Journal of Response to Writing

Volume 3 | Issue 1 Article 4

2017

The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures

Ahsan Pashazadeh
University of Tehran

Follow this and additional works at: https://scholarsarchive.byu.edu/journalrw

Part of the Arts and Humanities Commons, and the Education Commons

Recommended Citation

Pashazadeh, Ahsan (2017) "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures," *Journal of Response to Writing*: Vol. 3 : Iss. 1 , Article 4.
Available at: https://scholarsarchive.byu.edu/journalrw/vol3/iss1/4

This Featured Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Journal of Response to Writing by an authorized editor of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

www.journalRW.org

RW
JOURNAL OF RESPONSE TO WRITING

The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures

Ahsan Pashazadeh
University of Tehran

Studies that have reported delayed positive effects for written corrective feedback (WCF) have typically targeted the use of articles for first- and subsequent-mention functions, using narrowly focused corrections that lack ecological validity. Not much is known about how different grammatical features react to mid-focused and unfocused WCF options, which enjoy more ecological validity. This study investigates the delayed effect of different types of WCF on English as a foreign language (EFL) learners' accurate use of three features of English grammar (articles, infinitive, and unreal conditional). Four groups of participants (N = 77) were treated with different feedback options (mid-focused corrections, unfocused corrections, unfocused corrections plus revision, and no corrective feedback). WCF did not produce lasting accuracy gains, nor did it help corrected students outperform uncorrected students on a delayed posttest.

Keywords: accuracy, mid-focused WCF, unfocused WCF, revision

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.

Introduction

In a review of WCF studies, Truscott (1996) claimed that primary feedback studies showed correction to be harmful to accuracy, and he built a case for the abandonment of WCF in second language (L2) writing courses. Several researchers reacted to Truscott's review and questioned the validity of his radical thesis. Ferris (1999), for instance, maintained that empirical evidence was too limited to warrant the kind of conclusions drawn by Truscott. A few years later, there was a surge in experimental studies that set out to shed light on the efficacy of WCF. The results of the early attempts (e.g., Bitchener, Young, & Cameron, 2005; Chandler, 2003; Fazio, 2001) were not very encouraging for proponents of correction (see Truscott, 2007). Subsequently, researchers began to utilize highly focused corrections targeted at a couple of grammatical features, which showed that focused correction can produce significant gains in the accurate use of English articles for first- and subsequent-mention functions (Bitchener, 2008; Bitchener & Knoch, 2008, 2010a, 2010b; Sheen, 2007). Other studies showed that less focused corrections also have potential to help improve the accuracy of the two functions of English articles (Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007), but there is very little empirical evidence that WCF raises accuracy when targeted at other grammatical features.

One question that remains unanswered is whether there is a middle ground between correcting a range of grammatical errors (i.e., unfocused WCF), which seems to be ineffective in producing delayed accuracy gains, and correcting one or two error categories only (i.e., focused WCF), which lacks ecological validity. Is it possible to incorporate the effectiveness of highly focused corrections and the ecological validity of unfocused corrections into "mid-focused" WCF that targets a manageable number of error categories? If WCF is targeted at a manageably small number of error categories, will it be effective in helping L2 writers improve the accuracy of those structures on a delayed posttest? The aim of the present study was to investigate whether mid-focused corrections produce better delayed effects than (a) unfocused corrections, (b) unfocused corrections coupled with an opportunity to revise in class, and (c) writing practice alone. For this purpose, the accuracy of three English grammatical structures (i.e.,

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.

the use of articles for first and subsequent mentions, the infinitive, and the unreal conditional) was tracked using a pretest, an immediate posttest, and a delayed posttest administered one month after the treatment.

Literature Review

This section first discusses focused and unfocused corrections in L2 writing and details the studies that have experimented with these two types of corrective feedback, attempting to shed light on the factors that might help explain the conflicting results of the past and recent studies. The criteria for including a study in this part of the literature review were as follows: inclusion of a control group in the design, inclusion of a delayed posttest to measure delayed effects, use of reasonably authentic and valid measures of writing accuracy rather than such formal measures as multiple-choice grammar tests, and inclusion of new pieces of writing rather than revision tasks. Afterwards, revision and its significance in L2 writing is discussed. Finally, a separate section is devoted to the discussion of "dynamic WCF" studies, despite the lack of delayed posttests and true control groups that did not receive corrective feedback in dynamic WCF studies. The reason is that dynamic WCF studies share the concerns of the present study about the ecological validity of recent WCF studies that have reported positive outcomes for focused feedback.

Focused and Unfocused Feedback

In terms of focus, corrections are normally categorized into two types: focused and unfocused. Focused corrections, which have been utilized in a large number of studies in recent years with positive results, are offered on only one error category. Highly focused studies, such as Bitchener and Knoch (2010a, 2010b), have provided corrections on only two aspects of an error category (i.e., first- and subsequent-mention uses of English articles). Unfocused correction, which is also known as comprehensive correction, involves corrective feedback on a range of error categories. Liu and Brown (2015) have introduced the term *mid-focused* to refer to WCF that targets two to five error categories. If attention and noticing are factors that facilitate learning grammatical forms, as Schmidt (1994) argues, then corrections must become less useful to students as they lose focus. Thus, compared

to corrections that are offered on a wide variety of error categories, highly focused corrections have a better chance of being noticed by language learners. For this reason, recent scholarship has strongly advocated the use of focused corrections in writing courses.

Considering the results of recent feedback studies, one may be tempted to conclude that WCF is effective when it is focused on a simple and rule-governed aspect of grammar. However, it has been demonstrated by Shintani and Ellis (2013) that highly focused corrections might not help improve the accuracy of late-acquired features of grammar, even if they are simple and rule-governed. Shintani and Ellis provide evidence that although highly focused corrections increase the combined accuracy index of first- and anaphoric-reference functions of English articles, when the effects are teased apart, it can be demonstrated that it is only the anaphoric-reference function that benefits from corrective feedback and that the first-mention function is not affected by correction. Also, Fazio (2001) reported that focused corrections targeted at French grammatical spelling produced negative outcomes in terms of writing accuracy. Nevertheless, except for these two studies, the majority of the studies that have experimented with focused corrections have reported big delayed advantages for corrected students (see Bitchener & Knoch, 2010a for an overview).

One question that remains to be answered satisfactorily is whether WCF can prove effective when the focus of corrections is obscured with feedback that is less focused. The studies that have reported positive results for WCF have generally highlighted and focused on two functions of the article system ("a/an" for first mention and "the" for subsequent mentions) beyond what is considered reasonable in real language classrooms. It appears that in their efforts to prove the efficacy of WCF, researchers have experimented with types of feedback that lack ecological validity and potential for pedagogical application.

Even if there is compelling and irrefutable evidence to support the efficacy of focused corrections, there are at least two major reasons to explore the efficacy of less focused corrections. First of all, the actual culture of error correction in real-world classrooms demands the kind of correction that is not nearly as focused as the correction techniques utilized in recent WCF research. Although unfocused corrections are not recommended in recent

feedback scholarship, there is evidence that students demand corrections on all of their errors (Leki, 1991; Komura, 1999) and that language teachers normally offer corrections on a range of grammatical errors on the same essay (Ellis, Sheen, Murakami, & Takashima, 2008). Thus, despite all the recent recommendations, comprehensive correction seems to be the option that enjoys ecological and face validity, and the use of focused corrections is limited to the context of second language acquisition (SLA) studies.

The second reason is that early work in SLA shows that the course of acquisition is U-shaped (Kellerman, 1983). Focused feedback seems to be based on a linear concept of language acquisition: That is to say, students develop perfect or near-perfect command over certain linguistic structures and proceed to other structures in an additive fashion. Given the small number of essays produced during an instructional course, it would be impractical to use highly focused corrections to target a particular error category on more than a couple of occasions, whereas mid-focused corrections can be used to target a particular error category at different points in time. Thus, with mid-focused corrections, which do not assume a linear course of acquisition, L2 writing instructors can attempt to treat errors that keep reappearing in student compositions.

Of the studies that have experimented with unfocused corrections targeted at grammatical categories other than first- and subsequent-mention functions of English articles, only a few have reported positive outcomes for WCF (i.e., Chandler, 2003; Lalande, 1982; Sheen, Wright, & Moldawa, 2009; Van Beuningen, De Jong, & Kuiken, 2012). The positive results of Chandler (2003) and Lalande (1982) have been discussed and rejected by Truscott (2007), who showed that both authors had promoted their negative results as positive. Similar points can be made about the other two studies, which will be detailed below.

Sheen, Wright, and Moldawa (2009) set out to provide evidence for the efficacy of focused feedback (targeted at first- and subsequent-mention uses of English articles) and comprehensive corrections offered on a range of error categories. Unlike the studies reviewed by Truscott (1996), the study tracked the structures that had been corrected in earlier student essays. It can be argued that comprehensive feedback studies of the 1980s and 1990s failed in terms of accuracy improvement simply because they

did not track specific structures and reported a general index of accuracy. Challenging the conclusions of the Truscott and Hsu (2008) study, Bruton (2009) maintains that if the study had tracked the structures that had been corrected instead of reporting a general index of accuracy, signs of improvement might have been detected. This is a legitimate point. After all, it is not logical to correct a certain set of structures and then attempt to measure the effect on those structures as well as a number of other unrelated structures. Despite the claims by the authors to the contrary, the results of the Sheen et al. (2009) study show that unfocused corrections are not effective even when a select set of structures are corrected and tracked over time. In fact, the students who had received writing practice alone improved as much as or more than the students who had received either focused or unfocused corrections, both in the use of articles for first and subsequent mentions and in the use of the five grammatical categories targeted in the study. One important point is that the group Sheen et al. refer to as "control" was disadvantaged in terms of writing practice, as it did not perform two of the tasks that the other groups completed. For this reason, and in order to be consistent in the use of the term *control*, it is best if the "writing practice" group is considered the true control group.

Van Beuningen, De Jong, and Kuiken (2012) included four feedback conditions in their design: comprehensive direct corrections, comprehensive indirect corrections, self-correction, and writing practice. As with the Sheen et al. (2009) study, despite what the authors have claimed, because the students in the self-correction group were made conscious of their errors and were required to correct them, they cannot be considered a true control group (see Hyland & Hyland, 2006, for a discussion of different correction options). The only genuine control group was the one that received writing practice alone. The results showed that only the direct correction group outperformed the writing practice group on the delayed posttest and that there were no other significant differences between any of the groups. The important point about the study is that the "positive" effects assigned by the authors to WCF were, for the most part, the result of the writing practice and self-correction groups declining drastically in accuracy on the delayed posttest, rather than the direct and indirect correction groups making accuracy gains, as it was only the direct

correction group that made small gains on the delayed posttest, with the other three groups making declines.

Thus, apart from a few studies (e.g., Ellis et al., 2008) that have found positive effects for unfocused corrections on the accurate use of English articles for first and subsequent mentions, there is no evidence that comprehensive corrections are a useful tool to promote accuracy, particularly in the use of grammatical structures other than English articles. Clearly, in order for WCF to have a place in writing courses as an accuracy-promoting instructional practice, there is a need to demonstrate the efficacy of corrective feedback (especially, ecologically valid types of WCF such as mid-focused and unfocused corrections) on the acquisition of other grammatical structures that vary in terms of complexity.

Revision and Its Significance

After "process writing" became popular in the 1980s and 1990s, editing for sentence-level errors was deemphasized and delayed until the very last stage of writing, a practice which subsequently came under criticism as form-focused instruction gained ground in language teaching circles, and it was argued that sentence-level errors also deserve attention in the process of drafting and redrafting (Ferris & Hedgcock, 2005). It has been argued that revision frees up students' attentional resources to an extent and enables them to focus on form as well as on content. Bruton (2009) maintains that corrective feedback cannot be expected to have positive effects if the corrected students are not required to engage with it in some way. Ferris (2003, 2010) has argued that the feedback studies reviewed by Truscott (1996, 2007) failed to obtain positive results partially because they for the most part disregarded the common practice in writing courses—namely, prewriting, writing, feedback, and revision.

If, as Ferris (2003) argues, the failure of feedback in the studies reviewed by Truscott (1996) can be explained by the lack of revision in those studies, an interesting area for research would be whether requiring corrected students to revise their essays translates into any delayed accuracy gains in the use of corrected structures. It should be reiterated that the studies reviewed by Truscott (1996) utilized comprehensive corrections and did not track improvement in the use of the structures that had been corrected,

reporting only a general measure of accuracy. It would be interesting to see whether eliminating this shortcoming by tracking specific structures makes a difference. The effect of feedback and revision on specific structures (as opposed to feedback alone) is an area that has not been researched properly.

One way in which students can be forced to pay attention to corrections is to require them to revise their essays by incorporating the corrections they receive, thereby increasing the potential of corrective feedback for consciousness-raising. It is interesting to see whether and to what extent different grammatical structures are affected by feedback options that differ in terms of consciousness-raising potential. If the extra time that is spent on revision pays off in the form of improved delayed accuracy, language teachers cannot afford to ignore revision as an accuracy-promoting technique. On the other hand, if revision does not help improve linguistic accuracy on a delayed posttest, the extra time that is spent on it is not justifiable in terms of language acquisition (although it might still have value as a tool that helps students produce better drafts in the short term).

Dynamic WCF

Concerns about ecological validity of WCF studies have led to the development of *dynamic WCF*, an innovative approach to written feedback (Evans, Hartshorn, & Strong-Krause, 2011; Hartshorn et al., 2010; Hartshorn & Evans, 2012; Hartshorn & Evans, 2015). The results of dynamic WCF studies have been encouraging, although the methodology is more labor intensive compared to regular WCF practices. Hartshorn et al. (2010), whose design is typical of dynamic WCF studies, experimented with 47 advanced-level students in a U.S. university who were divided into two groups, one experimental and one contrast. The experimental group received dynamic WCF, involving coded feedback, instruction on how to interpret the error codes, error logs, short 10-minute essays written almost every day of the 15-week course, revision, assigning of grades based on holistic assessments, and personalized corrections. The contrast group, on the other hand, received regular process-oriented writing instruction, consisting of multidraft essay writing coupled with corrective feedback and revision. At the end of the experiment, the contrast group regressed in terms of overall accuracy, whereas the experimental group improved

significantly. The students in the experimental group were harmed slightly in terms of fluency and complexity. Based on the findings, the authors argue that in order for feedback to be effective, it has to be manageable, timely, constant, meaningful, and personalized (i.e., the components of dynamic WCF). Similar results have been reported by the other dynamic WCF studies which were mentioned above.

In summary, previous research has shown that WCF that is narrowly focused on the use of articles for first and subsequent mentions can help with accuracy in terms of both immediate and delayed effects. However, the reported delayed effects for unfocused corrections are not very positive. In the error correction literature, the general assumption has been that correction produces positive delayed effects, and two arguments have been presented to explain the conflicting results of focused and unfocused studies. The first argument is that generally unfocused correction studies did not track the structures that had been corrected, and the second argument is that some of these studies did not require the corrected participants to revise their essays on the basis of the corrections, thereby undermining the potential of WCF for raising accuracy levels. The present study attempted to address these two shortcomings of previous unfocused WCF studies. The following research questions were formulated:

1. Do different WCF options (i.e., mid-focused, unfocused, and unfocused plus revision) help experimental groups outperform the control group on a delayed posttest?
2. Does the provision of WCF or content comments help students significantly improve the accuracy of their writings on a delayed posttest?

The first research question is concerned with group differences and the performance of each group in comparison with other groups, whereas the second deals with absolute gains or declines on the delayed posttest in comparison with pretest scores.

Method

Participants and Setting

This study included 116 students from nine "pre-intermediate" classes in an all-male language school in Tehran, Iran. The language school had four main proficiency levels for placement purposes: elementary, pre-intermediate, intermediate, and advanced. It took around six months for the students to complete each proficiency level at the language school, assuming that they did not fail any of the 6-week "terms." The participants (almost all undergraduate and graduate students) had received at least six years of formal instruction in middle and high school; however, due to inefficiency of secondary education in Iran and very limited opportunities to interact with native speakers, students barely learn any English in secondary school. Generally, Iranian students who learn English in language schools are highly motivated to acquire English for academic and occupational success. Data from 39 of the students were excluded because of absenteeism, dropout, and (in the case of four students) failure to produce the targeted grammatical structures. The teachers who helped with data collection described the rate of absenteeism and dropout as "normal" for the language school in which the research was carried out. In terms of mother tongue, the participants were from Persian, Azeri, Kurdish, and Talyshi linguistic backgrounds. Thirty-seven of the 116 students said that they had been raised bilingual, speaking Persian (the official language) plus one of the other three languages. The ages of the participants ranged from 17 to 46, with a mean age of 25.8. The courses that the participants were enrolled in lasted for about six weeks and had speaking, listening, reading, and writing components. The textbooks that were used for the courses included in- and out-of-class writing activities. Most homework assignments were writing activities. The students met four times a week, and each class period was slightly more than one and a half hours. The courses were conducted almost exclusively in English by three teachers who agreed to help with data collection. All the teachers had over five years of language teaching experience.

Treatment

Students in the four feedback conditions received different treatments. The mid-focused group ($n = 23$) received corrections on all errors in the use of first- and anaphoric-reference functions of articles, the infinitive, and the unreal conditional. In the case of the unfocused group ($n = 17$), all errors in the use of the three grammatical structures (i.e., first- and anaphoric-reference functions of articles, the infinitive, and the unreal conditional) as well as a range of other error categories were corrected by the researcher: No attempt was made to correct each and every error. The students in the revision group ($n = 16$) received the same kind of corrections as the unfocused group, and they were also required to revise their essays on the basis of the corrections they had received. The students in the control group ($n = 21$) only received a couple of general comments (such as "the sentences about the 'cage' and 'famous people' are missing here!") relating to the content of their writings: No corrective feedback was provided on the use of any grammatical structure, and care was taken not to provide any content comments that would draw attention to grammatical errors.

The following are examples of errors that the participants made in the use of each of the three targeted error categories:

1. Errors in the use of articles: If I saw **bird*** in **cage***, I would free the bird from **cage***.
2. Error in the use of the infinitive: I always wanted **live*** in different places.
3. Error in the use of the unreal conditional: If I were a bird, I **fly*** to a new city, town or village every week to live there.

Instruments

The instruments that were utilized for this study consisted of one demographic information questionnaire that asked for such information as age and mother tongue, and three different reading passages coupled with three writing tasks. The three reading passages had been constructed carefully to include instances of the three structures targeted in the study (namely, first- and anaphoric-reference functions of articles, the unreal conditional, and the infinitive). The three writing tasks included instructions that asked the students to reproduce the reading passages as closely as they

could remember, a measure which was adopted in order to make sure the students would produce at least a portion of the structures included in the reading passages. The scores from the first, second, and third writing tasks were used as pretest, immediate posttest, and delayed posttest, respectively. The scores for accuracy in the use of the three targeted structures were calculated using the following formula for Target-Like Use (see Pica, 1994):

$$Accuracy = \frac{Correct\ instances\ provided\ in\ obligatory\ contexts}{All\ obligatory\ contexts + suppliance\ in\ wrong\ contexts} \times 100$$

For example, if the three structures were used correctly 15 times in 23 obligatory contexts and were oversupplied in 7 contexts in which their use was not necessary, the overall accuracy score would be 50%.

Procedure

Prior to the onset of the study, the teachers were briefed in a meeting on the kind of instruction they needed to avoid (specifically, instruction on first- and anaphoric-reference functions of English articles, the unreal conditional, and the infinitive, none of which were the focus of grammar lessons in the textbook, although the infinitive and the two functions of articles had been covered in earlier textbook units), and they were provided with general information regarding the study and its purpose. Having obtained the students' agreement to participate in the study, the researcher asked the participants to fill out a biographical data questionnaire. Afterwards, the participants were given the first reading task. The instructions at the top of the reading passage informed the participants that they would be required to reproduce the same passage after 10 minutes without being able to consult the passage. The students were told at this point that they were free to write down key words from the reading passage and that they would be able to ask the researcher for help with remembering the facts from the reading passage. In the majority of the cases, the researcher helped the students who needed prompts to remember the content of the reading passage by giving them information in Persian. The researcher walked around in the classrooms to make sure nobody was writing down large stretches of the reading passages, which would have defeated the whole purpose. The reading passages were

collected 10 minutes after distribution. All three of the reading passages that had been designed were less than 180 words long.

After the reading passages were collected, the students were given 20 minutes to complete the first writing task, a decision that was adopted on the basis of a small-scale piloting of one of the writing tasks with a group of students at the same pre-intermediate proficiency level. The student essays were collected and analyzed for accuracy in the use of the three structures that were targeted in the study, and this yielded the pretest scores. The errors were underlined, and the correct forms were written in red pen above or near the underlined errors. The control group received a couple of comments relating to content.

The essays displaying the researcher's corrections or content comments were returned to the students in the next class meeting. The students were asked to look at the corrections or content comments (and in the case of the revision group, the students were given 20 minutes to rewrite their essays on the basis of the corrections). Afterwards, the second reading and writing tasks were administered in the same way as the first reading and writing tasks. The second essays were analyzed later on for accuracy in the use of the three structures, which yielded the immediate posttest scores. The second writing task was returned to the students a few days later with no corrections or content comments. The researcher revisited the classes after four weeks to administer the third reading and writing tasks, which were analyzed for accuracy in the use of the three targeted structures and yielded the delayed posttest scores. Except for the corrections, the procedures that were adopted in the administration of all the three reading and writing tasks were the same.

Results

Having scored 20 essays randomly selected from among the essays written during the first writing session, the researcher asked a second rater to rescore the same essays while comparing the writings to the reading passage. The purpose of this step was to clear possible misunderstandings caused by the large number of mistakes in some essays and to make scoring more reliable and accurate. The inter-rater reliability score was .94 for the three structures combined. Given the high rate of inter-rater

reliability and the amount of work involved in scoring the essays, the rest of the compositions were scored by the researcher alone. Also, in order to determine intra-rater reliability, the researcher rescored the same 20 essays around three months after initial scoring, and a reliability of .96 was obtained (for the three structures combined). Further analysis revealed that inter-rater and intra-rater reliability values for each of the three structures were similar to the overall reliability values, although the scores relating to the accurate use of articles were slightly less consistent than the scores relating to the use of the other two grammatical structures.

The descriptive statistics for the overall accuracy scores (which are reported as percentages) are displayed in Table 1 for the experimental and control groups at the three testing points. The overall accuracy scores from the three testing points (i.e., pretest, immediate posttest, and delayed posttest) were fed into Statistical Package for the Social Sciences (SPSS), and a mixed within-between subjects test of analysis of variance (ANOVA) was conducted in order to determine whether the students in the four feedback conditions performed differently over time. The dependent variable was the overall accuracy scores in the use of the three structures targeted in the present study, and the independent variables were the feedback conditions and the testing times. Tests of statistical significance do not reveal the magnitude of an observed effect or the *effect size*. Therefore, in the present study, partial eta-squared (η_p^2) will be reported as the measure of effect size, and in line with the guidelines presented by Cohen (1988), the value of partial eta-squared (which is always positive and varies between zero and one) will be interpreted in the following way: The value of .01 will be interpreted as a small effect, and the values of .06 and .14 will be interpreted as medium and large effect sizes, respectively.

Table 1

Descriptive Statistics for Overall Accuracy

	Pretest	Immediate posttest	Delayed posttest
	Mean (SD)	Mean (SD)	Mean (SD)
Mid-focused (n = 23)	17.58 (12.53)	33.41 (15.65)	22.12 (8.57)
Unfocused (n = 17)	16.36 (8.51)	23.64 (17.60)	18.19 (9.26)
Revision (n = 16)	13.01 (7.44)	28.03 (14.68)	15.08 (7.40)
Control (n = 21)	14.94 (8.69)	17.14 (7.90)	17.42 (9.38)

There was a significant interaction between the feedback conditions and the testing times with a medium effect size, $F_{(6, 146)} = 2.56$; $p < .05$; $\eta_p^2 = .09$. As is illustrated graphically in Figure 1, this means that the four groups performed differently over time. That is to say, whereas the students in the three feedback conditions improved and declined substantially on the immediate and delayed posttests respectively, the control group improved on the immediate posttest and continued improving on the delayed posttest, although the gains in accuracy were negligible both times. As for the main effects, there was a significant effect for the different testing times

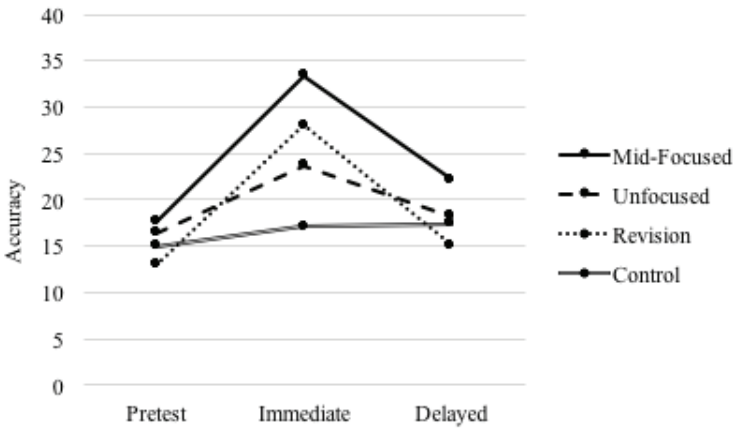

Figure 1. Overall Accuracy Over Time

with a very large effect size, $F_{(2, 146)} = 24.66$; $p < .05$; $\eta_p^2 = .25$. Also, the effect of the grouping variable (namely, the different feedback conditions) turned out to be significant as well with a medium effect size, $F_{(3, 73)} = 3.49$; $p < .05$; $\eta_p^2 = .12$.

A separate test of one-way ANOVA (with the feedback conditions as the independent variable and the overall scores from the three testing points as the dependent variable) was conducted to specify the exact location of significant group differences. Significant differences were detected on the immediate posttest only, $F_{(3, 73)} = 5.03$; $p < .05$. The immediate posttest will not be highlighted in the present study because the focus is on the delayed effects of correction. There were no significant differences between any of the groups on the delayed posttest, $F_{(3, 73)} = 2.24$; $p > .05$. That is to say, the delayed effect of mid-focused corrections proved to be similar to that of unfocused corrections (with or without revision) and content comments alone.

As Figure 1 shows, compared with their pretest scores, all four groups made small gains in overall accuracy from the beginning to the end of the experiment. In order to determine statistical significance of the absolute accuracy gains for each of the groups on the delayed posttest, separate paired samples *t* tests were run for each group on the overall pretest and posttest scores. The results showed that there were no statistically significant changes in accuracy scores from pretest to posttest in any of the groups. The strength of the change in accuracy scores was small for all groups ($\eta_p^2 = .043$, .01, .019, and .018 for the mid-focused, unfocused, revision, and control groups, respectively). Changes in accuracy scores of each of the three structures targeted in the present study were similar to the overall accuracy scores, in that the experimental groups made substantial gains in accuracy on the immediate posttest in the case of every one of the structures, but these gains were generally lost on the delayed posttest. However, for lack of space, the three structures will not be analyzed and discussed separately in this article.

Discussion

All four groups slightly improved their overall accuracy from the beginning of the study to the end, but this improvement was too small to

reach statistical significance. In the case of the experimental groups, there were substantial gains in accuracy on the immediate posttest, but these gains were lost by the delayed posttest administered one month later. The fact that there were substantial gains in the mean accuracy rates among the students whose essays had been corrected signifies that the students did indeed notice the corrections. In contrast, the control group showed little progress on the immediate posttest. Data analysis using a within-between ANOVA showed that the interaction effect between feedback and writing practice was significant. This means that as far as the overall accuracy scores are concerned, the treatment had an effect, but further analysis revealed that this effect was due to the drastic improvement in the accuracy scores of the correction groups on the immediate posttest and that there was no significant effect for correction on the delayed posttest.

The results of the within-between test of ANOVA also showed that the effects of feedback alone and writing practice alone were statistically significant as well, although these "main effects" are hard to interpret. The reason is that in the case of the main effect for feedback, the statistical procedure spreads out the effect of treatment (i.e., the four WCF options in the present study) over all testing times, including the pretest point. This is problematic because technically pretest scores are not affected by treatment. Also, the within-between test of ANOVA averages across groups in the case of the main effect for time (i.e., writing practice in the present study). In other words, all treatment and control groups are collapsed into one big group, which makes the main effect for writing practice meaningless because we are rarely interested in in the average performance of all treatment and control groups (for full elaboration of how to interpret interaction and main effects in within-between ANOVA tests, see Huck & Mclean, 1975).

One of the criticisms against comprehensive WCF studies reviewed by Truscott (1996) is that the measure of accuracy used in those studies did not exclude the structures on which the students did not receive corrections. That is to say, those studies only reported a general measure of overall accuracy, normally based on error rates. The findings of the present study show that as far as overall accuracy in the use of the three structures is concerned, comprehensive correction does not produce

delayed advantages in accuracy (which are widely interpreted as signs of acquisition and genuine improvement) even if the corrected structures are tracked over time.

The results of this study run counter to the results obtained by Van Beuningen, De Jong, and Kuiken (2012) and Ellis et al. (2008), both of which showed delayed effects for unfocused corrections. However, the positive results obtained by Van Beuningen et al. were mostly because of the accuracy declines in the control group, not because of the improvement in the experimental groups. The results of different corrective feedback studies show that the control group might decline or improve during the course of a study. This decline or improvement is considered a natural part of the dynamic process of L2 acquisition. When there are declines in the control condition and little or no absolute gains in experimental conditions, it is not a good idea to promote group differences as signs of the success of correction because it would imply that writing practice alone hurts students in terms of delayed accuracy, which is an untenable position that nobody has promoted. In the case of the Ellis et al. study, the measure of accuracy was based on the use of first- and subsequent-mention uses of English articles alone, which might be less complex than the infinitive and the unreal conditional (i.e., two of the three components of the overall accuracy index used in the present study). For this reason, the positive results of the Ellis et al. study might have resulted from the relatively simple focus of the unfocused and focused corrections.

Of particular interest in the present study is the behavior of the group that received unfocused feedback and was required to revise on the basis of the corrections: It showed improvement in overall accuracy on the immediate posttest but ended up at almost the same level of accuracy as at the beginning of the study. This is while the group that only received unfocused corrections and was not required to revise ended up making slightly more gains than the revision group. If we assume that correction is useful because it aids noticing, it stands to reason that revision helps corrected students pay further attention to the corrections they have received. However, in the present study, revision did not have a positive effect on delayed accuracy. Although it is completely possible that the non-significant results were due to chance or to factors uncontrolled in

the study, it is also possible that noticing is not positively correlated with increased levels of accuracy. In other words, it is possible that noticing might be harmful in the case of certain structures, particularly if there is reason to believe that noticing does not lead to understanding in the case of those structures. Of the studies that have compared revision with lack of revision, some have reported positive results for revision and some have reported no effects or negative effects. For example, Chandler's (2003) revision group made accuracy gains, while Polio, Fleck, and Leder's (1998) revision group declined in accuracy, although none of these studies tracked specific structures. The present study is the only attempt in WCF literature to compare the delayed effect of revision with corrective feedback alone while tracking specific structures.

That all the corrected groups declined from immediate posttest to delayed posttest is consistent with the vast majority of the studies that have been carried out in the area of L2 acquisition in general, although the declines in the present study were sharper than most studies. Norris and Ortega (2000) note in their meta-analysis that the effect of treatment tends to wear off over time, which is what was observed in this study. The students who had received corrections on their pretest essays were allowed to consult the corrections on the immediate posttest—this could have resulted in improved accuracy without necessarily understanding the grammatical logic behind the corrections. Also, the one-month time lapse could have undermined the possible memory factors, leading to decreased levels of accuracy on the delayed posttest.

In terms of absolute gains in accuracy, at the end of the experiment, the unfocused and revision groups were practically indistinguishable from the control group, and the group that received mid-focused corrections was only marginally better than the control group. These results are inconsistent with the majority of the studies that have experimented with focused and unfocused corrections over the past decade. Studies such as Bitchener (2008) that have used focused corrections have reported large gains in accuracy for corrected students. However, unlike the present study in which three different grammatical features were corrected, those studies have only targeted first- and anaphoric-reference functions of the English article system. The findings of the present study show that as far as combined

effects are concerned, mid-focused corrections are only marginally and non-significantly better than unfocused WCF and no correction.

In summary, the first research question asked whether corrected students outperform uncorrected students in terms of delayed accuracy, and the answer to this question is negative. The second research question asked whether the provision of WCF or content comments helps students significantly improve the accuracy of their writings over the course of the study, the answer to which is also negative. The kind of WCF utilized in the present study neither helped corrected students outperform uncorrected students, nor did it help them significantly improve their accuracy scores from the beginning to the end of the study.

Conclusion

In general, the present study points to the inefficacy of the kind of WCF that was used in this study. The results showed that although unfocused corrective feedback (with or without revision) and mid-focused feedback targeted at three different grammatical categories cause short-lived gains in accuracy rates among corrected students compared with uncorrected students, these gains do not last. If we come to the conclusion that corrective feedback is ineffective in producing positive delayed effects, or if the improvement comes at the cost of enormous time and effort on the part of L2 teachers, we might question the instructional value of correction, as the time and effort can be spent more productively on another area, such as preparing for class, strategy training, and grammar lessons, especially if ensuring the efficacy of corrective feedback requires extra time and effort in the form of revision, reformulation, error logs, and other such techniques that have been promoted by L2 writing scholars.

The first and most important limitation of this study is that there was only one round of correction. It might be argued that correction should be offered on numerous occasions in order to prove effective. However, many of the recent studies that have reported significant positive results for both focused and comprehensive corrections (e.g., Bitchener & Knoch, 2010b; Van Beuningen et al., 2012) have utilized only one round of corrections and have assigned a constructive role to corrective feedback on the basis of their positive results. In order to make the results of the current study

comparable to these recent studies and for reasons of practical convenience, it was decided that only one round of correction would be offered to the participants, which makes the results difficult to apply to normal writing courses in which teachers offer several rounds of correction, often involving the same error categories on several occasions, although probably not with the same level of consistency or intensiveness that was utilized in the present study.

The second important limitation of the present study is that the participants were instructed to reproduce the reading passages as closely as possible, which cannot be considered authentic writing. Nevertheless, because it is extremely difficult to design authentic writing tasks that force students to use certain structures in their writing, as students can almost always find ways of avoiding the structures that they are not comfortable with, the reproduction task, which has been utilized in other WCF studies (e.g., Sheen, 2007), can be considered a reasonably authentic option. In other words, although the kinds of feedback that were used in the present study (i.e., mid-focused and unfocused WCF) were more ecologically valid than the focused feedback of recent studies, the nature of the reproduction tasks would limit generalizability of the results.

The third limitation of the present work is that only accuracy scores were considered, and other measures such as fluency and complexity were not included. It is possible that these other measures might have shown some positive effects for WCF. Accuracy is an important consideration in L2 writing courses, but it is not the only consideration. It is necessary that the results be interpreted with this shortcoming in mind, although the nature of the reproduction writing tasks used in the present study would have limited the implications of fluency and complexity measures if they had been used.

Finally, the study was conducted with a limited number of participants at particular proficiency levels and was focused on a limited number of grammatical categories. Therefore, it is not possible to generalize the results to other participants and grammatical structures. Moreover, if the participants had been tested with other testing options, particularly formal grammatical tests, different results might have been obtained. This means that even if correction fails to produce accuracy gains in authentic or semiauthentic writing, this should not be considered as evidence to

reject correction altogether, as, for example, its value for training students to take formal multiple-choice tests would still be unknown (see Cardelle & Corno, 1981, for evidence of how corrective feedback can contribute to formal knowledge of language). It is best to consider the current study as a snapshot of a particular learning situation at a specific point in time, and thus the results could have been different if the participants had been tested in different ways or at different points in time, or if a different student population had been selected. For this reason, generalizing the results to other contexts would involve a degree of uncertainty.

Despite the vast amount of research that has been carried out over the past half century in the area of corrective feedback, there is a need to carry out further research to shed light on the remaining WCF questions, particularly in the area of selective feedback. For example, there is a distinct possibility that focused corrections on different error categories may interfere with one another over time. For example, corrective feedback on the use of past perfect may interfere with learners' knowledge of past simple. This problem becomes particularly serious when corrections are offered on grammatical structures that are closely related. For instance, it would be very difficult for L2 writers to distinguish the passive suffix *–ed* from the past tense *–ed* on the basis of written corrections. Because of the complexity of grammar, these problems may result in student confusion, thereby reducing the efficacy of selective corrections in the long run. Similarly, corrections on different aspects of a certain grammatical category—such as the English article system—at different points in time are likely to lead to the same problems. So far, no study has investigated this possibility.

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.

References

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing, 17*(2), 102–118.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research, 12*(3), 409–431.

Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31*(2), 193–214.

Bitchener, J., & Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19*(4), 207–217.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 14*(3), 227–258.

Bruton, A. (2009). Designing research into the effect of error correction in L2 writing: Not so straightforward. *Journal of Second Language Writing, 18(*2), 136–140.

Cardelle, M., & Corno, L. (1981). Effects on second language learning of variations in written feedback on homework assignments. *TESOL Quarterly, 15*(3), 251–261.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12*(3), 267–296.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36*(3), 353–371.

Evans, N., Hartshorn, K. J., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System, 39*(2), 229–239.

Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- and majority-language students. *Journal of Second Language Writing, 10*(4), 235–249.

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

Ferris, D. R. (2003). *Response to student writing: Implications for second language students.* Mahwah, NJ: Erlbaum.

Ferris, D. R. (2010). Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition, 32*(2), 181–201.

Ferris, D. R., & Hedgcock, J. S. (2005). *Teaching ESL composition: Purpose, process and practice* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Hartshorn, K. J., & Evans, N. W. (2012). The differential effects of comprehensive corrective feedback on L2 writing accuracy. *Journal of Linguistics and Language Teaching, 3*(2), 217–247.

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.

Hartshorn, K. J., & Evans N. W. (2015). The effects of dynamic written corrective feedback: A 30-week study. *Journal of Response to Writing, 1*(2), 6–34.

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly, 44*(1), 84–109.

Huck, S. W., & McLean, R. A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potential confusing task. *Psychological Bulletin, 82*(4), 511–518.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal, 99*(1), 1–18.

Kellerman, E. (1983). Now you see it, now you don't. In S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112–134). Rowley, MA: Newbury House.

Komura, K. (1999). *Student response to error correction in ESL classrooms* (Unpublished master's thesis). California State University, Sacramento.

Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal, 66*(2), 140–149.

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*(3), 203–218.

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30*, 66–81.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493–527.

Polio, C., Fleck, N., & Leder, N. (1998). If only I had more time: ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing, 7*(1), 43–68.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Consciousness in Second Language Learning, 11*, 237–326.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255–283.

Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System, 37*(4), 556–569.

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22*(3), 286–306.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46*(2), 327–369.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16*(4), 255–272.

Truscott, J., & Hsu, A. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292–305.

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

Copyrights
© JRW & Authors.
Copyright for this article is retained by the author(s), with first publication rights granted to the Journal. This is an open-access articles distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Pashazadeh, Ahsan. (2017). "The Effect of Mid-Focused and Unfocused Written Corrections on the Acquisition of Grammatical Structures." *Journal of Response to Writing, 3*(1): 56–82.