The effectiveness of providing second language (L2) writers with on-line written corrective feedback

Professor John Bitchener, AUT University
Dr Martin East, University of Auckland
Helen Cartner, AUT University

April 2010

AOTEAROA
NATIONAL CENTRE FOR TERTIARY TEACHING EXCELLENCE

An Ako Aotearoa publication. Support for this work was provided by Ako Aotearoa through its Regional Hub Project Funding scheme

Abstract

The value of written corrective feedback for second language development is controversial (Ferris, 1999; Truscott 1996, 1999). The only way to determine whether or not it is effective is to empirically investigate its effect when learners write new texts over time. In recent years, a growing number of studies have investigated the effectiveness of providing learners with targeted feedback on certain linguistic error categories (e.g. past tense, article use) and with different types of direct and indirect feedback. While positive findings have emerged for the tested error categories, additional research is required to examine the extent to which feedback can help learners improve the accuracy with which they use other linguistic categories. The available research on the effectiveness of different types of feedback has been inconclusive for a variety of reasons including poor research design and analysis. This pilot project investigated the effectiveness of providing advanced learners with feedback on their two most frequent error categories (singular and plural nouns and subject-verb agreement). Some students received direct error correction while others received indirect coded meta-linguistic feedback. This paper reports the key findings of the study and includes recommendations for classroom application and further research.

The aims of the study

The aim of the study was to investigate the extent to which second language writers at a pre-degree level benefit from written corrective feedback on their writing when delivered on-line. In doing so, the study examined the effectiveness of two different types of feedback (direct error correction in blog entry - the correct version was provided beside each targeted error - and a meta-linguistic code below the blog entry e.g. DA for definite article error) and the level of retention over time (five weeks after the immediate post-test text had been written).

The background and rationale of the study

There are two key reasons that a focus on accuracy is important for international and migrant writers of English as a second language. First, it has been well documented that individuals who reveal shortcomings in their written expression may not only be stigmatized as 'outsiders' but also treated unequally when seeking employment opportunities in areas they may have been trained or educated in before coming to New Zealand. As the TEC (Tertiary Education Commission) Education Strategy (2007-2012) from the Ministry for Tertiary Education explains, success for all New Zealanders is vitally important so that 'they can contribute fully to our economy and society' (p.21). To this end, the report adds that 'tertiary educational organisations need to address the disparities that exist for populations such as .... migrants and refugees ....' (p.21). It identifies the need for strong foundation skills, 'especially

literacy, numeracy and language' because 'a lack of literacy, numeracy and language skills in the workforce impedes productivity and will, in the long run, impede economic growth'(p.22). Another reason for focusing on accuracy is the number of students choosing to complete English language programmes of study before applying for or being admitted to degree level programmes because their written work reveals shortcomings that may limit the progress they make at higher levels of tertiary study.

Research in recent years (e.g. Bitchener, 2008; Bitchener & Knoch, 2008, 2009; Ellis et al, 2008; Ferris, 2006; Sheen, 2007) has reported that learners can improve the accuracy of their written expression in specifically targeted areas but debate continues about the relative effectiveness of different types of feedback over time for more advanced L2 writers for whom some linguistic forms and structures may have begun to fossilize. Additionally, it is also unknown what effect written corrective feedback can have when it is provided on-line rather than off-line.

Much of the early research on written corrective feedback focused on its value for helping L2 writers revise draft texts accurately (see Ferris, 2004). While the findings from this literature are instructive for language and writing teachers, they do not reveal whether or not written corrective feedback can also play a role in language development or language acquisition, measured by the level of retention revealed by writers when new texts are written and assessed over time. More recently, though, this key question has begun to be investigated. The studies referred to above have all revealed that written corrective feedback has the potential to effect change in written accuracy when certain linguistic forms and structures are targeted (e.g. past simple tense, some functional uses of the English article system) but to date research has not been carried out to examine the extent to which written corrective feedback can also treat other linguistic error categories. As Truscott (1996, 1999) and others have noted, it is important to find out whether or not it is limited to the treatment of only certain linguistic errors. Thus, this study sought to investigate whether or not written corrective feedback could also be effective in targeting other problematic error categories in the texts of relatively advanced L2 writers about to enter the first year of university degree study.

Traditionally, teachers of L2 writers have provided handwritten off-line feedback and commentary on written texts (Ferris, 2006). More recently, track change feature has provided teachers with an additional means of response. Students are now used to working and thinking on-line and so are well prepared for using computer and internet options for further study and for communicating with English-speaking communities in the workplace and in society generally. This study was therefore designed to investigate the effectiveness of providing L2 writers with on-line corrective feedback on their computer-written texts.

Most of the early research on written corrective feedback was unfocused in the sense that a wide range of error categories (that were equally wide and loose in definition) were investigated. This meant that it was difficult to know the exact cause of a learner's difficulty (see Bitchener, 2008). More recently, written corrective feedback research has followed the example of oral corrective feedback research and focused its investigations on specifically targeted linguistic error categories. Most often the categories chosen have been reported in the literature as areas of general concern for teachers and learners. This approach has meant that the specific needs of learners identified in the types of texts they are asked to write at the time of investigation have usually been ignored. To address this shortcoming, this study identified all the linguistic errors that were made in the first written text of the participants and from this analysis an investigation was made of the effectiveness of written corrective feedback for treating the two most frequently occurring linguistic error categories. This analysis found that a large number of participants made errors in the use of singular and plural nouns and in subject-verb agreement. For this reason, the study focused on the group performance of those making these types of error by measuring the extent to which written corrective feedback helped these learners improve their use of these forms/structures in the writing of two new texts.

The research questions

1. Does targeted written CF help L2 writers reduce the percentage of errors they make when using singular and plural nouns in new pieces of writing?

2. Is direct error correction more effective than coded meta-linguistic feedback for reducing the percentage of errors made in the use of singular and plural nouns in new pieces of writing?

3. Does targeted written CF help L2 writers reduce the percentage of errors they make when making subject-verb agreements in new pieces of writing?

4. Is direct error correction more effective than coded meta-linguistic feedback for reducing the percentage of errors made when making subject-verb agreements in new pieces of writing?

5. Does targeted written CF help L2 writers reduce the percentage of total errors (combining singular/plural noun and subject-verb agreement scores) in new pieces of writing?

6. Is direct error correction more effective than coded meta-linguistic feedback for reducing the percentage of total errors made in new pieces of writing?

Participants

The participants in this study (n = 20) were recruited from the advanced level programme in English as an additional language in a New Zealand tertiary institution. A slightly larger number of students were invited to take part in the study. The students were enrolled in a programme comprising four papers: Listening and Note-taking, Writing and Research Skills, Reading and Vocabulary Development and Oral Interaction and Presentation Skills. In the context of the programme students were required to write several short pieces on-line in response to a variety of listening stimuli. They were instructed that they would be making written contributions as part of an on-line 'corrective feedback blog' to which they would have the opportunity to contribute at various times throughout the programme. They were told that they would receive corrective feedback on their writing. Participants were recruited from this programme through an invitation, at the start of the course, to take part in the study.

Design

The study used a pre-test / post-test / delayed post-test design. On three occasions participants would be asked to complete a 30-minute written task on-line in response to a listening stimulus. The following stimulus sources were used:

Pre-test: Listening to lecture 1: The process of lecture comprehension (PP3-4) in Lebauer, R.S. (2000). *Learn to listen, listen to learn.* New York: Addison Wesley Longman.

Immediate post-test: Listening to public transport level 4 listening in
Kaufmann, H., & Westwood, V. (1996). *Issues in English.* Melbourne: Protea Textware.

Delayed post-test: Listening to the environment level 4 listening in
Kaufmann, H., & Westwood, V. (1996). *Issues in English.* Melbourne: Protea Textware.

Pre-test

The pre-test was conducted in week 3 of the course. Participants were asked to complete a short essay, in class and within 30 minutes, in response to the following task:

> What problems do you have when you listen to a lecture in English? Do you have any of these problems when you listen in your native language? In your opinion what are the sources of these problems?

Immediately after participants had completed the task, they were randomly assigned to two groups, each containing ten participants. It was planned that each group would be provided with a different type of on-line written corrective feedback.

Participants in Group 1 would receive direct error correction. That is, the error would be corrected, in red, in the blog entry. The correct form would be noted in brackets after the error. No explanation of the error would be provided. For example, a

student who had written “I think I have two main problem when i listen to the lecture” received the correction: “I think I have two main ~~problem~~ (problems) when i listen to the lecture”

Participants in Group 2 would receive direct feedback on the type of error they had made, but the error would not be corrected. Instead, participants would receive an abbreviated meta-linguistic code. An explanation of the code would be provided below the blog entry. For example: “Yes, I do have a lots (n/s) of problem (n/pl) in listening during class” code n/s = noun singular; n/pl= noun plural.

Two of the researchers read the blog entries. One researcher read and annotated the Group 1 essays, and the other worked with the Group 2 essays. All error categories were noted and were then ranked to determine the two most frequent error categories. These two error categories became the focus of the corrective feedback. The error categories that were identified from this exercise are listed in Table 1.

Table 1: Ranked error categories

Error category	Type of error
Noun plural/singular	Participants used a singular noun when it should have been plural, and vice versa
Subject-verb agreement	There was an error of agreement between subject and verb
Tense error	Participants used an inappropriate tense in the context
Infinitive verb	Participants failed to conjugate the verb correctly, and instead wrote an infinitive
Article error	Participants used an article where none was required, and vice versa or used an incorrect article
Collocations	Participants failed to use correct collocations
Wrong part of speech	Participants used, for example, an adverb where an adjective was required in the context
Prepositional error	An inappropriate preposition was used

Immediate post-test

Feedback on the pre-test was provided to the participants a week after completing the first task (week 4). The participants were asked, in class, to look at the errors on which they had received feedback. Participants in Group 1 were asked to read and take note of the errors; those in Group 2 were asked to do the corrections beside each bracketed code. Participants then completed Task 2 (the immediate post-test), for which they had 30 minutes:

> What do you know about Auckland’s public transport? Do you use it? Why do so many people prefer to use cars?

After the participants had completed this task the data were analysed by the two researchers who noted, and gave feedback on, the two most frequent error categories as identified from Task 1. The researchers worked with the same Group as for Task 1. The feedback was made available to the participants a week later (week 5). On this occasion they were able to access and read the feedback, but were not required to produce another sample of writing.

Delayed post-test

In week 9 of the course, five teaching weeks after completing Task 2, participants were given a third task:

> Do you believe that reducing garbage is a good idea? What are some of the ways you can reduce your garbage?

A similar procedure was followed for Task 3 to the one that had been followed in Task 2. The researchers read the blog entries for one of the two groups (the same group as for Tasks 1 and 2). Errors were noted and recorded for the two most frequent error types. Subsequent to data collection, participants received written feedback on Task 3, and the frequencies of error across all three tasks were analysed.

Findings of the study

Research questions 1 and 2

Table 1 below reveals the descriptive statistics for research questions 1 and 2 - error rates in the use of singular and plural nouns by the two feedback groups (group 1 received direct error correction and group two received meta-linguistic feedback in the form of meta-linguistic coding) across the three testing times (pre-test, immediate post-test and delayed post-test).

Table 1: Error rates by group and time for noun plural/singular

Groups	Time 1		Time 2		Time 3	
	M	SD	M	SD	M	SD
1	17.33	8.234	7.33	7.76	5.33	7.53
2	13.25	4.5	2	1.41	4.75	3.78

Before analysing the results of each research question, we were interested in confirming whether or not the two group scores were similar or different. If they were significantly different, this would have meant that our two groups were on an unequal playing field. The one-way ANOVA test on the pre-test scores revealed in fact that there was no difference in the error rates of the two groups with respect to singular/plural noun usage: $F(1) = .800$, $p = .397$. In seeking an answer to research

question 1, we examined the error rates of the immediate and delayed post-test pieces of writing and found the difference to be statistically significant for group 1, the direct error correction group: F (2) = 4.07, p = .051. Significance was also found for group 2, the meta-linguistic feedback group: F (2) = 9.293, p = .015. As Figure 1 below shows, the difference occurs for both groups between the pre-test and the immediate post-test. This result provides clear evidence of the effectiveness of written corrective feedback on this targeted linguistic error category.

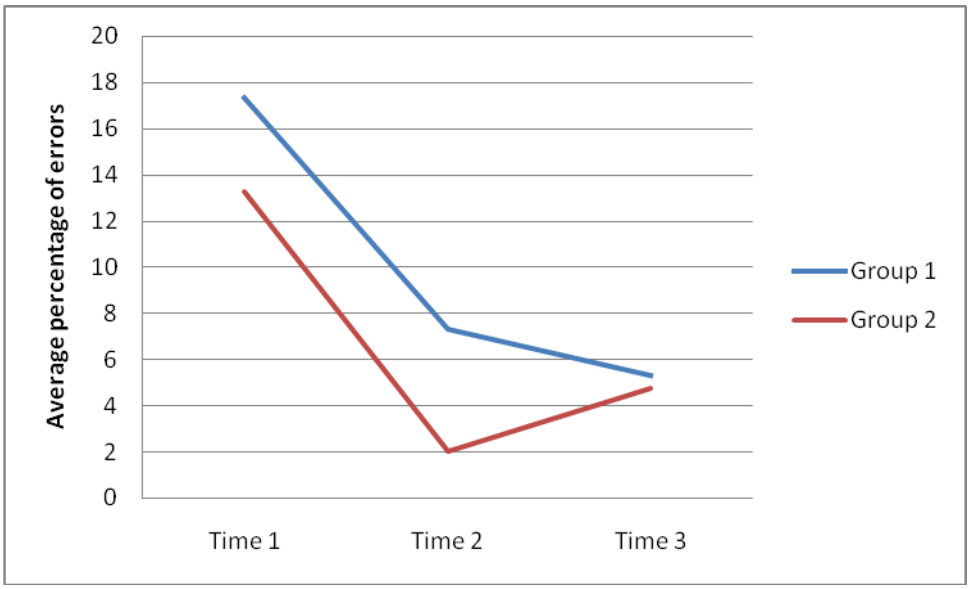

Figure 1: Error rates by group and time for singular/plural nouns

The second research question investigated whether or not there was a differential effect between groups 1 and 2. The repeated measures ANOVA test for group (treatment groups 1 and 2) x time (the 3 writing occasions) revealed that there was no significant difference in effect between the two treatment groups (df = 1, F = 1.91, p = .204). This can be seen in the visual representation of Figure 1 above.

Research questions 3 and 4

Table 2 below reveals the descriptive statistics for research questions 3 and 4 - error rates in the use of subject-verb agreement by the two feedback groups (group 1 received direct error correction and group two received meta-linguistic feedback in the form of meta-linguistic coding) across the three testing times (pre-test, immediate post-test and delayed post-test).

Table 2: Error rates by group and time for noun subject-verb agreement

Groups	Time 1	Time 2	Time 3

	M	SD	M	SD	M	SD
1	8.67	5.54	4.5	3.89	5.17	6.21
2	18	15.55	8.43	9.71	2.43	4.16

Before analysing the results for each research question, we were again interested in confirming whether or not the two group scores were similar or different. The one-way ANOVA test on the pre-test scores revealed in fact that there was no difference in the error rates of the two groups with respect to subject-verb agreement: F (1) = 1.931, p = .192. In seeking an answer to research question 3, we examined the error rates of the immediate and delayed post-test pieces of writing and found the difference to not be statistically significant for group 1, the direct error correction group: F (2) = .861, p = .452. However, there was a statistically significant difference across the three testing times for group 2, the meta-linguistic feedback group: F (2) = 6.879, p = .010. Figure 2 below shows clearly this reduction in error rate.

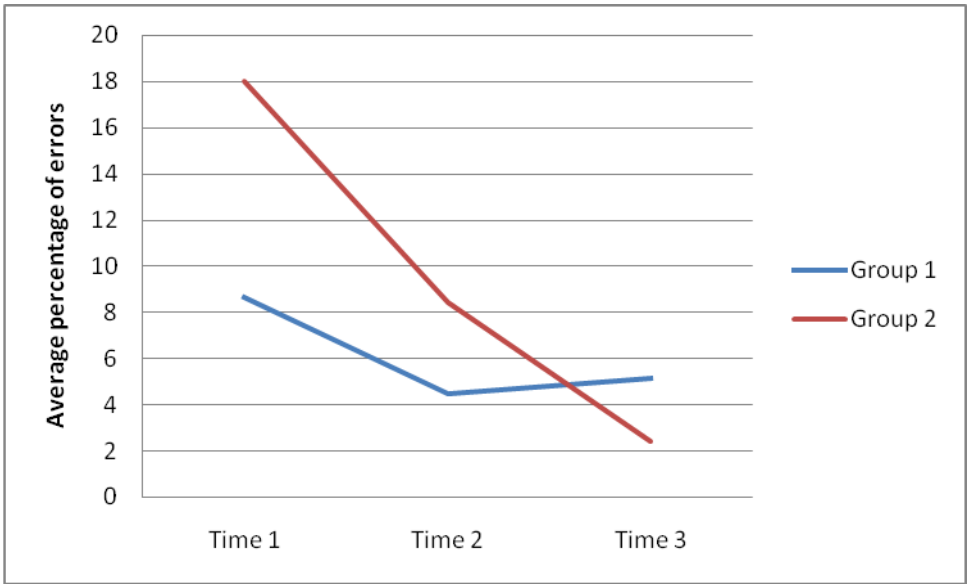

Figure 2: Error rates by group and time for subject-verb agreement

Thus, it can be concluded that the provision of meta-linguistic feedback in the form of coded feedback had a more significant effect than direct error correction on treating subject-verb agreements. It would seem therefore with this linguistic error category that learners benefitted from feedback about the nature of the error (revealed by means of a meta-linguistic code) more than they did from direct error correction.

Research question 4 investigated whether or not there was an effect between groups 1 and 2. The repeated measures ANOVA test revealed no difference between groups (df = 1, F = .912, p = .36) and this can be seen in Figure 2 above.

Research questions 5 and 6

Table 3 below reveals the descriptive statistics for research questions 5 and 6 – total error rates (combining both linguistic error categories) by the two feedback groups (group 1 received direct error correction and group two received meta-linguistic feedback in the form of meta-linguistic coding) across the three testing times (pre-test, immediate post-test and delayed post-test).

Table 3: Error rates by group and time for combined scores

Groups	Time 1		Time 2		Time 3	
	M	SD	M	SD	M	SD
1	15.6	7.26	7.1	8.67	6.3	9.21
2	22.38	15.06	8.38	8.98	4.5	4.07

Before analysing the results of each research question, we were again interested in confirming whether or not the two group scores were similar or different. The one-way ANOVA test on the pre-test scores revealed that there was no difference in the error rates of the two groups with respect to combining their total scores: $F(1) = 1.583$, $p = .226$. In seeking an answer to research question 5, we examined the error rates of the immediate and delayed post-test pieces of writing and found a statistically significant difference for group 1, the direct error correction group: $F(2) = 6.103$, $p = .009$. A statistically significant difference across the three testing times for group 2, the meta-linguistic feedback group, was also revealed: $F(2) = 9.944$, $p = .002$. Figure 3 reveals the reduced error rates for both groups between the pre-test and immediate post-test and as such further attests to the value of written corrective feedback over time.

Research question 6 investigated whether or not there was a differential effect between groups 1 and 2. The repeated measures ANOVA test revealed that there was no significant difference in effect between the two groups ($df = 1$, $F = .385$, $p = .544$). Again, Figure 3 reveals this absence of effect visually.

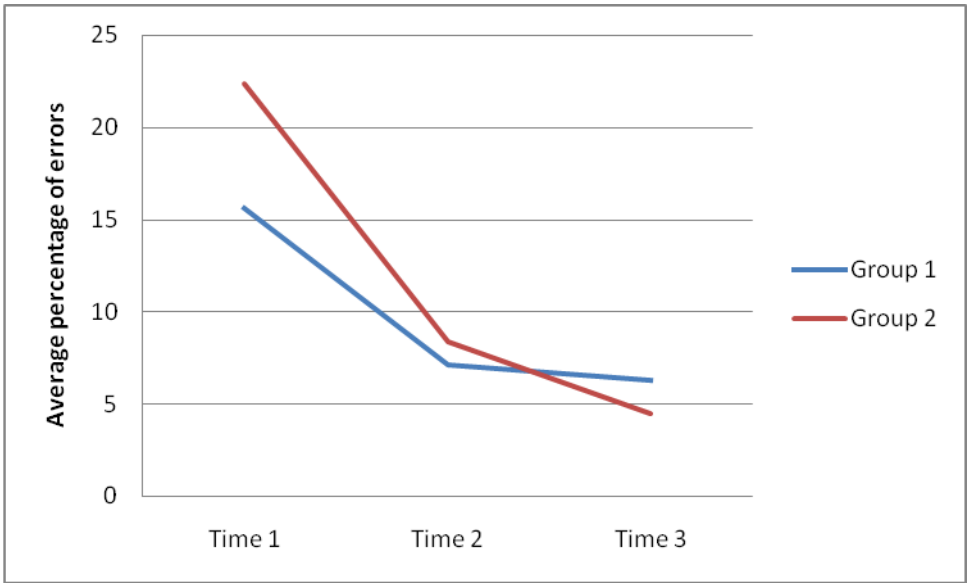

Figure 3: Error rates by group and time for combined scores

Summary of findings

RQ 1 and 2	Targeted written CF helped learners reduce their error rate in using singular/plural nouns over time but there was no difference in effect between the two types of feedback.
RQ 3 and 4	Only coded meta-linguistic feedback (group 2) helped learners reduce their error rate in subject-verb agreements over time.
RQ 5 and 6	Targeted written CF helped learners reduce their total error rate (combination of singular/plural noun and subject-verb usage) over time but there was no difference in effect between the two types of feedback.

Conclusions of the study

1. From the results of this study it would seem that written corrective feedback is effective in helping advanced L2 learners/writers improve the accuracy with which they use two relatively simple, partially acquired linguistic forms/structures - singular/plural nouns and subject-verb agreement. However, it needs to be realised that the small sample size of this pilot investigation may be considered a limitation. A larger sample may reveal additional findings of significance.
2. The study also shows that these benefits are maintained over time – in this case, over a 6 week period.

3. These findings add to the growing evidence (see Bitchener and Knoch, 2009 for a review of this research) in support of the effectiveness of written corrective feedback for treating partially acquired rule-based discrete linguistic forms/structures.

4. The study shows that written corrective feedback can be effectively provided on-line but further research, comparing on-line and off-line provision within a single study design, would be required to find out if one medium is more effective than the other.

5. Targeting two linguistic error categories rather than many categories and loosely defined categories is again shown (see Bitchener, 2008; Ellis et al, 2008) to be an effective way of treating linguistic error difficulties and helping writers retain the level of mastery and control demonstrated immediately after feedback has been provided.

6. The findings of this study are applicable to advanced proficiency L2 writers. Further research would be required to find out if the forms/structures targeted in this study can also be successfully treated with writers at other proficiency levels.

7. The findings of the study suggest little difference in effect between direct error correction and coded meta-linguistic feedback. The only difference between the two was found in their effect on subject-verb agreement. It may be that some forms/structures such as this are more effectively treated when the nature/cause of the error is revealed. On the other hand, in the case of singular/plural nouns, error correction was sufficient. That there is a difference in effect between the two types of feedback indicates that further research should be done to explore the extent to which explanation (i.e. nature/cause of error) is more helpful than unexplained correction.

Acknowledgements

The researchers are grateful to the Northern Regional Hub of Ako Aotearoa for its funding of this project and to the students who took part in the study.