Should Feedback be Direct or Indirect? Comparing the Effectiveness of Different Types of WCF on L1 Arabic Writers' Use of English Prepositions

Khaled Karim*, Martin J. Endley

Department of Linguistics, United Arab Emirates University, UAE

Abstract

This study investigated the effects of direct and indirect WCF on students' revision accuracy as well as on new pieces of writing over time. Intermediate level pre-faculty university students were divided randomly into four groups: direct, underlining only, underlining metalinguistic, and a control group. They produced two texts from two different picture prompts and revised these over a two-week period. They also produced two new texts three and four weeks after the last treatment. Both WCF significantly reduced learner errors in subsequent revision tasks. All three treatment groups also made significant reduction of errors after 3 weeks of CF treatment.

Keywords: *Written corrective feedback, Direct and indirect feedback, Comprehensive corrective feedback, Focused feedback, Unfocused feedback*

Introduction

While corrective feedback (CF) is a central component of second language (L2) writing programs around the world, there continues to be considerable controversy regarding its value and effectiveness in the literature. Most L2 writing researchers will be familiar with the debate initiated by the publication of Truscott's paper in 1996, which strongly argued against the effectiveness of grammar correction in L2 writing, calling CF a waste of time; and Ferris' (1999) response to Truscott's argument, providing both theoretical and empirical evidence in support of error correction in L2 writing. Since then, an increasing number of L2 writing studies have investigated the effectiveness of written corrective feedback (WCF) in various contexts. These

khaledk@uaeu.ac.ae

studies have examined not only the efficacy of WCF itself, but also the differential effects of different types of WCF (e.g., Baker & Bricker, 2010; Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009a; Bitchener & Knoch, 2009b; Bitchener, Young & Cameron, 2005; Ellis et al., 2008; Ferris, 2006; Hartshorn et al., 2010; (Author XXX), 2018; Liu, 2008; Mirzaii & Aliabadi, 2013; Sheen, 2007; Sheen, Wright & Moldawa, 2009; Storch & Wigglesworth, 2010; Truscott & Hsu, 2008; Van Beuningen, De Jong & Kuiken, 2008; Van Beuningen, De Jong & Kuiken, 2012). Among these studies, several have focused on comparing the effectiveness of direct and indirect WCF, exploring also the effectiveness of different subtypes of direct feedback, such as direct WCF only, direct WCF with underlining, and direct WCF together with oral or written metalinguistic WCF (e.g., Bitchener, 2008; Bitchener et al., 2005; Bitchener & Knoch, 2009a, b; Sheen, 2007; Sheen et al., 2009). While some of these studies have shown an advantage for indirect feedback (e.g., Ferris, 2006; Sheen, 2007) over direct feedback, others have found the reverse (e.g., Bitchener et al, 2005; Mirzaii & Aliabadi, 2013; Van Beuningen et al., 2008, 2012). Thus, there is disagreement among researchers regarding which type of CF is more effective.

A further point worth noting is that most of the studies that found WCF to be effective (e.g., Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009 a & 2009b; Bitchener et al., 2005; Ellis et al, 2008; Sheen, 2007; Sheen et al., 2009; Shintani & Ellis 2013, Shintani, Ellis & Suzuki, 2014) have targeted either a single grammatical structure or only a limited number of grammatical structures. In a recent synthesis of research, Liu and Brown (2015) drew attention to a major methodological limitation of such 'one-shot' treatments in that while they may be easier to implement, their pedagogical validity is questionable. Moreover, as Storch (2010) points out, "learning requires extensive and sustained meaningful exposure and practice" (p. 42). Therefore, there is a need for studies that investigate the effects of focused CF over a longer period of time. The study presented here was designed to address some of these concerns. First, it compared the effects of direct CF with two types of indirect CF; second, it investigated the effect of CF on both revision of the same piece of writing and on new pieces of writing. To this end, the study involved multiple feedback sessions on multiple pieces of writing, tracking the revision effects of feedback on individual pieces of writing and also its putative learning effect on subsequent writings overtime.

Literature Review
Studies of Corrective Feedback in L2 Writing
Ever since the publication of Truscott (1996) researchers have debated the effectiveness of corrective feedback in L2 acquisition and L2 writing development. As Hyland and Hyland (2006) put it some twelve years ago, "while feedback is a central aspect of L2 writing programs across the world, the research literature has not been unequivocally positive about its role in writing development" (p. 83).

A notable feature of some of the earliest WCF studies was that they adopted a quite general perspective, comparing the provision of feedback with no feedback (e.g., Ashwell, 2000;

Fathman & Whalley, 1990; Ferris & Roberts, 2001; Kepner, 1991; Polio, Fleck & Leder, 1998; Semke,1984). Very few early studies considered the relative effectiveness of different types of feedback. Moreover, those which did often reported strikingly contradictory results. Thus, Robb, Ross and Shortreed (1986) found no differences based on different types of feedback; similarly, Semke (1984) found no significant differences in accuracy across two treatment groups, one receiving direct feedback and one receiving indirect feedback. On the other hand, Lalande (1982) found that learners who had been provided with indirect feedback showed greater improvement than learners who had been provided with direct feedback. Again, in Sheppard's (1992) comparison of the effects of holistic comments versus direct CF on grammar and linguistic complexity, both groups demonstrated improvement on the correct use of verbs; however, in terms of grammatical accuracy and linguistic complexity the group that was provided with holistic comments outperformed the group that received direct feedback.

In addition to the somewhat general approach that tends to characterise these early studies, several critics (e.g., Bitchener, 2008; Ferris, 2008; Storch, 2010) have pointed out that these studies also display various methodological limitations. Among these we may note the lack of a control group; the failure to measure improvement of accuracy in new pieces of writing; and (more contentiously) the provision of feedback on all the errors (that is, unfocused feedback) rather than on one or only a small number of types of error (focused feedback). It has been noted that the failure to measure improvement of accuracy in new pieces of writing is especially unfortunate. Without such a measure there is no way of knowing whether the same errors were repeated in subsequent pieces of writing or whether errors not found in the initial writing sample appear subsequently. In other words, there is no way of knowing whether writing development has taken place (Bruton, 2009).

A feature of various more recent studies is the attempt to address these design flaws, including a control group and measuring improvement of accuracy in new writing, while examining the effectiveness of different types of feedback, both direct and indirect. Again, however, the findings have tended to be noticeably mixed, with little sign of any consensus emerging among researchers. On the one hand, some recent investigations have reported an advantage for indirect feedback (e.g., Ferris, 2006; Sheen, 2007) and metalinguistic feedback (e.g., Bitchener et al. 2005; Bitchener 2008; Ellis et al., 2006; Gholaminia, Gholamnia & Marzbanc, 2014; (Author XXX), 2018; Sheen, 2007; Sheen 2010a) over direct feedback. On the other hand, some other studies have reported direct feedback to be more effective (e.g., Bitchener et al., 2005; Mirzaii & Aliabadi, 2013; Van Beuningen et al., 2008, 2012). Notable here are two studies by Van Beuningen et al. (2008) and (2012), both of which found more significant accuracy gains for direct feedback on new texts, although both direct and indirect feedback were effective for revised texts. Again, Bitchener's (2008) study indicated that the two groups that received written or oral metalinguistic feedback along with direct feedback outperformed the group which received direct feedback only. Similarly, Bitchener et al. (2005) reported that a combination of direct oral metalinguistic feedback and direct written feedback was more effective than direct written feedback alone. However, Bitchener and Knoch's (2008) study

reported no difference between three treatment conditions (direct WCF, written and oral metalinguistic explanation, and direct WCF only). Sheen (2007) found that the group receiving indirect metalinguistic feedback outperformed the group receiving direct feedback group in a delayed posttest. Gholaminia, Gholaminia, and Marzbanc (2014) also found that their metalinguistic code-correction group outperformed their direct feedback group. Shintani, Ellis and Suzuki (2014) found a more durable effect for direct WCF than for metalinguisctic explanation. In sum, then, the picture with regard to WCF is a complex one. Indeed, in their recent meta-analysis, Kang and Han (2015) concluded that there is no clear-cut evidence in the research literature about the efficacy of direct vs. indirect WCF.

A further issue debated in the literature is the question of whether feedback is more effective if focused on selected structures or whether comprehensive feedback (i.e., correction of all errors) is better. Of those studies that have reported positive effects for feedback, most have targeted only a limited number of grammatical structures rather than providing comprehensive feedback on all errors. For example, in Bitchener, Young and Cameron (2005) feedback was targeted on the definite article, the simple past tense, and the use of prepositions. Other researchers (e.g., Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009a; Bitchener & Knoch, 2009b; Bitchener et al., 2005; Ellis et al., 2008; Sheen, 2007; Sheen et al., 2009; Shintani & Ellis, 2013) have focused only on the English article system. Shintani, Ellis and Suzuki (2014) focused on the hypothetical condition and indefinite articles. By contrast, few researchers (e.g., Ferris, 2006; Hartshorn et al., 2010; Liu, 2008; Truscott & Hsu, 2008 and Van Beuningen et al., 2008; 2012) have provided comprehensive feedback on all errors.

A few studies have compared the effectiveness of corrective feedback on single errors versus feedback on all errors. An example is Sheen, Wright and Moldawa (2009), who investigated the differential effects of what they termed focused and unfocused feedback on the accuracy of definite and indefinite articles together with four other grammatical structures (copula 'be', regular past tense, irregular past tense and prepositions). They reported no difference was found between the two types of feedback. Still more recently, Frear and Chiu (2015) also investigated the effects of indirect WCF on the accuracy of a limited number of grammatical structures (i.e., weak verbs) as well as on all errors. They report that both the focused and the comprehensive groups were successful in accuracy gains

Investigating the effect of feedback on a limited range of linguistic targets has obvious methodological advantages. However, some researchers (e.g., Storch, 2010) have expressed a concern that although focusing the treatment on certain errors may show that the feedback is effective in improving learners' linguistic accuracy of the selected features, it can tell us little about the effect of feedback on learners' overall writing ability. That said, most researchers have tended to favor focused WCF. The main reason for this is the pedagogical assumption that, as Ellis et al. (2008) put it, "Learners are likely to attend to correction directed at a single (or limited number of) error type(s) and more likely to develop a clear understanding of the nature of the error and correction needed" (Ellis et al. 2008; p. 356). The present study followed this tendency in providing focused feedback on one target structure, namely, prepositions.

A further issue raised in recent investigations has to do with the length and duration of the feedback treatment. Storch (2010) noted that most of the recent studies (e.g., Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2009a; Bitchener & Knoch, 2009b; Sheen, 2007; Van Beuningen, De Jong & Kuiken, 2008) used 'one-shot' designs, in which feedback was provided on one occasion and on a single text. Furthermore, in some studies where direct feedback was provided with or without written explanations (e.g., Bitchener & Knoch, 2008), participants were given only a few minutes to review the feedback provided before being asked to write a new text in an immediate post-test. It is scarcely necessary to draw attention to the questionable nature of such a procedure. Indeed, it is difficult to disagree with Storch's observation that while such treatments "may be easier to implement and control when conducting experimental studies", they lack "theoretical and pedagogical validity" insofar as "learning requires extensive and sustained meaningful exposure and practice" (2010, p. 42).

Current Study

The study reported here was designed to address some of the issues noted in the literature review above. It investigated the effect of direct and indirect focused WCF on students' grammatical accuracy, both in the revision of the same piece of writing and on new pieces of writing over time. The research design encompassed a number of unique aspects. First, unlike other recent major studies (e.g. Truscott & Hsu, 2008; Shintani & Ellis, 2013; Shintani et al. 2014; Van Beuningen et al., 2012) the students in the treatment groups were provided with feedback on more than one occasion (i.e., they received two feedback sessions); second, the students produced a new piece of writing three and four weeks after they received the final corrective feedback treatment, thereby providing insight into the effect of feedback over time. Finally, the study explored the effectiveness of two different subtypes of indirect feedback that differed in terms of feedback explicitness: indirect feedback involving underlining only and indirect feedback with metalinguistic cues. While some previous studies have examined the various forms of direct feedback such as direct correction only or direct correction with oral or written metalinguistic explanation (e.g., Bitchener, 2008; Bitchener et al., 2005; Bitchener & Knoch, 2009a,b; Sheen, 2007; Sheen et al., 2009), few studies have compared the effectiveness of different types of indirect feedback. Thus, the present study addressed the question of the degree of explicitness of indirect feedback as well. In addition to the above, no previous studies have investigated the effects of WCF in an Arab EFL context. In particular, the effect of WCF on the use of prepositions by the Arab EFL students has not been investigated previously. The choice of target structure was further motivated by the fact that, as Al Yaari and Almaflehi (2013) have pointed out, prepositions are widely acknowledged to be a problematic feature of English for learners whose first language is Arabic. In short, then, the current study not only addressed the existing gaps in the research literature, its findings are valuable additions to the current knowledge of effectiveness of WCF.

Research Questions

The study addressed the following research questions:

1. What are the short-term effects of differing types of WCF (direct, indirect with underlining, and indirect with underlining and metalinguistic explanation) on the L2 writing of Arabic-speaking EFL learners?
2. What are the delayed effects of differing types of WCF (direct, indirect with underlining, and indirect with underlining and metalinguistic explanation) on the L2 writing of Arabic-speaking EFL learners?
3. Does the effect of WCF depend on the type of feedback provided (direct vs. indirect)?

Methodology

Background, Participants and Instructional Setting

The study was conducted at United Arab Emirates University (UAEU), a government-sponsored, English-medium university located in the United Arab Emirates. UAEU serves predominately (circa 95%) UAE nationals, with an approximate 70:30 female to male student ratio. All the participants in this study were female students attending the University Foundation Program (UFP) within the university. Foundation programs, as their name indicates, comprise a sequence of proficiency-based courses designed as a bridge between high school and university for those learners unable to meet the minimum proficiency requirements for direct entry into undergraduate degree programs. In order to gain admission to an undergraduate degree program at UAEU students must attain a minimum level of English language proficiency (either a score on the TOEFL iBT of 61 or IELTS 5.0) on completion of the program.

Selection of participants for this study was conducted by means of purposive sampling. After discussion with senior management of the foundation program, four intact classes were identified as suitable. In total 86 female UFP students participated in the study. All were Emirati nationals with Arabic as their first language. Their ages ranged from 18 to 20 years.

Targeted Linguistic Error

The target structure for this study was selected English prepositions of place. The decision to target prepositions of place was taken for two reasons. First, the effect of WCF on the use of prepositions by the Arab EFL students has not been investigated previously; second, prepositions are widely acknowledged to be a problematic feature of English for learners whose first language is Arabic (Al Yaari & Almaflehi, 2013).

Design

The three treatment groups received different types of focused WCF as follows. Treatment Group A received direct feedback on their errors. That is, incorrect prepositions were crossed out and the correct preposition was written in their place. Treatment Groups B and C received different types of indirect feedback. For Group B incorrect prepositions were underlined to

indicate an error had been made but no further information was provided. For Group C incorrect prepositions were underlined and an explanatory note ("incorrect preposition") was written in the margin of the script. The control group completed the writing tasks without receiving any WCF. Table 1 below summarises the arrangement of the groups:

Table 1
Groups

Group	Type of treatment
Treatment Group A ($n = 24$)	Direct corrective feedback
Treatment Group B ($n = 23$)	Underlining only
Treatment Group C ($n = 18$)	Underlining with metalinguistic explanation
Control Group ($n = 21$)	No corrective feedback

Writing Task

The method chosen to elicit the target structure was picture prompts. Participants produced two descriptions based on different picture prompts and revised these over a two-week period. The two prompts took the form of pictures of a living room in a house in which various familiar items (TV, lamp, coffee table etc.) could be seen. The items in the picture were numbered with the corresponding names of the items provided in a list below the picture.

In Week 1 Day 1 participants were instructed to write sentences describing the location of various items (e.g., "The lamp is on the desk"). A sample sentence was provided as a model. Students were instructed to produce a minimum of ten sentences. They were given thirty minutes to write the description. Feedback was provided on their sentences. On Day 2, participants revised their writing. They were given forty minutes to read the feedback provided and to revise the descriptions. Week 2 followed precisely the same procedure. To examine the delayed effects of feedback on students' writing skills, each group was also asked to produce two new pieces of writing, in Weeks 5 and 6 (i.e., three and four weeks after the final treatment). The above procedures are summarised in Table 2.

Table 2
Procedures

Week 1	Week 2	Week 5	Week 6
Day 1: Writing 1	Day 1: Writing 2	Writing 3	Writing 4
Day 2: Revision	Day 2: Revision		

Accuracy Measure

Every text was scored for writing accuracy. Like previous studies of the effectiveness of WCF (e.g., Bitchener et al., 2005), the writings were measured on a quantitative variable. Accuracy was calculated as the percentage of incorrect usage for each specific preposition. For example, three incorrect uses of a specific preposition from ten obligatory occasions would give an error rate of 30%. Intra-rater and inter-rater reliability of the scoring were also calculated. In order to

check the intra-rater reliability, 50% of the writings were graded a second time by one of the researchers of the current study (a native speaker of English) 3 months after the initial scoring and analysis of all data. Pearson correlation coefficients for the scores at two times for the four writing tasks were: 0.96 (Writing 1), 0.97 (Writing 2), 0.94 (Writing 3), and 0.97 (Writing 4). The Pearson correlation coefficients for the scores at two times for the two revision tasks were: 0.97 (Revision 1) and 0.98 (Revision 2). To check the inter-rater reliability, another scorer, an ESL teacher and a native speaker of English, scored 20% of the writings individually. Pearson correlation coefficients for the two scores in the four writing tasks were: 0.98 (Writing 1), 0.97 (Writing 2), 0.94 (Writing 3), and 0.93 (Writing 4). The Pearson correlation coefficients for the two scores in the two revision tasks were: 0.98 (Revision 1) and 0.97 (Revision 2).

Data Analysis

Analyses were carried out on students' writing performance both within and across sessions. To examine the effect of WCF on students' subsequent revisions, a comparison was made between the four groups' error rates in the two revision tasks. To examine the effect of WCF on new writings, a comparison was made between the groups' error rates from Writing 1 to Writing 2, from Writing 2 to Writing 3, and from Writing 3 to Writing 4. Both one way and two-way analyses of variance (ANOVA) were used to identify the effects of feedback. The two-way ANOVA model contained treatment (type of feedback) as a between-subject variable and time as a within group variable. Seperate analyses were conducted for the revision effect and the transfer effect of feedback (i.e., the effect of feedback on new pieces of writing). Separate one-way ANOVAs were conducted to determine where the differences actually were. Additionally, if a test revealed statistical significance, post hoc tests to evaluate differences among specific means were also conducted.

Results
Revision Effects
In order to examine the effects of WCF on revisions, a two-way ANOVA was used on the revision gain scores. The results demonstrated that time had a significant effect [$F (1, 82) = 8.3$; $p= .005$], suggesting that there was an improvement from Revision 1 to Revision 2. The results also showed that condition (i.e. feedback types) also had a significant effect on the revisions [$F (3, 82) = 3.7$; $p= .014$]. There was also interaction between time and condition [$F (3, 82) = 3.2$; $p = .025$], which suggests that the difference in the effectiveness of WCF varied for the three feedback types over the three feedback sessions. Consequently, one-way ANOVA tests were conducted individually for each condition to find out which feedback type was more effective at each time. The results are presented in Tables 3 and 4. The one-way ANOVA test of the error rates revealed that reduction of errors was not significant in Revision 1 [($F (3, 82) = 2.54$, $p< .062$]; however, in Revision 2 there was a significant reduction of errors: [$F (3, 82) = 5.5$, $p< .002$]. Post hoc test with Bonferroni revealed a significant difference between the Underline only CF and the No CF group in Revision 2 task. The Underline only CF group was most successful

in reduction of errors in Revision 1, and they outperformed the other two treatment groups as well as the control group in both revision tasks.

Table 3
Descriptive statistics: Overall accuracy gains in revision 1 and 2 by group

Groups	Reduction of errors:			Reduction of errors:		
	Revision 1			Revision 2		
	M	S.D.		*M*	S.D.	
1. Underline+metalinguistic (*n*=18)	12.55	18.97		13.4	19.11	
2. NoCF (*n*=21)	22.00	18.49		20.8	18.9	
3. DirCF (*n*=24)	24.16	17.25		9.4	10.80	
4. Underline only (*n*=23)	11.21	22.03		3.5	6.5	
Total 17.74 19.80 (N=86)			11.47	15.52		

Table 4
One-way ANOVA for accuracy gains: Writing 1 and 2 to their revisions

		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	2834.681	3	944.894	2.541	.062
Revision1	Within Groups	30493.691	82	371.874		
	Total	33328.372	85			
	Between Groups	3474.865	3	1158.288	5.580	.002
Revision2	Within Groups	17020.588	82	207.568		
	Total	20495.453	85			

Transfer Effects

In order to examine the effects of WCF on new writings, both short-term (i.e., after 1 week) and delayed (i.e., after 3 weeks [comparing week 3 and week 5] and 6 weeks [comparing week 1 and week 6]) transfer effects were examined. As with the revision effects, a two-way ANOVA was used to test within and between-group differences from one piece of writing to a new piece of writing and over time. Subsequently, one-way ANOVAs and also pairwise comparions were conducted to find out where the differences lay.

The initial two-way ANOVA with time as a within-subject variable and feedback condition as a between subject variable demonstrated that time had highly significant effect (i.e., there was significant improvement from Writing 1 to Writing 2, Writing 2 to Writing 3, and Writing 3 to Writing 4) irrespective of the types of WCF [$F (3, 246) = 20.41$; $p = .000$]. There was no main effect of conditions [$F (3, 82) = 1.81$; $p = .150$); however, there was significant interaction of time and feedback condition [$F (9, 246) = 4.92$; $p = .000$]. This shows that learners' improvement over time depended on the type of feedback. One-way ANOVAs were conducted individually to find out where the effect was. These analyses were conducted regarding both the short-term and delayed transfer effects.

Short-term Effects

In order to examine the short-term effects of CF treatments (from Writing 1 to Writing 2), a comparison was made between the percentages of errors produced in each of the four groups by using a one-way ANOVA. There was noticeable decrease of errors from Writing 1 to Writing 2 (Week 1 to Week 2). Table 5 provides the descriptive statistics for overall reduction of errors for all groups, itemized per session.

Table 5

Descriptive statistics: Short-term error reduction by group (Week one vs. Week 2)

Groups	Error rates in Writing 1		Error rates in Writing 2	
	M	S.D.	*M*	S.D.
1. Underline+metalinguistic (*n*=18)	28.22	19.41	35.50	23.73
2. NoCF (*n*=21)	24.52	20.82	24.61	18.79
3. DirCF (*n*=24)	32.75	20.68	24.95	16.00
4. Underline only (*n*=23)	40.95	18.75	18.52	13.18
Total	31.98	20.56	25.36	18.56
(N=86)				

The one-way ANOVA showed that the four groups varied in reduction of errors in the Writing 2 [$F(3, 85) = 3.05$, $p = .033$]. In order to determine where the significant differences in accuracy gains lay between groups, post-hoc multiple comparisons were conducted using Bonferroni tests.

In the case of the reduction of errors from Writing 1 to Writing 2, a post hoc test with Bonferroni revealed that in Writing 2 there was a significant difference in error rates between the Underline + metalinguistic ($M = 35.50$; $SD=23.73$) and the Underline only ($M = 18.52$; $SD =13.18$) groups. The mean error rate difference between these two groups was 16.97. In case of the other two treatment groups, the differences in mean scores indicated that students who received Direct WCF performed better than students who received Underline + metalinguistic WCF; however, the difference in accuracy gain scores for these two groups was not significant. Students who received Underline only WCF produced fewer errors than students in the Underline + metalinguistic group in Writing 2. Comparison of means between the Direct WCF group and the No WCF group further demonstrated that the Direct WCF group and No WCF group performed almost identically in terms of reduction of errors in Writing 2.

Delayed Effects

In order to examine the delayed effects of WCF treatments, a comparison was made between the error percentage scores (from Week 1 to Week 5; Week 1 to Week 6; Week 2 to Week 5 and Week 2 to Week 6) in all four groups by using a one-way ANOVA. Results revealed that all three treatment groups were highly successful in reducing errors in Week 5 and in Week 6. Table 6 provides the descriptive statistics for delayed overall accuracy gains for all four groups.

Table 6

Descriptive statistics: delayed effect of WCF by group and session

Groups			Error rates: Writing 3 (Wk. 5)		Errors rates: Writing 4 (Wk. 6)	
M	S.D.	*M*	S.D.			
1. Underline+metalinguistic (*n*=18) 19.88	13.49				30.33	18.86
2. NoCF (*n*=21)		21.23	18.07		21.52	17.93
3. DirCF (*n*=24)		11.50	11.72		16.20	14.33
4. Underline only (*n*=23)		8.65	9.69		11.78	20.63
Total			14.87	14.27	19.27	18.91
(N=86)						

According to the one-way ANOVA, the difference in error rates from Writing 1 [*F* (3, 85) = 2.75, *p* = .048] and Writing 2 [*F* (3, 85) = 3.05, *p* = .033] to the delayed writing (Writing 4) between the four groups was significant [*F* (3, 85) = 3.93, *p* = .011]. Tables 7 and 8 provide the ANOVA results of the reduction of errors from Writing 1 and Writing 2 to Writing 4.

Table 7

ANOVA: Comparing error rates of Writing 1 and Writing 4

		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	3289.183	3	1096.394	2.753	.048
Writing1	Within Groups	32655.806	82	398.242		
	Total	35944.988	85			
	Between Groups	3824.193	3	1274.731	3.933	.011
Writing4	Within Groups	26579.109	82	324.135		
	Total	30403.302	85			

Table 8

ANOVA: Comparing error rates of Writing 2 and Writing 4

		Sum of Squares	Df	Mean Square	F	Sig.
	Between Groups	2941.676	3	980.559	3.050	.033
Writing2	Within Groups	26360.150	82	321.465		
	Total	29301.826	85			
	Between Groups	3824.193	3	1274.731	3.933	.011
Writing4	Within Groups	26579.109	82	324.135		
	Total	30403.302	85			

ANOVA tests also revealed that the difference in error rates from Writing 1 and 2 to Writing 3 (Week 5) also reduced significantly [*F* (3, 85) = 4.53, *p* = .005]. Post-hoc pairwise comparison using Bonferroni tests revealed that in Writing 4, the Underlining + metalinguistic (*M* = 30.33; *SD=18.86*) and Underline only (*M* = 11.78; *SD* =20.63) groups had significant differences in

error reduction rates. The mean error rate difference between these two groups was 18.55. In addition, the differences in mean error rate scores indicated that students who received Underline only WCF, outperformed Direct CF and No CF groups. Post-hoc pairwise comparison of the mean error rate scores of Writing 1 and Writing 2 with Writing 3 (Week 5) using Bonferroni also revealed that all three treatment groups were highly successful in reducing errors in writing 3. However, in Writing 3, the No WCF group ($M = 21.23$; $SD =18.07$) and Underline only ($M = 8.65$; $SD =9.69$) group had significant differences in error reduction rates. The mean error rate difference between these two groups was 12.58. The Direct WCF group was able to utilize the feedback more effectively than students who received Underline + metalinguistic WCF and those who received no WCF. To sum up, the study found there were durable or delayed effects of WCF treatments on error reduction. Participants successfully reduced errors in new writing 5 weeks after the WCF treatment was provided, and, the effects of WCF treatment were retained in week 6.

Discussion

This study investigated both the revision and transfer effects (effects on new pieces of writing) of focused WCF. It also investigated whether the effect depends on the type of feedback. The findings showed that the feedback improved learners' revision accuracy and that both direct and indirect feedback significantly reduced learner errors in subsequent revision tasks. These findings are consistent with those of previous research that has examined and found effects of feedback on revision (e.g., Ashwell, 2000; Bitchener, 2008; Bitchener & Knoch, 2008, 2009a, Ferris & Roberts, 2001; Sachs & Polio, 2007).

As for the revision effects of different feedback types, the results showed that indirect feedback was more effective than direct feedback. There were also differences in the different types of indirect feedback and these effects varied over time. The Underline only feedback group outperformed both the control group and the two other treatment groups in reduction of errors in Revision 1 with the reduction in in Revision 2 being significant. Direct WCF, however, outperformed the Underline + metalinguistic group in Revision 2.

Interestingly, both types of indirect WCF was very successful in this study. Indirect feedback may be effective in subsequent writing tasks if the feedback is repeated because repetition of the feedback can help learners recall their understanding of the grammatical structure and at least understand the type or nature of the errors made. In the current study, only preposition of place errors were targeted for feedback. It is possible, then, that WCF involving underlining only might prove sufficient if repeated.

All three treatment groups retained accuracy and made significant reduction of errors after 3 weeks of treatment. In terms of short term accuracy, there was a noticeable decrease of errors from Writing 1 to Writing 2 (Week 1 to Week 2). However, reduction of errors from Writing 3 to Writing 4 (Week 5 to Week 6) was not significant. These accuracy gains are encouraging. They differ from Truscott and Hsu's (2008) and Liu's (2008) findings, which only found accuracy gains on the revised texts, not on new pieces of writing. One possible reason for this difference

could be that in the above-mentioned studies, the feedback was provided only one time and therefore learners might not have had enough time to process the feedback, whereas, in the present study the feedback was provided over time. This again suggests that when feedback is provided on multiple occasions, it might be more effective for both revision and L2 development.

In addition to short term accuracy gains, delayed effects of WCF were also found in this study. All three treatment groups displayed significant decrease of errors from Writing 1 (Week 1) to Writing 3 (Week 5) and Writing 1 (Week 1) to Writing 4 (Week6). Also, there were decrease of errors from Writing 2 (Week 2) to Writing 3 (Week 5) and Writing 2 (Week 2) to Writing 4 (Week 6). In Writing 3 (Week 5), Direct WCF was most successful in reducing errors followed by Underlining only. The success of Direct WCF is in line with the findings of Bitchener et al. (2005), Mirzaii and Aliabadi (2013), and Van Beuningen et al. (2008 & 2012). In Writing 4 (Week 6), Underlining only WCF was most successful in reducing error. Underlining only WCF had revision as well as short term and delayed effects in this study. This finding is interesting because the more implicit form of WCF gained higher success than the more explicit type of WCF (i.e., Underline + metalinguistic WCF). A possible explanation for this is that since in this study only prepositions of place errors were targeted for correction, participants might not have had to think about other grammatical features while correcting errors. However, among the three types of WCF, direct feedback was also very successful on both revision and new piece of writing. This could have been because direct feedback provided learners with sufficient information to resolve their errors during revisions and this might have helped some of the learners effectively internalize the correct form.

Apparently, there was variation in the effectiveness of different types of feedback (direct and two types of indirect WCF). For example, in revision 1, underline only WCF was more effective than direct WCF. On the other hand direct WCF was more successful than both types of indirect WCF in revision 2. Similarly, while in short term writing tasks both direct and indirect WCF were equally successful, in delayed writing tasks direct WCF was more successful than indirect WCF. The mixed findings or variation in results about the effect of feedback is an evidence for the complexity of corrective feedback and the various factors that may influence its effectiveness. Indeed, in a recent meta-analysis, Kang and Han (2015) found that the efficacy of written corrective feedback was influenced greatly by factors such as context as well as learners' level of linguistic knowledge. Bitchener (2017) further added that within a single context, for example, the foreign language learning context, "there may be a complex interaction of external and internal motivational factors and while these may determine whether or not a learner is likely to respond positively or negatively to written CF" (p. 136).

Conclusion, Implications and Direction for Future Research
In summary, the findings of the present study show that both direct and the indirect WCF significantly improve revision accuracy. The findings are in line with previous studies that investigated the effects of focused WCF (e.g., Bitchener, 2008; Bitchener & Knoch, 2008;

Bitchener & Knoch, 2009a; Bitchener & Knoch, 2009b; Bitchener et al., 2005; Ellis et al., 2008; Sheen, 2007; Sheen et al., 2009; Shintani & Ellis, 2013). However, no previous studies investigated the effects of WCF in an Arab EFL context. In addition, while almost all of the previous studies provided treatment on errors only one time, in the current study treatment on errors was provided two times. Thus, the findings of the present study are a valuable addition to the current knowledge of effectiveness of WCF.

Based on the findings of the present study, it can be suggested that WCF is a useful editing tool insofar as it helped learners improve the accuracy of their initial writings during revisions. Teachers, therefore, should allow learners the opportunity to revise their written texts based on the WCF they provide. Moreover, providing students with opportunities to revise their writing could foster second language learning, as production of the correct form might help learners to automatize their L2 production (Loewen, 2004). In addition, successful error correction in revisions might lead to improved motivation in L2 writing, which might lead to improved accuracy in new texts and L2 acquisition (Van Beuningen, 2011). Apart from revision effects, WCF displayed both short term and delayed accuracy gains in this study. All three treatment groups retained accuracy and made significant reduction of errors after 3 weeks of treatment. And Underlining and Direct WCF proved to be very effective in improving grammatical accuracy on selected error types (i.e., preposition). Like previous studies, the findings of the present study demonstrated that focused WCF has the potential to help improve accuracy in writing.

The present study has certain limitations that should be kept in mind when interpreting its results. First, the sample size was relatively small. Therefore, future research should be conducted using larger sample sizes. However, although the sample size could be considered small, each participant in this study produced four pieces of writing and revised two of those writings. In consequence, the amount of data collected was by no means negligible.

To assess learners' level of L2 proficiency, the present study, like some previous research (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009a, 2009b; 2010; Ellis et al., 2008; Sheen, 2007; Sheen, 2010b; Truscott & Hsu, 2008) used participants who had already been assigned to a class on the basis of a placement test. However, placement tests might not be fully accurate in terms of assessing language proficiency. Studies using the same design but using more robust language proficiency tests to measure learners' initial linguistic and writing abilities would be helpful.

The present study measured the delayed effects of WCF. However, there was only a two-week gap between the third writing, and the final writing. Such a gap or interval is not sufficient to provide evidence for long-term accuracy gain. Therefore, to determine long-term effects, future studies could use longer intervals between writings. Finally, although attempts were made to have writing tasks with similar levels of difficulty, since there were different prompts, there is a possibility that participants responded to these tasks differently. In other words, there is the possibility of a task effect, which could explain some of the differences in the effectiveness of feedback on both revision and the accuracy of new writings.

References

Al Yaari, S., Almaflehi, N. (2013). The problem of translating the prepositions *at, in* and *on* into Arabic: An Applied Linguistic Approach. *Journal for the Study of English Linguistics, 01,* 256-273.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft

composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing,* 9, 227–57.

Baker, W., & Bricker, R. H. (2010). The effects of direct and indirect speech acts on native English and ESL speakers' perception of teacher written feedback. *System, 38,* 75-84.

Bitchener, J. (2017). Why some L2 learners fail to benefit from written corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language teaching and learning* (pp. 129-140). New York: Routledge.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal ofSecond Language Writing, 17,* 102–18.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrantand international students. *Language Teaching Research Journal, 12,* 409–31.

Bitchener, J., & Knoch, U. (2009a). The relative effectiveness of different types of direct written corrective feedback, *System, 37,* 322–29.

Bitchener, J., & Knoch, U. (2009b). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31,* 193-214.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2writers with written corrective feedback. *Journal of Second Language Writing, 19,* 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, 9,* 227–58.

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.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28,* 339-368.

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

Fathman, A. K., & Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom,* (pp. 178-190). Cambridge: Cambridge University Press.

Ferris, D. (1999). The case for grammar correction in L2 writing classes: a response to Truscott (1996). *Journal of Second Language Writing, 8,* l-l1.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues,* (pp. 81-104). Cambridge: Cambridge University Press.

Ferris, D. (2008). Feedback: Issues and options. In P. Friedrich (Ed.), *Teaching academic writing,* (pp. 93-124). London, UK: Continuum.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing, 10,* 161–184.

Frear, D., & Chiu, Y. H. (2015). The effect of focused and unfocused indirect writtencorrective feedback on EFL learners' accuracy in new pieces of writing. *System, 53,* 24-34.

Gholaminia, I., Gholaminia, A., and Marzbanc, A. (2014). An investigation of meta-linguistic corrective feedback in writing performance. *Procedia - Social and Behavioral Sciences, 116*, 316-320.

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*, 84-109.

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

Karim, K., & Nassaji, H. (2018). The revision and transfer effects of direct and indirect comprehensive corrective feedback on English-as-a-second-language (ESL) students' writing. *Language Teaching Research.* https://doi.org/10.1177/1362168818802469

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal, 7*, 305-313.

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

Liu, Y. (2008). The effects of error feedback in second language writing. *Arizona working papers in SLA and teaching*, *15*, 65-79.

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

Mirzaii, M., & Aliabadi, R. B. (2013). Direct and indirect written corrective feedback in the context of genre-based instruction on job application letter writing. *Journal ofWriting Research, 5,* 191-213.

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

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, *20*, 83-95.

Sachs, R., & Polio, C. (2007). Learners' use of two types of written feedback on a L2writing revision task. *Studies in Second Language Acquisition, 29,* 67-100.

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, *17*, 195–202.

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

Sheen, Y. (2010a). The Role of Oral and Written Corrective Feedback in SLA. *Studies in Second Language Acquisition*, *32*, 169– 179.

Sheen, Y. (2010b). Differential effects of oral and written corrective feedback in the ESLclassroom. *Studies in Second Language Acquisition, 32*, 201-234.

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

Sheppard, K. (1992). Two feedback types: do they make a difference? *RELC Journal, 23*, 103-l10.

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,* 286–306.

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of Written Feedback and Revision on Learners' Accuracy in Using Two English Grammatical Structures. *Language learning, 64,* 103-131.

Storch, N. (2010). Critical feedback on written corrective feedback research. *IJES, 10*, 29-46.

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition, 32*, 303-334.

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

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

Van Beuningen, C. G. (2011). *The effectiveness of comprehensive corrective feedback in second language writing.* Unpublished doctoral dissertation. University of Amsterdam, Netherlands. Retrieved from http://dare.uva.nl/en/record/374645.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and Indirect corrective feedback on L2 learners' written accuracy. *International Journal of AppliedLinguistics, 156*, 279–296.

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