Research into Corrective Written Feedback: An EFL Teacher's Perspective

Darby McGrath

Kwansei Gakuin University

This paper focuses on four important questions regarding written corrective feedback (CF), and to what extent the research can help us to answer these questions. The conclusion is that the research has yet to offer compelling reasons for writing teachers to adopt any particular approach to CF, beyond a broad indication that some CF is better than no CF, and that students are better served by feedback that identifies the type of error they have made than by less direct feedback. Beyond this, we know very little about what kinds of written feedback and how much feedback is most effective. There are two reasons for this. First, what researchers mean when they talk about corrective feedback is so varied that it is difficult to discern a consistent pattern of what approach to feedback works best. Furthermore, recent research has become so limited in scope that it is of little practical value to teachers and students. For research into CF to inform teaching practices in any meaningful way, these two factors will need to be addressed.

この論文では、添削に纏わる修正を促すフィードバック（CF）に関する4つの大きな疑問について、どの程度学術研究がこの疑問に答えてくれているのかについて論述しています。結論としては、学術研究は、これが最も有効なCFであるといったものを提示せず、ライティングの教師たちにCFをしないよりも、したほうが良い。また、学習者には間違えの場所だけを教えるよりも、どのような間違いをしているのかを提示するほうがいいといった広域な示唆を与える以外には、どのような添削を行えば良いのか、どれくらいのフィードバックが学習者にとって一番有効なのかといったことについてはあまりふれられていないのが現状です。これには二つ理由があります。一つ目は、研究者達がCFについて述べるとき、修正すべき間違えが多様すぎ、これが全ての間違えに対して一番有効であるといった一貫した方法を提示する

ことが難しいということがあげられます。二つ目に、近年の研究では、ある一つのことに焦点を当てすぎており、実際の教室ではあまり実践性をもてないということがあげられます。実際に使える修正を促すフィードバック、CFを研究するためにはこの二つが改善される必要があるのではないのでしょうか。

Since Truscott's (1996) paper calling for the abandonment of written corrective feedback (CF), there has been disagreement over the efficacy of CF. At issue is the question of whether CF has any positive effect on students' long-term writing accuracy. It is arguably one of the more important unresolved questions for EFL writing teachers. Instructors in many contexts spend dozens of hours every year giving CF with the assumption that it will be beneficial. Students are also asked to devote substantial time attending to this feedback. As Truscott points out (1996; 1999), if CF is ineffective, this represents a tremendous waste of time and resources. It is therefore in the interests of stakeholders that we develop a greater understanding of the following questions surrounding written corrective feedback.

Is Some Corrective Feedback Better than No Feedback at All?

In 1996 Truscott claimed that "None of the studies that purportedly support the practice of grammar correction actually do so" (p. 341). This was based on two issues. First, he makes the valid point that the early studies in support of the efficacy of CF made no effort to measure learning. These papers only looked at the effects of CF on revisions of existing texts, not the extent to which CF brought about gains in accuracy in new texts. Two examples of this are Fathman and Whalley's (1990) study of 72 U.S. university students and Haswell's (1983) study of 24 U.S. university students. In both studies the treatment groups showed an increase in accuracy in revised drafts after receiving CF. The limitation here is that this is uncontroversial and not particularly useful. Even Truscott (1996, 1999) accepts that CF contributes to accuracy in revisions. The more interesting question is whether the effect lasts for any length of time. None of the early

studies points to a longer-term effect on written accuracy based on CF. Semke (1984), Robb, Ross, and Shortreed (1986), Kepner (1991), and Sheppard (1992) all found no significant effect for written CF in new writing tasks. At this point Truscott's case looks conclusive: the evidence we have for an effect for CF on written accuracy is only for text revisions, and there is no evidence for long-term gains.

Proponents of CF, however, take a different view. Ferris (2004) and Bitchener, Young, and Cameron (2005) point out the that studies showing no effect for CF from this time all lack a control group. Without adequate controls as a point of comparison, it is difficult to draw any conclusions about what effect any feedback technique has. The implication is that the pre-1996 studies indicating no long-term effects from CF were too flawed to inform any conclusions.

Truscott's (1996) second point is that CF is inherently implausible as a learning tool. The primary reason he gives for this is that written CF is not and cannot be sensitive to developmental sequences in grammar acquisition. He suggests that any learning that extends from CF is at best "pseudolearning."

Truscott's point here is intuitively appealing, but the important question is the extent to which later research supports his concerns. Since 1996, the research has tended to be more rigorous. In that time, all notable studies bar one (Truscott & Hsu, 2008) have supported the efficacy of CF to some degree. This is significant, but much of the evidence is somewhat equivocal. Many of these studies still have problems with methodology or generalisability.

Guénette (2007) argued that it is often difficult to make sense of what the data from CF research say because of the variations in methodology that are employed. She makes four important points. First, the research has often failed to account for student variation, in particular L2 proficiency. There is a least a suspicion that L2 proficiency will play some part in a student's capacity to internalise CF. Second, it has often failed to make use of adequate control groups, so we can not always be sure of what is being measured. In addition, research has often not controlled for variation between the classroom activities to which the treatment and control groups are exposed. Finally, she also pointed to a need for more longitudinal research if we are to make claims about the long-term benefits

of CF.

Ferris and Roberts' (2001) study of 72 university ESL students found a significant effect for CF when students were asked to self-edit their texts. However, while this study was bolstered by the use of a genuine control group, it primarily investigated the relative merits of explicit and non-explicit CF on revision, not new texts. As such, it lacked a longitudinal element and cannot be said to provide any evidence of long-term learning from written CF. Ferris (2004) later reiterated Guénette's concern that a lack of longitudinal studies had undermined the research base.

Chandler (2003) investigated the effects of CF on the writing of 31 Asian students at an American conservatory. She found significant effects for CF on new work written weeks after the treatment. Although this is promising, the study unfortunately lacked an adequate control group, which makes it difficult to be confident of what was measured.

Bitchener et al. (2005) also found a significant effect for CF on both short-term revision and longer-term learning. However, this study clearly contravenes Guénette's call to limit the variation in classroom activities between treatment and control groups. In this study, the two treatment groups received either 20 or 10 hours of intensive instruction per week. The control group received only four. Not only does this equate to a difference of nearly 200 classroom hours over the 12 weeks of the study, but because the students chose which class to attend, there is at least a suspicion that less motivated students would have chosen the shorter, 4-hours-a-week class, and thus formed the lower-performing control group.

Truscott and Hsu (2008) published the one notable study since 1996 that finds no significant effect for CF on learning. However, Truscott makes the unsupported claim that there is no substantive difference between different types of feedback. For this study, he and Hsu limited the feedback to implicit CF in the form of underlined errors only, an approach which places the onus on students to both identify and repair their errors, and is seldom employed in recent CF research. A second criticism that has been made of this study is that it likely suffered from a ceiling effect, whereby there was very little room for students to improve in their second piece of writing because they performed so well in the

first (Bruton, 2009; Bitchener & Ferris, 2012).

Another study from Bitchener (2008) is among the first significant studies of CF not to fall foul of Guénette's (2007) objections. In a study of 75 students at two New Zealand language schools, Bitchener (2008) found that the accuracy of the students who received CF outperformed controls and that "this level of performance was retained two months later" (p. 102). However, this study has a focus that undermines its value to teachers. Bitchener's article signalled the beginning of a trend of highly selective CF research that evaluated CF in terms of its effect on only one language point. For example, Sheen (2007), Bitchener and Knoch (2008), Sheen, Wright, and Moldowa (2009), Farrokhi and Sattarpour (2012), and Shintani and Ellis (2013) have also published recent articles which show an effect for CF on accuracy in the use of a single language point. All five of these articles limited their investigation to students' use of articles in writing. From a teacher's perspective, there is some question of how useful such a limited focus is. Three of the papers (Bitchener, 2008; Bitchener & Knoch, 2008; Shintani & Ellis, 2013) also broadened the definition of CF to include metalinguistic feedback. In the cases of Bitchener (2008) and Bitchener and Knoch (2008), this feedback even included extensive oral elements. This is not cause to doubt the findings, but this is clearly leaving the realm of written CF and is perhaps moving beyond what can be expected of most teachers in real-world contexts.

One recent piece of CF research was called by Ellis and Shintani (2014) "one of the best studies carried out to date" (p. 268). Van Beuningen, De Jong, and Kuiken (2012) investigated the effectiveness of CF on post-test accuracy in Dutch high school students. Where this departed from many of the more recent studies is that it investigated comprehensive feedback (i.e., CF for all errors), something which arguably adds to its real-world relevance. It also found a significant lasting effect for CF on accuracy.

At this point we can say that the literature points at some effect for written CF on accuracy in student writing. Since 1996, only one significant study (Truscott & Hsu, 2008) has contradicted this. However, some uncertainty remains over the extent to which the research supports an approach to CF that is feasible in real-world settings. A number of the studies are methodologically flawed, while

others limit the range of feedback to a degree that would not be useful in a classroom setting. Some of the studies even purport to show an effect from CF when in fact what they measure is CF plus extensive oral feedback. The issue is also significantly more complex than simply a question of whether CF works. CF takes many different forms, all of which can affect uptake in different ways.

What Types of Corrective Feedback Should Be Used?

There are two main considerations here. The first is reasonably straightforward: whether implicit CF (in which the student is made aware of an error, but not what kind of error it is) or explicit CF (in which the type of error is identified) is more effective. Of the major studies published since 1996, the only one which focused entirely on implicit feedback was also the only study to find no effect for CF on learning (Truscott & Hsu, 2008). One other study (Chandler, 2003) found that implicit feedback (underlined errors only) had a greater long-term effect than explicit feedback (underlined errors with codes identifying error types). All of the others looked at either explicit feedback only or compared explicit and implicit CF and found explicit to be superior. On this point the research seems clear: most of the evidence points to the value of explicit feedback over implicit feedback.

Less clear is the question of direct feedback (in which the student is supplied with the correction) or indirect feedback (in which the student is given the error type, but is expected to repair it herself). Very little research has compared the relative merits of indirect and direct CF, with most researchers investigating direct CF only. Of two studies that have made the comparison, the effect on learning is inconclusive. Chandler (2003) found that direct feedback had a greater effect on both revision and new writing, while Ferris (as cited in Bitchener & Ferris, 2012) found that direct feedback had a greater effect on the accuracy of revision, whereas indirect feedback was more beneficial to long-term learning. One possible reason for Chandler's findings is that students sometimes have difficulty understanding the codes often used by teachers in indirect CF (Hyland, 1998). The question of which type of feedback is more effective is important, as indirect CF is much

more time consuming for teachers and students, but at this point there has not been enough research into the matter for teachers to make a confident decision either way.

How Much Corrective Feedback Should Be Given?

Although this question has not received much attention, it makes intuitive sense that too much corrective feedback may be confusing and dispiriting for students. On the other hand, if teachers elect to focus on a small number of language points, how many are still too many? Additionally, how are teachers to decide which are most useful to correct? The tendency in recent research has been to use highly selective treatments.

Four recent studies have focused entirely on the treatment of articles in student writing (Bitchener, 2008; Bitchener & Knoch, 2008; Farrokhi & Sattarpour, 2012; Shintani & Ellis, 2013). Although these studies seem to show an effect for CF on learning, what they represent is so far removed from the classroom reality that they are effectively useless for teachers. The suggestion of Ellis, Sheen, Murakami, and Takashima (2008) that teachers simply treat a different language point in each paper they set is no solution in contexts in which students only write one or two papers a semester. For this reason, there is a need for more research into comprehensive or semi-comprehensive CF that treats most or all tokens of many or all types of grammatical errors. Such treatments have been found, in the Van Beuningen et al. study (2012), to have a significant effect. It is tempting to agree with Van Beuningen et al. when they state that "most important to us is the fact that the comprehensive approach most closely resembles the correction method used in actual teaching practice" (p. 6). In recent years, too little of the research has taken this approach.

Can Corrective Feedback Have a Negative Effect on Students' Writing?

One of Truscott's (1996, 1999) less examined claims is the notion that CF can be counter-productive: that it can can have a negative impact on student writing.

One reason he suggests this is that CF makes students fearful of making mistakes, so they avoid complex language in which errors are more likely, and they write less. Truscott makes reference to three earlier works which appear to provide evidence for this (Sheppard, 1992; Semke, 1984; Kepner, 1991). However there has been little investigation of this in the past 20 years. Chandler (2003) found no negative effect for CF on written fluency, although her lack of a genuine control group undermines the reliability of this claim. Van Beuningen et al. (2012) found no effect for CF on grammatical complexity or on lexical complexity. This is the best recent evidence we have for or against Truscott's claim. We are left with on the one hand an intuition that CF has a "chilling" effect on complexity in students' writing, and on the other a single robust study indicating otherwise. There is a need for further research to clarify this.

Conclusion

Since 1996 a majority of the research has suggested that some CF is probably better than no CF at all, and that explicit feedback is probably more effective than implicit. Beyond this qualified statement, little is settled in any meaningful sense. Much of the early research was of questionable validity. The methodology of recent studies has improved, but their practical usefulness often has not. Beginning with Bitchener (2008), much of the research has been concerned with the treatment of a single error type. Some of these later positive studies have even added metalinguistic training and oral feedback into the mix. While these studies are theoretically significant, this approach bears very little resemblance to what teachers can do in real contexts.

To become more useful, the research needs to follow the example set by Van Beuningen et al. (2012). Not only should it be rigorous to satisfy Guénette's (2007) call for controls and consistency, but the parameters need to be sufficiently inclusive as to include the kinds of feedback that teachers can and do use, and students expect. If these standards become established, we will begin to have a much better understanding not only of whether written CF is effective, but also of how much and what kinds of feedback to provide. Until such time, we will continue to offer written CF with no clear idea of how effective it is.

References

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

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing.* New York, NY: Routledge.

Bitchener, J., & Knoch, U. (2008). The value of a focused approach to written corrective feedback. *ELT Journal, 63*(3), 204-211.

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.

Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, *18*(2)*,* 136-140.

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)*,* 297-296.

Ellis , R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research.* New York, NY: Routledge.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, *36*(3), 353-371.

Farrokhi, F., & Sattarpour, S. (2012). The effects of direct written corrective feedback on improvement of grammatical accuracy of high-proficient L2 learners. *World Journal of Education*, *2*(2), 49-57.

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, England: Cambridge University Press.

Ferris, D. (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*(1), 49-62.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How

explicit does it need to be? *Journal of Second Language Writing, 10*(3), 161-184.

Guénette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing, 16*(1), 40-53.

Haswell, R. H. (1983). Minimal marking. *College English, 47*, 600-604.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing, 7*(3), 255-286.

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, 5*, 305-313.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly, 20*(1), 83-93.

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals, 17*(3), 195-202.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255-283.

Sheen, Y., Wright, D., & Moldowa, A. (2009) Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37*(4), 556-569.

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal, 23*(1), 103-110.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286-306.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327-369.

Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing, 8*(2), 111-122.

Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, *17*(4), 292-305.

Van Beuningen, C., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, *62*(1), 1-41.

Author Bio

Darby McGrath *is an Associate Lecturer of English at the School of Policy Studies, Kwansei Gakuin University, Hyogo. He has a Master of Applied Linguistics (TESOL) and is currently working on his Ph.D. His research interests include academic writing and the bilingual mental lexicon. darbymcg@gmail.com*