Peer Corrective Feedback on L2 Writing: Does It Help Improve Written Accuracy and L2 Explicit Knowledge Over Time?

Maryam Homayounzadeh, Shiraz University, maryam.homayounzadeh.shiraz@gmail.com

Saeed Mehrpour, Shiraz University, smehrpur@rose.shirazu.ac.ir

Mahboobeh Saadat, Shiraz University, mahsaadat@gmail.com

Peer Corrective Feedback on L2 Writing: Does It Help Improve Written Accuracy and L2 Explicit Knowledge Over Time?

Maryam Homayounzadeh [1,] Saeed Mehrpour[2], Mahboobeh Saadat[3]

ARTICLE INFO

ABSTRACT

The purpose of the study was to explore the effect primarily of written corrective feedback (WCF) provided by peers on enhancing learners' internalization/retention of grammatical structures and further of task type in determining learners' engagement with WCF. Participants in the study were a group of EFL students (n = 36) who completed a number of controlled and free-writing tasks and provided metalinguistic feedback on the errors in the tasks accomplished by their peers. The grammatical structures, which the participants were to focus on included primarily errors of verb use and after some sessions of receiving WCF from their teacher, errors of noun, pronoun, and article use as well. The results of RM ANOVA suggested significant improvement in not only the accuracy with which learners used these structures after involvement in peer WCF, but also in the precision of the metalinguistic feedback they provided.

Written corrective feedback (WCF) refers to the feedback provided on the linguistic errors in student writings to promote their acquisition of target-like structures, demonstrated in their writing of new texts (Bitchener, 2012). This is often distinguished from the feedback provided in composition classes generally on the macro-level aspects of student writings, such as content and organization, to improve the overall quality of their essays (Ferris, 2010). A significant question that has so far received extensive attention in the literature concerns how WCF is to be provided to the learners for it to be effective and lead to learning. Various options have been identified for responding to the errors in the students' writings (Ellis, 2009; Ferris, 2002) and numerous studies have been conducted to compare their effectiveness (Bitchener, 2012; Bitchener & Knotch, 2010; Sheen, 2007) and to consider the effect of other intervening variables, such as

[1] Shiraz University, maryam.homayounzadeh.shiraz@gmail.com, +98-917-187-1128

[2] Shiraz University, smehrpur@rose.shirazu.ac.ir, +98-917-711-7011

[3] Shiraz University, msaadat@rose.shirazu.ac.ir, +98-917-316-5785

the context of learning and the learners' attitude, motivation and learning style, influencing their uptake of the feedback (Ferris, Liu, Sinha & Senna, 2012; Rahimi, 2015).

Nevertheless, in search for a more effective way to give WCF, the studies conducted thus far have focused merely on 'how' WCF is to be provided, disregarding the significant issue of 'who' provides the feedback. They have all exclusively considered the teacher as the only source for WCF, while a plethora of research, though carried out in composition classes, demonstrates that feedback provided by peers can be more advantageous than that given by teachers in helping learners improve the quality of their compositions (e.g. Lundstorm & Baker, 2009; Min, 2005; Rahimi, 2013; Ruegg, 2015). Indeed, peer feedback has been demonstrated to be beneficial not only to the receiver of feedback but to the feedback provider as well (Strijobs, Narciss, & Dünnebier, 2010). However, no study has yet taken into account the efficacy of peer feedback for language acquisition purposes. Although the research findings on peer feedback in composition classes suggest that it is beneficial mostly to the macro-level aspects of writing, these findings are inconclusive as the training they provided the learners with to give feedback had its focus upon the macro-level features of writing, rather than on linguistic errors. Therefore, it is not still known whether by training the learners to provide WCF, they can improve their own acquisition of the given structures.

The current investigation embodies an attempt primarily to examine the effect of systematic training for giving WCF on the learners' noticing and acquisition of target-like linguistic structures. The aim is to explore whether as a result of engaging in peer corrective feedback, learners can enhance both their noticing of errors in their peers' writings and the accuracy of their own written production. The results will be compared against the gains they acquire as a result of exposure merely to the WCF given by the teacher to determine which one, in case of any differences, can better lead to learning. The study will further take into account the extent to which task characteristics can affect learners' engagement with the WCF they give/receive. The results will be significant in providing useful insights for language teachers regarding how best to respond to the linguistic errors in the learners' writings to promote their retention and long-term acquisition. The findings could shed further light on the types of tasks that are most appropriate to be used in language learning classrooms to promote the students' learning.

2. Literature Review

The language learning potential of writing and WCF has been corroborated by various theoretical and empirical research studies within both the cognitive and sociocultural paradigms (Manchón, 2011). Theoretical support within the cognitive framework comes from the Noticing (Schmidt, 2001) and Output Hypotheses (Swain, 1985, 1995) which insist on the significance of attracting the learners' attention to language as an object while they are engaged in communication for L2 development. Accordingly, WCF is believed to provide the learners with even a greater learning opportunity than oral corrective feedback in that it is always explicit in directing the learners' attention to the gap(s) in their interlanguage (even if provided indirectly, e.g. by underlining or circling the error) and allows the learners time to give their error(s) more attention (Bitchener, 2012). Within the sociocultural paradigm, which emphasizes the collaborative nature of learning, the language learning potential of writing is seen in languaging (meaning making and knowledge construction through language) via collaborative writing. Accordingly, giving WCF is regarded as the joint construction of a Zone of Proximal Development (ZPD), i.e. assisting language learners to use structures, which they cannot employ individually (Erlam, Ellis, & Batstone, 2013).

Within these two theoretical paradigms, empirical research has been conducted exploring, on the one hand, the extent to which theoretical predictions are actually met in practice (e.g. Truscott, 1996) and on the other hand, the relative effectiveness of various types of WCF (Bitchener, 2010; Bitchener & Knotch,

2009; Erlam, Ellis, & Batstone, 2013; Ferris, 2010) as well as the mediating factors which affect the learners' acquisition from the input provided (e.g. Ferris et al., 2012).

2.1. Research Findings About the Language Learning Potential of WCF

Given Truscott's (1996) contention that WCF is ineffective and can even be detrimental to learning, the early studies on WCF explored its efficiency for language learning in practice. Drawing a distinction between feedback for accuracy and feedback for learning, depending on the increases in learners' accurate production over time (Manchon, 2011), the plethora of research conducted recently evince the effectiveness of WCF for improving the learners' accurate production of a text (Bitchener, 2008; Bitchener & Knotch, 2010; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007). As these studies suggest, the learners did notice, as a result of the feedback they received, the gaps in their language use and in some cases they retained their levels of improvement over time (Bitchener & Knotch, 2010). In fact, as Bitchener (2012) indicates, such unfavorable results about WCF, observed by Truscott (1996, 1999, 2007) and others (e.g. Kepner, 1991; Semek, 1984), were due to the lack of a rigorous methodological design, rather than the inefficiency of WCF.

However, this does not entail that the mere reception of feedback by the learners fosters their language acquisition. In fact, while WCF can enhance the students' noticing of the gaps in their interlanguage, it is the quality of noticing that determines the learners' retention of the feedback and long-term acquisition of the linguistic point (Qi & Lapkin, 2001). In this regard, Qi and Lapkin (2001) draw a distinction between two types of noticing substantive and perfunctory, with the former referring to the type of noticing in which the learners understand as well the reasons underlying the particular feedback provided. In the case study they conducted, Qi and Lapkin (2001) found that substantive noticing led to grater improvements in the revised texts than perfunctory noticing. Accordingly, the significant question that concerns L2 writing pedagogy is how the quality of noticing can be improved in providing WCF. Hence, numerous studies are being carried out to identify the variables that can promote the language learning potential of WCF.

2.2. Factors Affecting the Effectiveness of WCF

2.2.1. Types of Feedback

Among the factors whose impact on the language learning potential of WCF has been investigated throughout the literature is the particular type of feedback provided. In this regard, the question of concern has been whether feedback should be focused or comprehensive, explicit or implicit, and if explicit, how much metalinguistic information is necessary (Bitchener, 2008; Bitchener & Knotch, 2008; Ellis et al., 2008; Ferris, 2010; Sheen, 2007). Although research findings in this regard are contradictory, in general, it is suggested that focused WCF, provided for specific error types, is more valuable than the unfocused one (Bitchener & Knotch, 2009, 2010; Sheen, 2007), particularly for the learners of lower intermediate levels since it places a lower cognitive load on their processing capacity (Bitchener & Ferris, 2012).

However, with respect to the dichotomy of explicit versus implicit feedback, it is still controversial which one can better lead to acquisition. The advantage of implicit feedback is that it engages the learners in guided learning and problem-solving and thereby enhances the probability of their reflection upon their currently internalized knowledge and of their deeper processing and learning (Bitchener, 2012; Ferris et al., 2012). On the other hand, explicit feedback is recommendable in that it reduces learners' confusion and the number of hypotheses they need to test and further provides them with the information they require to solve the more complex errors (Bitchener, 2012). Indeed, explicit

feedback is considered "the quickest, most appropriate and most useful way of helping" (Scrivener, 2005, p.3). However, as the study by Erlam, Ellis, and Baston (2013) suggests, neither of these WCF types per se can contribute to learners' acquisition of the forms and correct use of them with less assistance over time. Therefore, the question is still open of how feedback is to be provided for better retention and long-term learning of language structures to take place.

2.2.2. Individual and Contextual Variables

In trying to identify the factors that enhance the learning potential of WCF, researchers have further attended to the influence of the individual and contextual variables. Although research in this area is still in its infancy, there is evidence that individual factors, such as the learners' goals, attitudes, beliefs, educational background, confidence and motivational levels may also affect their uptake of the WCF provided (e.g. Ferris et al., 2012; Hyland, 1998, 2003; Rahimi, 2015; Storch & Wigglesworth, 2010; Swain, 2006; Swain & Lapkin, 2003). Hyland (1998, 2003), for instance, through retrospective interviews and case studies found that learners' response to feedback and the strategies they adopt depend to a large extent on the importance they attribute to the grammatical accuracy of their writing. In this regard, Hedgcock and Lefkowitz (1994) contend that in foreign language contexts, where language is acquired mainly as a qualification, accuracy might not be a concern for the language learners and this could be significant in determining their retention of feedback. Rahimi (2015) further found learners' motivation and cognitive style (field dependence/independence) significant in determining their retention of WCF.

Besides such cognitive and affective factors, the differences found in the learners' written accuracy across tasks (Bitchener, 2012; Diab, 2016; Eckerth, 2008) suggest that task characteristics could also be influential in determining the students' learning from the corrective feedback they receive. Indeed, as Noris and Ortega (2012) indicate, "tasks or other elicitation devices may themselves constrain or facilitate the use of certain language forms … so the apparent lack of sought-after L2 forms in production may not necessarily indicate the lack of learner knowledge of them" (p. 581). However, no empirical study has yet been conducted to verify this in practice.

2.2.3. Source of feedback

Another significant factor influencing the effectiveness of feedback could be the source of feedback (Diab, 2016). A plethora of studies have been conducted in composition classrooms to investigate the extent to which peer feedback can affect the students' revising skills and as a result, develop in them the habit of self-editing (e.g. De Guerro & Villamil, 1994; Min, 2006; Peterson, 2003; Rahimi, 2009; Tsui & Ng, 2000). The majority of these studies suggest that learners are more likely to include in their revisions the feedback they receive from their peers, which they find more compatible with their own proficiency level and more manageable to apply, as compared to those of their teachers (Caulk, 1994). Moreover, it has been suggested that by reviewing their peers' writings and providing, rather than merely receiving feedback, learners can make more gains in promoting their writing ability (Lundstorm & Baker, 2009). Nevertheless, with the focus of these studies being on academic writing ability, rather than language acquisition, questions exist of the extent to which long term acquisition of linguistic structures can take place if WCF is provided by peers.

Of course, the current literature does make suggestions regarding how peer feedback might affect the linguistic accuracy of student writings. Ruegg (2015), for instance, compared the effect of peer and teacher feedback on the improvement of both the students' writing quality and their writing accuracy. Based on the findings, no significant difference was found between the improvement in writing quality, either as a result of teacher or peer feedback; however, the group that received feedback from the teacher received a significantly higher grammar score. Hence, based on the findings, the researcher suggested that

31

teachers provide comments on the grammatical aspects and students on the content of writing. However, Ruegg (2015), similar to other studies of peer feedback and writing quality (Berg, 1999; Hu, 2005; Min, 2006; Rahimi, 2013; Zue, 1995), provided the participants with no training for giving WCF. If there was any training, it was concerned mainly with the global aspects of writing, rather than its linguistic accuracy. With the learners not possessing the necessary skills for providing WCF, no significant results can be expected in this respect.

To explore the extent to which peer feedback can promote learners' language acquisition and written accuracy for the long run, the current investigation will adopt a pre-post-and delayed post-test design and will systematically train a group of EFL undergraduates to provide WCF on their peers' writings. The aim is to verify whether this way they can improve their learning of particular linguistic items, evinced in the accuracy both of the comments they provide and the revisions they make. Given the concerns in the literature (Bitchener, 2012), as reviewed above, about the potential effect of task type on the effectiveness of WCF, the study will take that into account by comparing the results across different task types.

2.3. Research Questions

The present study seeks answer to the following questions:
1. Can training peer reviewers help learners improve their writing accuracy in the short/long run?
2. Does providing peer WCF help learners improve their explicit knowledge of language structures?
3. Does the learners' engagement with the WCF vary across task types?

3. Research Context

The research took place in an institute of higher education in Iran in a four-credit course of grammar and writing for EFL undergraduates. As Hedgcock and Lefkowitz (1994) suggest, in foreign language contexts the participants might be less willing to attend to WCF because they are acquiring the language as a qualification, rather than with the purpose of becoming a member of the target language community. However, in the present study, the nature of the course, which had its focus on the fundamentals of grammar and writing, as well as of the participants' major (Teaching English as a Foreign Language and English Translation) could overcome this effect of the context, by requiring them to develop, not only a fluent but also an accurate knowledge of English. Moreover, the participants, had already passed a similar course the previous semester, and therefore had some exposure to the explicit knowledge of the target language, another factor which Hedgcock and Lefkowitz (1994) found significant in determining the learners' engagement with the WCF.

4. Methodology

4.1. Participants

The study used single-subject repeated measures design, in which the performance of the same group of individuals is measured under different conditions. Therefore, participants in the study were a group of thirty-six English majors (with an average age of 25), pursuing their BA degrees at an institute of higher education in Shiraz, Iran. They had enrolled in the four-credit course of the *Fundamentals of English Grammar and Writing II* (class meetings being two days a week) and had already passed a similar course the previous semester. Nevertheless, the group was not homogeneous in language proficiency, as their performance suggested on a test of structure and written expression, retrieved from a paper-based Test of English as a Foreign Language (TOEFL PBT). The sample indeed represented a range of proficiency levels from lower intermediate (31-40), to intermediate (41-50) and upper-intermediate (51-68).

4.2. Targeted Structures

Given the focus of the class, the structures whose acquisition was investigated in the study included four of the five error categories identified by Ferris and Roberts (2001), i.e. verb errors, noun ending, pronoun, and article use errors and errors of mechanics. The categorization and criteria in each section, however, have been modified according to the purpose of the study. Table 1, below provides an elaboration on each category and what the students were expected to focus upon.

Table 1
Description of error categories used for feedback and analysis

Verb errors	Errors in verb tense and form as well as the relevant errors of subject-verb agreement
Noun ending/pronoun errors	Incorrect, omitted, or unnecessary possessive/plural noun endings, incorrect use of pronouns, as well as the relevant errors of subject-verb and noun-pronoun agreement
Article errors	Incorrect use/omission of articles and other determiners

4.3. The Tasks

For the purpose of the study, two types of tasks were used, from the either ends of the continuum of language-based and meaning-based tasks (Manchón, 2011). The language-based tasks were in the form of cloze-tests, which the participants were to complete by filling the gaps with appropriate forms of the verbs, nouns, pronouns, and/or articles, depending on the instructional focus of the class. Six such tasks were designed, three focusing specifically on each of the given error categories (verbs, nouns/pronouns, and article errors), one on both verbs and nouns and one on all the three grammatical issues.

The meaning-based tasks, on the other hand, did not have their focus on any particular grammatical forms; rather, they provided the participants with a topic and required them to express their ideas within a narrative paragraph of 120-180 words. The topics were selected based on the majority agreement. The aim in using these tasks was to assess the learners' accuracy in using the stated grammatical structures and their ability to distinguish their erroneous use in uncontrolled contexts, wherein students are known to attend less to their use of grammatical structures (Manchón, 2011). Moreover, the use of free writing tasks along with the grammar-oriented ones would make it possible for the study to investigate the legitimacy of Bitchener's (2012) contention regarding the impact of task type on the learners' engagement with WCF.

Besides the stated tasks, the study used two structure-and-written expression tests, whose items were selected, based on the focus of the course, from a set of TOEFL PBT tests. The tests were used as a pretest and a posttest to assess the participants' pre-experimental and long term acquisition of the targeted structures.

4.4. Procedure and Training

The aim of the course was to familiarize the learners with the fundamental issues in English grammar and writing. The topics covered in the course were verb tenses, nouns, pronouns, noun-pronoun and subject-verb agreement, and articles. At the beginning of the term, the students received a structure and written expression test of TOEFL PBT, which was intended to assess their proficiency in English grammar and classify them into groups accordingly. Nine groups were formed, each consisting of two lower-intermediate, one intermediate, and one upper-intermediate student, who would function as the head of the group. They were then instructed on how to work in groups and provide WCF on the grammatical errors in the tasks performed by their peers.

After each grammatical point was taught, the students were provided with a grammar-oriented and a free writing task to complete. Then within groups, they exchanged their papers with their peers for correction. They were instructed to number each error and provide metalinguistic comments at the end of the texts, explaining why, for instance using that particular tense in that context was wrong and what could have been a better option. Having corrected their peers' papers, they handed them in to the head of the group to check the accuracy of the corrections made and provide the necessary explanations for the group mates in case of any problems. The instructor (one of the researchers) then received the papers to check for any unnoticed errors by the groups and to keep record of each student's scores on the tasks. The scores were given based on the accuracy not only of the responses to the task, but also of the feedback provided.

The first grammar-oriented task the students accomplished had its focus on verb tenses and subject-verb agreement. Following that, as their first free-writing task, they wrote a paragraph describing one of their most memorable trips. Since focused WCF is generally considered more conductive to learning (e.g. Ellis, 2009), in correcting this writing of their peers, the participants were encouraged to concentrate mainly on their use of verbs. The scores the students obtained on these two tasks served as a pre-experimental account of their accuracy in using this grammatical point and the precision of their explicit knowledge in this respect.

In the second set of tasks the participants completed, the concern was with their accuracy in using not only verb tenses but also nouns/pronouns and articles in a sentence. To observe the impact of peer feedback on the learners' acquisition of verb tenses, the instructor herself assessed the grammar-oriented task and provided metalinguistic feedback where necessary. The free-writing task, on the other hand, was corrected once by the peers to provide feedback on the use of verbs and once by the instructor to comment on the use of nouns and pronouns. The scores were calculated separately for each category of errors. The learners' accurate use of verbs as well as the precision of the feedback they offered was to be considered a clue to their mastery of the verb tenses, having twice received feedback from their peers on this grammatical point in the context of an uncontrolled task. Their scores for the accurate use of nouns/pronouns and articles, on the other hand, was to be taken as an account of their improvement in acquiring the point having received WCF from the teacher. The aim was ultimately to compare the two types of WCF in terms of their effectiveness in promoting the students' acquisition of grammatical structures.

Nevertheless, to further examine the learners' retention of the given grammatical issues in a more controlled context after exposure to more peer and teacher feedback, they were provided with a grammar-oriented task, focusing on the three points. They assessed their peers' tasks based on their use of verbs and received feedback from the teacher on their use of nouns, pronouns and articles. Similar to the previous tasks, the learners were scored both based on their correct answers and the accuracy of the feedback they gave for their peers on their use of verbs.

Following that, there was another controlled task whose focus was on nouns, pronouns, and articles. The learners then wrote their third writings about a blessing in disguise in their lives, i.e. their experience of something bad which turned out to be good. Nevertheless, instead of the teacher, the learners corrected and commented on their peers' papers. In assessing the writings, their focus was to be not only on the errors of verb use but on those of the nouns/pronouns and article use as well; the accuracy of their feedback could be indicative of their long-term acquisition of the given grammatical items.

The ultimate tasks the participants responded to in this experiment were, on the one hand, a controlled one, whose concern was with all the three error categories and on the other hand, a composition in which they described an experience that led them to appreciate things they had taken for granted. The tasks were to serve as delayed post-test, demonstrating the examinees' long-term acquisition of the grammatical items. In assessing these tasks, the learners were again asked to focus on the errors of

all the three categories and provide metalinguistic explanation for the errors they specified. They were then scored for the accuracy of both the verbs, nouns, pronouns, and articles they used and the comments they provided. The learners' accuracy in using the forms was to be interpreted as an indication of their long term acquisition of the stated structures, manifested in the context of both a controlled and an uncontrolled task (Lantolf, 2000). The results could further provide information on whether giving WCF makes the learners more sensitive to errors than teacher feedback and whether this has any effect on their accurate use of the forms.

4.5. Data Analysis

The participants' performance on the tasks was scored according to the number of errors of each type and the accuracy of the feedback they provided. To assess error counts of each type in the free writing tasks, the writings were rewritten accurately to estimate the total number of verbs, nouns, and articles needed to be used in the text. Error mean was then calculated for each category of errors separately, dividing the error counts by the total number of verbs, nouns/pronouns, or articles needed to be used in the text, multiplied by the average number of verbs, nouns/pronouns, or articles in the whole sample. A similar approach was followed to score the learners' feedback. The obtained scores were then transformed into t-scores, to make them comparable for the purpose of the study. One-way repeated measures ANOVA (RM ANOVA) was then run to investigate the impact of WCF provided by peers on language acquisition and written accuracy. Paired samples t-test was further used to compare the learners' improvement across tasks of different types to verify whether this was a significant factor affecting the learners' engagement with the corrective feedback (CF) provided.

5. Results

The main concern of the study was to investigate the extent to which by providing WCF on their peers' writings, learners can improve their writing accuracy and acquisition of linguistic structures. Hence, the first set of analyses involved examining whether there existed a reduction in the number of errors in student writings as a result of exposure to peer corrective feedback. Table 2 below presents the descriptive statistics for the errors of verbs in the pre-task.

Table 2
Descriptive statistics for errors of verb use

Tasks	Minimum	Maximum	Mean	Std. deviation
Controlled	34.62	68.83	50.0001	10.00001
Free Writing Task	34.77	82.71	50.0000	9.99999

5.1. Peer WCF Effect on the Acquisition of Verb Use

Considering verbs as one of the error categories, the results of the RM ANOVA, suggest significant differences between the error means of the pretest, post-test and the delayed post-tests (F = 3.8502, p < .05). The effect size for this difference is .987, which according to Cohen (1988) is a large one. Apparently, peer feedback has had quite a noticeable effect on the reductions observed in the number of errors in the students' use of verbs.

Table 3

Results of RM ANOVA for the difference in the number of errors of verb use in pre, post, and delayed post-tests

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilk's Lambda	.013	3.8502	6.000	30.000	.000	.987

The table of pairwise comparisons below demonstrates the trend of change in the number of errors in the students' use of verbs in the tasks. Throughout the results, it is possible to detect the trend of change in the error means and verify the effect of task type on the learners' engagement with the WCF.

Table 4

Pairwise comparisons for the errors of verb use

		Mean difference	Std. Error	Sig.	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	1.458	.913	1.000	.347	-1.346
	3	4.916	.123	.000	3.152	5.879
	4	8.289	.182	.000	-6.500	-1.928
	5	3.668	.991	1.000	6.688	8.648
	6	10.833	.213	.000	-7.361	-14.014
	7	14.340	1.015	.000	-11.095	-17.585
2	3	1.289	.182	1.000	.597	.597
	4	3.725	.947	1.000	1.173	6.969
	5	4.016	.309	1.000	3.004	5.027
	6	21.189	.812	.000	18.523	23.854
	7	14.340	.991	.000	11.095	17.585
3	4	1.285	.278	1.000	.375	1.194
	5	-6.525	.245	1.000	3.503	6.290
	6	18.904	.840	.000	16.125	21.656
	7	10.383	1.015	.000	7.361	14.014
4	5	-6.840	.914	1.000	-8.874	-4.826
	6	6.333	.113	.000	1.702	9.604
	7	-12.188	.946	.000	-15. 288	-9.708
5	6	17.173	.878	.000	14.298	20.042
	7	10.652	.186	.000	8.042	12.263
6	7	13.521	.913	.000	10.531	16.511

Based on estimated marginal means

 *The mean difference is significant at .05

 a. Adjustment for multiple comparisons: Bonferroni

1. Controlled pretest   2. Free writing pretest   3. Free writing post-test   4. Controlled post-test   5. Free writing post-test   6. Controlled post test   7. Free writing post-test

As the results suggest, learners began to improve in their use of verbs, both in the controlled and free writing tasks, from the first post-tasks (3 and 4) they accomplished (p < .05). Some regression is found in Task 5, a free writing task, wherein no significant improvement is observed in the learners' performance, as compared to the pre-tasks (p > .05). Nevertheless, the learners seem to resume their improvement in the accurate use of verbs in the following two tasks, which show a significant difference not only with Task 5, but with all the preceding ones (p < .05). Nevertheless, in general, error mean differences with the pretest tasks, seemed to be more noticeable in the controlled, as opposed to the free writing tasks. To verify whether this apparent difference was significant and thus whether task type affected the learners' engagement with WCF, paired samples t-test was run to compare the degree of improvement in controlled and composition tasks.

Table 5

Results of paired samples t-test for the difference in the number of errors of verb use in the controlled and free-writing tasks

	Paired Differences							
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair1 C1C4 – F2F3	2	3.41666	.56894	-1.15507	2.15496	.000	35	1.000
Pair2 C4C6 – F5F7	.681	2.17952	.36352	44.26378	45.73767	123.881	35	1.000
Pair3 C1C6 – F2F7	3.507	4.14157	.69029	-1.40130	1.40143	.000	35	1.000

As the results indicate, although the number of errors in controlled tasks was evidently smaller, the degree of improvement in the learners' written accuracy from one controlled task to another was not any significantly different from that found between two composition tasks (p > .05). Therefore, with respect to this particular grammatical item, i.e. verbs, task type did not appear to play a significant role in affecting the learners' engagement with the WCF provided.

The last question that needed to be investigated about the learners' acquisition of verbs concerned the precision of the feedback they offered. To investigate the impact of WCF on the improvement of the learners' explicit knowledge, RM ANOVA was run between the t-scores estimated for the accuracy of the learners' metalinguistic explanations to determine if any improvement had occurred. Significant results were found (F = 41.113, p < .05) with a high effect size for the dependent variable (.895).

Table 6

Results of RM ANOVA comparing the precision of metalinguistic explanations across tasks

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilk's Lambda	.271	41.113	4.000	27.000	.000	.895

As the table of pairwise comparison presented below suggests, learners' precision in assessing their peers' use of verbs suggests a constant improvement across tasks (p < .05). The regression observed in Task 5 in the learners' accurate use of verbs is not observed here in their precision in noticing their peers' errors. Apparently, peer feedback has had a significant effect on the learners' sensitivity to errors.

Table 7

Pairwise comparison of metalinguistic feedback on errors of verb use across tasks

		Mean difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	.898	.913	1.000	.115	1.346
	3	-4.101	.237	.000	2.126	6.879
	4	-7.649	.130	.000	3.850	7.970
	5	-5.328	.463	.000	4.521	6.353
	6	-12.931	.200	.000	10.511	14.750
	7	-9.343	.115	.000	6.998	10.095

2	3	-4.455	.117	.000	2.290	-5.597
	4	-7.254	.732	.000	5.433	8.990
	5	-4.916	.260	.000	3.154	6.022
	6	-14.100	.970	.000	10.523	15.664
	7	-9.040	.703	.000	9.085	12.625
3	4	-5.228	.198	.000	3.124	7.151
	5	-7.221	.405	.000	3.503	6.290
	6	-12.540	.840	.000	10.111	14.610
	7	-9.305	.989	.000	10.281	13.985
4	5	5.220	.914	.000	2.874	-6.840
	6	-12.271	.110	.000	10.102	14.005
	7	-11.188	.946	.000	9. 288	13.708
5	6	-8.103	.468	.000	6.158	10.002
	7	-7.999	.186	.000	5.042	9.263
6	7	4.012	.763	.000	3.641	6.120

Based on estimated marginal means

 *The mean difference is significant at .05

  a. Adjustment for multiple comparisons: Bonferroni

1. Controlled pretest   2. Free writing pretest   3. Free writing post-test   4. Controlled post-test   5. Free writing post-test   6. Controlled post test

7. Free writing post-test

The results of the paired samples t-test conducted suggest that the enhancement in the learners' sensitivity to errors of verb use was consistent across tasks and was not affected by the particular type of tasks used.

Table 8

Results of paired samples t-test between controlled and free-writing tasks for the precision of metalinguistic feedback

	Paired Differences								
					95% Confidence Interval of the Difference				
	Mean	**Std. Deviation**	**Std. Error Mean**		**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair1 C1C4 – F2F5	-2.835	8.78494	2.54276		1.15507	3.15496	1.178	35	1.000
Pair2 C4C6 – F5F7	-4.272	6.17702	2.78042		3.26378	7.73767	5.459	35	1.000
Pair3 C1C4 – F2F3	-3.194	6.17446	1.49750		1.34612	5.02586	1.038	35	1.000
Pair3 C1C6 – F2F7	-3.891	6.31853	2.82715		-1.40130	1.40143	5.602	35	1.000

Nevertheless, to assure that the findings obtained were a result of the peer feedback learners provided and were not merely a consequence of natural development or practice effect, it was necessary to study their performance in contexts where no peer feedback was provided to see whether such an improvement would take place. Hence the learners were provided with similar tasks focusing on the use

of nouns, pronouns, and articles; however, instead of the peers, the instructor provided them with feedback on their errors. The following sections present the results.

4.2. Teacher vs. Peer Feedback Effect on the Errors of Noun and article Use

Table 9 below provides an account of the learners' proficiency in the use of nouns, pronouns and articles in both a controlled and a free writing task prior to receiving any feedback from the teacher.

Table 9.
Descriptive statistics for the errors of noun, pronoun, and article use

Tasks	N	Minimum	Maximum	Mean	Std. Deviation
Controlled	36	1	14	5.8889	4.38033
Free writing	36	5.70	36	15.8289	8.75577

To investigate whether the learners' use of nouns/pronouns and articles would similarly improve merely with exposure to teacher feedback, RM ANOVA was run between their scores across tasks. Significant results were found ($p < .05$) with a moderate effect size ($r = .48$).

Table 10
Results of RM ANOVA comparing errors of noun, pronoun, and article use across tasks

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilk's Lambda	.002	3.610	5.000	31.000	.000	.480

However, the results of pairwise comparisons across tasks suggested that the significance of the findings was merely due to tasks 5 and 6, which showed a significant difference with all the other tasks ($p < .05$). Respectively, these were the ultimate controlled and composition tasks learners performed having once provided and received peer feedback from their peers on the use of nouns, pronouns and articles. Although they did get similar metalinguistic corrective feedback from the teacher, no significant advancement was noticeable in their performance on any of the other tasks whether controlled or composition ($p > .05$). Besides the efficacy of peer WCF, this information per se demonstrates that learners' engagement with WCF might not be task-dependent.

Table 11
Pairwise comparison of metalinguistic feedback across tasks

(I) Nouns	(J) Nouns	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	4.633E-5	.376	1.000	-1.183	1.183
	3	2.462E-5	.848	1.000	-2.670	2.670
	4	-2.537E-5	.908	1.000	-2.862	2.862
	5	-50.000*	1.527	.000	45.189	54.811
	6	-48.889*	5.477	.000	-66.143	-31.635
2	1	-4.6335	.376	1.000	-1.183	1.183
	3	-2.1715	.953	1.000	-3.001	3.001
	4	-7.1705*	5.531	1.000	-66.311	-31.467
	5	98.889*	6.984	.000	76.887	120.891
	6	50.000*	1.533	.000	45.169	54.831
3	1	-2.4625	.848	1.000	-2.670	2.670
	2	-4.9995	.184	1.000	-.580	.580
	4	2.1715	.953	1.000	-3.001	3.001

	5	50.000*	1.501	.000	45.270	54.730
	6	48.889*	5.745	.000	30.790	66.988
4	1	2.5375	.908	1.000	-2.862	2.862
	2	7.1705	1.017	1.000	-3.203	3.203
	3	4.9995	.184	1.000	-.580	.580
	5	50.000*	1.500	.000	45.274	54.726
	6	48.889*	5.787	.000	67.120	30.658
5	1	50.000*	1.527	.000	45.189	54.811
	2	98.889*	6.984	.000	76.887	120.891
	3	50.000*	1.500	.000	45.274	54.726
	4	50.000*	1.533	.000	45.169	54.831
	6	50.000*	1.501	.000	45.270	54.730
6	1	-48.889*	5.477	.000	31.635	66.143
	2	-98.889*	6.984	.000	76.887	120.891
	3	-48.889*	5.787	.000	30.658	67.120
	4	-48.889*	5.531	.000	31.467	66.311
	5	-48.889*	5.745	.000	30.790	66.988

Based on estimated marginal means

 *The mean difference is significant at .05

  a. Adjustment for multiple comparisons: Bonferroni

1. Controlled pretest   2. Free writing pretest   3. Free writing post-test   4. Controlled post-test   5. Free writing post-test   6. Controlled post test

The advantages of peer CF were further evident in the enhancement observed in the learners' heightened sensitivity to and noticing of the errors in their peers' performance and the precision of the metalinguistic comments they offered, and thus their explicit knowledge about the given structural items. Similar to their accuracy of use, as the results of the RM ANOVA suggested, the learners tended to give more on-target comments on their peers' performance in the last two tasks (P < .05), having done so in the previous two tasks.

5. Discussion of the Results

The purpose of the study was to investigate the efficacy of peer CF in improving the learners' internalization or long-lasting acquisition of grammatical structures, so far as their accurate use in writing texts was concerned. In this regard, the study focused upon verb, noun, pronoun, and article use as a sample of grammatical items and explored the learners' acquisition of them as a result of giving and receiving WCF to and from their peers. Further significant to the study was, on the one hand, the learners' noticing and sensitivity to errors, which has been emphasized as being significant to learning (Schmidt, 2001) and on the other hand the effect of task type on the learners' acquisition, as Bitchener (2012) and Norris and Ortega (2012) suggest.

With respect to its impact on the long-term acquisition of grammatical structures, the results of the study supported peer feedback as an appropriate approach. This was revealed in the learners' developing accuracy in using verb tenses after their involvement in giving and receiving peer corrective feedback. This effect could be attributed to the explicit feedback the learners themselves were required to provide on their peers' errors, which, in Qi and Lapkin's (2001) terms, had given them the chance to understand the reasons behind their errors and had thereby enhanced their substantive noticing because, when similar feedback on noun, pronoun, and article use was provided by the teacher, no significant

improvement was observed in the learners' use of the forms. It was after their engagement in peer feedback that significant advancement could be discerned. The finding, on the one hand, corroborates Lundstorm and Baker's (2009) contention that giving can do learners more good than merely receiving feedback. As Ohta (2000) and other advocates of sociocultural theory (e.g. Lantolf, 2000) contend, it is learner engagement in the process of learning that makes a difference in the results and giving WCF to their peers' errors provides such an opportunity for the learners.

The finding, on the other hand, negates Ruegg's (2015) argument that feedback on the grammatical errors is better to be provided by the teachers, given the learners' disregard of them in turn for the content in providing feedback to their peers' writings. In the results of the current investigation, it was revealed that if adequately trained, learners not only notice the errors in their peers' writings but can consequently improve their own written accuracy as well. Indeed, in the results of the present study, along with the increase in the learners' writing accuracy, there was a growth in the precision of the feedback they offered and the metalinguistic explanations they provided. In other words, as a result of peer feedback, there appeared to be a parallel growth in the learners' substantive noticing of the gaps in their own interlanguages, and in their explicit knowledge about the given grammatical items, which ultimately led to their own learning of the structures.

Further significant in the results of the study was the realization that the learners' engagement with WCF might not be dependent on the type of the tasks involved. So far as the results of the current investigation are concerned, learners' written accuracy equally grew in both controlled and free writing tasks. The same was also true of the accuracy of the CFs learners gave to their peers. Nevertheless, this is not to ignore Eckerth's (2008) claim that the learners' number of grammatical errors in controlled tasks generally tends to be fewer than that in free writing tasks; indeed this was the case in the current study as well, so far as the mean of errors was concerned. However, what was found almost similar between the two task types was the extent of the learners' growth in accuracy from one controlled or composition task to another and this could indeed be due to the quality of their engagement with feedback and their increased substantive noticing

The results, nonetheless, apparently contradict Diab's (2016) contention regarding the significance of task type in determining language development from the WCF provided. In fact, as the findings of the current investigation suggest, although the number of errors learners make might differ across tasks depending on their focus and the cognitive load they place on the learners, it might not be so significant in determining their engagement with the feedback they receive and their consequent learning. Indeed, in the results of the current study, it was the students' substantive noticing and their realization of the reasons for the CFs they gave/received which made a difference in their acquisition.

6. Conclusion

The results of the study highlight the significance and efficacy of peer corrective feedback in language instruction. In spite of the preceding research, which suggested peer feedback to be merely effective for the global aspects of writing (Ruegg, 2015), the current investigation found it quite effective in enhancing the learners' long term acquisition of linguistic structures and thus their enhanced writing accuracy as well as noticing of grammatical errors. The finding is important, not only in further highlighting the significance of learner engagement in the process of language learning, but in finding a more efficient and perhaps even more effective approach toward assessing the learners' writings for grammatical accuracy, as compared to teacher feedback. As Erlam et al. (2013) indicate, efficiency is a critical variable to be taken into account.

The study has some implications for the future research to consider. Primarily, from among the various options identified in the literature for providing WCF (Ellis, 2009), the current investigation trained the learners to provide metalinguistic comments as a variety of focused, direct, explicit WCFs; it

was assumed that this way the participants would be encouraged to reflect upon the rules, consider alternatives and with reason decide which one best suits the context, thereby promoting their long-term acquisition. Notwithstanding, it is still open to further investigation the question, on the one hand of which approach to WCF is more appropriate for peer feedback and how much feedback is enough for the peers to provide, for the internalization of grammatical structures to take place. For WCF provided by the teachers, Bitchener and Knotch (2009) suggest that the mere marking of errors would suffice; however, further studies are needed to consider this issue with respect to peer corrective feedback.

Furthermore, the study dismissed a consideration of learner characteristics and how they interact with the students' engagement with the corrective feedback provided. However, as Ferris et al. (2012) contend, this is quite a significant issue to be considered in research on WCF. Moreover, in examining the learners' acquisition in the study, merely their accurate use in writing tasks was considered. Future research could explore this issue by examining the extent to which learners maintain this accuracy in speaking as well.

References

Aljaafreh, A., & Lantolf, J. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78, 465-483.

Berg, B. (1999). The effect of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8, 215–241.

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

Bitchener, J. (2012). A reflection on 'the language learning potential' of written CF. *Journal of Second Language Writing*, 21, 348-363.

Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System*, 37(2), 322-329.

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

Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly*, 28, 181–188.

De Guerrero, M., & Villamil, O. (1994). Social cognitive dimensions of interaction in L2 peer revision. *The Modern Language Journal*, 78, 484–496.

Diab, N. M. (2011). Assessing the relationship between different types of student feedback and the quality of revised writing. *Assessing Writing*, 16, 274-292.

Diab, N. M. (2016). A comparison of peer, teacher and self-feedback on the reduction of language errors in student essays. *System*, *57*, 55-65.

Eckerth, J. (2008). Investigating consciousness-raising tasks: pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics*, 18, 2, 119-145.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63, 97-107.

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.

Elram, R., Ellis, R., & Batstone, R. (2013). Oral corrective feedback on writing: Two approaches compared. *System*, 41, 257-268.

Elwood, J. A., & Bode, J. (2014). Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System*, 42, 333-343.

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

Ferris, D. (2010). Second language writing research and written corrective feedback: Intersections and practical applications. *Studies in Second Language Acquisition*, 32, 181-201.

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, *22*(3), 307-329.

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.

Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing*, 1(3), 255–276.

Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9, 321–342.

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*, 75, 305–313.

Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of second language writing*, *18*(1), 30-43.

Manchon, R. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. John Benjamins Publishing Company.

McMartin-Miller, C. (2014). How much feedback is enough?: Instructor practices and student attitudes toward error treatment in second language writing. *Assessing Writing*, 19, 24-35.

Min, H. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15, 118–141.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375-401.

Norris, J. M., & Ortega, L. (2012). Assessing learner knowledge. In S. Gass & A. Mackey (Eds). *The Routledge handbook of second language acquisition* (pp. 573-589). London and New York: Routledge, Taylor and Francis Group.

Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. *Sociocultural theory and second language learning*, *4*, 51-78.

Peterson, S. (2003). Peer response and students' revisions of their narrative writing. *L1-Educational Studies in Language and Literature*, 3, 239–272.

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of second language writing*, *10*(4), 277-303.

Rahimi, M. (2009). The role of teacher's corrective feedback in improving Iranian EFL learners' writing accuracy over time: Is learner's mother tongue relevant? *Reading and Writing: An Interdisciplinary Journal*, 22, 219–243.

Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research*, 17, 1, 67-89.

Rahimi, M. (2015). The Role of Individual Differences in L2 Learners' Retention of Written Corrective Feedback. *Journal of Response to Writing*,*1*(1).

Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, *29*, 73-82.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Scrivener, J., 2005. *Learning Teaching: A Guidebook for English Language Teachers*. Macmillan Education, Oxford.

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. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32, 203-234.

Strijobs, J. W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and students' competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency?. *Learning and Instruction*, 20, 291-301.

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

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in the study of language: Studies in honour of H. G Widdowson*. Oxford: Oxford University Press.

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

Truscott, J. (1999). The case for ''the case for grammar correction in L2 writing classes": a response to Ferris. *Journal of Second Language Writing*, 8, 111–122.

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

Tsui, A., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9, 147–170.

Unlu, Z., & Wharton, S. M. (2015). Exploring classroom feedback interactions around EAP writing: A data-based model. *Journal of English for Academic Purposes*, 17, 24-36.

Zue, W. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication*, 1, 492–528.

Appendix
List of words used to perform the three tasks in the pre-test and post-test
l'outarde, le pêcheur, le caribou, les préparatifs, le dépeçage, le poteau, la tranche, paisiblement, serviable, débrouillard, fendre, transporter, étirer, camper, chasser (the goose, the fisherman, the caribou, preparations, butchering, the post, the slice, peacefully, helpful, resourceful, split, transport, stretch, camp, hunt)