Is Corrective Feedback in the ESL Classroom Really Effective? A Background Study

Gaston Bacquet*
University of Reading, UK

Corresponding Author: Gaston Bacquet, E-mail: g.bacquet@pgr.reading.ac.uk

Article history
Received: February 19, 2019
Accepted: March 17, 2019
Published: May 31, 2019
Volume: 8 Issue: 3
Advance access: April 2019

ABSTRACT
The use, range and efficacy of different types of corrective feedback in the ESL classroom have been widely researched, especially over the past 30 years. This paper attempts to dissect such research and to provide language teachers with an understanding of the background literature and theoretical views surrounding the subject. By summarizing different, often contrasting views on what corrective feedback is, how it works and how effective it is (or not), I strive to provide ESL teachers with a deeper understanding of the available literature and different perspectives that can hopefully better inform and shape our practice. The paper first provides definition of key terms related to corrective feedback, it proceeds to explore a range of theories and views and to look into the core studies regarding both written and spoken feedback and it ends with a short discussion on how these finding are relevant for ESL teachers.

Key words: Feedback, Fluency, Recasts, Metalinguistic Clues, Elicitation, Explicit, Implicit, Correction, Grammar Correction, Esl, Direct Feedback, Indirect Feedback

INTRODUCTION
The role and effectiveness of corrective feedback has been an area of keen interest for ESL/EFL specialists, teachers and researchers, which can be ascertained by the large amount of literature on the subject, especially since 1997, date in which the number of studies increased by 15% compared to the previous decade (Li, 2010).

Previous studies have covered areas as diverse as the classification of feedback in the classroom (Sarandi, 2016), its effectiveness in SLA (Li, 2010) and the types of corrective feedback that yield best results in terms of corrected output (Norris and Ortega, 2000).

However, oral and written feedback have been explored, researched and documented separately; this paper will attempt to dissect the literature in both areas, explore the question “Is feedback in the ESL classroom actually effective”? and if so, which type of feedback seems to yield the most improved linguistic output both in spoken and written form.

This essay will draw on that previous research, review the existing narrative on the subject of oral and written corrective feedback and draw its own conclusions as to how effective such feedback actually is, as well as the type of feedback that according to accurate measurements produces evidence that learning has occurred if that is indeed the case.

DEFINITION OF TERMS
Lightbown & Spada (1991) define corrective feedback as “any indication to the learners that their use of the target language is incorrect” (p.171).

According to Lyster & Ranta (1997), and using terms that have come to be generally accepted in SLA terminology, there are six generally accepted types of feedback:

a. Recasts, a reformulation of the learner’s output that leaves the error out for the student to notice.
b. Repetition, where the teacher draws the learner’s attention to a specific linguistic feature in an utterance by repeating what the learner said while adjusting intonation.
c. Explicit correction, where the teacher provides the correct form after it’s become clear the learner’s output was incorrect.
d. Clarification requests, where the teacher uses questions to show that an utterance has not been understood, incorrectly formulated or it needs to be uttered again.
e. Metalinguistic clues, where the teacher provides further information or comments regarding what’s been said by the learner.
f. Elicitation, where the teacher asks a questions that requires more than a simple yes or no, and gives learners time to think about it and reply to it.
(Source: Tedick & Gortari, 1998)

Recasts and repetitions are generally considered to be implicit forms of CF, since the error is not directly pointed out to the learner, whereas the remaining four are commonly called explicit forms of CF. It must be pointed out, however, that a clear distinction between explicit and implicit is not always present and some techniques for CF flow interchangeably between both depending largely on how the teacher uses them. Using a study by Chaudron (1977), Salica found three different types of repetitions which were at different times...
of a corrective nature or used for agreeing, appreciating or understanding (1981).

As far as CF on written output, the most widely used terms are:

a. **Direct feedback**, where the specific location of the error is pointed out and the correct form is provided (Sheen, 2007)

b. **Indirect feedback**, wherein the instructor points out that there has been an error but without providing the correct form (Ferris, 2002)

Rod Ellis (2009) further broke this down into **metalinguistic feedback**, focused and unfocused CF (the former focusing on one specific type of error and the latter pointing out to whatever error has been made), electronic CF and reformulation.

**THEORETICAL VIEWS AND BACKGROUND**

Long (1996) view on feedback is that the learner is faced with positive and negative evidence of language use. By positive evidence is meant models of what is considered acceptable form, whereas negative evidence refers to information provided to the learner which show what is unacceptable form. The terms defined above represent ways for teachers to provide this type of evidence.

In 1982, Stephen Krashen put forth what is known in the SLA field as the Input Hypothesis, which argues that lengthy exposure to the target language (implicit learning) is enough to acquire a second language. He further argued that explicit instruction or feedback would not bring about proficiency in the use of L2, thus rendering the idea of CF irrelevant.

In can be argued that the process of first language acquisition, where little feedback is provided and children are exposed to large amounts of input, supports this theory.

Chaudron (1988), on the other hand, suggests that feedback might be the single most important source of improvement and development available to learners. It is unknown, however, the degree to which this information actually leads towards acquisition. He also sustains that it is inevitable that students derive some kind of information from teachers, be it explicit correction, their behavior or facial expressions when questioning, and that the role of feedback should be to provide learners with tools they can use to modify their behavior (Chaudron, 1988).

In her seminal study with French Canadian immersion students, Swain (1995) demonstrated that at that scale, input by itself did not bring about acquisition as it might have been expected with younger learners, and that adult language learning is a much more conscious process.

Thus, and as a reaction to this emphasis on output, two other relevant hypothesis gained importance: one was the Output Hypothesis (Swain, 1995) and the second was the Noticing Hypothesis (Schmidt, 1990).

The first argues that learning occurs when a learner notices a gap in their linguistic knowledge and by noticing it and trying to fill that gap, he learns something new about the language.

The second proposes that the learning of specific linguistic items cannot take place unless they are noticed, and that although noticing by itself is not enough to acquire language it provides with a solid starting point (Schmidt, 1990, 1995).

Both of these hypothesis would indeed require either explicit or implicit feedback as an awareness-rising tool.

Gass (1991) also pointed out that for acquisition to take place, learners must be able to notice the mismatch between their own interlanguage and the input they are receiving, and that this noticing must be a conscious process in which feedback plays an important role.

Gass and Varonis, in their 1994 paper, as well as Ellis (1991), argue in favor of CF as a tool that might trigger changes in the learners’ knowledge and use of L2 which “may show up at a later point in time” (Gass & Varonis, 1994).

Several of these studies (Lightbown & Spada, 1991, Swain, 1995 and subsequent studies by them) have been carried out in Canada, in instructional settings that favor a communicative approach, where meaning-based exchanges take precedence over form-focused instruction. It is a telling fact that what led to these studies was the observation that students were not becoming more proficient under that approach, thus making it necessary to conduct research to find out why. In fact, all the studies mentioned before were conducted under a form-focused instructional model, where feedback does become necessary.

A number of researchers, however, have held discordant views not only concerning the type of feedback that works best in SLA, the context in which it should be provided and finally whether or not CF is useful at all.

Hendrickson (1978) suggests that feedback should be constrained within the boundaries of form-focused tasks that require grammatical accuracy, and to leave meaning-focused tasks (where communication is the main goal) alone.

He further suggests the three types of errors that should be treated:

a. Those that impair communication to a great extent
b. Those that have stigmatizing effects on the recipient
c. Those that are occur frequently in both speaking and writing

Hillocks (2006), sustains the notion that meaning-based instruction gives students the power to work with their own ideas and thus CF (best suited and effective in form-focused instruction) becomes unnecessary. Some time earlier, Schafer (2004) held the view that form-based teaching and CF provided students with a “false and anachronistic view of the language”, in reference to the Audio-Lingual method and the unnatural characteristics of the instruction-based setting.

Finally, there are those proponents of self-correction or peer feedback. Makino (1993), suggests that self-correction has two advantages: the first is that it gives learners the chance to think about their own output, be it spoken or written, as well as pay closer attention to structural form. The second one is that it helps activate linguistic competence on part of the learners.

All of the above refers to oral CF, where most of the research has been done. As I mentioned earlier, This essay will also look into the use and role of written CF; although research in this area is less abundant, there is some relevant information worthy of mention which ties into the aforementioned studies.
In an influential paper from 1996, John Truscott called for grammar correction to be abandoned at a time where there was little to no evidence as to the effectiveness of correction to students’ written output.

Part of his hypothesis was based on the work of other researchers (Krashen, 1992; Leki, 1991; VanPatten, 1986a, 1986b), who had also concluded that grammar correction was ineffective. He further cites a study done by Semke (1984) which ascertained that students who received content-only feedback outperformed those who did receive grammar correction in post-tests.

His main contention was that grammar correction did not result in better writers and that in some cases it turned learners into worse writers, in which case it not only did not help but it was also harmful (Truscott, 1996). He further sustained that one of the main reasons for it being harmful was an affective one: students perform better when not under stress and when they feel relaxed and confident, elements which are disturbed when corrective feedback is given (specifically grammar correction). From an language-acquisitional point of view, he pointed out that teaching practices that depend on transferring knowledge rather than focusing on the developing the language don’t show very positive prospects, and that grammar correction falls into the “transfer of knowledge” slot.

His review paper led to a myriad of subsequent studies by other scholars. The most immediate response came from Dana Ferris, who contested Truscott’s paper in 1999 and argued, first of all, that the effectiveness of feedback depended largely on the type being provided and second, on the manner in which is provided.

Later, less skeptical studies such as those done by Sheen (2007) and Ellis (2008) have tried to ascertain if in fact acquisition is possible and have attempted demonstrate that there is a value to CF and that it can lead to acquisition by following certain parameters.

**REVIEW OF STUDIES**

As pointed out earlier, Krashen (1982) considered CF to be harmful when learners were engaged in tasks that favored communication (meaning) over linguistic accuracy. He theorized that providing correction during this type of task activated the affective filter and made it more difficult for students to communicate, given their possible fear or embarrassment.

Tomasello and Herron’s Garden Path theory shed some light on the possible effects of corrective feedback through their 1988 research project, where learners were divided into two groups and taught various “exception to the rule” items. One group was taught explicitly and the other was led into a process of induction and deduction (“down the garden path”). The project aimed at investigating the process of overgeneralization (like adding and ed ending to verbs when using them in their simple past form, even irregular ones) and it proved through test results that those students who had been exposed to both the rules and the negative evidence of their output outperformed those who had been received explicit instruction by up to 24% (Tomasello and Herron, 1988). They replicated their study with other groups, obtaining similar results. Although not conclusive (by their own comments), it does provide some evidence that corrective feedback plays an important role in second language learning. It must be pointed out that their study was carried out in a classroom setting where focus-on-form was the objective. It remains debatable if such treatment would have produced the same results in a more naturalistic setting or a classroom where the focus of instruction was on communication.

Further to Tomasello and Herron, Lightbown and Spada (1990) conducted a study with French-speaking students in Canada over a period of 5 months. These were 5th and 6th grade students immersed in an ESL program after having been exposed evidence that the learners’ syntax and morphology were clearly below what could have been expected after years of such immersion (Harley & Swain, 1984; Swain 1984, 1985).

The purpose of their study was to look into the effects of form-focused instruction and corrective feedback in the context of CLT (Communicative Language Teaching). As I pointed out earlier, the instructional model in Canada provided plenty of opportunity for meaning-based, communicative-type tasks as instructed by the Canadian Ministry of Education, with teachers providing very limited correction to students’ output (if at all), and Lightbown and Spada suggested that the inclusion of form-focused instruction may be useful in improving the development and accuracy of such output. Their study posed the following question:

**“Are there clear differences in learner language outcomes that may be related to difference in instruction?”**

The research was conducted in four different classes which underwent different types of instruction, but where at least 70% of if was meaning-based. They analyzed inflectional morphemes (plural “-s”, progressive “-ing”), adjectival/noun placement and possessive prepositions.

Results of their research reveal that students in the class where the teacher devoted little to no time to grammar instruction showed the least accurate output, and inversely, the class where the most form-focused instruction was provided showed the most improved use of the forms being taught, including presentational forms (“there is...” versus “I have...”).

Notwithstanding the fact that there are self-admitted limitations to their study (size, number of subjects, setting and context), the results do confirm that corrective feedback can be useful with certain aspects of the language form (Lightbown & Spada, 1990).

Susanne Carroll and Merrill Swain produced a study in 1993 which sought out to demonstrate whether or not grammatical learning did in fact occur as a result of feedback, by investigating the effects of both explicit and implicit forms of CF. They argued that although the best way to provide feedback on form was through a thorough and detailed explanation, this would be very difficult to understand for untrained learners. One the other hand, too implicit a correction, by failing to indicate the source of the error, might prove useless unless limited in scope and thus reducing the cognitive load on the learner, who could therefore draw some conclusions as to the (in)correctness of form (Pinker 1989, Carroll & Swain 1993). Their study involved 100 adult ESL learners who
spoke Spanish as a first language and it focused on English dative alternation, which according to their own previous research was a difficult item for adults to grasp. In fact, other studies (Hawkins 1987; Mazurkevich, 1984a, 1984b) show that even highly proficient, native-like speakers tend to overgeneralize the rules governing this linguistic item. Learners were divided into five groups of 20 students each, each of which received a different feedback treatment and group Z (the comparison group) receiving no treatment at all.

Results were measured by having learners take two recall tests within a week of each other, after the experimental sessions were concluded (these sessions were preceded by a listening test and a background questionnaire).

This table, extracted from their original study, shows the mean results per feedback item obtained.

The most striking finding remains the fact that group A, which was the only one receiving explicit metalinguistic correction, outperformed the other four and that according to their own results and findings, providing indirect feedback did not help learners as much in producing the targeted structure.

This of course, and according to their admission, raises the possibility that the claim that learning about the language does not promote language learning might be incorrect (Carroll & Swain, 1993)

Further, well-reputed studies (White, 1991; DeKeyser, 1993; Ellis, 2006; Lyster & Ranta, 1997; Ortega, 2001) have explored various aspects of corrective feedback as given by the teacher and yet they all delved into the oral production aspect of the language with little attention paid to aspects such as writing output, or even self or peer correction. They additionally focused on form and accuracy as their main benchmark and have measured acquisition by assessing the use of specific grammatical features.

There are important differences between oral and written CF: while the former happens (generally speaking) almost immediately after the error has been made, the latter is delayed. It also imposes less of a cognitive load on the learners, while oral CF demands a greater use of students short-term memory (Sheen, 2007).

In a controversial paper from 1996, Truscott made the now well-known statement that grammar correction should be altogether abandoned from writing courses due its ineffectiveness and possible harm. This posed an obvious challenge to researchers in the field, and yet, as controversial as it was, his theories, as pointed out by Ellis (2008), put forth two valid points: one is that several studies done up to that point which supported the notion and validity of CF lacked a control group, and thus, there was no way to prove whether or not acquisition had happened as a result of such feedback; the other point being that the research did not require for students to produce a piece of writing anew, but just a new draft of work previously done.

Ferris’ response to Truscott in her 1999 paper makes a case FOR grammar correction, though agreeing with Truscott on the fact that no single form of correction can possibly be effective in treating syntactic, morphological and lexical errors simultaneously (1999).

In defending her stance, she points out that indirect correction works better than direct correction and that this correction works better when it relates to word patterns or rule-governed linguistic items.

She also points to the unwillingness of teachers to provide thorough grammar explanation as a possible source of failure, and to unpreparedness or lack of strong theoretical knowledge on the part of the teachers as another one. She details her experience in training students in the use of editing techniques and further explains a system she created to provide error correction on non-idiomatic, idiosyncratic errors, encouraging instructors to move away from the “one size fits all” approach of using a uniform coded system for error marking (VT for “verb tense”, WC for “word choice”, SV for “subject-verb agreement”, and so on).

In 2007, Dr. Younghee Sheen conducted a study in the US to determine whether or not corrective feedback improved written accuracy. Her research was carried out with 111 community college students from diverse ethnic/language backgrounds, age 21 – 56, all of whom had the same level of proficiency and received the same amount and type of instruction that involved identical materials and tasks, which made it easier to ascertain the effects of corrective feedback alone without any other variables at play.

Through the use of four types of tests (speeded dictation, writing test, error correction and language analytical ability) and two different types of CF, direct-only and direct metalinguistic, the study concluded that the latter did in fact produce a higher degree of accuracy on the one linguistic feature being worked on (definite and indefinite articles). In fact, the groups the received direct metalinguistic feedback scored higher than the control group and the one which received direct-only correction with no metalinguistic correction given by up to 13 percentage points.

All the above fits with earlier research and assertions by Stakes (1975), Fanselow (1977b) and McTear (1975), all of whom pointed to the need for:

a. feedback to be consistent within one type of error. 

b. a verification of the learner’s perception, and 

c. evidence of the learner’s understanding.

Another relevant study conducted by Rod Ellis et al in 2008 also produced some evidence that narrow, direct, focused feedback does lead to a higher degree of accuracy, although not as high as the established criteria of 90% set out by Brown in 1973. His research was carried out in a national university in Japan with a much smaller sample: 49 students,
all of whom had received about 800 hours of classroom instruction over the course of 6 years. The study called for all of them to write exactly the same narrative and receive feedback from the same teacher.

Unlike previous studies on the role of written CF and taking Truscott's point into consideration, this study did require students to produce an entirely new narrative and not just a redraft.

The methodology was similar to the one used by Sheen (pretest-treatment-posttest) but the latter only considered two types of test: narrative and error correction.

Results show that although there was no significant difference between those groups treated with direct focused correction (only the targeted feature was corrected) and direct unfocused correction (different types of errors were made evident, including the targeted form), there was a very obvious improvement of those groups compared with the control group which received no treatment at all.

Later studies (Bitchener, 2008; Bitchener & Knoch, 2010; Sheen, 2010) have gathered evidence to prove that CF (especially direct, focused and metalinguistic feedback) does help learners produce more accurate narratives in subsequent attempts (and not just in redrafts, as Truscott contended).

One hindrance of such research, however, is the limited linguistic scope of the studies, most of which focused on the use of articles, past simple forms and prepositions. In fact, according to Bitchener and Knoch (2010), although research subjects scored considerably higher than the control groups when using the first two, there was no marked improvement on accuracy when it came to use the latter. More research is necessary to ascertain the effectiveness of these methods with more complex grammar/linguistic items.

One last element to consider and which has been widely researched over the past 20 years is the use of peer-feedback in ESL writing classes, especially (but not only) when involved in using the process approach (Meriwether, 1997).

The perceived advantages of peer feedback have been documented in studies by Clifford (1981), Purves (1986), Berlin (1988) and Fox (1990), and range from feeling less judgmental and confusing than that of the teacher's, to more helpful (1988) and Fox (1990), and range from feeling less judgmental and confusing than that of the teacher's, to more helpful (Leaver, 1987; Murphy, 1996).

It is interesting to note that all 16 subjects chose peer feedback over teacher feedback, noting, however, that they LEARNED more through the latter but were more comfortable with the former.

Further analysis of this study show that all students expressed a higher degree of stress when faced with their instructor's comments compared to those of their peers.

It is important to note that all of the above research points to affective factors and none have measured elements such as improved accuracy or greater acquisition, so further research needs to be conducted to measure the effectiveness of peer feedback in terms of improved output beyond the learners' own perceptions of its usefulness, and the research-supported fact that peer feedback creates a social contract that enhances students' attitude towards writing and brings about a beneficial and advantageous learner setting (Chaudron, 1984; Mendonca and Johnson, 1994; Nelson and Murphy, 1996).

FINDINGS AND DISCUSSION

This study, which aimed at investigating the effectiveness of oral and written feedback in the ESL classroom according to existing literature, began with the question “Is feedback in the classroom actually effective? And if so, which type of feedback seems to yield the most improved linguistic output?”. The following conclusions can be drawn:
**Oral Corrective Feedback**

As we’ve seen previously, research into this area has produced some shared results but is hardly conclusive. There is enough evidence to sustain that a non-corrective learning environment, set up for communicative language learning, does improve students fluency and their meaning-based output (Krashen, 1982, 1985; Hendrickson, 1978, 1980; Shafer, 2004). It is also safe to say that unless some type of corrective feedback is provided, learners are unlikely to become more accurate language users or to acquire specific form-based features (Lightbown&Spada, 1990; Swain, 1995; Ortega, 2000).

As to the type of feedback being used and its effectiveness, studies have shown that recasts are in fact the most widely used of error correction, but not necessarily the most effective (Lightbown&Spada, 1993). Although some scholars had held the view that metalinguistic feedback did not lead to improved and more accurate output, Carroll and Swain’s 1993 study shows evidence that this assertion might not be entirely accurate and that it does in fact help learners become more competent language users. So although implicit error correction is greatly favored in the current ESL world, metalinguistic feedback has its own place in the classroom (Carroll&Swain, 1993; Sheen, 2010).

Of course, how this works depend greatly on the instructor and the type of instruction they choose to provide (meaning-based, form-focused or an integrated model such as the one attempted by Lighbown and Spada in 1990) and how consistent they are with the feedback they provide within a specific type of error (McTear, 1975)

**Written Corrective Feedback**

As has been explored here, there’s also enough evidence showing that, although John Truscott had some valid points concerning the shortcomings of available research studies on the effectiveness of grammar correction (no control group, “one size fits all” correction, no follow-up to ascertain whether or not acquisition had taken place), some type of feedback is necessary if one wants learners to write more accurately, at least in a product-approach writing setting (Sheen, 2007; Ellis, 2008, Bitchener&Knoch, 2010).

An additional element to consider is the type of feedback given to learners and how this feedback is provided (Ferris, 1996, 2007). Several researchers have gathered evidence that direct, metalinguistic, focused correction help learners produce more accurate narratives on subsequent tasks and not only in redrafts, as Truscott had suggested.

Further research needs to be conducted to explore if corrective feedback has its use beyond acquiring targeted language items beyond those that have been tested (articles, prepositions, verb tenses), but what has been done so far does show that students have been able to acquire those linguistic features (Sheen, 2007, Ellis 2008).

On the issue of peer feedback, there is not enough evidence to sustain that it aids either acquisition or accuracy, but research does support its use on lowering anxiety, affective barriers and stress that might be present when faced with correction from the instructor (Zhang, 1995; Rollinson, 1998, 2005), and as such it becomes a valid instructional tool in ESL writing classes, especially while engaged in process writing.

**CONCLUSION**

In addition to the research that has been done, a further conclusion that can be drawn from this study is that in the end, it is largely dependent on each teacher to analyze their individual teaching context (students’ ethnicity, educational background, proficiency level, learning aims, instructional focus) to determine which type of feedback, if at all, is to be given. Zhang’s research, for instance, was entirely conducted on students of Asian background, whose educational model differs from Western setting.

None of the research referenced here is prescriptive and it has been done across a variety of settings and L2s, from English and French to Spanish and Japanese, which is partly the reason previous studies are indeed helpful as a guideline but hardly conclusive as a rule of thumb (no that any of the researchers mentioned here have ascribed to that notion).

Is our aim to improve fluency for communicative purposes? Is it to help learners achieve a higher degree of accuracy? Are our students writing for academic purposes, or to become better writers in general? What type of learners are we working with? Are they college students, school children, adult immigrants? These are some of the questions that fare beyond the scope of this particular study and which further research will need to answer; it is also up to each instructor to delve deeper into those factors and hopefully find further answers that will help learners achieve their aims, not only from the point of view of feedback but the type of instruction we are providing.

**REFERENCES**


UNANNOTATED REFERENCES
