Academic Writing and Grammatical Accuracy: The Role of Corrective Feedback

Sasan Baleghizadeh\textsuperscript{2*} & Yahya Gordani* \\
\textit{Shahid Beheshti University, G.C. Irán}

Abstract

Investigations into the effect of providing corrective feedback on L2 writing have often produced contradictory results. This study, following a line of research concerned with the role of corrective feedback in writing, contributes to this line of research by analyzing different feedback types in an EFL academic writing context. 45 graduate university students enrolled in an academic writing course were provided with different types of feedback (direct feedback; student-teacher conference; no corrective feedback) over a course of 12 weeks (24 sessions). The study found significant immediate and delayed effects for the student-teacher conference type of feedback on students’ overall accuracy improvement. It is suggested that improvements in writing accuracy could at least in one respect be attributed to the type of feedback provided.

\textit{Keywords:} academic writing, writing accuracy, corrective feedback, EFL learners.

Resumen

Las investigaciones sobre el efecto de proporcionar retroalimentación correctiva sobre escritura en L2 a menudo han producido resultados contradictorios. Este estudio se desarrolla bajo una línea de investigación que busca indagar sobre el papel que juega la retroalimentación correctiva en el proceso de escritura. Así mismo, se realiza un gran aporte a la línea de investigación al analizar los diferentes tipos de retroalimentación existentes en un contexto de escritura académica de inglés como lengua extranjera. La población objetivo del estudio fueron cuarenta y cinco estudiantes de posgrado matriculados en un curso de escritura académica el cual tenía una duración de doce semanas (24 sesiones). Estos estudiantes durante el curso recibieron los siguientes tipos de retroalimentación: retroalimentación directa, reunión docente – estudiante y sin retroalimentación correctiva. El estudio reveló importantes hallazgos en

\textsuperscript{1} Received: March 23rd, 2012 / Accepted: May 15th, 2012

\textsuperscript{2} Email: sasanbaleghizadeh@yahoo.com
The application of accurate grammar is an important aspect of any good piece of writing. In addition, students can advance their level of English by producing written work that employs the grammatical structures they have learned. Although it is unrealistic for nonnative students to expect to reach 100% accuracy (and many native English speakers may have similar difficulty), they should aim to continuously improve their writing accuracy, in order to make their work as readable and efficient as possible.
Many studies have made it clear that after having studied English as well as academic writing for years, non-native students experience a great deal of difficulty in their writings. For example, Johns (1997) found that many non-native speaking graduate and undergraduate students, after years of ESL training, often fail to recognize and appropriately use the conventions and features of academic written prose. Researchers have pinpointed many reasons that the academic writing of even highly advanced and trained non-native students continues to exhibit numerous problems and shortcomings (See Hinkel, 2002; Johns, 1997; Jordan, 1997; Leki & Carson, 1997; Prior, 1998). Such shortcomings include, among others, the ineffectiveness of writing courses in preparing students for academic writing tasks and the disparity between the existing teaching and assessment practices in academic writing contexts. The effectiveness of writing courses in preparing NNS students for actual academic writing in universities is discussed by Leki and Carson (1997). They found that, “what is valued in writing for writing classes is different from what is valued in writing for other academic courses” (p. 64).

However, it should be noticed that academic writing is different from personal or creative writing in a few respects. Thaiss and Zawacki (2006) asked professors at George Mason University what they thought academic writing was and what its standards were. They came up with three characteristics: 1. Clear evidence in writing that the writer(s) have been persistent, open-minded, and disciplined in study. 2. The dominance of reason over emotions or sensual perception. 3. An imagined reader who is coolly rational, reading for information, and intending to formulate a reasoned response.

What seems obvious to assert is that in academic writing one should always follow rules of grammatical accuracy since the end-user or consumer of the writing is likely to be an expert in particular fields. Hence, it is vital that writing is clear through compliance with rules of punctuation and the conventions of grammar to maintain clarity and avoid ambiguity in expression. Leki and Carson (1997) emphasized that the teaching of writing in ESL and EAP programs needs to provide students with linguistic and writing skills that can enable the learners to be involved with and make sense of the new information.

On the other hand, in teaching L2 writing to academically bound learners, rhetorical and discourse features of written English have often been overemphasized. What has become of lesser importance, as Hinkel (2004) puts it, is the language tools (i.e., the grammar and vocabulary that L2 writers must have to construct academic text, which
in turn can be organized into a coherent written academic discourse). In other words, no matter how well discourse is organized or how brilliant the writer’s ideas may be, it would be hard to understand them if the language is opaque. When it comes to assessment, though, raters seem to be more concerned with the linguistic errors made by the writers than the rhetorical and discourse features of texts. This incongruity between teaching and assessment practices might therefore be a potential reason behind student’s failure in academic writing.

Accuracy in Academic Writing

Many researchers have reasonably argued that for academically oriented and advanced L2 learners, grammar instruction is essential if they are to achieve their educational and professional goals (Celce-Murcia 1991; Schmidt 1994; Shaw & Liu 1998). Celce-Murcia (1991), for instance, emphasized the importance of a reasonable degree of grammatical accuracy in academic writing. She mentioned that high frequency of grammatical errors in nonnative speaker’s academic writing (an average of 7.2 errors per 100 words) most probably makes their writings unacceptable to the University faculties.

A large number of extensive and detailed studies have demonstrated that mere exposure to L2 vocabulary, grammar, discourse, and formal written text is not the most effective means of attaining academic L2 proficiency (e.g., Ellis, 1990; Hinkel, 2002; Laufer & Nation, 2001; Norris & Ortega, 2000; Schmidt, 2000). In other words, exposure to the input is not a guarantee for language acquisition. Schmidt (2000) proposes the noticing hypothesis to emphasize that only items in linguistic input that are attended to by language learners are likely to be acquired.

Chang and Swales (1999) investigated specific discourse and sentence-level writing skills of highly advanced non-native speaker students. They indicate that even in the case of advanced and highly literate non-native speakers, exposure to substantial amounts of reading and experience with writing in academic contexts does not ensure their becoming aware of discourse and sentence-level linguistic features of academic writing and the attainment of the necessary writing skills. Chang and Swales concluded that explicit instruction in advanced academic writing and text is needed. Similarly, Ellis (1990) believed that formal classroom teaching with its emphasis on linguistic accuracy will engage the learner in planned discourse and develop the corresponding type of competence.
However, as mentioned, the predominant method of instruction in the teaching of L2 writing has mainly remained focused on the writing process (Johns, 1990; Reid, 1993; Zamel, 1983) and the product of writing is seen as secondary to the writing process. Therefore, as Hinkel (2004) mentions, issues of L2 grammar, lexis, and errors are addressed only as needed in the context of writing, and L2 writers with proficiency levels higher than beginning are exposed to text and discourse to learn from them and, thus, acquire L2 grammar and lexis naturally. She goes on to suggest that the assessment of L2 writing skills by ESL professionals on standardized and institutional placement testing has largely remained focused on the writing product without regard to the writing process (Vaughan, 1991) and concludes that “the disparity between the teaching methods adopted in L2 writing instruction and evaluation criteria of the quality of L2 writing has produced outcomes that are damaging and costly for most ESL students, who are taught brainstorming techniques and invention, prewriting, drafting, and revising skills, whereas their essential linguistic skills, such as academic vocabulary and formal features of grammar and text, are only sparsely and inconsistently addressed” (p.6).

Corrective Feedback and Writing Accuracy

Previous research has made it clear that in order to be able to write successfully in a second language and, in particular, to be able to learn the formal L2 academic prose crucial in nonnative speakers’ academic and professional careers, students need to develop a basic linguistic threshold, without which they simply do not have the range of lexical and grammar skills required in academic writing (Berkenkotter & Huckin, 1995; Byrd & Reid, 1998; Chang & Swales, 1999; Hinkel, 2002; Horowitz, 1991; Johns, 1997; Kroll, 1980; Paltridge, 2001; Read, 2000).

Xudong, Cheng, Varaprasad, and Leng (2010) investigated the impact of English for Academic Purposes course on the development of academic writing abilities of ESL/EFL graduate students. The study found that not much progress had been made by these students in terms of grammar accuracy. In addition students’ responses to the questionnaire indicated that they felt the course did not help them improve their grammar accuracy. As Hinkel (2004) mentions, intensive and consistent instruction in L2 grammar is essential for academically bound nonnative speakers. Consistent grammar instruction has been shown to be effective in improving the quality of L2 production (Cumming, 1990; Ellis 2001; Fotos 2002; Norris and Ortega 2001; Schmidt 1994).
In their meta-analysis, Norris and Ortegga (2000) conclude that focused instructional treatment of any sort is far better than nonintervention and is durable over time. What remains controversial, however, is the nature of this instructional treatment which includes the unresolved issue of the provision of corrective feedback to student writings.

Feedback as viewed by Furnborough and Truman (2009) entails the existence of a gap between what has been learned and the target competence of the learners, and the efforts undertaken to bridge these gaps. In a study Treglia (2009) posits that students understood and were able to address corrections irrespective of the type of feedback provided, assuring writing teachers that student writers are able to benefit strongly from teacher feedback.

In a series of studies, Truscott (1996, 1999, and 2007) pointed to the fact that there was no sufficient research in favor of grammar correction. He referred to many studies which couldn’t actually support grammar correction for different reasons such as the absence of control groups and delayed posttests or the use of grammar exercises as their only writing tasks. This claim caused criticisms on the part of the proponents of grammar correction (Ferris, 1999, 2004), and some researchers tried to generate research to counter the conclusions (Ferris & Roberts, 2001; Chandler, 2003; Bitchener 2008; Bitchener et al., 2005, 2008, 2009, 2010; Rahimi, 2009; Sheen et al., 2009), and after each of their attempts Truscott has responded with critiques claiming that their work fails to demonstrate that error correction has any benefit (Truscott 2004; Truscott and Hsu, 2008).

Many researchers investigated the efficacy of different feedback types on students’ writing accuracy. For example, Bitchener, Young, and Cameron (2005) found that direct oral feedback in combination with direct written feedback did not only have a greater effect than direct written feedback alone on improved accuracy over time, but it also found that the combined feedback option facilitated improvement in some error categories but not others. Moreover, they believe that upper intermediate L2 writers can improve the accuracy of their use of rule-governed linguistic features if they are regularly exposed to oral and written corrective feedback.

In addition, Bitchener and Knoch (2008) analyzed the extent to which different written corrective feedback options (direct corrective feedback, written and oral metalinguistic explanation; direct corrective feedback and written meta-linguistic explanation; direct corrective feedback only; no corrective feedback) improve students’ accuracy in the use of two functional uses of the English article system. Their
findings revealed that students who received all three written corrective feedback options outperformed those who did not receive written feedback and that students’ level of accuracy was retained over seven weeks, while there was no difference in the extent to which students improved the accuracy of their writing as a result of written corrective feedback.

Objective of the Study

Improving students’ writing accuracy is an essential factor in effective writing. Effectiveness of a piece of writing will be determined in part by its accuracy. This is the reason why grammar correction has received so much attention on the part of researchers, teachers, teacher educators, and students in the recent decades. In the writing classroom, teacher feedback on grammar may be a useful pedagogical device to enhance the accuracy of writing. The present study aims to analyze different feedback types (direct feedback; student-teacher conference; no corrective feedback) in an EFL academic writing context. It is hypothesized that the provision of different feedback types or no feedback at all will have different effects on student writings’ overall grammatical accuracy.

The study

Participants

A total of 45 students took part in this study. They came from among 56 students enrolled in three intact classes of academic writing in Tehran’s Shahid Rajaee Teacher Education University. The participants were homogeneous in terms of their age, gender, major and English learning backgrounds. They were at the age of 21 to 37, with an average of 28.4. The ratio of male to female students was also controlled to be 22 female and 23 male students so as to avoid issues of gender bias. The classes were academic writing classes consisting of graduate students majoring in Teaching English as a Foreign Language (TEFL). Their learning backgrounds were similar due to the fact that they had studied English solely within the educational system of Iran. Besides, none of them had the experience of studying or living in English-speaking countries. The two classes were randomly assigned to one experimental (group A) and one control group (group B).
Instrument

A writing test package consisting of four academic topics together with instructions on the number of words needed and the allotted time was used in this study. The first writing task was applied to ascertain the homogeneity of the participants in terms of their writings’ grammatical accuracy. The other three writing tasks (a pretest, an immediate posttest and a delayed posttest) were administered at different phases during and after the course.

Participants’ writings at each phase of the study were evaluated and scored by two raters for the matter of inter-rater reliability. To estimate the inter-rater reliability of the tests, the correlation coefficients between the two raters were calculated. All the four tests were shown to have very high inter-rater reliability with an average of 0.913 which were found statistically significant at p<0.1 level of significance.

Procedures

Out of the 56 students who took the first writing test, the eligible ones (those whose scores ranged from one standard deviation above and below the mean on the test; n=45) were selected to serve as the participants of the study. The participants were then randomly assigned into three homogeneous groups each with 15 participants. All the condition for the groups was exactly the same except for the method used for the provision of corrective feedback.

The three groups were administered the pretest at the beginning of the first educational semester (fall 2011) and during the first session of their course. Before the test was administered, participants were provided with an explanation of the purpose of the study and assured that the results would have no influence on the course outcomes.

In the second phase, the groups were taught the course for 24 sessions (12 weeks). However, only the two experimental groups received the treatment. The treatment for the first experimental group involved direct corrective feedback. Every week in the first experimental group the students were given a topic to write as an assignment for the next session. Their writings were then collected and the instructor provided each student with corrective feedback on grammatical points of his/her writing product in the next session. Grammatical errors were underlined and suggestions for alternatives were provided as much as possible. In the second experimental group this procedure was supplemented with a teacher-student conference on the grammatical errors in which the class discussed and gave their
ideas about the problems with the errors made by themselves or their peers and they suggested alternatives for erroneous items. The control group, however, received no feedback on the grammatical accuracy of their writings. Their writings were returned to them with the teacher’s comment on the content.

In the final phase, the posttests were administered to both experimental groups. The immediate posttest was administered right at the end of the semester and during the last session of the course (the 12th week). Six week later when the class met again, the delayed posttest was administered about which there was no prior notice to the students. The collected data --the scores obtained from the pretest and the two posttest administrations-- were statistically analyzed using the SPSS program.

Results

The present study investigated the effectiveness of different feedback types on grammatical accuracy of the students’ academic writing. In this section, the results of the study will be presented. Table 1 presents the results of an academic writing test to 45 students assigned to three groups. This test was administered to ascertain the homogeneity of the three groups.

Table 1. Descriptive statistics for the first academic writing task

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>15</td>
<td>13.4700</td>
<td>3.17588</td>
<td>.71082</td>
<td>13.5122</td>
<td>16.4187</td>
<td>09.00</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>12.9500</td>
<td>3.08650</td>
<td>.69506</td>
<td>14.1057</td>
<td>16.5543</td>
<td>09.00</td>
</tr>
<tr>
<td>C</td>
<td>15</td>
<td>13.1000</td>
<td>2.73188</td>
<td>.61087</td>
<td>13.8214</td>
<td>16.3786</td>
<td>11.00</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>13.1733</td>
<td>2.96557</td>
<td>.38260</td>
<td>14.4511</td>
<td>15.9822</td>
<td>09.00</td>
</tr>
</tbody>
</table>
Table 2. One way ANOVA to assess the homogeneity of the three groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3.483</td>
<td>2</td>
<td>1.717</td>
<td>.270</td>
<td>.923</td>
</tr>
<tr>
<td>Within Groups</td>
<td>514.750</td>
<td>42</td>
<td>9.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>518.133</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 presents the results of a one way ANOVA which was run between the mean scores of the three groups. With an alpha level decided at .05 level of significance, there was not any significant difference observed between the three groups (F=.270) which along with the randomization showed that the three groups were homogeneous. In other words, the results of this test showed that there was no significant difference in the grammar accuracy of the students’ writing prior to the experiment.

In order to examine the performance of the three groups in the immediate posttest of writing accuracy a one way analysis of ANOVA was conducted. The results are presented in the table 3 below.

Table 3. One-way ANOVA for the immediate posttest of writing accuracy

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>345.578</td>
<td>2</td>
<td>14.723</td>
<td>7.213</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>72.951</td>
<td>42</td>
<td>.770</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>292.476</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the table shows, the differences between the groups are significant (Sig=.000). Therefore the participants in the three groups differ in their performance in the immediate posttest. To specify exactly which two groups are different from each other, post hoc analysis was conducted through the Scheffe test. This is presented in the table 4 below.
As the table shows, experimental group B (student-teacher conference group) is significantly different from the other two groups with regard to their performance on the immediate posttest. Mean differences make it clear that group B has a significantly better performance than the other two groups. However, there is no significant difference in the performance of the experimental group A and the control group.

To compare the performance of the three groups in the delayed posttest, a second ANOVA was applied. The results from the application of this procedure tell us whether or not there are differences among the three groups in their performance on the delayed posttest. These are presented below in Table 5.

*Table 5. ANOVA*

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>669.409</td>
<td>2</td>
<td>34.704</td>
<td>12.662</td>
</tr>
<tr>
<td>Within Groups</td>
<td>90.247</td>
<td>42</td>
<td>1.037</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>759.656</td>
<td>44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The mean difference is significant at the 0.05 level.*
As the table shows, the differences between the groups are significant (Sig=.000). Therefore the participants in the three groups differ in their performance in the delayed posttest. To specify exactly which two groups are different from each other, post hoc analysis was conducted through the Scheffe test. This is presented in the table 6 below.

Table 6. Multiple Comparisons
Dependent Variable: delayed posttest

<table>
<thead>
<tr>
<th>(i) groups</th>
<th>(j) groups</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>A</td>
<td>1.52069*</td>
<td>.26523</td>
<td>.000</td>
<td>.9613</td>
<td>2.2813</td>
</tr>
<tr>
<td>C</td>
<td>B</td>
<td>-.639488*</td>
<td>.26312</td>
<td>.000</td>
<td>5.7396</td>
<td>7.0502</td>
</tr>
<tr>
<td>C</td>
<td>A</td>
<td>-.18747</td>
<td>.29376</td>
<td>.749</td>
<td>-.9032</td>
<td>.3398</td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>.18747</td>
<td>.29376</td>
<td>.749</td>
<td>-.3358</td>
<td>.9032</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the 0.05 level.

The results show a significant difference among two of the paired groups with regard to their performance in the delayed posttest. A look at the mean differences from the table makes it clear that just like the immediate posttest, group B has performed significantly better than groups A and C on the delayed posttest.

Discussion and Conclusion

The objective of this study was to compare and contrast two types of corrective feedback namely the direct correction versus the student-teacher conference feedback. As discussed before, the student-teacher conference feedback group was found to be significantly better than the direct feedback group and the control group on both immediate and delayed posttests suggesting the immediate and lasting effectiveness of this type of feedback over the direct feedback or the provision of no feedback at all. This finding confirms the findings of Rahimi, 2009, Hartshorn (2008), Chandler (2003) and Fazio (2001) who found corrective feedback as a way of improving the structural accuracy of
L2 student writing. However the present findings run counter to that of Truscott (1999, 2007) who claimed that correction has a small negative effect on learners’ ability to write accurately and that we can be 95% confident that if it has any actual benefits, they are very small. These results are also inconsistent with findings of Bitchener, Young, and Cameron (2005) who found a combination of written corrective feedback and conference feedback to improve accuracy levels in some structures, but found no overall effect on accuracy improvement.

From a theoretical perspective, perhaps there is place to claim a strong bond between providing students with error feedback and students’ writing accuracy. In line with Schmidt’s (1990, 1994) noticing hypothesis, only items which are noticed by the learners will be likely to be acquired. Thus, error feedback will push the students towards noticing the linguistic problems they are struggling with and that sometime they take for granted. In other words, providing corrective feedback will prompt the students to try and modify their developing interlanguage system in line with the feedbacks provided. The important issue, however, is the most effective type of feedback possible.

The effectiveness of student-teacher conference feedback type and the failure of direct feedback in this study may be attributed to previous research evidence suggesting that when it comes to written corrective feedback students often do not understand the meaning of much of the feedback on their papers and also do not know what they are expected to do with them (see Ferris 1995 and Hyland 1998). The oral student-teacher conferences held every session helped clarify the student errors and the errors made by their peers to them. By the third or fourth session the students had got used to most of the common errors discussed in the classroom and they had come to terms with the teacher’s assumptions of the structural errors they made. With regard to the direct feedback, however, this was not the case. Students were left with corrected papers and in some cases they might not be able to make sense of the feedback provided.

Another explanation for the effectiveness of the provision of feedback in the form of conferences might be the nature of the course. Students preparing themselves for doing academic writings in the future are well aware of the importance of structural accuracy of their writings. In fact, feedback on the structure of their writings is usually what they expect the teacher to provide. Content, on the other hand, is what they are already familiar with and they have usually read extensively on areas of their interest to be able to satisfy the needs of their readers. This supposed agreement between students and teachers on the feedback issue adds to the possibility that students try to learn
and make use of the provided feedback in their future writings (See Lee, 2005; Jeon & Kang, 2005; for more on this issue).

Deciding on the type of the corrective feedback is an important pedagogical issue since it requires different amounts of time and teaching skill. Direct provision of error feedback in the form of underlining and labeling errors by type may be less time-consuming for teachers than holding student-teacher conferences of discussion groups. It is certainly much easier to just underline or circle errors. Hence the direct method may seem to be a more handy option for teachers. On the other hand, holding student-teacher conferences on errors will necessarily call for sufficient metalinguistic knowledge possessed by students as well as teachers. Therefore the usage of this type of feedback may be suggested for the adult language learners and/or higher levels of language learning.

Future research may benefit from examining the effects of corrective feedback: (1) on new pieces of academic writing (2) on a wide range of linguistic error categories which was not considered in this study, (3) with students of different proficiency levels (4) in ways that involve different other feedback strategies and combinations of strategies.

References


**Authors**

*Sasan Beleghizadeh* has a Ph.D in TEFL from the University of Tehran and an M.A. in TEFL from Allameh Tabataba’i University. He is currently working as an assistant professor of TEFL at Shahid Beheshti University, G.C. in Iran, where he teaches courses in applied linguistics, syllabus design, and materials development. His published articles have appeared in journals like TESL Reporter, ELT Journal, and The Teacher Trainer. E-mail: sasanbaleghizadeh@yahoo.com

*Yahya Gordani* is current Ph.D. candidate of TEFL at Shahid Beheshti University, G.C. in Iran. He has vast experience in teaching English as a foreign language and has published in Asian EFL Journal.