Written Corrective Feedback: A Review of Studies since Truscott (1996)

Yuan-Yuan Meng[1]

INTRODUCTION

Linguistic errors are pervasive in second language (L2) students' writing. Depending on their gravity, the errors may cause a minor degree of irritation to the reader or even lead to total communication breakdown. As such, errors have always been a major concern to both students and teachers, and error correction has also assumed a central position in language teaching. Students generally expect that their errors will be pointed out and dealt with by their teachers. For instance, in a study on students' attitudes toward corrective feedback (CF) in college-level English writing classes, Leki (1991) surveyed 100 English as a Second Language (ESL) students, asking them such questions as how concerned they were with their written errors, what they thought were the most important features in their writing that the teacher should attend to, and what they looked at when receiving a graded paper from the teacher. The results of the survey indicated that the students believed that good writing should be error-free, and the majority wanted all their written errors to be corrected. For L2 teachers, providing written CF on student writing has long been an essential practice. In fact, "grammar correction is something of an institution" (Truscott, 1996, p. 327) in L2 writing courses. Despite the fact that correcting students' written errors is a time-consuming ordeal, and the endeavor is "fraught with uncertainty about its long-term effectiveness" (Ferris, 1999, p. 1), most L2 teachers have continued to slave over students' errors in one form or another. As confirmed by a recent study on practitioners' perspectives, the majority of teachers believe that students need CF and that written CF is overall an effective pedagogical practice (Evans et al., 2010).

In the field of Second Language Acquisition (SLA), the positive effect of oral CF in facilitating L2 acquisition is generally supported by research (Li, 2010; Lyster & Saito, 2010; Russell & Spada, 2006), but research on the role of written CF in SLA has produced mixed results, and a consensus on the positive role of written CF has not yet been obtained (Ferris, 2010; Guenette, 2007; Hyland and Hyland, 2006; Van Beuningen, 2010). Hyland and Hyland (2006) commented that due to the varied populations, treatments, and research designs of many of the earlier studies on written CF, it was difficult to draw valid conclusions and generalizations from them. Likewise, Guenett (2007) attributed the conflicting research findings on written CF to methodological inconsistencies between the earlier studies and the presence of confounding external variables.

[1]Yuan-Yuan Meng is currently an EdM student in Applied Linguistics at Teachers College, Columbia University, where she obtained an MA in TESOL and an MA in Applied Linguistics. She has been teaching Chinese as a foreign language at Columbia University since 1993. Her research interests include formulaic expressions, literacy education, and the stylistic features of Chinese journalistic language. Correspondence should be sent to: ym11@columbia.edu.

Can written CF help L2 learners improve their grammatical accuracy in writing? Is it worthwhile for L2 teachers to devote countless hours to written error correction? These questions not only concern L2 teachers; they have also become a source of controversy among SLA researchers, who, for over a decade now, have vigorously debated the effectiveness of written CF in facilitating L2 development (Ferris, 1999, 2004; Truscott, 1996, 1999, 2004, 2007). The debate was famously triggered by Truscott's (1996) review article, "The Case against Grammar Correction in L2 Writing Classes," in which he made the strong claim that grammar correction is ineffective and potentially harmful and should therefore be abandoned. Responding to Truscott, Ferris (1999) argued that when correction is "selective, prioritized, and clear" (p. 4), it can be beneficial for some student writers. She also argued that research evidence up to the late 1990s had been too limited to support Truscott's strong position. Furthermore, many of those earlier studies suffered such design flaws as the lack of a control group or generalizability, so she contended that the results obtained from them could not serve as a basis for abolishing the long-valued tradition of error correction. In 2004, Ferris argued again in a review article that the existing research findings were still inconclusive in determining the efficacy of written CF and called for more robust investigation into the effect of error correction, which ideally should address the design flaws found in many previous studies. Specifically, she called for replicable longitudinal studies, carefully designed with a control group that does not receive feedback. Following this call, there has emerged a good number of experimental and quasi-experimental studies (e.g., Bitchener, 2008; Ellis et al., 2008; Sheen, 2007; Truscott & Hsu, 2008), which made an effort to address the shortcomings of prior research by including a control group and examining the effect of written CF on new pieces of writing. Apart from these tightly controlled studies, descriptive research has also been carried out (e.g., Evans et al., 2010; Vyatkina, 2011).

In this paper I review 15 empirical studies on written CF conducted since Truscott's (1996) call for the abandonment of grammar correction. In the sections below, I first summarize the arguments presented in Truscott's (1996) review and then provide the definitions for several key terms related to the studies on written CF. Following that, I review the empirical studies and discuss their findings and implications. I conclude this paper by making suggestions for pedagogy and future studies.

SUMMARY OF TRUSCOTT (1996)

In "The Case against Grammar Correction in L2 Writing Classes," Truscott (1996) advanced the position that grammar correction is ineffective, counterproductive, and should have no place in writing courses. By grammar correction, he meant "correction of grammatical errors for the purpose of improving a student's ability to write accurately" (p. 329). Citing research evidence from first language (L1) writing studies (e.g., Knoblauch & Brannon, 1981) and L2 writing studies (e.g., Kepner, 1991; Semke, 1984; Sheppard, 1992), and arguing on both theoretical and practical grounds, he repeatedly emphasized that grammar correction does not work. His major arguments are summarized as follows:

1. Second language development involves complex learning processes rather than the simple process of passing information from teacher to student. Error correction is a form of information transfer. To harbor favorable intuitions about error correction is thus to subscribe to the simplistic "information-transfer" view of learning, which expects, unrealistically, that students will take in everything the teacher imparts to them.

2. Related to the notion of complex learning processes, research has shown that in acquiring certain grammatical features, L2 learners follow a natural pre-determined acquisition order, which is unaffected by instruction. It follows then that error correction, which functions as a form of instructional intervention and which often fails to match students' current stage of interlanguage development, is not likely to have much value.

3. Not all learning is equal. Learning which does not affect learners' underlying, developing system and which only results in a superficial, transient form of knowledge, is "pseudo-learning." Grammar correction only produces pseudo-learning because it is "typically done in terms of isolated points and without reference either to the processes by which the linguistic system develops or to the learner's current developmental stage" (p. 347).

4. From a practical perspective, grammar correction is unlikely to be successful as it requires that many other conditions be met. For one, the teacher must have the ability to recognize errors and be able to explain the problems well. But grammar is often too complex to be adequately explained by any single linguistic theory or teacher, and even if the teacher is capable enough to recognize and explain all errors, he may still lack the time and patience required for high-quality correction. On the students' part, they may fail to understand the teacher's corrections or only achieve a surface understanding of the teacher's comments without grasping the general principles involved. In the latter case, they are likely to forget the new knowledge quickly and repeat the errors later. Students may also lack the motivation to process the teacher's feedback thoughtfully.

5. The inherent unpleasantness of correction triggers much anxiety in students and discourages them from experimenting with the language. In order to avoid errors and corrections, students tend to shorten and simplify their writing, which hurts the development of their linguistic complexity.

6. The time spent on providing and processing low-level grammar feedback diverts the attention of both teacher and student from other high-level aspects of writing such as organization and content. When effort directed at grammar correction diverts class resources from other important activities like additional writing practice, the correction harms students' interlanguage development.

DEFINITION OF TERMS

Direct vs. Indirect Written CF

Direct written CF involves the provision of the correct form on students' work, including such treatments as crossing out an unnecessary element, inserting a missing element, and writing the correct form above or near the error (Ellis, 2009). Indirect written CF only requires singling out the errors without actually providing the corrections. It can be done by underlining or circling errors, inserting a caret (^) sign to indicate a missing word, or placing a question mark alongside a confusing phrase (Hendrickson, 1980). It can also include the use of error codes.

Metalinguistic CF

Metalinguistic CF involves providing explicit explanations of the nature of the student's errors (Ellis, 2009). The explanations can be directly written down on the student's paper or given through face-to-face conferencing.

Focused vs. Unfocused Written CF

The teacher can choose to treat students' errors selectively (focused approach) or comprehensively (unfocused approach). The focused approach targets only one or a limited range of error types (e.g., the use of English articles), whereas the unfocused approach involves correcting all errors in the student's text (Van Beuningen, 2010).

REVIEW OF STUDIES

At the start of the 2000s, two studies investigating the effect of written CF on revised texts have provided support for the short-term effectiveness of CF (Ashwell, 2000; Ferris & Roberts, 2001). Ashwell (2000) compared the effect of different feedback patterns on the writing by 50 Japanese college students in the context of a process-oriented writing instruction. He wanted to find out if the content-then-form feedback sequence as suggested by Zamel (1985) was superior to other sequences in improving students' writing. The students were assigned to one of four different groups: the content-then-form feedback group, the form-then-content feedback group, the form-and-content feedback group, and the non-feedback group. Form feedback was provided by underlining or circling linguistic and mechanical errors, and content feedback addressed discourse-level issues such as organization, paragraphing, and cohesion. The students produced three drafts for a writing assignment. The first group received content feedback on their first draft, then form feedback on the second draft; the second group received feedback in the reverse order; the third group received mixed feedback on both drafts. The results showed that the recommended feedback sequence was not superior to the other two feedback patterns, and the three groups that received feedback outperformed the non-feedback group. When revising their texts, the students acted on three quarters of the form feedback, which helped them to improve the linguistic accuracy of their writing.

Ferris and Roberts's (2001) study, also involving the provision of indirect CF (codes and underlining) and text revision, likewise showed that ESL college students who received CF produced significantly better revised texts than those who did not receive feedback. This study consisted of a four-stage cycle (pre-test, writing, feedback, editing) and targeted five error categories. Prior to the writing, a "Grammar Knowledge Questionnaire" and a "Grammar Knowledge Pretest" were administered to ensure that the students' knowledge of the five error categories was comparable. Students were assigned to one of three groups: Group A had their errors underlined and coded by error type; Group B only had their errors underlined; and Group C did not receive any feedback. Two weeks after the original writing, the students self-edited their texts in class based on the papers — corrected or uncorrected — that they received. Apart from providing support for the value of CF for text revision, this study also showed that there were no significant differences in revision performance between the two feedback groups.

Notwithstanding the short-term effectiveness of written CF demonstrated by the above revision studies and an earlier study by Fathman and Whalley (1990), Truscott (2007) contended that because these studies did not produce evidence of learning, which in his term refers to improved accuracy on new pieces of writing, the investigation was of limited value.

In the same year as Ferris and Roberts's study, Fazio (2001) conducted an experiment investigating the effect of different feedback on the writing accuracy of 112 fifth-graders. These fifth-graders were L1 and L2 learners of French at a French-language school in Montreal,

randomly assigned to one of three feedback conditions — form-focused, content-based, and the combination of form and content feedback. For this study, the students produced one piece of journal writing each week over the course of four months, with complete freedom to topic selection and length of entry. For each journal entry, the form-focused group received direct CF on two targeted grammatical features; the content group received a content-related comment; and those in the combination group received both types of feedback. Accuracy of writing was measured as the total number of targeted errors committed divided by the total number of uses of the targeted structures. The findings showed that neither the form-focused nor the content-based feedback exerted much positive influence on the L1 and L2 students' grammatical accuracy in the targeted features.

Chandler (2003) reported positive findings in two studies on ESL music students, who wrote five autobiographical essays over one semester for a writing course. In the first study, students were assigned to an experimental group (n = 15) and a control group (n = 16), with the former required to correct the errors marked by the teacher and the latter not. The experimental group's papers were treated in two steps: First, the teacher underlined all the errors in the students' first draft. Then, after students rewrote the paper based on the initial feedback, the teacher provided direct correction on their revised text. The control group, who were not required to rewrite the first draft, only received the initial CF from the teacher. Students' errors were classified into 23 different categories, including but not limited to word form, word choice, verb tense, and word order. Error rate was calculated as number of errors per 100 words, and writing accuracy was measured by calculating error rate on the first and fifth papers. The results revealed that the experimental group showed significant improvement in grammatical and lexical accuracy on the fifth paper, but no significant difference was found between the control group's first and fifth papers. Between the experimental and the control groups, there was also significant difference in their improvement in accuracy over the 10 weeks. Furthermore, both groups showed marked increase in writing fluency, measured by the time spent on writing each piece.

Chandler's (2003) second study investigated the differential efficacy of four types of feedback methods on students' writing accuracy. It also examined change in students' writing fluency over one semester and looked into students' attitudes to the different feedback methods through surveys. Participants were 20 students enrolled in the same ESL writing course in a different year, who were also required to produce five autobiographical essays. For each of the first four essays they wrote, the students had to revise based on the teacher's feedback, which was provided in four ways: direct correction, underlining with description, description of error type only, and underlining. All students received these four types of feedback but in varying order. Error rate was calculated in the same way as for the first study, and improvement in writing accuracy was assessed by looking at the change in error rate in the revised and subsequent new papers. The overall writing quality of the first draft of the first paper and that of the fifth paper was rated holistically, on a scale of 1-6, by two trained raters. In addition, writing fluency was measured by using students' self-reports of the time spent on writing each paper. The study yielded five main findings: (1) that students' writing accuracy and fluency both improved significantly over the semester; (2) that concerning the overall writing quality, there was only a small difference between the ratings for the first paper and the last; (3) that with respect to accuracy in revision, students produced fewest errors with the direct correction method, second fewest errors with underlining plus description, and most errors with description alone; (4) that in terms of accuracy in subsequent writing, the direct correction and underlining methods led

to more accurate writing on the next assignment; and (5) that most students considered direct correction to be the easiest and the most preferred feedback method.

Although Chandler's (2003) two studies suggest that the provision of CF, when combined with revision requirement, helps learners improve their writing accuracy over time, some researchers (Guenette, 2007; Truscott, 2004, 2007) argued that due to the absence of a non-feedback control group in both studies, their findings failed to provide solid evidence about the superiority of CF. To answer what Truscott (2004) called the "Big Questions" — whether written CF facilitates the development of grammatical accuracy and whether teachers should provide written CF — Ferris (2004) stressed that studies with more robust design were needed.

As mentioned previously, Ferris's (2004) call for more robust investigation into the role of written CF has given rise to a growing body of new studies, the most noteworthy strand of which investigated the effect of focused written CF on learners' development of grammatical accuracy (Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009; Bitchener & Knoch, 2010; Bitchener et al., 2005; Ellis et al., 2008; Sheen, 2007; Sheen et al., 2009). This body of research shares several methodological features. First, all employed a quasi-experimental design with a non-feedback control group for comparison. Second, most targeted only the use of two functions of the English article system: the indefinite article *a* for the first mention and the definite article *the* for subsequent mention. These studies also went beyond examining short-term improvement in revisions and sought to measure accuracy gain on new pieces of writing. The research questions that they explored generally fall into two categories: whether written CF improves accuracy in the use of the targeted features on new pieces of writing, and whether there is a difference in the effect of the different feedback strategies. Below I describe each of them in more detail.

Bitchener et al.'s (2005) 12-week study examined the effects of written CF and face-to-face conference on the development of three English grammatical features: prepositions, simple past tense, and articles. Participants were 53 ESL immigrant learners, who completed four writing tasks, each in the form of an informal letter, and were divided into three treatment groups. Group 1 received direct CF and a 5-minute student-researcher conference; Group 2 received direct CF only; and Group 3 received no CF. Accuracy was measured as the percentage of correct suppliance of each targeted linguistic feature. The study found that Group 1 outperformed Group 2 and the control group. The combination of direct CF and individual conferences helped learners improve their use of the simple past tense and the definite article, but such effect was not found for the use of prepositions. The researchers attributed the positive results with simple past tense and definite article to the fact that they are more rule-based, and the negative results with prepositions to their idiosyncratic nature. Also, the meta-linguistic explanations students received during the one-on-one meeting helped them notice the gap between their incorrect usage and the correct form, and thereby facilitated their understanding and acquisition of the rule-based features.

Following Bitchener et al. (2005), Bitchener (2008) and Bitchener and Knoch (2008, 2009, 2010) conducted a series of narrowly focused studies, all targeting the use of two functions of the English article system as described above. In Bitchener (2008), 75 intermediate-level ESL students participated in an experiment involving a pre-test, an immediate post-test, and a delayed post-test. The delayed post-test was administered two months after the immediate post-test. For each of these tests, the students completed a 30-minute in-class writing task describing what was happening in a given picture, and they were assigned to four different conditions: direct CF plus written and oral meta-linguistic explanation; direct CF and written meta-linguistic explanation;

direct CF only; and no feedback. Oral meta-linguistic feedback was delivered to Group 1 via a 30-minute mini-lesson. Obligatory contexts for the targeted article uses were identified, and accuracy was calculated as percentage of correct uses in obligatory contexts. The results demonstrated that all the treatment groups receiving CF outperformed the non-feedback control group on both the immediate and delayed post-tests, suggesting that written CF is not only effective in the short-term, but the level of effectiveness can be sustained for a longer period. The study also showed that of the three feedback methods, the combination of direct CF plus written and oral metalinguistic explanation produced the best results in helping students improve the accuracy of their use of articles. Interestingly, though, Group 2 who received direct CF plus written metalinguistic explanation performed worse than Group 3 who only received direct CF.

Extending Bitchener's (2008) study with the same design, Bitchener and Knoch (2008) carried out a two-month experiment with 114 low-intermediate ESL learners and obtained slightly different results. While the study also showed that students who received written CF performed better than those who did not, and that the level of accuracy was sustained for seven weeks, nonetheless no significant difference was found in the efficacy of the different feedback methods.

In Bitchener and Knoch's (2009) ten-month investigation on the effect of written CF, 52 low-intermediate ESL students were assigned to the same three treatment groups and a non-feedback control group. The major difference between this study and the previous two lies in its duration. After the immediate post-test, the students were evaluated on three delayed post-tests at three different points of time. The study found (1) that all feedback groups outperformed the control group; (2) that there was no difference in effectiveness among the three feedback options; and (3) that the accuracy-gain was durable, retained over a 10-month period.

Still another study by Bitchener and Knoch (2010) applied similar methodology to compare the effects of three feedback strategies on 63 advanced ESL learners' use of the two articles. The three treatment conditions were: written meta-linguistic explanation, indirect circling of errors, and written meta-linguistic explanation plus oral form-focused instruction. The study showed that all experimental groups outperformed the control group in the immediate post-test, and only the two meta-linguistic groups retained the accuracy-gains across the 10-week period.

In a similar vein, Sheen (2007) examined the differential effect of two types of written CF — direct CF and direct CF plus meta-linguistic comments — on the acquisition of English articles by 111 intermediate ESL students. Through a correlational analysis, she also investigated how language aptitude affected the results. The students took a language analytic ability test two weeks prior to the experiment, which was then followed by the pre-test, two CF treatment sessions, the immediate post-test, and the delayed post-test. During each of the two treatment sessions, the experimental groups wrote a narrative based on a story they read. Then the narrative was corrected by the researcher, who focused mainly on article errors, and returned to the students 2-4 days later. The control group received no CF treatment. To measure acquisition, three kinds of tests (a speeded dictation test, a narrative writing test, and an error correction test) were given to all three groups during the pre-test, immediate post-test, and delayed post-test. This two-month study reached three conclusions. First, the CF groups outperformed the control group on both post-tests. Second, there was no difference between the two feedback methods in the immediate post-test. Third, in the delayed post-test, there is an advantage for the combination of direct CF and written metalinguistic explanation over direct CF alone. This third conclusion supports Bitchener et al.'s (2005) findings that when meta-linguistic explanation accompanies

direct CF, learners are more likely to achieve enduring improvement. The study also showed that learners with stronger language aptitude benefitted more from the CF.

Building on Sheen (2007), Sheen et al. (2009) examined the differential effects of three treatments – direct focused CF, direct unfocused CF, and writing practice alone – on the use of the English articles by 80 ESL learners. Students in the focused CF group only had their article errors corrected, whereas the unfocused CF group had their errors in five categories (articles, copula *be*, regular past tense, irregular past tense, and preposition) corrected. Like those in the Sheen (2007) study, these two CF groups followed the same procedure to complete two written narrative tasks and receive CF during the treatment sessions. The writing-practice group only did the narrative tasks without receiving CF; the control group received no CF. All four groups completed the pre-test, immediate post-test, and delayed post-test, and acquisition of the articles was measured by three versions of a narrative writing test, which asked the students to write a story based on a series of pictures. The results showed that all three treatment groups outperformed the control group. For the acquisition of the articles, the focused group showed an advantage over the unfocused group in both the immediate and the delayed post-tests, suggesting that focused CF is more effective than unfocused CF. What is surprising, though, is that in terms of the overall accuracy in the five targeted features, the focused group, who did not receive corrections on features beyond the articles, also outperformed the unfocused and the control groups. The unfocused group, who in fact received CF for the targeted features, did not demonstrate significantly greater accuracy than the control group. Sheen attributed this to the possibility that when CF addresses a range of errors, learners might be less able to process the feedback effectively. Another reason might have to do with the manner in which the CF was provided — the feedback which the focused group received was systematic, but that which the unfocused group received was much less so. Although the writing-practice group also outperformed the control group in both the article uses and in the other grammatical features, suggesting that accuracy can improve through the kind of writing practice used in this experiment, the effectiveness of focused CF was still greater than that achieved by writing practice alone.

Adopting several methodological features from Sheen (2007), Ellis et al. (2008) also compared the effects of focused and unfocused written CF on Japanese EFL learners' use of the two English articles. On three written narratives, students in the focused group received corrections of article errors, the unfocused group received corrections of all errors, and the control group received no CF. Accuracy-gain on the article uses was measured by narrative writing tests and error correction tests. The study found that both CF groups performed significantly better than the control group on the delayed post-test. The researchers contended that contrary to Trucscott's claim (1996, 1999), written CF can indeed facilitate acquisition. But different from the findings revealed in Sheen et al. (2009), the results indicated that there was no statistically significant difference between the effects of focused and unfocused CF options.

The next three studies represent a refreshing break from the above "article craze," in that they all employed a true experimental design and treated errors comprehensively.

Truscott and Hsu (2008) tested the effectiveness of unfocused CF on students' revision and new writing. Participants were 47 university students in Taiwan, who were randomly assigned to either the experimental group that received CF from the teacher, or the control group that did not receive any CF. Both groups were given 30 minutes to write a narrative essay in class during the 12^{th}, 13^{th}, and 14^{th} weeks of the semester. In week 12, all students wrote a narrative story based on a sequence of eight pictures. One week later, they received the first draft

back from the teacher and revised it. The experimental group rewrote the story based on the teacher's feedback, delivered via underlining of errors without codes. The control group simply rewrote the draft without the benefit of CF. In week 14, the two groups wrote another narrative story based on a new sequence of eight pictures. In this study, the treatment of errors focused on grammatical and spelling errors only. Statistical analysis revealed that for the rewrites, there was a significant difference between the two groups in error reduction. However, for the new writing task, the error rate between the two groups was virtually identical, suggesting that CF did not exert a strong influence on students' writing development. On the one hand, these findings corroborated previous research that CF does have a positive effect on reducing errors in revision. On the other hand, they failed to show that success with the revision task can translate into improved accuracy on a new writing task. Truscott and Hsu thus concluded that successful error reduction on the revision cannot be taken as evidence of learning.

Van Beuningen et al.'s (2008) study, however, provided evidence to support the beneficial role of unfocused CF. The study was carried out on 62 secondary-school L1 and L2 learners of Dutch, who were randomly assigned to one of four groups: direct unfocused CF, indirect unfocused CF, practicing writing (practice), and self-correction. All four groups participated in a three-session experiment and completed two writing tasks for Session 1 and another two tasks for Session 3. The two writing tasks for each session were of different topics, both requiring that the students write an email explaining the topics to a classmate. Each topic was accompanied by a series of pictures, and students were allowed 20 minutes to perform each task. The purpose of using two task topics was to control for topic influences and increase generalizability. Students in the CF groups received either direct corrections or coded CF on their writing. The Practice group and the self-correction group received no CF. During Session 2, the CF groups revised their texts based on the corrections they received; the Practice group did not revise their uncorrected texts but instead, completed two additional writing tasks; and the Self-correction group was required to revise their texts without the benefit of CF. The results revealed that the two feedback groups significantly outperformed the self-correction group on the revised texts. With regard to performance on the new tasks given at Session 3, only direct CF resulted in improved accuracy. Of the four groups, the practice group performed worst, displaying no improvement from Session 1 to Session 3. This is different from the results obtained by Sheen et al. (2009), which showed that writing practice alone was more helpful than receiving no feedback. It also challenged Truscott's (1996) suggestion that having students do additional writing practice might be more worthwhile than giving them CF.

Following up on their 2008 study, Van Beuningen et al. (2012) investigated the effect of direct and indirect unfocused CF on 268 L1 and L2 Dutch learners. This study is unmatched by previous experiments in that it not only dealt with a language other than English, but also involved a large number of secondary-school participants representing two educational levels. It also explored eight research questions, some of which had not been addressed or only touched upon briefly before. For instance, the study sought to test Truscott's (2007) hypotheses that CF might only have value for non-grammatical errors and that CF compels students to simplify their writing. Like in the previous study (Van Beuningen et al., 2008), the students were randomly assigned to one of four conditions: direct CF, indirect CF (via error codes), Self-correction, and Practice. The experiment consisted of four sessions: a pre-test session (Session1), a treatment/control session (Session 2), a post-test session (Session 3), and a delayed post-test session (Session 4). At Session 1, all four groups completed the first writing task. One week later at Session 2, the two CF groups revised their first written texts based on the CF they received;

the Control group self-corrected their original texts without the help of CF; and the Practice group completed a new writing task. During Session 3 (week 3) and Session 4 (week 6), all four groups produced a new text based on a new topic. All writing topics were related to biology (e.g., honeybees, ladybugs), and each writing task required that the students wrote an email to a classmate explaining the topic.

The results showed that unfocused CF led to improved accuracy in both the revised texts and the new texts. The positive effect of unfocused CF observed at the revision stage was retained four weeks later. In terms of the differential efficacy of direct and indirect CF, the study found that only direct CF facilitated "durable grammatical accuracy improvements of a medium size" (p. 32), and that indirect CF had a greater effect on non-grammatical errors. The findings also rebutted Truscott's (1996) hypotheses that CF harms the complexity of students' writing and that additional writing practice may be more beneficial than the provision of CF. Lastly, the study found no significant interaction between feedback effectiveness and the participants' educational level. In short, this study demonstrated that comprehensive treatment of errors can help students improve their grammatical accuracy over time, and the findings do not support Truscott's claim that grammar correction may bring about harmful effects.

DISCUSSION

Effects of Written CF on Text Revision

The findings of two revision studies (Ashwell, 2000; Ferris & Roberts, 2001), together with Truscott and Hsu (2008) and the two studies by Van Beuningen et al. (2008, 2012), have provided clear evidence that written CF can serve as an effective editing tool — students who received both direct and indirect CF outperformed those who received no CF on the revision task. By now, a consensus regarding the short-term effectiveness of CF is apparently established.

Effects of Written CF on New Texts

But what is of greater concern is whether the short-term effect can be transferred to new tasks and help students acquire grammar over time. Most of the studies reviewed above have striven to address this issue. Even though a unanimous answer to the question has yet been reached, there is growing evidence that written CF can help learners improve their linguistic accuracy for the longer term. Of the 11 recent quasi-experimental and experimental studies, ten have shown that both selective and comprehensive correction of errors can result in learners' improved accuracy in new texts, and the effect could be as durable as up to ten months. Most notably, with its robust research design (i.e. true experimental design using a large sample), Van Beuningen et al.'s (2012) study provided compelling evidence of the effectiveness of written CF. Nevertheless, problems still exist in some of these studies, which I will turn to shortly.

Focused Written CF vs. Unfocused Written CF

If written CF can be useful for the acquisition of certain linguistic features for the long term, it is worth asking then whether focused written CF (selective correction) is more effective than unfocused feedback (comprehensive correction) and whether direct written CF is more

effective than indirect feedback. For the focused-unfocused dichotomy, the limited findings so far do not provide a clear picture. Ellis et al. (2008) and Sheen et al. (2009) both compared the relative effectiveness of focused and unfocused CF, but the former found the two approaches to be equally effective, while the latter showed that the focused approach yielded better results. To complicate the picture a little, Van Beuningen et al.'s (2008, 2012) studies showed that unfocused CF helped their Dutch learners improve written accuracy both for the short- and the longer-term.

Direct Written CF vs. Indirect Written CF

With respect to the relative effectiveness of direct and indirect CF, four studies that compare the two approaches (Bitchener & Knoch, 2010; Chandler, 2003; Van Beuningen et al., 2008, 2012) demonstrated that direct CF led to greater accuracy gain than indirect CF. Bitchener and Knoch's (2010) and Van Beuningen et al.'s (2008, 2012) particularly showed that while both direct and indirect CF were effective for the short-term, only direct CF yielded a more significant long-term effect. From a theoretical point of view, some researchers (e.g., Ferris, 2002; Lalande, 1982) had suggested that indirect CF has the greatest potential to facilitate learning because it engages learners in deeper cognitive processing and "promotes the type of reflection on existing knowledge or partially internalized knowledge that is more likely to foster long-term acquisition and written accuracy" (Bitchener & Ferris, 2011, p. 65). Chandler (2003), however, argued that the benefit created by indirect CF may be cancelled by delayed access to the correct form. Direct CF, on the other hand, allows learners prompt access to the target form, enabling them to confirm or abandon their hypotheses about the language soon after they write, and thereby helps them to internalize the corrections better. Other benefits of direct CF may include reducing confusion resulted from ambiguous indirect CF and providing learners with information to solve complex errors involving syntactic structure or idiomatic expressions (Bitchener & Ferris, 2011).

Amenability to Written CF

Not all errors are equally treatable — different types may respond to correction differently (Ferris, 1999; Hyland & Hyland, 2006). According to Ferris (1999), "treatable errors" are those that occur in "a patterned, rule-governed way" (p. 6) such as verb-form errors and the aforementioned article use, whereas untreatable errors (e.g., lexical errors, sentence construction) are more idiosyncratic in nature. Truscott (2001) also suggested that some error types might be more correctable than others. To him, the more appropriate targets for selective correction are discrete items involving simple principles (e.g., lexical items, spelling, punctuation), and the uncorrectable errors are those "stemming from problems in a complex system, particularly the syntactic system" (p. 93). In his view, grammar errors are generally poor candidates for CF, but non-grammatical errors can fare well with correction. On this point, Van Beuningen et al.'s (2012) study demonstrated that both complex grammatical problems (e.g., word order) and non-grammatical errors (e.g., spelling) are susceptible to CF. Regarding the correctability of lexical errors, it is interesting to see how Truscott's view contradicts Ferris's. In my opinion, lexical items may just be as complex as articles. Some lexical items are indeed simple and easily correctable, but others that involve multiple meanings or complicated selectional restrictions may not respond to CF well. Future research could investigate whether lexical errors would be good targets for correction since the extant research has not shed much light on that.

Limitations of the Current Research

Methodological flaws are apparent in some of the studies. In Bitchener et al.'s (2005) experiment, the participants, who were divided into three feedback groups, received different amount of instruction from their respective classes — 20 hours a week for Group 1, ten hours a week for Group 2, and four hours a week for Group 3. Though the authors claimed that all three groups received the same amount of grammar instruction and attention to writing, such dramatic differences in instructional exposure would most likely have confounded the results. It is possible that the positive effects that Group 1 obtained did not actually result from the written CF but rather were attributable to the longer hours of instruction they had received. Given this flaw, Truscott (2007) argued that the study did not qualify as a controlled experiment. A closer examination of the three groups' performance on the writing tasks also led him to question the validity of the study's findings.

As noted earlier, the majority of the quasi-experimental studies post 2005 only targeted one single aspect of the English article system. With improved research design, this body of work has consistently shown that written CF can facilitate the acquisition of the targeted feature. Despite the optimistic findings, however, these studies have come under criticisms chiefly because of their extremely narrow focus. Bruton (2009b) and Storch (2010) both criticized these studies for lacking ecological validity, having little pedagogical relevance to real classrooms. Bruton (2009b) commented that with the single-feature focus, these studies give the impression that the writing tasks more resembles "the third stage in a Present-Practice-Produce (PPP) sequence" (p. 608) than the different stages of a communicative writing process. Similarly, Storch (2010) questioned the generalizability of these studies' findings, arguing that the provision of CF on only one grammatical feature in controlled environments will not be of much value to language teachers. Sympathetic to these legitimate concerns, Ferris (2010) also noted that

> L2 writing researchers and practitioners might wonder if, in the interest of empirical rigor, some of the SLA research efforts on written CF have been so narrowly focused that it would be difficult to transfer their approach and findings to a real writing classroom or to a diverse group of students. (p. 186)

Another problem with this body of research, which relates more to grammar instruction than methodology, has to do with how the English article system was presented to the students. As is well known, the article system is highly complex. Certainly there is more to the articles than what was presented in these studies. By limiting the article usage to the simple contrast of "first mention" and "anaphoric second mention," students might be misled into thinking that that was the whole picture of the article system, and might even feel confused when they encounter other types of article uses that are beyond the simple scope of "first mention vs. second mention".

As an avid critic of empirical research on written CF, Truscott himself has only conducted one study on the subject with Hsu in 2008, and this study did not escape scrutiny, either. Truscott and Hsu concluded that learning did not occur as a result of the provision of indirect written CF. On the basis of the surface evidence, this conclusion seemed justified. It is true that compared to the control group, the experimental group showed significant improvement on their revision texts after receiving correction, but on the subsequent new writing task, they

produced virtually the same amount of errors as the control group. However, upon careful examination of the data provided by Truscott and Hsu, Bruton (2009a) pointed out two problems that called into question the author's claims. First, the experimental group's lack of improvement from the pre-test to the post-test might have been due to a "ceiling effect," which refers to the fact that the experimental students did not make that many errors on their first writing piece to begin with. These students were thus left with little room to make significant improvement on the subsequent text. But more crucially, the way accuracy gain was measured has seriously undermined the validity of the study's findings.

In this experiment, each piece of the students' writing was assigned an error rate, calculated as the total number of errors divided by the total number of words produced, and the mean error rates for both the experimental and the control groups were also calculated. The authors then compared the global error rates of the two groups without taking into account what specific errors were made on the initial text and the new text. This approach is problematic because to truly capture the changes in the student's writing, one would need to examine the specific details in each of the texts. To accurately gauge the effect of written CF on error reduction, one would also need to consider if the student's original errors in the initial text, for which he receives CF, actually reappear or disappear in the new one. If what shows up in the new writing are mostly new errors which have never before been treated, the high error rate on the new text cannot be taken as evidence of non-learning or ineffectiveness of correction.

Such is precisely the case with a particular participant, whose three sample texts (initial text, revision, second text) were included in the study, and on which Bruton (2009a) conducted a detailed qualitative analysis. It was found that the student's most frequent error in the initial text was the past tense of *to be*, which was actually successfully reduced in the subsequent new text by 75%. As for other error types such as lexical items and verb prepositions, none of those occurred in the new text bore any relation to those in the original one. These valuable details and insights were lost in the global error rate comparison that Truscott and Hsu (2008) had opted for. If not for Bruton's (2009a) meticulous analysis of the "sequential details that reveals the underlying inconsistency in the design and the argument" (p. 139), the reader would have been left with only one interpretation of the results as presented by the authors and perhaps would have even gone on to accept Truscott's view that error correction is indeed unhelpful and should thus be avoided.

CONCLUSION

Since Truscott (1996) called for the abandonment of grammar correction in L2 writing, the past 16 years has seen a proliferation of studies on written CF. This paper reviews the empirical evidence provided by 15 recent studies, with the aim to gain more insights into the state of affairs regarding the written CF debate. Overall, these studies have shown that error correction is effective for short-term revision, and increasing evidence also suggests that both focused and unfocused CF can facilitate the acquisition of either a single or a wider range of grammatical features. The findings of some studies also challenged Truscott's hypothesis that grammar correction may bring about detrimental effects such as reducing the complexity of students' writing or that mere writing practice would be more valuable than providing CF.

Despite the promising results, there are still lingering concerns. For one, most of the studies dealt with English learners in an L2 context, and many of them only focused on the simple usage of the English article system. It is therefore difficult to generalize the results to other contexts, learners, or linguistic features. Now is time that the research community puts the

narrow article investigation to rest and turns its attention to other syntactic and lexical features. Given the scarcity of research on unfocused CF and given the positive findings revealed by Van Beuningen et al. (2012), future studies should continue to explore the potential of comprehensive error correction. In addition, to better understand how learners process CF and how the depth of their processing affects their interlanguage development, more qualitative research examining learners' uptake of CF is also needed.

Drawing on the available research evidence and discussions of written CF, I would like to make the following pedagogical suggestions. First of all, written CF can help learners develop greater control over the use of grammar, so the teacher should not hesitate about providing it. But s/he should also know that, however great its potential might be, CF will not cure all of the learners' language problems. Its success will be determined by a host of other factors such as error types, feedback characteristics (e.g., explicitness, clarity, consistency), the learner (e.g., language aptitude, proficiency level, motivation, developmental readiness), the teacher (e.g., personality, teaching methods), and the learning context. It is also vital that the teacher make sure that learners attend to the CF by taking such measures as requiring them to revise the original text. Having learners process feedback is essential because as Schmidt's Noticing Hypothesis (1990) suggests, linguistic elements in the input cannot be learned unless they have been noticed by the learner. Last but not least, it is important for the teacher to remember that "second language acquisition is slow, gradual, and often arduous, and that corrective feedback is only one of the many factors that contribute to that process" (Guenette, 2007, p. 52). Sometimes it may take an extended period of time before one begins to see CF bear fruit. Patience and persistence are key.

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*, 227–257.

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

Bitchener, J. & Ferris, D. (2011). *Written Corrective Feedback in Second Language Acquisition and Writing*. New York, NY: Routledge.

Bitchener, J. & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, *12*, 409–431.

Bitchener, J. & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, *31*, 193–214.

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

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

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

Bruton, A. (2009b). Improving accuracy is not the only reason for writing, and even if it were... *System, 37*, 600-613.

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.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97-107.

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, 353-371.*

Evans, N., Hartshorn, K., & Tuioti, E. (2010). Written corrective feedback: Practitioners' Perspectives. *International Journal of English Studies, 10*, 47-77.

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

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

Ferris, D.R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing, 8*, 1–10.

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

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. (2010). Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition, 32*, 181–201.

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.

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

Hendrickson, J. M. (1980). The treatment of error in written work. *Modern Language Journal, 64*, 216-221.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*, 83-101.

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.

Knoblauch, C. H., & Brannon, L. (1981). Teacher commentary on student writing: The state of the art. *Freshman English News, 10*(2), 1-4.

Lalande, J. (1982). Reducing composition error: 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.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60*, 309-365.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition, 32*, 265-302.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.),

Synthesizing research on language learning and teaching (pp. 133-164). Amsterdam: Benjamins.

Schmmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 128-158.

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

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

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

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

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies*, *10*, 29-46.

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 against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*, 111–122.

Truscott, J. (2001). Selecting errors for selective error correction. *Concentric: Studies in English Literature and Linguistics*, *27*, 93-108.

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, *13*, 337-343.

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

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

Van Beuningen, C. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, *10*, 1-27.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *International Journal of Applied Linguistics*, *156*, 279-296.

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

Vyatkina, N. (2011). Writing instruction and policies for written corrective feedback in the basic language sequence. *L2 Journal*, *3*, 63-92.

Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, *19*, 79-101.