistics for the accuracy scores of each group on each of the three test occasions are presented in [Table 4](#) and are illustrated in [Fig. 1](#).

The first research question investigated the effect of targeted corrective feedback on three pieces of writing at different times by means of an immediate post-test and a delayed post-test. Because the sphericity assumption on Mauchly's Test of Sphericity was violated, *F* values were corrected using the Greenhouse–Geisser adjustment. The ANOVA test revealed a significant difference in accuracy scores across the three writing tests ($F = 60.028$; *d.f.* = 2; $p = .000$). Pairwise comparisons further revealed significant differences between the pre-test and the immediate post-test scores ($p = .000$), between the pre-test and delayed post-test scores ($p = .000$) but no difference between the immediate post-test and delayed post-test scores

Table 4
Descriptive statistics: accuracy by group and time

Groups	Pre-test		Immediate P-T		Delayed P-T	
	<i>M</i>	S.D.	<i>M</i>	S.D.	<i>M</i>	S.D.
1. CF, W, O	67.59	19.26	82.53	14.33	82.65	15.47
2. CF, W	53.11	21.73	77.50	16.07	76.78	20.67
3. CF	59.35	18.39	81.25	13.00	80.05	12.10
4. Control	51.90	28.32	52.75	23.27	63.90	18.90

P-T (post-test); CF (corrective feedback); W (written meta-linguistic explanation); O (oral meta-linguistic explanation).

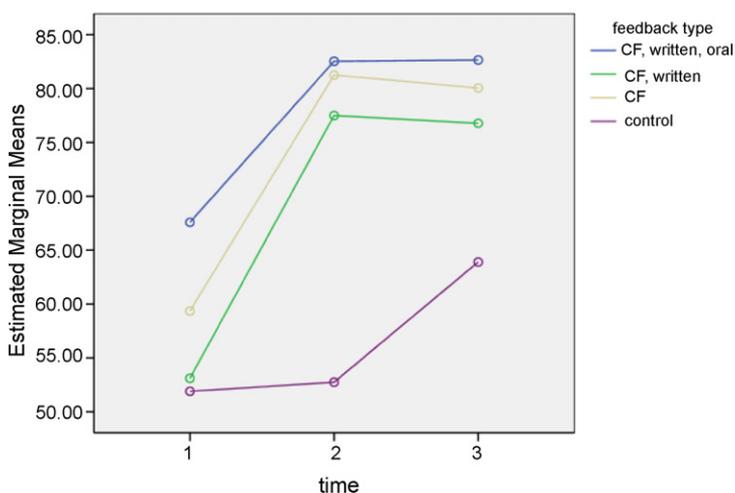


Fig. 1. Accuracy by group and time.

($p = .229$). There was a significant improvement in accuracy immediately after the treatment had been provided, and this level of accuracy was retained in the third piece of writing.

This will come as a pleasing result for researchers and teachers who, despite Truscott's (1996, 1999, 2004) claims that corrective feedback should be abolished because it is ineffective, have assumed that it is facilitative of improved written accuracy and have believed that what they see as improvements in their students' writing is worth the time and effort. Having said this, it is nevertheless important to realise that a finding such as this was not the result of an unfocused, unsystematic, random treatment of diverse error difficulties. It was the result of a moderately intensive, targeted focus on two functional uses of one problematic linguistic domain for ESL learners. It should also be acknowledged that the participants in the study (international students from predominantly East Asian countries where English is most often studied as a foreign language in formal instructional settings and the focus is usually form- and structure-based as opposed to competency-based) are likely to have had some earlier instruction in the use of the targeted functions, but that their mastery or acquisition was still being established. Thus, further research is needed to determine the extent to which corrective feedback helps learners develop accuracy in the use of completely new linguistic forms and structures.

The findings of this question not only indicate the immediate effect of written corrective feedback on a new piece of writing but also the extent to which the level of accuracy was retained

over a 2-month period without additional corrective feedback and classroom instruction. While performance on the delayed post-test reveals the learners' level of retention, it does not mean that accuracy in this test was necessarily and only the result of the treatment provided on the day of the immediate post-test. In any longitudinal study, it is not possible to control for the effect of intervening variables such as additional instruction that may have been received outside of class time or additional self-study engaged in by highly motivated students. Theoretically, further research could investigate whether or not students receive such input by means of a self-report questionnaire or interview.

An important contribution of this finding to the existing evidence in support of corrective feedback (Ashwell, 2000; Fathman & Whalley, 1990; Ferris & Roberts, 2001) is its measurement of accuracy on new pieces of writing rather than on the revision of single texts. This marks a step forward because, as Truscott (1996, 1999, 2004) and Ferris (1999, 2004) point out, the effectiveness of written corrective feedback can only be assessed when accuracy is measured on new texts. It is interesting to note that this finding is also supported by a recent SLA study (Ellis et al., 2006) of the effectiveness of oral corrective feedback. Ellis et al. investigated the effect of two types of corrective feedback on the acquisition of past tense -ed by low intermediate ESL students. One group received implicit corrective feedback in the form of recasts, a second group received explicit corrective feedback in the form of meta-linguistic explanation, and a third group, acting as the control group, received no corrective feedback. Two post-test scores revealed a clear advantage for students who received corrective feedback.

The second research question investigated whether or not there was a differential effect on accuracy for the different treatment options. The ANOVA test revealed a statistically significant difference in accuracy scores for the four groups ($F = 5.779$; $d.f. = 3$; $p = .001$). Pairwise comparisons further revealed that the difference lay between the control group score and those of treatment groups one ($p = .002$) and three ($p = .011$) but not group two ($p = .127$). As Fig. 1 illustrates, participants in group one who received direct corrective feedback as well as written and oral meta-linguistic explanation and those in group three who received direct corrective feedback and no meta-linguistic explanation outperformed the control group who did not receive corrective feedback.

Following Bitchener et al. (2005) and Sheen (2006), it was not surprising to find that students who received direct corrective feedback as well as written and oral meta-linguistic explanation outperformed those who received no corrective feedback. What was interesting and is therefore worthy of some discussion is the fact that the accuracy of group two (direct corrective feedback and written meta-linguistic explanation) was lower than that of group three (direct corrective feedback and no meta-linguistic explanation). As one considers the pre-test scores for the four groups, one might be inclined to suggest that this finding is not surprising because the mean performance of participants in group three was higher than that of group four (control group). However, a post hoc ANOVA revealed that the differences between the pre-test scores of the four groups were not statistically significant ($p = .148$). In other words, they can be considered equal. Because the targeted feature was unlikely to have been a new linguistic form, it is possible that the limited detail and the single provision of written meta-linguistic explanation may not have been sufficient for it to have had a significant effect.

In two respects, this finding makes an important contribution to earlier work. First, it corroborates the finding of Bitchener et al. (2005) who found that the addition of written and oral meta-linguistic explanation to direct corrective feedback significantly helps learners improve the accuracy of their writing. Second, it demonstrates that oral meta-linguistic explanation in the form of a clearly focused mini-lesson (30 minutes) may be as effective as the more

time-consuming one-on-one conferences that were included as the oral meta-linguistic explanation in Bitchener et al. (2005). However, further research is needed to see if there is a difference in the effectiveness of the two types of oral meta-linguistic explanation and in the separation of oral and written meta-linguistic explanation and direct error correction.

The third research question investigated the interactional effect of feedback type and test time. The ANOVA test revealed a significant effect for this interaction ($F = 4.828$; $d.f. = 5$; $p = .001$). The test of within-subjects contrasts further revealed a significant difference for their interaction between the immediate post-test and delayed post-test ($F = 5.441$; $d.f. = 3$; $p = .002$) but not between the pre-test and delayed post-test ($F = 1.703$; $d.f. = 3$; $p = .174$). As Table 4 reveals, for the delayed post-test (time 3), there was a significant difference in accuracy between the control group and groups one and three but no difference between the control group and group two. In other words, there was still a significant difference in accuracy between learners who received direct corrective feedback as well as written and oral meta-linguistic explanation and those received only explicit error correction.

Even though the performance of the control group improved in the delayed post-test, it was significantly different to that of groups one and three. One explanation for this improvement might be that some members of the group sought input on the targeted feature during the weeks between the immediate and delayed post-tests. It is always possible that students from the feedback groups may have passed on information about what they were receiving feedback on or that students in the control group sought instruction from out-of-class sources. Studies with additional delayed post-tests would be needed to see whether variations in control group scores are retained or whether they are indicative of an idiosyncratic performance on a single test.

Conclusion

The aim of this study was to investigate whether targeted corrective feedback on ESL student writing results in improved accuracy in new pieces of writing over a 2-month period and to see whether there is a differential effect on accuracy for different corrective feedback options. In doing so, the study endeavoured to address a number of design issues that have prevented us from obtaining clear answers to these issues. Most importantly, the study found that written corrective feedback had a significant effect on improving accuracy in the use of two functional uses of the English article system (the use of “a” for first mention and “the” for subsequent mentions) and that this level of accuracy was retained 2 months later without additional feedback or instruction. Compared with the majority of earlier studies that had measured the effectiveness of corrective feedback on text revisions, this study demonstrates its effectiveness on new pieces of writing. The study also found that students who received direct corrective feedback on the targeted features as well as written and oral meta-linguistic explanation (group one) and those who received direct corrective feedback but no meta-linguistic feedback (group three) outperformed the control group (group four) who did not receive corrective feedback.

Although every attempt was made to eliminate design and analytical flaws, there were inevitably a few minor limitations, and these should be taken into account when further research is designed. On the one hand, the focus on only two functional uses of the English article system may be seen as both a limitation and a strength. In adopting this focus, it has been possible to measure the accuracy of a specific problematic linguistic area for many ESL students. However, further research is needed to investigate the extent to which the positive findings of the study apply to other linguistic error categories. Because the focus of this study was on partially learned linguistic features, further research is also needed to examine the effectiveness of corrective

feedback in helping students develop accuracy in the use of completely new linguistic forms and structures. Further research might also find it useful to investigate whether instrument differences, for example, different genre tasks, have similar or different effects on accuracy. The population focus (mainly East Asian students; all international students with an EFL background; all at a low intermediate level of proficiency) in further research could also be extended to include students from other L1 and ethnic backgrounds (international and migrant) and other proficiency levels of English. This study measured accuracy retention over 2 months, but further research would do well to extend this scope to include several additional post-tests over a longer period of time (in preparation) so that the ultimate value of written corrective feedback for acquisition can be determined.

This study investigated the effect of direct corrective feedback options and, in doing so, has filled another gap in the research, but further studies are now needed to examine the relative merits of different types and combinations of direct and indirect feedback (in preparation). While there is a need for further research in these areas, there are a number of pedagogical applications that can be made from the findings of this study.

The study has demonstrated that significant improvements in accuracy can result from the provision of written corrective feedback on errors that are made in the use of the referential indefinite article “a” (first mention) and the referential definite article “the” (subsequent mentions). It has also shown that a focused approach to the treatment of recurrent linguistic errors does not have to involve extensive amounts of class time. A one-on-one conference will take more class time than small group conferences which, in turn, will take more time than plenary mini-lessons. Therefore, teachers may find it helpful to provide mini-lessons to all students on a small range of recurrent error categories and follow these up with small group meta-linguistic sessions on particular error categories with those students having the most difficulty with a particular form or structure. One-on-one conferences could then be offered to those requiring more attention. Additionally, teachers would do well to negotiate with their students the features that will be focused on so that they buy into this targeted focus. Several students who participated in this study said informally, after they had completed the final piece of writing, that they thought the single targeted approach was more helpful than the unfocused approach that is often adopted by classroom teachers and that they would recommend it be followed by other teachers. Some teachers and students might consider a targeted approach over time an indulgence but, if it is adopted for a period of time that is agreed upon by both teachers and students, for example 6–8 weeks or until noticeable improvements in accuracy are evident, it does not have to dominate the students’ learning programme. Further research is needed, of course, to confirm the merits of these approaches.

Acknowledgements

The author would like to thank the editors and their reviewers for a number of insightful comments on an earlier draft of this article.

References

- Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227–257.
- 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, 191–205.

- Butler, Y. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, 24, 451–480.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357–386.
- Chandler, J. (2000, March). The efficacy of error correction for improvement in the accuracy of L2 student writing. *Paper presented at the AAAL conference*.
- 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, 267–296.
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York: Cambridge University Press.
- Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Ellis, R. (1998). Teaching and research: Options in grammar teaching. *TESOL Quarterly*, 32, 39–60.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.
- 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.
- Ferris, D. R. (1995). Can advanced ESL students be taught to correct their most serious and frequent errors? *CATESOL Journal*, 8, 41–62.
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, 31, 315–339.
- Ferris, D. R. (1999). The case for grammar correction in L2 writing classes. A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1–11.
- 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.), *Perspectives on response*. Cambridge: Cambridge University Press.
- Ferris, D. R., Chaney, S. J., Komura, K., Roberts, B. J., & McKee, S. (2000, March 14–18). Perspectives, problems, and practices in treating written error. *Colloquium presented at International TESOL Convention*.
- Ferris, D. R., & Helt, M. (2000, March 11–14). Was Truscott right? New evidence on the effects of error correction in L2 writing classes. *Paper presented at the American Association of Applied Linguistics Conference*.
- 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.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543–572.
- Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25, 1–36.
- James, C. (1998). *Errors in language learning and use*. White Plains, NY: Addison-Wesley Longman.
- Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, 75, 305–313.
- Komura, K. (1999). *Student response to error correction in ESL classrooms*. Master's thesis, California State University, Sacramento.
- Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, 66, 140–149.
- Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals*, 24, 203–218.
- Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82, 357–371.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Mackey, A., & Philp, J. (1998). Conversational interaction in second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338–356.
- Master, P. (1995). Consciousness raising and article pedagogy. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy*. Norwood, NJ: Ablex Publishing Corporation.

- 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.
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20, 83–93.
- Roberts, B. J. (1999). *Can error logs raise more than consciousness? The effects of error logs and grammar feedback on ESL students' final drafts*. Master's thesis, California State University, Sacramento.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition*, 15, 147–163.
- Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, 17, 195–202.
- Sheen, Y. (2006). *Corrective feedback, individual differences, and the acquisition of English articles by second language learners*. Unpublished PhD thesis, University of Nottingham, United Kingdom.
- Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, 23, 103–110.
- 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(2), 111–122.
- Truscott, J. (2004). Dialogue: Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, 13(4), 337–343.

Appendix A Controlled practice exercise

Fill in the gaps with **a**, **an**, or **the**.

1. This morning I bought Newspaper and magazine. newspaper is in my bag but I don't know where I put magazine.
2. I saw accident this morning. car crashed into tree. Driver of Car wasn't hurt but car was badly damaged.
3. There are two cars parked outside: blue one and grey one. blue one belongs to my neighbours; I don't know who owner of grey one is.
4. My friends live in old house in small village. There is beautiful garden behind house. I would like to have garden like that.