Feedback Complexity and Writing Skill: The Relative Effectiveness of No Feedback, Error-flagging and Response-contingent Feedback Types

Dönüt Karmaşası ve Yazma Becerisi: Dönüt Yokluğu, Teşhis Edici Dönüt ve Cevaba Bağlı Dönütünün Karşılaştırmalı Geçerliliği

Maryam ZEİNOLABEDİNİ∗ Javad GHOGLAMI∗∗

Abstract: Despite controversies about the value of presenting written corrective feedback, research suggests that feedback has an undeniably significant role in improving performance in skill learning. However, little extant research has attended the different types of feedback ranged along a dimension of complexity. Thirty participants participated in this study in order to compare their writing accuracy under the no feedback, error-flagging, response-contingent and control conditions. The results of repeated measures, ANOVA revealed a significant main effect and interaction effect for feedback types on the accuracy improvement of learners from pre-test to post-test. Overall, these results suggest that unelaborated feedback is of limited pedagogical value, whereas elaborated feedback can contribute to linguistic accuracy in L2 writing.

Keywords: Written corrective feedback, elaborated feedback, no feedback, error-flagging, response-contingent


Anahtar sözcükler: Tazılı yapıcı dönüt, ayrıntılı dönüt, dönüt yokluğu, teşhis edici dönüt, cevaba bağlı dönüt

Introduction

In recent decades, foreign or second language instructors and administrators have changed their emphasis from a strictly product-oriented approach to learners’ writing toward a process-oriented approach. In fact, a more dialectic approach is encouraged to include both process and product (O’Sullivan, 2007; Warschauer, 2002). One significant element of a process-oriented

∗ PhD. Candidate of TEFL at Urmia University, Experienced EFL Teacher, Iran, maryam_shimi_ir@yahoo.com
∗∗ Assist. Professor, TESOL from Urmia University, Department of English, Iran, gholamij@gmail.com
approach to L2 writing is the issue of written corrective feedback (WCF). Literature in this regard has been barely introduced concerning the best and most appropriate methods of feedback, yet language teachers are faced with the real necessity of making decision about the content and different aspects of the learners’ writing pieces. In reality, how different teachers deal with the writing problems of different learners remains an area of concern to be inquired into in different contexts. Therefore, it is important to have more studies conducted in foreign language writing classrooms, as this type of research is definitely limited, when compared with the quantity of research in programs on writing in English as a second language.

To reach this point, many researchers try to find ways and develop strategies that help teachers to improve the EFL/ESL learners’ writing (e.g., Ashwell, 2000; Fazio, 2001; Chandler, 2003). Leki (1990) believed that learners need to get responses to their writing. One method of these communications between teacher and learner is through writing. Weigle (2002) argued that writing is a standardized system of communication and also is an essential tool in learning a system. In fact, the way teachers respond to the learner’s writing for the purpose of improvement is of essential importance. Therefore, in order to reduce unpleasant effects upon learner’s view towards writing, the way the teacher provides learners with a suitable type of feedback has to be accounted for within the pedagogical system. It needs to be noted that writing in foreign language education is one of the skills that demands greater attention due to its supportive role as reinforcement for learning other aspects of the language. Based upon which are supported by proponents of Grammar Translation Method, in EFL and ESL language classes, in order to learn other aspects of language such as grammar and vocabulary, writing can support and reinforce learning through oral communication.

There are, some research articles published on written feedback until the late 1990s (Lalande, 1982; Semke, 1984; Zamel, 1985). Since Truscott (1996) published his controversial article considerable debate followed concerning if and how to give L2 learners feedback on their written errors, among SLA researchers and language educators. This issue result in two opposing viewpoints regarding the usefulness of error correction in enhancing the learner’s writing ability. The first group believed that error correction not only is of no use in teaching writing classes, but also in some cases can be deleterious and may damage the act of learning (e.g., Krashen, 1982; Truscott, 2007). Some scholars (e.g., Kepner, 1991; Sheppard, 1992; Truscott, 2007) believed that corrections do not have a positive effect in improving learners L2 writing in courses. However, the second group (e.g., Carroll, Swain, & Roberge, 1992; Bitchener, 2008; Van Beuningen, De Jong, & Kuiken, 2012) believes that learners’ writing ability can improve through correcting errors in writing courses.

On different grounds, Truscott (1996, 328) claimed that “grammar correction has no place in writing courses, and should be abandoned” (see also Bitchener, Young, & Cameron, 2005). By analyzing previous research (Hendrickson, 1978; Semke, 1984; Zamel, 1985; Robb, Ross, & Shortreed, 1986; Frantzen, & Rissel 1987; Leki, 1990; Kepner 1991; Krashen, 1992; Sheppard, 1992), Truscott concluded that in both L1 and L2 writing courses, grammar correction is not helpful. For two major theoretical and practical reasons, he explained this finding in a more logical fashion. On the one hand, he believed that correction has a negative effect on the natural sequence of acquisition and understated the gradual and complex process of acquiring the forms and structures of a second language. On the other hand, he pointed to a list of practical problems related to the ability of teachers to give correction, and of a learners’ desire to receive error correction.

Although there are proponents and opponents of feedback provision in the classroom, what
seems evident is that for a learner to remain interested and motivated, there needs to be a close relation between the learner’s goals and the expectation that these goals can be met (Fisher, & Ford, 1998; Ford et alii, 1998). The replace, if the feedback is goal-directed and supplys learners with information about their improvements towards a goal, it is assumed to be certainly more effective than when this information is provided in discrete responses (i.e., answers to individual activities). For the feedback to be more effective and adapted to the learners’ goals, the issue of complexity needs to be addressed too. The provision of either elaborated or a more simple type of feedback is an important concern in the success of correction which is discussed in the section below.

**Feedback Complexity**

Whereas more specific feedback is usually better than less specific feedback under special conditions, a pertinent issue to consider in producing feedback is one of the length or the complexity of the information. For instance, if the correction is too long or complex, many learners’ attention can be distracted making it ineffectual. Lengthy feedback may also disperse and extenuate the information. Feedback complexity, therefore, relates to the quantity and the quality of information included in the feedback message.

There have been numerous studies that have examined the issue of feedback complexity, but only a few attempted to capture the fundamental factors along side a dimension of complexity (Dempsey, Driscoll, & Swindell, 1993; Mason & Bruning, 2001; and Narciss, & Huth, 2004).

If feedback aims to provide a corrective function, even in its simplest form, it should confirm as to if the learner’s response is correct or incorrect and supply information to the learner as to the correct answer (either directive or facilitative). Studies that have investigated the type and content of information in feedback, have however generated inconsistent findings (Kulhavy, 1977; Mory, 2004). They have examined different feedback types specified according to complexity including no feedback, verification, correct response, try again, error-flagging, elaborated, attribute isolation, topic contingent, hints/cues/prompts, bugs/misconceptions, and informative tutoring.

For the effect of feedback complexity, Schimmel (1983) carried out a meta-analysis on feedback as employed in computer-assisted pedagogy and programmed instruction. He analyzed the findings of 15 empirical researches and concluded that the amount of information, namely the feedback complexity, was not significantly related to feedback impact. He also found that feedback impact was significantly higher in computer-assisted instruction than in programmed instruction.

Sleeman et alii, (1989) investigated the controversial results in the literature with respect to the identification and remediation of learners’ errors. They perceived that few studies have systematically confronted the effects of different styles of error-based feedback and for those that have, the results are indeterminate.

Sleeman et alii, (1989) compared the relative effectiveness of bugs/misconceptions and the topic-contingent as elaborated feedback types. Although the bugs/misconceptions employ more complex methods of correction based upon formal error analysis, while the topic-contingent method is the re-teaching of the specific problematic content. The results showed that feedback based upon the re-teaching of content was as influential as feedback, based upon more complicated error analyses.

Generally, elaborated feedback addresses the right response and might additionally explain the reason why the chosen answer is incorrect and highlight the correct answer. There seems to
be growing compliance that the response-specific feedback as a type of elaborated feedback has the potential to improve the learner’s attainment more than other types of feedback (e.g., Gilman, 1969; Corbett, & Anderson, 2001; Mory, 2004). However, since there are still inconclusive results regarding feedback complexity and its types as reviewed above, additional studies seem to be indispensable in resolving this issue.

The purpose of the present study is to examine the differential role that feedback complexity plays in writing ability. The following research question is addressed:

1. To what extent do the types of feedback concerning linguistic errors determine accurate performance in learners’ post-test writings?

Methodology

Participants

The study was conducted at a language institute in Urmia, Iran. One complete upper-intermediate class of EFL learners was invited to participate. The First Certificate in English (FCE) test was administered as a pre-test to learners, to determine their writing ability. The language institute in the present study focuses upon the communicative approach to language teaching and devotes an equal attention to reading, writing, speaking and listening skills. Learners at the institute received 3 hours of instruction two days a week. As a result of the pre-test 30 learners were selected as participants. The teacher herself who was an experienced (years of teaching = 10) 35 years old teacher provided the correction types to the learners’ errors.

Procedure

The design of the present study is a pre-test post-test experimental group only design. The treatment (different written corrective feedback (WCF) options on the linguistic errors) lasted for in total 17 sessions. For the first five sessions, the first session was devoted to the pre-test and the three sessions included the control condition in which the learners did not receive any feedback on their writing performance. Only the writing section of the FCE test was included as appropriate to the purpose of this study. The writing section has two parts which require the writers to compose two different pieces of writing within a time of one hour and twenty minutes. For the first writing task, writers are asked to write an answer to one compulsory question using up to 160 words. For part 2, writers need to select one writing task from a choice of five and are allowed to use 120-180 words. The FCE writing tasks were used as both the pre-test at the beginning of the study and as post-tests after the fulfillment of the treatment sessions. In all classroom sessions, the learners composed three writing pieces according to the writing tasks of the Summit 1 A book (Saslow, & Ascher, 2012). The book introduces all the competencies of B1 Level described in the Common European Framework of Reference, and introduces competencies at the B2 level. The writing tasks included a variety of writing activities such as descriptions, biographies, articles, letters to newspapers, summaries, and critiques of articles.

In the sixth session, the no feedback strategy was employed which is a condition in which the learner is provided with a question and is asked to answer, but there is no mention of the correctness of the learners’ answer. After the completion of three sessions, the FCE post-test was administered.

The tenth session began with the application of error-flagging type of feedback, also known as location of mistakes, in which the errors are highlighted in a solution and the correct answer is not supplied. In other words, “error-flagging is a via media—it detects errors for the learner,
but leaves it to the learner whether and how to correct the errors” (Kumar, 2010). After the completion of three sessions of treatment devoted to error-flagging strategy, session thirteen included the FCE post-test. Session fourteen started with the response-contingent type of feedback provided on the learners’ linguistic errors. Response-contingent which is a type of elaborated feedback attends to the learner’s specific response. It can describe the reason why the response is incorrect and why the response is correct. However, in this type of feedback, no formal error analysis is performed. The post-test was again administered at session seventeen to evaluate the effect of response-contingent strategy upon participants’ writing performance.

Learners in all correction conditions were required to complete the writing tasks within 30 minutes and the corrected pieces of writing for the treatment groups were returned within a week of writing.

Error Marking
Whereas learners in the experimental conditions revised their writings based upon the feedback, those in the contro condition were given unmarked copies of their original writings. After the errors of the post-test writings (i.e., FCE writing test) were marked, each piece of writing was assigned an accuracy rate. This was achieved through determining the total number of words written divided by the total number of errors. Marking was carried out by the researcher. A trained teacher/researcher in the institute conducted an inter-rater reliability check on this analysis. The initial index of reliability was 89%. Collaborative analysis of the other 11% of instances resulted in complete agreement.

The choice of error types to be marked was according to the need for both consistency and broad inclusion. All linguistic and grammatical errors were entailed, but errors in word choice were not, except when they comprised the choice of a function word and could thus be counted as grammatical problems. Errors of mechanics were excluded, except for the cases that were fragments and run-on sentences. Spelling errors were also included.

Results
To examine the research question, a repeated measures ANOVA was run on the four groups’ pre-test and post-test scores with time (pre- vs. post-test) as within and each feedback type (no feedback, error-flagging, response-contingent, and control condition) as between subject factor. Table 1 below shows the descriptive statistics for the learners’ pre-test and post-test writing performance. The mean scores refer to the mean of accuracy.

Table 1. Descriptive Statistics for Writing Performance in Four Conditions

<table>
<thead>
<tr>
<th></th>
<th>feedback</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest</td>
<td>no feedback</td>
<td>7.6167</td>
<td>1.85994</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>error-flagging</td>
<td>7.6167</td>
<td>1.85994</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>response-contingent</td>
<td>7.6167</td>
<td>1.85994</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>control condition</td>
<td>7.6167</td>
<td>1.85994</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>7.6167</strong></td>
<td><strong>1.83634</strong></td>
<td><strong>120</strong></td>
</tr>
<tr>
<td>posttest</td>
<td>no feedback</td>
<td>8.1500</td>
<td>2.05321</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>error-flagging</td>
<td>13.4500</td>
<td>2.51283</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>response-contingent</td>
<td>14.1733</td>
<td>2.38991</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>control condition</td>
<td>7.8167</td>
<td>2.34507</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>10.8975</strong></td>
<td><strong>3.73389</strong></td>
<td><strong>120</strong></td>
</tr>
</tbody>
</table>
The results of descriptive statistics show that the mean scores of no feedback group (M = 8.15, SD = 2.05) and the control group (M = 7.81, SD = 2.34) were similar to each other in the post-test. Moreover, the participants’ accuracy rate in the response-contingent condition (M = 14.17, SD = 2.39) was slightly higher than that of the error-flagging condition (M = 13.45, SD = 2.51).

In order to compare the test scores across the treatment conditions, a one-way repeated measures ANOVA was computed. Table 2 below shows the results of this analysis.

**Table 2. Repeated Measures ANOVA Results for Writing Performance in Four Conditions**

<table>
<thead>
<tr>
<th>Source</th>
<th>Type of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>Linear</td>
<td>1</td>
<td>645.832</td>
<td>147.004</td>
<td>.000</td>
<td>.559</td>
</tr>
<tr>
<td>time * feedback</td>
<td>Linear</td>
<td>3</td>
<td>171.433</td>
<td>39.021</td>
<td>.000</td>
<td>.502</td>
</tr>
<tr>
<td>Error(time)</td>
<td>Linear</td>
<td>116</td>
<td>4.393</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The repeated measures ANOVA indicated statistically significant differences between the four conditions, F(1, 116) = 147.004, p = .000. The results of this analysis revealed a significant effect of time (F(1,116) = 147.004, p = .000) and a significant feedback-time interaction (F(3,116) = 171.433, p = .000), suggesting that only error-flagging and response-contingent feedback conditions led to an improvement from pre-test to posttest in some aspects of the writing. In other words, time did have an effect on the two types of feedback.

The results of the Tukey post-hoc test are reported in Table 3 to isolate the exact points where the differences between the conditions occurred.

**Table 3. Tukey Test Results for Writing Performance in Four Conditions**

<table>
<thead>
<tr>
<th>(I) feedback</th>
<th>(J) feedback</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>no feedback</td>
<td>error-flagging</td>
<td>-2.6500*</td>
<td>.38736</td>
<td>.000</td>
<td>-3.6597</td>
<td>-1.6403</td>
</tr>
<tr>
<td></td>
<td>response-contingent</td>
<td>-3.0117*</td>
<td>.38736</td>
<td>.000</td>
<td>-4.0214</td>
<td>-2.0020</td>
</tr>
<tr>
<td></td>
<td>control condition</td>
<td>.1667</td>
<td>.38736</td>
<td>.973</td>
<td>-.8430</td>
<td>1.1764</td>
</tr>
<tr>
<td>error-flagging</td>
<td>no feedback</td>
<td>2.6500*</td>
<td>.38736</td>
<td>.000</td>
<td>1.6403</td>
<td>3.6597</td>
</tr>
<tr>
<td></td>
<td>response-contingent</td>
<td>-.3617</td>
<td>.38736</td>
<td>.787</td>
<td>-1.3714</td>
<td>.6480</td>
</tr>
<tr>
<td></td>
<td>control condition</td>
<td>2.8167*</td>
<td>.38736</td>
<td>.000</td>
<td>1.8070</td>
<td>3.8264</td>
</tr>
<tr>
<td>response-contingent</td>
<td>no feedback</td>
<td>3.0117*</td>
<td>.38736</td>
<td>.000</td>
<td>2.0020</td>
<td>4.0214</td>
</tr>
<tr>
<td></td>
<td>error-flagging</td>
<td>.3617</td>
<td>.38736</td>
<td>.787</td>
<td>-.6480</td>
<td>1.3714</td>
</tr>
<tr>
<td></td>
<td>control condition</td>
<td>3.1783*</td>
<td>.38736</td>
<td>.000</td>
<td>2.1686</td>
<td>4.1880</td>
</tr>
<tr>
<td>control condition</td>
<td>no feedback</td>
<td>-.1667</td>
<td>.38736</td>
<td>.973</td>
<td>-1.1764</td>
<td>.8430</td>
</tr>
<tr>
<td></td>
<td>error-flagging</td>
<td>-2.8167*</td>
<td>.38736</td>
<td>.000</td>
<td>-3.8264</td>
<td>-1.8070</td>
</tr>
<tr>
<td></td>
<td>response-contingent</td>
<td>-3.1783*</td>
<td>.38736</td>
<td>.000</td>
<td>-4.1880</td>
<td>-2.1686</td>
</tr>
</tbody>
</table>
Based on observed means. The error term is Mean Square(Error) = 2.251.
* The mean difference is significant at the .05 level.

The results of Tukey corroborate the findings of descriptive statistics, indicating that there are not statistically significant differences between the no feedback and the control condition (p=0.97) and error-flagging and response-contingent conditions (p=0.78). However, there are significant differences between the no feedback and error-flagging conditions (p=0.000), no feedback and response-contingent conditions (p=0.000), control condition and error-flagging conditions (p=0.000), control condition and response-contingent conditions (p=0.000).

Figure 1 provides a schematic representation of the mean percentages for the four conditions in the pre-test and post-test.

Discussion
Truscott’s (1996) review of studies by Kepner (1991), Semke (1984), and Sheppard (1992) contended that feedback cannot have an influence on enhancing L2 learners’ writing. With respect to the restricted range of studies, the present study attempted to expand the base by examining the impact of three different types of feedback, distinct in complexity on the accuracy performance of learners in new pieces of writing. The results of the statistical analysis showed that the feedback types had a significant impact on the accuracy rate of learners’ post-test writing. In other words, there were significant differences between the levels of complexity such that the more elaborated and complicated type of feedback (i.e., the response-contingent) could lead to a higher accuracy rate than the simpler and less elaborated types of no feedback and control conditions. This finding was not particularly surprising as one would tend to expect that clarifying the rules and identifying the errors and the provision of explanation for them in the response-contingent condition would help learners to notice the difference between their errors and the corrections they receive. Noticing such distinctions is now confirmed in the SLA literature as essential to uptake and long-term learning (Schmidt, 1994; Gass, 1997). By comparison, learners in the error-flagging condition who received the identification and highlight of errors were also given the opportunity to notice their problematic areas and could thus improve their writing, albeit to a lesser degree. And, the learners in the no feedback condition who were not made aware of their errors and those in the control condition who were not provided with any written or oral feedback on the linguistic features showed a similar performance on the post-test, obtaining lower rates of accuracy when compared with that of other groups.

The results of this study can be elaborated further from sociocultural theory. The fact that the elaborated type of feedback resulted in a higher accuracy rate than the other feedback strategies implies that the provision of opportunities for learners to see the problematic areas according to their specific response scaffolds their performance. In this way, the reviewer (i.e.,
the teacher) provides feedback for the writer that falls within the writer’s zone of proximal development (ZPD). Therefore, the noticing of the gap in the writing performance scaffolds the learning process (Nassaji, & Swain, 2000).

According to the results, it is suggested that classroom L2 writing teachers provide their learners with written feedback in form of the identification of errors according to the writer’s specific response on a regular basis. In order for learners to buy into the learning process, it is even suggested teachers discuss with their learners which linguistic errors should be attended most and instruct them about the central attributes of the linguistic feature or skill being studied. However, it should be mentioned that in all the elaborated feedback types, the correct answer is not provided by the teacher, while an attempt is made to guide the learner in the right direction.

This finding adds to the growing body of research that has examined the impact of different feedback types on accuracy performance. For example, it has already been observed that elaborated feedback is more effective than less elaborated feedback in helping learners enhance their writing accuracy. Since little, to no research has specifically investigated the effect of different feedback options regarding complexity, the results of the present study are noteworthy.

Finally, the results of this study have demonstrated that L2 learners can enhance the accuracy of their use of linguistic features if they are exposed to written feedback on a regular basis. Further research would need to be conducted to test if this finding holds true for other categories of elaborated feedback and whether it applies for specific linguistic form in which the rules of usage are more sophisticated and more idiosyncratic.
REFERENCES


