
Abstract

Given the fact that the incorrect use of grammar negatively affects one's writing performance, most EFL Iraqi university-level students suffer from different grammatical problems, which in turn, lead to inefficient writing output. Accordingly, the researchers find that it is important to identify and fix the most crucial items that hinder students to perform well while putting pen on paper. One of the approaches to reduce the number of students' grammatical errors is by means of supporting them with feedback. The current study tackles two types of feedback, i.e. traditional and metalinguistic feedback; following Ellis (2008) as the theoretical model of the study. This experiment was conducted on 60 EFL Iraqi university-level students (3rd year – evening classes). The participants were randomly put in two main groups, and a subgroup: 20 in a control group and 40 in an experimental group. The latter was divided into two smaller groups: 20 in a descriptive group, and 20 in a code group. The results of the study found out that participants in the control group made a very moderate development in their writing performance, while those in the experimental group made a noticeable improvement. The study also found out that the descriptive group scored slightly higher marks than the code group in Posttest 1; however, the latter participants outperformed those in the former group in Posttest 2.

Keywords: Writing Performance, Teacher's Corrective Feedback, Metalinguistic Feedback, Traditional Feedback, EFL Iraqi University Students

الملخص

بات من المسلمات أن الاستخدام الخاطئ للنحو يؤثر سلباً على أداء الكتابة حيث يعاني معظم طلبة الجامعات العراقية الدارسين للغة الانكليزية بوصفها لغة أجنبية من مشاكل نحوية شتى. وفقاً لذلك، وجد الباحث أن من الضروري تحديد المشاكل الأكثر تأثيراً على أدائهم الكتابي ومساعدة الطلبة لحل تلك المشاكل. تتمثل إحدى تلك الطرق في دعمهم بالتغذية الراجعة كونها تمثل مفتاحاً أساسياً لمساعدة الطلبة بمعرفة مدى تقدمهم ومنحهم كذلك الإحساس بالإنجاز والتقدم في الأداء. تتناول هذه الدراسة نوعين التغذية: التقليدية والميتا لغوية، متبنيه اليس (2008) كإطار نظري لتحقيق اهدافها وذلك من خلال اختبار ستين طالباً عراقياً جامعياً الدارسين الانكليزية بوصفها لغة أجنبية (المرحلة الثالثة/ الدراسة المسائية). قسم الباحثان هؤلاء وبشكل عشوائي إلى مجموعتين رئيسيتين وأخرى ثانوية: عشرون منهم كانوا في المجموعة الضابطة وأربعون في المجموعة التجريبية. قسمت الأخيرة إلى مجموعتين: عشرون طالباً كانوا في المجموعة الوصفية و عشرون آخرون في المجموعة الرمز. أجرى الباحثان اختباراً تمهيدياً لتحديد المشاكل النحوية التي تسهم سلباً على جودة كتابتهم كما أجرى أيضاً أربعة اختبارات بعد ذلك لغرض مراقبة مدى تحسن الطلبة ومدى استجابتهم لتصحيحات الباحث أثناء فترة الاختبار ومن ثم قام الأخير بعمل اختبارين آخرين بعد الانتهاء من فترة الاختبارات: كان الاختبار الأول لغرض التحقق من مدى استفادة الطالب من الاختبارات السابقة أما الآخر فكان أيضاً لاختبارهم بعد فترة لغرض التحقق أي الاختبارين أكثر ديمومة من غيره. توصلت نتائج هذه الدراسة إلى أن طلبة المجموعة الضابطة كانوا قد أحرزوا تقدماً طفيفاً جداً، بينما أحرز الطلبة في المجموعة التجريبية تقدماً ملحوظاً في أدائهم الكتابي. وتوصلت الدراسة أيضاً إلى أن الطلبة في المجموعة الوصفية كانوا قد حققوا نتائج أفضل بقليل من أولئك في مجموعة الرمز في الاختبار النهائي الأول بينما حققت المجموعة الأخيرة نتائج أفضل بكثير من أقرانهم في المجموعة الوصفية في الاختبار النهائي الثاني.

الكلمت المفتاحية : اداء الكتابة، التغذية الراجعة للتدريسين ، التغذية الميتالغوية ، التغذية التقليدية ، الطلبة العراقيين ، الجامعين الدارسين للغة الاجنبية.

1. Introduction

Providing students with feedback on writing assignments of English as a foreign language is an indispensable aspect of classroom pedagogy. Being a debatable issue in the field of second language acquisition (SLA), Hyland and Hyland (2006) viewed this literature as troublesome; a point that addresses an inevitable systematic corrective feedback (henceforth CF) that encompasses different types of written corrective feedback (henceforth WCF).

Educators and researchers identified different strategies to address students' writing performance. Ellis (2008) exploited two major strategies for correcting students that teachers and researchers are recommended to consider: strategies for providing feedback such as direct or indirect metalinguistic feedback; and the students' response to the feedback such as revision and attention to the given correction. These options are essential for both determining whether or not WCF is effective and if it is so, what kind of CF is most effective. Earlier, Ferris (2002) categorized feedback as either direct corrective feedback (henceforth DCF) (explicit) or indirect corrective feedback (henceforth ICF) (implicit). In the case of the former, the teacher provides students with the correct form of correction by crossing out unnecessary words, phrases, or morphemes; inserting a missing word or morpheme and writing the correct form above or near the erroneous form. In the same vein, the latter can be used as a means of providing learners with explicit guidance about how to correct their errors, particularly those who do not know how to tackle their writing sufficiency. According to Ferris and Roberts

(2001), DCF is probably better than ICF for student writers of a low level of proficiency.

A study by Sheen (2007) suggested that DCF could be fruitful in promoting learners' acquisition of particular grammatical features. Comparing DCF to ICF, she found that both types of feedback were effective in increasing students' awareness of the use of articles in subsequent writing assignments carried out right after the CF treatment. She also proved that ICF is more effective than DCF in the long term.

In the case of the latter, Ferris and Roberts (2001) stated that ICF has been made in some way or another. They asserted that it can be provided by either underlining or circling the potential error(s). These errors should be recorded in the margin beside the number of errors in a given line, or the teacher/researcher can use particular codes to show exactly where the error, as well as its type, has occurred. Similarly, Ellis (2008) added that the teacher may point out an error without offering any supplementary correction which "takes the form of underlining and using cursors to show omissions in the student's text".

Metalinguistic CF (henceforth MCF) is another strategy suggested by Ellis (2008:100). In this literature, he stated that teachers provide learners with a particular form of explicit comment regarding the nature of the error they make. He also divided this type of comment into two forms:

The use of error codes. These consist of abbreviated labels for different kinds of errors. The labels can be placed over the location of the error in the text or in

the margin. In the latter case, the exact location of the error may or may not be shown. In the former, the student has to work out the correction needed from the clue provided while in the latter the student needs to first locate the error and then work out the correction

Mostly advantageous, Nicole and Macfarlane-Dick (2006) believed that MCF increases students' self-regulated learning when they receive this type of feedback from their teacher. They, as well, are required to fix their errors based on the written feedback provided.

Trupe (2001,1) emphasizes:

Teaching writing must involve both process and product. Teachers should first focus on the organization of the writing. As the next step, they should deal with grammatical problems seen in writing. When students are not good at organizing their ideas, the teacher should deal with this before moving on to grammatical mistakes

He pinpointed that the former appears to be the mechanical aspects of writing given that it focuses on grammar and syntax wherein learners should imitate a particular writing model while putting pen on paper. This approach primarily requires the teacher to focus on correctness and form of only the final students' products without teacher's intervening during the initial stages of writing. On the contrary, the latter approach focuses on how ideas are comprehensively developed and

formulated in writing. It further gives students room to comprehend the meaning and guides them to self-discovery. This approach requires the teacher to intervene and guide students during the initial processes of writing without correcting their errors until the final product.

In this work, the researchers amalgamated the product-oriented approach and the process-oriented one. These two approaches often lead to the reduction of errors and clear expression of ideas. Therefore, they are both required to assess the students' writing performance.

2. Review of Related Literature

This section is dedicated to reviewing studies on different types of CF with a special focus on the MCF since it effectively helps students understand the study topic, guides them to know how to run over their errors, and enhances their learning performance. Lalande (1982) conducted an experiment to include a group of US students of German. He divided the participants into two groups: the first group received direct correction, while the second one was provided with the error code to rewrite. The result of this experiment showed that the former group made more errors in their required assignments at the end of the semester, while the latter group improved their grammatical accuracy on subsequent writing assignments. But, the difference between the two groups had not recorded a statistically significant difference.

Another study was conducted by Ferris and Roberts (2001) whose participants were divided into three

groups. The participants, in the first two groups, received CF either on the type of the errors they made or on location. The participants in the third group got no feedback at all. Accordingly, the findings showed that the first two groups outperformed the group with no feedback. However, neither group showed significant development in their writing performance over the period of the study.

Bitchener (2008) examined the efficacy of WCF to 75 international ESL students in Auckland, New Zealand. He put his subjects into two groups: experimental and control groups. The former was divided into three subgroups: the first one received oral metalinguistic explanations; the second group was given written metalinguistic explanations; group three got DCF. The results of his study showed that the writing accuracy of students who received written metalinguistic explanations in the immediate Posttest outperformed the rest of the groups.

In a recent study, Gholaminia et al (2013) conducted a study to examine two groups of EFL Iranian learners. The first one received traditional corrective feedback (henceforth TCF), while the second one got MCF. The study revealed that the experimental group subjects, who were assessed with implementing metalinguistic code correction, achieved better results compared with those who got the traditionally instructed feedback in their Posttest.

3.The Aims of the Study

This paper aims to examine the effect of teachers' TCF and MCF (direct and indirect) on EFL students' writing performance. Moreover, it sheds light on a long-

term CF which will be carried on through two un-sequential Posttests. Given that this tendency has not been processed so far in the Iraqi context, it motivates the researchers' interest to articulate this procedure to help validate which CF approach is apt to enhance students' writing accuracy over time.

3.1 Statement of the Problem

The present study highlights the importance of feedback to students to improve their writing performance. This can be done through a number of tests to measure which feedback is more useful: the TCF or MCF, and any of the latter: the code or descriptive feedback. So that, teachers may take advantage of the recommendations of the study when evaluating their students' writing performance. Through the related literature involved in this paper, it is found that the Iraqi context, particularly university level students, has not been tested so far. Thus, it generates a gap deserved to be bridged to reduce the targets' grammatical writing errors.

4. Methodology

The study exploits the difference in the effect of the TCF and teachers' MCF on EFL students' writing performance. This section presents the research design (section 4.1) which outlines the approach of this paper. It also includes the research questions (section 4.2) which tackle particular gaps that are needed to be bridged. This section also presents the participants engaged in this study (section 4.3) as well as the instruments used (section 4.5) and the procedure followed (4.4) to accomplish this work.

4.1 Research Design

The present work is quantitative in nature since it delves in depth into the difference between teachers' TCF and MCF. Furthermore, the data of this study was gathered through the use of instruments that were used to estimate the participants' English proficiency level and measure their language potentials. A multivariate analysis of variance (MANOVA) was conducted to assess if there were significant differences in the linear combination of Pretest and Posttest 1 between the levels of Feedback. In addition, an analysis of variance (ANOVA) was conducted for each dependent variable to examine the effects of Feedback on Pretest and Posttest 1.

4.2 Research Questions

This work seeks to answer the following research questions:

1. Does MCF have a noticeable improvement in the Iraqi EFL university students' writing performance if compared to the TCF?
2. Which MCF is more effective in assessing students' writing performance across time, the descriptive or code written feedback?

4.3 Participants

The current study was conducted in the Iraqi educational context. The number of participants is 60 students (19 males and 41 females). They were studying

at the University of Basrah, College of Arts, Department of English, Third-year (evening studies). This choice was based on the fact that one of the researchers was himself the participants' instructor. Their ages ranged between 21 to 40 years old. All of them are Arabic native speakers. Their selection was based on Nelson English Language Proficiency Test. This test helped the researchers find out whether or not the participants were at the same level of proficiency in the English language. In the first place, they were 82 (class A and B); however, after processing the intended test, only 60 students were found legible to undertake the current experiment, being the most homogenous ones in both classes. In other words, the highest and lowest scores were not included in this paper. But, they were later provided with a list of the summary of the study after each processed session so as to let them take advantage of the results obtained.

4.4 Instruments

Four instruments were involved in the present paper:

1. Nelson English Language Proficiency Test was adopted before the Pretest to estimate the proficiency level of the study subjects. This test serves the purpose of evaluating the students writing ability.
2. Pretest (Writing Pretest) was the second instrument conducted in this study. The key purpose of this test is to measure the students' linguistic performance, and what they know and can do before enrolling in the very experiment. A prominent way to do this is to use a writing Pretest that assesses students' proficiency in the targeted skill. The participants were allocated to write a

narrative essay on a topic selected from their coursebook (Essay and Letter Writing- by L.G Alexander, 1965); the time given was 30 minutes; the number of words required was between 150 to 200.

3. Four in-tests were also conducted in classes to trace students' responses and to monitor their writing performance. Notable are the various selected writing topics as all of them were chosen from their coursebook, referred to earlier, so as to cover the already designed syllabus items during the course.

4. Two un-successive Posttests were used in this paper. The importance of these tests is to determine how much students have progressed and overcome the deficiencies in their writing in the short and long time periods by the virtue of feedback correction.

After the treatment sessions were made, the participants were given another topic to write about (Posttest 1). Similarly, the number of words, the time allocated, the instructions, and the topic, were all adopted from the previously mentioned coursebook so as to keep the targets close to the course objectives. Besides, the reliability of the tests and their selected topics and contents were consulted with *three university professors at the Department of English, College of Arts, University of Basrah, Iraq to validate and evaluate the context of the selected items.

Name	Narrow Field	Years of Experience
Assist. Prof. Dr. Ali Qasim Juda	English Language Teaching	27
Assist. Prof. Dr. Samir Talib Dawood	Renaissance Drama	12
Lecturer Dr. Mohammed Qasim Zboon	Critical Thinking and Pedagogy	11

4.5 Procedure

After the Pretest, the participants were randomly divided into two main groups: the control group with TCF (either by underlining or circling their errors) and the experimental group with metalinguistic WCF, and were again randomly divided into two subgroups: the first group was provided with a description of their errors in the margins, while the second one received error codes above their mistakes. The latter group participants were provided with a list of error codes and explanations, assessed with examples extracted from their test sheets, to help them respond properly to their errors. Notable is that each group equally contained 20 students.

It is worthy to mention that the study subjects were aware of the writing skills as part of the first and second-year courses in their university studies. At the beginning of the course, the researchers rapidly reviewed the previously studied course clauses to make sure they could match with the intended experiment. Afterward, they were introduced to the types of WCF which helped them follow up the codes and explanations. For further clarification, the students with the experimental explanation treatment were introduced to the explanation of the meaning of each CF, and those with the error code experimental group were also presented to what each given code meant. The criterion under which

the researchers evaluated the participants' test sheets was by marks from 0 to 10. However, these marks were not given to the students so as not to confuse them with the given grades. But, they were grouped and had various opportunities to discuss their errors in general.

5. Data Collection

The students' writing sheets were collected and scored, adopting the Writing Rating Scale which was developed by Gassner et al (2007) (see appendix). This scale rates the students' products depending on four criteria: task achievement, coherence and cohesion, grammar, and vocabulary. The scoring ranged from 0 for no task achievement and not enough assessable language; and 7 for complete task achievement, cohesive and coherent text, a reasonable range of grammatical structures, and vocabulary. This scale has been modified in terms of (0-10) rating, and in terms of focus, where grammar issue is the main concern in this paper.

Posttest 1 and Posttest 2 were successively processed in two periods: The first one was run in November 2017. During this period, the researchers collected the data and put them in tables to facilitate understanding the results clearly. The second stage was conducted in March 2018 where they tested the same participants and followed the procedures shown above so as to answer the second research question "*Which MCF is more effective in assessing students' writing performance across time, the descriptive or code written feedback?*"

6. Descriptive Statistics

Summary statistics were done for each interval and ratio variables. Frequencies and percentages were calculated for each nominal variable. The most frequently observed categories of Feedback were Direct and Indirect, each with an observed frequency of 20 (34%). Frequencies and percentages are presented in Table 1.

Table 1 Frequency Table for Nominal Variables

Variable	<i>n</i>	%	Cumulative %
Feedback			
No	20	33.90	32.20
Direct	20	33.90	66.10
Indirect	20	33.90	100
Missing	0	0	100

Note. Due to rounding errors, percentages may not equal 100%.

6.1 Summary Statistics

The observations for Posttest 1 had an average of 5.93 ($SD = 1.94$, $SE_M = 0.25$, $Min = 1.00$, $Max = 9.00$, $Skewness = -0.64$, $Kurtosis = -0.57$). The observations for Pretest had an average of 3.71 ($SD = 1.84$, $SE_M = 0.24$, $Min = 0.00$, $Max = 8.00$, $Skewness = 0.21$, $Kurtosis = -0.30$). The observations for Students_Numbers had an average of 31.00 ($SD = 17.18$, $SE_M = 2.24$, $Min = 2.00$, $Max = 60.00$, $Skewness = 0.00$, $Kurtosis = -1.20$). When the skewness is greater than 2 in absolute value, the variable is considered to be asymmetrical about its mean. When the kurtosis is greater than or equal to 3, then the variable's distribution is markedly different than a normal

distribution in its tendency to produce outliers (Westfall & Henning, 2013). The summary statistics can be found in Table 2.

Table 2 Summary Statistics Table for Interval and Ratio Variables

Variable	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE</i> <i>M</i>	Mi n	Ma x	Skew ness	Kurt osis
Posttest 1	5.9 3	1.9 4	5 9	0. 25	1. 00	9.0 0	-0.64	-0.57
Pretest	3.7 1	1.8 4	5 9	0. 24	0. 00	8.0 0	0.21	-0.30
Students_Nu mbers	31. 00	17. 18	5 9	2. 24	2. 00	60. 00	0.00	-1.20

Note. denotes the sample size is too small to calculate statistics.

6.2 Data analysis

To answer the research questions, a multivariate analysis of variance (MANOVA) was conducted to assess if there were significant differences in the linear combination of Pretest and Posttests between the levels of Feedback. Table 3 below summarizes the tests taken by all participants and the grades they got. The participants have been organized according to their corresponding groups. Participants 1 through 20 constitute the control group students who were given TCF; participants 21 through 40 constitute the experimental group of students who were provided with direct MCF, while participants 41 through 60 are those participants who received indirect MCF.

Table 3 Students' Progress through the Tests

No.	Pretest	In-test 1	In-test 2	In-test 3	In-test 4	Posttest 1	Posttest 2
1.	3	3	2	3	2	4	3
2.	0	1	1	1	2	1	2
3.	2	2	2	3	3	3	3
4.	4	3	4	3	5	3	4
5.	4	4	5	3	4	4	5
6.	3	2	2	3	1	2	3
7.	2	3	2	2	2	3	3
8.	5	6	5	6	5	5	6
9.	2	2	3	3	4	3	3
10.	4	3	5	5	4	5	4
11.	1	2	1	1	2	2	2
12.	3	2	3	2	3	3	3
13.	7	7	8	7	7	8	8

14.	5	4	5	5	5	6	6
15.	2	2	2	3	4	4	3
16.	6	6	6	6	5	5	6
17.	2	3	2	2	4	3	3
18.	4	4	4	4	4	4	4
19.	5	5	5	6	5	5	5
20.	2	4	4	3	4	4	4
21.	2	4	4	3	4	5	5
22.	3	4	4	5	6	6	9
23.	3	5	5	5	7	6	7
24.	4	5	5	6	6	7	7
25.	7	6	6	6	6	7	6
26.	8	8	7	6	<u>7</u>	8	8
27.	7	8	9	8	8	9	7
28.	7	7	7	8	8	8	8
29.	3	4	7	8	8	8	7

30.	2	5	6	6	9	7	6
31.	3	4	6	7	7	7	4
32.	3	4	7	7	7	7	5
33.	4	4	5	7	8	8	6
34.	5	6	7	8	7	7	7
35.	4	5	6	6	6	8	6
36.	5	6	6	7	8	7	8
37.	6	6	5	6	8	8	7
38.	3	3	5	6	7	8	6
39.	3	4	5	5	6	6	9
40.	3	4	6	6	7	8	4
41.	4	4	5	5	6	5	5
42.	2	3	3	4	6	7	8
43.	2	2	3	5	7	5	5
44.	3	3	4	4	6	7	7
45.	2	3	4	4	7	6	8
46.	4	5	4	5	6	7	8
47.	2	3	5	5	7	6	7
48.	3	4	5	5	6	8	7
49.	5	5	6	5	7	6	7

50.	4	5	5	6	7	7	9
51.	5	5	5	5	5	8	8
52.	3	4	5	5	6	7	8
53.	2	2	5	6	8	8	7
54.	2	2	4	6	7	8	8
55.	7	7	6	6	8	7	7
56.	6	6	7	7	7	8	9
57.	5	7	6	6	7	7	7
58.	5	6	6	7	7	5	5
59.	5	6	6	6	7	6	8
60.	4	5	4	6	5	4	8

Table 3 above shows that 5 participants got no advantage, 12 got slight progress, while 3 of them failed to enhance their writing ability from the Pretest to Posttest 1. The table also illustrates the progress that the same students made in Posttest 2, compared to the Pretest. Twelve participants developed very moderate modification, while 8 made no change in their writing performances. The table also displays the students' obtained results in Posttest 1 and Posttest 2. Six participants made inconsiderable improvement in enhancing their writing ability. Eleven ones got no change in their output, while the results of the rest declined in

Posttest 2. Clearly enough is the instability in the above displayed results which is back to the null input exposure referred to earlier.

Moreover, Table 3 demonstrates that there can be seen various changes made by the participants in this descriptive group. Thirteen students have got a noticeable modification found in Posttest 1 in their writing performance results, compared with the ones in the Pretest. It can also be noted that 5 students made a very slight change when results are again compared between the Pretest and Posttests, while 2 students made no change. The dissimilarity in the results was due to the students' ability to respond to the instructions provided during the test period. The majority of students in this group made fruitful advantages of the researchers' assessment.

As to the results students got in the Pretest and Posttest 2, table 3 shows that 9 of them achieved noticeable improvement in their output; 8 participants made a slight change; 2 students made no pace at all, and only 1 among the twenty failed to take any advantage of the researchers' instructions along the sessions. A comparison between Posttest 1 and Posttest 2 results is also presented in Table 3. Only two students obtained higher marks in Posttest 2 than theirs in Posttest1. Two additional participants got a slight advantage in their Posttest 2. Five of them made no change at all, and the results of the remaining 11 failed to score higher marks in Posttest 2, compared with their responses in Posttest 1.

There can also be seen that the majority of students in the experimental group (Code Group) had got a

noticeable improvement in the Pretest and Posttest 1 in their writing performance. In a further detailed description, 12 of them were successful to benefit from the researchers' scaffolded feedback presented to them. Five students made very moderate advantage, while the remaining 3 made neither positive nor negative reaction to the presented feedback. The variation in the students' responses was due to their poor potential reaction to the researchers' assessment, and/or they were not able to interact fully with the aforementioned coded errors. Note that these codes were explained and exemplified in oral and written forms during the pre-period of the current experiment.

Table 3 presents, as well, the development in the experimental group (code group) participants from an early testing stage (Pretest) up to the final one (Posttest 2). Fifteen participants had got noticeable enhancement in their writing performance. Only 3 students made a moderate change in their responses. Two of them achieved no advantage, and none of the participants declined their results in the final test. In addition, another important comparison can also be noticed in Table 3. Only 1 participant was able to score a higher mark in Posttest 2 than Posttest 1. The majority of students (9 students) were able to gain higher marks, though moderate and slight, in Posttest 2. Eight of them got equal marks in both targeted tests. While the remaining 2 students failed to change or positively improve their writing performance in Posttest 2 when compared with Posttest 1.

7. Results and Discussions

The main effect for Feedback was significant, $F(4, 112) = 12.43, p < .001, \eta^2_p = 0.31$, suggesting the linear combination of Pretest and Posttest 1 was significantly different among the levels of Feedback. The MANOVA results are presented in Table 4. In other words, the students' writing performance differed significantly between the groups. The experimental group students outperformed the control group students regarding their writing performance. These differences are due to the MCF which they received during instructions.

Table 4 MANOVA Results for Pretest and Posttest 1 by Feedback

Variable	Pillai	F	df	Residual df	p	η_p^2
Feedback	0.61	12.43	4	112	< .001	0.31

To further examine the effects of Feedback on Pretest and Posttest 1, an analysis of variance (ANOVA) was conducted for each dependent variable. An analysis of variance (ANOVA) was conducted to determine whether there were significant differences in Pretest by Feedback. The ANOVA was examined based on an alpha value of 0.05. The results of the ANOVA were not significant, $F(2, 56) = 1.96, p = .151$, indicating the differences in Pretest among the levels of Feedback were all similar (Table 4). The main effect, Feedback was not significant, $F(2, 56) = 1.96, p = .151$, indicating there were no significant differences of Pretest by Feedback levels. The means and standard deviations are presented in Table 5.

Table 5 Analysis of Variance Table for Pretest by Feedback

Term	SS	df	F	p	η_p^2
Feedback	12.81	2	1.96	.151	0.07
Residuals	183.29	56			

Table 6 Mean, Standard Deviation, and Sample Size for Pretest by Feedback

Combination	M	SD	n
No	3.11	2.05	20
Direct	4.25	1.83	20
Indirect	3.75	1.52	20

Note. indicate sample size was too small to calculate statistics.

Similarly, an analysis of variance (ANOVA) was conducted to determine whether there were significant differences in Posttest 1 by Feedback. The ANOVA was examined based on an alpha value of 0.05. The results of the ANOVA were significant, $F(2, 56) = 38.94$, $p < .001$, indicating there were significant differences in Posttest 1 among the levels of Feedback (Table 7).

Table 7 Analysis of Variance Table for Posttest 1 by Feedback

Term	SS	df	F	p	η_p^2
Feedback	126.65	2	38.94	< .001	0.58

Residuals 91.08 56

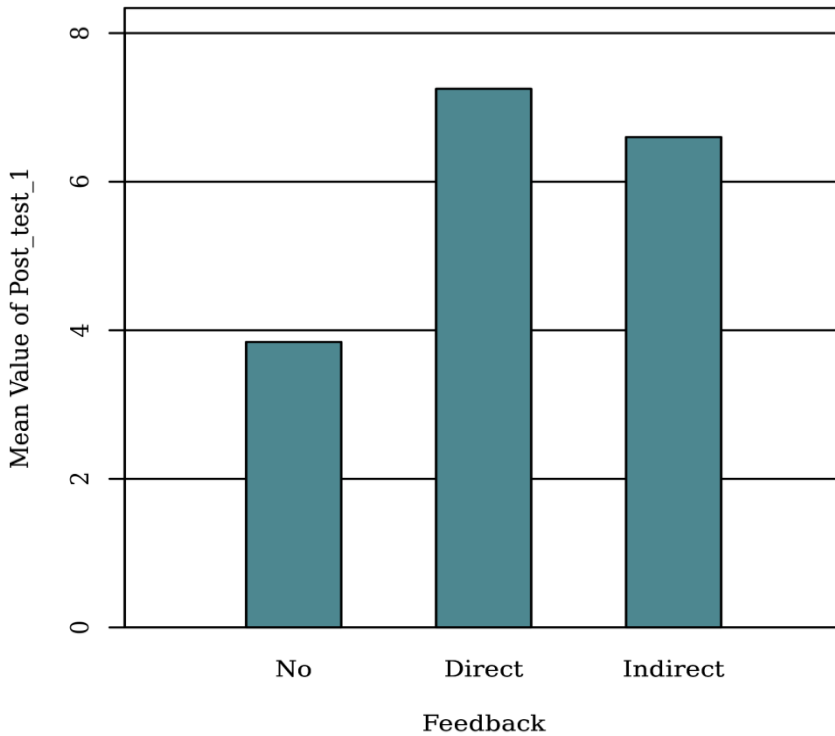


Figure 1. Posttest 1 Means by factors levels of Feedback

The eta squared was 0.58 indicating Feedback explains approximately 58% of the variance in Posttest 1. The means and standard deviations are presented in Table 8.

Table 8 Mean, Standard Deviation, and Sample Size for Posttest 1 by Feedback

Combination	<i>M</i>	<i>SD</i>	<i>n</i>
No	3.84	1.61	20
Direct	7.25	0.97	20
Indirect	6.60	1.19	20

Note. indicate sample size was too small to calculate statistics.

To answer the second research question, a simple linear regression analysis was conducted. A linear regression analysis was conducted to assess whether Feedback significantly predicted Posttest 2. The 'Enter' variable selection method was chosen for the linear regression model, which includes all of the selected predictors. The results of the linear regression model were significant, $F(2,56) = 28.79$, $p < .001$, $R^2 = 0.51$, indicating that approximately 51% of the variance in Posttest 2 is explainable by Feedback. The Direct category of Feedback significantly predicted Posttest 2, $B = 2.55$, $t(56) = 5.67$, $p < .001$. Based on this sample, this suggests that moving from the No to **Direct** category of Feedback increased the mean value of Posttest 2 by **2.55 units** on average. The **Indirect** category of Feedback significantly predicted Posttest 2, $B = 3.25$, $t(56) = 7.23$, $p < .001$. Based on this sample, this suggests that moving from the No to Indirect category of Feedback increased the mean value of Posttest 2 by **3.25 units** on average. Table 9 summarizes the results of the regression model. Therefore, based on the results discussed above, direct MCF has been found to be less effective than indirect MCF

regarding the students' writing performance in Posttest 2.

Table 9 *Results for Linear Regression with Feedback predicting Posttest 2*

Variable	B	SE	CI	β	t	p
(Intercept)	4.05	0.32	[3.41, 4.70]	0.00	12.61	< .001
Feedback Direct	2.55	0.45	[1.65, 3.45]	0.62	5.67	< .001
Feedback Indirect	3.25	0.45	[2.35, 4.15]	0.79	7.23	< .001

Note. CI is at the 95% confidence level. Results: $F(2,56) = 28.79, p < .001, R^2 = 0.51$

Unstandardized Regression Equation: Posttest 2 = 4.05 + 2.55*Feedback Direct +

9. Findings

According to the results found above, the participants in the control group, whose errors were treated in the traditional way, made insignificant improvement in their writing performance. Their poor performances are due to the fact that they were not given the opportunity to have their errors notified, described, and fixed during the test period. However, they were only exposed to the results obtained from the other two groups (descriptive and code groups). This exposure was processed by supporting them with handouts that contained the students' error corrections discussed in the classes.

As to the findings displayed earlier, the majority of students in the experimental group, particularly the descriptive group, made considerable enhancement in their writing performance. This substantial development was to the opportunity these students had when their errors were described, identified, and corrected continually along the test period. In other words, they were provided with specific written corrective information which helped them understand what they needed to do to enhance their writing performance.

The diversity of the findings obtained from the analyzed data answered the first research question "*Does metalinguistic corrective feedback have a noticeable improvement in the Iraqi EFL college students' writing performance if compared to the TCF?*". In that, the participants in the experimental groups who received different types of metalinguistic feedback made a recognisable improvement in their writing ability. Meanwhile, those in the control group who received no feedback made less advantage in their writing performance.

The findings of this study are compatible with Sheen's (2007) and Ellis's (2008), which were both displayed earlier in the introduction section. They together collectively indicate that DCF has the advantage to provide learners with explicit guidance about how to correct their errors. These results are also in accordance with Gholaminia, et al (2013) who found out that teaching different writing tasks through MCF can help the learners to make fruitful use of these tasks more effectively than those who receive no feedback.

As far as the analysis of the data shown above is concerned, it answers the second research question *“Which metalinguistic corrective feedback is more effective in assessing students’ writing performance across time, the descriptive or code written feedback?”*. It is essential to mention that during the gap months between Posttest 1 and Posttest 2, the participants, generally, were not given any new instructions or exposed to any modified input. This procedure was to test the significance and validity of the two given types of metalinguistic feedback. The analysed data proved that the results of the students in the code group outperformed those in the descriptive group. This finding is accredited to the students’ self-assessment while delving into decoding their error codes during the test period. Generally, when students know that the correction of their writing is to be based on code correction, they feel more responsible for their learning quality.

The above mentioned finding corresponds to Ferris and Roberts’s (2001) who believed that indirect (code) metalinguistic feedback brings greater benefit to learners if compared with the direct (descriptive) one. It also goes in line with Norrozizadeh’s (2009) finding who stressed that ICF motivates students to be more independent in the learning process, which by the end leads to long term learning.

10. Conclusions

The present study investigated the effect of different types of teachers’ written corrective feedback: (Traditional and Metalinguistic: Descriptive and Code) on

students' writing performance. The experiment tested 60 EFL Iraqi University level students' (3rd-year students-Evening studies) writing accuracy. The 60 participants were randomly put into two main groups. Twenty students were in the control group who had neither type of feedback but were only supported with handouts that contained the class discussions and the students' corrections. Forty participants were in the experimental group; 20 of them were in the descriptive, and the rest were in the code group. All the participants were subject to a Pretest to identify the most common grammatical items, which affect the quality of their writing accuracy.

The grammatical items covered in this study resembled the most critical ones that students encountered while writing and were also recommended by the Academic Writing and Drama Professors in the Department of English – College of Arts – University of Basrah- Iraq. Forty students in the experimental groups passed through 4 in-tests to monitor their progress. During the treatment period, the researchers put the participants into smaller groups and made open sessions to help the targets modify their errors after every single test.

The findings indicated that all types of feedback contributed to improving the students' writing accuracy, though the effect was varied. In other words, students in the control group got very poor results due to the reasons mentioned previously. However, the rest of them in the metalinguistic groups, who were supported with the researchers' feedback (direct and indirect), made noticeable development in their writing output.

The current study results also indicated that students in the descriptive group scored higher marks in Posttest 1 than those in the code group, though comparatively insignificant and very slight difference. However, the latter group outperformed the former in Posttest 2. Researchers are recommended to conduct more studies utilizing a larger sample size to arrive at more generalized findings in the future. They are, as well, advocated to consider different substantial obstacles that hinder the development of college students' writing. If so is done, they probably come up with different results. Furthermore, new studies may tackle lower or higher college students' levels to catalogue a teaching model for others to take advantage of, and have them applied in classes.

Bibliography

Alexander, L.G (1965) *Essay and Letter Writing*. London: Longmans.

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4: arXiv preprint arXiv, *Journal of Statistical Software*.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102-118.
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2(3), 292-307
- Ellis, R. (2008). A typology of written corrective feedback types. *ELT Journal*, 63 (2), 97-107.
- Eastwood, J. (1992). Oxford practice grammar: with answers (2nd edition). Oxford: OUP.
- Ferris, D. R. and 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.
- Ferris, D. R. (2002). Treatment of error in second language student writing. Ann Arbor: University of Michigan Press.
- Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). Thousand Oaks, CA: Sage.
- Gholaminia, I. et al. (2013). An investigation of meta-linguistic corrective feedback in writing performance. *Social and Behavioral Sciences*, 116, 316-320.

- Hyland, K. and F. Hyland. (2006). 'Feedback on second language students' writing'. *Language Teaching* 39: 83-101.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY: Guilford Publications.
- Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, 66, 140-149.
- Murphy, R. (2015). *English Grammar in Use: A self-study reference and practice book for intermediate students: with answers (4th edition)*. Cambridge: CUP.
- Newton, R. R., & Rudestam, K. E. (2012). *Your statistical consultant*. Thousand Oaks, CA: Sage Publications.
- Nicole, D., and Macfarlane-Dick, D. (2006). Formative Assessment and Self-regulated Learning. A Model and Seven Principles of Good Feedback Practice. *Studies in Higher Education*, 31 (2), 199-218.
- Norrozizadeh, S. (2009). Indirect Feedback: A plausible suggestion for overcoming error occurrence in L2 writing. *The Journal of Asia TEFL*, 6 (4), 245-262.
- Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2), 1-9.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language attitude on ESL learners' acquisition of articles. *TESOL Quarterly* 41, 255-283.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Mahwah, NJ: Routledge Academic.

Westfall, P. H., & Henning, K. S. S. (2013). *Texts in statistical science: Understanding advanced statistical methods*. Boca Raton, FL: Taylor & Francis.

Vince, M. (2008). *Macmillan English Grammar in Context. Advanced*. Oxford: Macmillan Education.

Websites:

- Gassner, O., Mewald, C. & Sigott, G. (2007). Testing Grammar: LTC Technical Report 2, Language Testing Centre, Alpen-Adria- Universität Klagenfurt. Retrieved on 3rd August 2014 from <http://www.uni-klu.ac.at/ltc/inhalt/346.htm>

- <https://www.issuu.com/rasvamiti/docs/nelson-language-proficiency-test-pd>

-Intellectus Statistics [Online computer software]. (2019). Retrieved from <https://analyze.intellectusstatistics.com/>

-Trupe, A. L. (2001). A process approach to writing. Retrieved on 29 July 2014 from <http://www.bridgewater.edu/WritingCenter/Resources/Process.htm>

Appendix

Writing Rating Scale (May 2009)

	Scales	Extended Scales	CEFR
7	<ul style="list-style-type: none"> • good range of structures • few inaccuracies 	<ul style="list-style-type: none"> • Shows a good range of structures for most communicative needs. • Communicates with few inaccuracies and a relatively high degree of grammatical control. 	<i>Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding.</i>
6			<i>Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influence. Errors occur, but it is clear what he/she is trying to express.</i>
5	<ul style="list-style-type: none"> • generally sufficient range of structures for familiar contexts • occasional inaccuracies • message clear 	<ul style="list-style-type: none"> • Shows a generally sufficient range of structures for familiar contexts; mostly uses a repertoire of 'routines' for predictable situations. • Communicates with occasional inaccuracies which can impair communication, but message is clear. • L1 influence is noticeable. 	<i>Uses reasonably accurately a repertoire of frequently used 'routines' and patterns associated with more predictable situations.</i>
4			
3	<ul style="list-style-type: none"> • limited range of simple structures • frequently inaccurate • generally without causing breakdown 	<ul style="list-style-type: none"> • Shows a limited range of simple structures. • Communicates frequently inaccurately with basic mistakes, e.g. tenses or verb-noun agreement, but generally without causing breakdown; message is usually clear. 	<i>Uses some simple structures correctly, but still systematically makes basic mistakes – for example tends to mix up tenses and forget to mark agreement; nevertheless, it is usually clear what he/she is trying to say.</i>
2			
	<ul style="list-style-type: none"> • extremely limited range of structures 	<ul style="list-style-type: none"> • Shows an extremely limited range of simple structures or patterns within a learnt repertoire. 	<i>Shows only limited control of a few simple grammatical structures and sentence patterns</i>