L2 Wreeσ'' rr cceiii ng ff Teccher ooo ooo pueσ-generated Feedback*

Shiva Kaivanpanah**
Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran.
Dr. Seyed Mohammad Alavi***
Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran.
Rose Meschi (Corresponding author) ****
Faculty of Foreign Languages and Literatures, University of Tehran, Kish
International Campus, Kish, Iran.

Abstract

Writing is thought as the most complicated skill in second language acquisition; therefore, L2 researchers have always been in pursuit of discovering an effective approach to improve it. One of the most debated ways is feedback which has a key role in improving the quality of writing. Much of the previous research on feedback has focused on analyzing different types of feedback and their effect on the learners' writings and few studies have examined the effectiveness of computer feedback. Therefore, the present study was conducted to 1) determine what aspects of students' writings receive computer feedback, 2) examine the difference in the effect of computer-generated feedback (CBF) and Teacher-based feedback (TBF) on improving the students' writing quality and 3) compare the differences in Depth of Processing (DOP) in processing computer and teacher feedback. The results indicated that content, style and organization of their essays received feedback from the teacher and the computer. Teacher feedback was more effective in terms of its impact on improving the quality of the writing of the students than computer-generated feedback and it resulted in deeper processing of lexical items, whereas computer-generated feedback invoked medium processing on grammar.

Key words: Computer-generated feedback, Depth of processing, Teacher-based feedback, Writing, Writing quality.

1. Introduction

Numerous studies on L1 and L2 writing have pointed to the role of feedback in writing (Beach, 1976; Bitchener, 2008, 2012; Bitchener, Young & Cameron, 2005; Chandler, 2003; Ferris 1995, 1999, 2004, 2010; Sommers, 1980). Feedback functions as an important feature in developing students' learning through assessment, which can boost their self-confidence and encourage them (Alavi, Kaivanpanah, Danesh, 2019). Nevertheless, researchers have not yet come to a consensus regarding the effectiveness of feedback on improving the quality of students' writings. Truscott (1996) stated in his review essay that: "…correction is harmful rather than simply ineffective" (p. 360). He noted that although students desire grammar correction, teachers should not provide them with corrective feedback as it is harmful. Ferris (1999) offered a rebuttal to Truscott's strong opposition against feedback and argued that evidence and research on feedback was not conclusive enough to arrive at such strong stance. She contended that Truscott had disregarded some helpful evidence of corrective feedback. In 1999, Truscott published another study, in which he strongly held his position on feedback. In this regard, Ferris (1999, 2004) argued that providing feedback and grammar correction on students' writings had positive effects in terms of improving the accuracy and quality of their writing. She argued that students desire to receive feedback on their errors could not be simply disregarded. Similarly, some declare WCF to be effective in decreasing the number of students' errors (Ferris, 1999; Chandler, 2003; Lalande, 1982) while others have argued that it should not be used due to its fruitlessness in the long run (Krashen, 1984; Truscott, 1996; Truscott, 1999).

Despite differences in their opinion about the usefulness of feedback, Truscott (1996, 1999) and Ferris (2004) agree on two things: first, that there is not enough research available on feedback and second, that "proof of burden" is on researchers who claim that feedback and error correction are necessary for improving writing quality. Reviewing previous studies on the effectiveness of feedback (Ashwell, 2000; Fathman & Whalley, 1990; Ferris & Roberts, 2001), Ferris (2004)

concluded that most studies are in favor of error correction, and reported that it is not only helpful but necessary

Several studies on L2 writing have proposed that feedback plays a crucial role in writing and have focused on studying and comparing different types of feedback which lead to improvement in writing (Carroll & Swain, 1993; Ferris, 1999, 2002, 2003, 2004, 2010; Ferris & Roberts, 2001; Han, 2002; Kim & Bowles, 2019; Mackey, Gass, & McDonough, 2000). Researchers and writing instructors have also had a strong interest in deciding what types of feedback can better improve students' writing skills especially after Ferris (2004) pointed put the positive effects of error correction on writing and due to the everlasting controversies among scholars about feedback and its types, in addition to discover the most effective type.

Written Corrective Feedback (WCF) is considered as one of the most commonly investigated types of feedback to aid language learners and to provide them with the correct forms (Long, 1991). Yet the lack of agreement among researchers has encouraged the researchers of the present study to examine how deeply learners process the feedback they receive when revising/ editing their writing. The notion of depth of processing (DOP) states that if the learner uses higher and greater cognitive efforts at the time of processing information while drawing from the prior background knowledge, the possibility of recalling the processed information increases.

Most studies conducted in WCF have not compared computer-generated feedback with teacher feedback with a focus on students' DOP using think-aloud protocols. This has encouraged the researchers to examine the writing aspects attended to in each type of feedback, find out whether there is a difference between computer-generated and teacher-based feedback in terms of their effect on improving the quality of writing, and find out how deeply learners process feedback.

2. Review of the Literature
2.1. Written Corrective Feedback

Written Corrective Feedback (WCF) is a reaction to language errors in a writing made by L2 learners. WCF is commonly given on linguistic errors and targets the accuracy issues of the L2 learners rather than content, stylistic or organizational errors (Bitchener & Storch, 2016). It allows students to become aware of their shortcomings and creates a teacher-student communication in L2 writing class (Ferris, Pezone, Tade, & Tinti, 1997). Yoshida (2010) states that "Teachers give feedback in the classroom to prompt learners to use more appropriate expressions or sentences, even when their utterances are grammatically correct" (p. 293). Researchers examining the role of corrective feedback have always been concerned with its effect on improving students' writing accuracy as well as its facilitative function (Ferris, 1995; Ferris, 2004; Ferris, 2010; Hyland & Hyland 2006;). To address controversies about the effectiveness, Chandler (2003), Ferris and Roberts (2001), Semke (1984), Frantzen (1995), Ferris (2010) and Kim and Bowles (2019) examined the differences among essays which had received WCF and those which did not. They found positive effects of WFC on writing quality. Their results were in line with Lalande (1982) indicating that WCF improved the L2 learners' writings. In a similar vein, a growing number of studies have found positive effects for WCF, focusing on factors such as the nature of writing assignment task, students' proficiency level and the context of the writing class (Ashwell, 2000; Fathman & Whalley, 1990; Ellis, Basturkmen, & Loewen, 2001; Ellis, Loewen, & Erlam, 2006; Ferris, 2010; Ferris & Roberts, 2001; Kim & Bowles, 2019; Norris & Ortega, 2000).

2.2. Studies on Feedback
2.2.1. Types of Feedback

WCF can be direct and indirect. If the instructor seeks to provide the correct linguistic form, they use direct CF (Bitchener & Storch, 2016). Direct feedback provides the learners with the correct lexical or grammatical form, and guidance on how to correct the errors. This type of feedback is appropriate for situations in which the learners are not

aware of the correct form (Ellis, 2008). Ferris and Roberts (2001) believed direct CF is more effective than indirect feedback for lower level students as it might be more difficult for them to find out the nature of the error in their writing due to their level of proficiency. Sheen (2007) noted that direct CF can be useful for the acquisition of specific grammatical features, especially when it targeted a specific linguistic feature. As this type of feedback does not require much cognitive processing from the learner, it may not result in long-term learning (Ellis, 2008). On the other hand, if the focus is on pinpointing the error with the purpose of encouraging the learner to use the correct form, teachers provide indirect CF (Bitchener & Storch, 2016). Indirect feedback indicates the error and informs the learner that a mistake has occurred without giving the correct linguistic form. This is usually done by underlining or using cursors to attract the learner's attention to the missing elements (Lee, 2004).

Previous research on direct and indirect feedback has indicated that for most learners with different proficiency levels indirect feedback is more effective than the direct type as it accommodates "guided learning and problem solving" (Lalande, 1982). Indirect feedback also has the advantage of encouraging students to process the linguistic structures, whereas in direct CF the disadvantage is that it fosters marginal reflection on the side of learners, hence it may not lead to long-term acquisition (Ellis, 2008; Lalande, 1982).

Numerous studies examining the usefulness of feedback have reached similar conclusions and noted that teacher feedback results in better improvements of students' writings (Enginarlar, 1993; Ferris, 1995, 1999, 2002. 2003. 2004. 2010; Ferris & Roberts, 2001; Hedgcock & Lefkowitz, 1996; Miao, Badger & Zhen, 2006). Nowadays in the fast pacing modern world of education the necessity to use technology mediated ways of providing feedback has risen more than the past. Therefore, the possible effective role of computer programs has gained the researchers in the field of feedback.

2.2.2. Computer-generated feedback

While studies in second language writing propose that teacher feedback has a positive effect on students' writings, providing students with such feedback on a large scale can be problematic (Ebyary & Windeatt, 2010). Teachers cannot give personal, fast, feedback on the content to a large number of students (Grimes & Warschauer, 2010; Lee et al, 2009). Given that, numerous researchers have searched for alternatives to teacher feedback. One potential solution is employing computer programs which provide computer-generated feedback (CFB) on students' writings.

Computer software that provide feedback on students' essays are cost-effective compared to teacher-based feedback. Computer feedback is provided by software that provides immediate feedback on students' writings (Warschauer, 2004). Most computer-generated feedback tools are web-based which enables them to provide a set of support and customer features that range from securing the learner's privacy by using a password to providing different writing practices, model essays and even a system of scoring or placement test (Ware, 2011). These programs provide feedback on a range of various elements of writing, such as grammatical errors, holistic evaluations regarding content, style and mechanical and organizational aspects of writing (Chen & Cheng, 2008; Ericsson & Haswell, 2006; Shermis & Burstein, 2003; Ware, 2005; Warschauer & Ware, 2006).

Computer-generated feedback is an economically feasible alternative to teacher feedback, as it provides immediate feedback to a large number of students. Using computer-generated feedback has dramatically decreased the hours instructors spend on commenting and correcting writings (Warschauer & Ware, 2006). However, research on CBF is controversial; whereas some scholars believe it to be effective after intensive use (Coniam, 2009; Hutchison, 2007; Lee, Gentile, & Kantor, 2010; Shermis & Burstein, 2003), others have reported no improvement or negative results after receiving computer feedback (Warschauer & Grimes, 2008; Lai, 2010; Lee et al, 2009; Tuzi, 2004). Deciding on whether computer-generated feedback helps students'

writings or not mainly depends on how writings is described, what is the purpose of writing and in what way computer is going to help. Researchers who describe writing as a discrete skill report improvement after long-term use of this type of feedback (Attali & Burstein, 2006; Lee, Gentile, & Kantor, 2010; Shermis & Burstein, 2003).

A number of computer software which provide feedback on a written work are *Automated Essay Scoring* (AES) (Shermis & Buretein, 2003), *Automated Writing Evaluation* (AWE) (Warschauer & Ware, 2006), *MY Access*, *Holt Online Scoring, BETSY*, *Criterion* and *Grammarly.* In order to test *AES* effectiveness in providing feedback, Lee et al (2009) compared a web-based essay evaluating system which they developed to provide immediate feedback on content and organization. They compared participants in experimental group who had feedback from the system and those in control group who wrote their essays on the computer without receiving any web-based immediate feedback. No significant difference was found between the two groups in terms of essay length or in the final scores given by two human raters.

Attali (2004) employed *Criterion*, which gives feedback on form, language use, mechanics, and style as well as a holistic essay score to find out more about computer-generated feedback. He counted 9275 essays submitted to *Criterion* by the learners. The software provided immediate feedback, on the error numbers, and feedback on grammar, usage, mechanics, and style errors for each essay. The participants were asked to submit a revised essay according to the feedback they had received. The data from the first and last submitted essays indicated improvement of total scores and increase in the length of the revised versions.

Lai (2010) examined learners' preferences for computer-generated feedback. the effect of the software, *MY Access* on students' writings was compared with peer feedback. The results showed that even though both types of feedback were helpful and effective, students were inclined to receive peer feedback. In another study, Tuzi (2004) investigated the relationship between student preferences and

effectiveness of different types of feedback. He aimed to discover if the most popular type of feedback among students was the most effective one leading to success in writing. Twenty L2 writers submitted their writings to a website which was specially designed for writing and receiving responses from friends and peers, and face-to-face feedback from writing instructors. Though students preferred teacher feedback, computer-generated feedback was more effective on improving the revised version of their work than oral feedback. Considering the value of computers, most scholars agree that now the time is ripe to inspect the potential of computer-generated feedback as a supplement for teaching writing(Warschauer & Ware, 2006; Chen & Cheng, 2008; Shermis & Burstein, 2003; Ware, 2005; Hyland, 2010).

The controversies regarding feedback has heated discussions regarding the types of feedback are most effective; hence researchers have always been concerned with questions as to whether they should provide feedback on learners' writing (Carroll & Swain, 1993; Chandler, 2003; Ferris, 1999, 2002, 2003, 2004, 2006, 2010; Ferris & Roberts, 2001; Han, 2002; Kim & Bowles, 2019; Mackey, Gass, & McDonough, 2000). In addition, they have always been concerned with the type of feedback which could result in significant improvement of EFL learners' writings (Bitchener, Young & Cameron, 2005; Guénette, 2007; Sheen, 2004; Yoshida, 2010). Nevertheless, there is lack of consistency among studies on the effect of feedback on writing. One possible reason might be lack of attention to feedback processing by learners (Kim & Bowles, 2019). Therefore, an examination of the notion of depth of processing (DOP) is necessary to shed light on how deeply learners process feedback.

2.3. Depth of Processing (DOP)

Depth of processing (DOP) is defined as the combination of cognitive effort, degree of analysis, and expansion of intake, together with employing background knowledge, hypothesis examining, and rule making used in analyzing and encoding same grammatical or lexical item in the input. DOP concentrates on participants' level of cognitive work and time spent processing the target item, than on accuracy of the

produced language. DOP measures how deeply the L2 learner takes part in processing the new data, though he may not find out the correct answer related to the original structural rule (Leow, 2015; Leow et al., 2008; Leow & Mercer, 2015; Morgan-Short et al., 2012). The notion of DOP has not been largely used in relation to writing or feedback, as it has mostly been investigated in reading tasks. Nevertheless, DOP can be exceptionally useful for gaining insight into the minds of students as they receive feedback, how they process it and how they implement it into their writings.

In one of the few attempts to use think-aloud protocols to discover how deeply language learners analyze different types of feedback in their writing, Kim and Bowles (2019) studied 22 adult learners in an academic writing course. One of the participants' argumentative essays received direct correction, whereas in the other one their incorrect sentences were reformulated in a counterbalanced fashion. Think-aloud protocols of participants revealed that they processed sentential and paragraph-level mistakes at a deeper level; however, they overlooked the superficial-level mistakes when they get reformulation feedback. In addition, they discovered that reformulations evoked a higher DOP when participants processed non-surface errors. They could not report reformulation as the superior type of feedback in comparison with direct corrections, as their participants were advanced level students and as they had no measures for learning. Hence they could not generalize their findings. Similar to Adrada-Rafael and Filgueras-Gómez (2019), they did not discover any differences between the performance of the think-aloud and silent groups in their writing tasks. They concluded that there might not be a one-size-fits-all kind of feedback for improving students' writing and that different errors require different kinds of feedback.

Literature review indicates that studies comparing the effect of computer and teacher feedback is scarce. Except for one study (Kim & Bowles, 2019), the role of DOP to in relation to feedback has not been examined. Therefore, the present study aims to examine how deeply

learners process different types of feedback. The following questions are addressed in the present study:

1. Which aspects of students' writing receive feedback in computer-generated or teacher-written feedback?

2. Is there a difference between computer-generated and teacher-based feedback in terms of their effect on improving the quality of writing?

3. Is there any difference between DOP in computer-generated and teacher feedback?

3. Method
3.1. Participants
The participants were 60 (38 females and 22 males) intermediate Persian-speaking English-as-a-foreign language learners who wrote an argumentative essay in class. They were assigned to two experimental conditions (one receiving teacher feedback and the other computer-generated feedback), each including 30 participants. In each group, the participants were divided into two groups; one had to think aloud while revising their writing and the students in the other group were silent. Both groups revised their essays after they receiving indirect/coded teacher feedback. The participants had an average of 5.5 years of study of English as their L2.

3.2. Materials
To make sure that participants were homogenous in terms of their general English proficiency, they took Oxford Placement Test (OPT). The OPT includes two sections: listening and grammar. There are 100 items in the listening section, which take approximately ten minutes to complete. The grammar section includes 100 items, which would take up to 50 minutes. The results revealed that from the original pool of 45 participants, 30 of them were at an intermediate level of proficiency.

Their writings were scored based on the ESL Composition Profile (see Appendix D) developed by Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981) which considers content, organization, vocabulary, language use, and mechanics; each one has four scoring ranks ranging from very poor, poor to fair, average to good, and very good to

excellent. The proficiency level of the writer is described through a series of components with clear definers, as well as a numerical scale. The content component scores range on a scale of 13-30, organization 7-20, vocabulary 7-20, language 5-25, and mechanics 2-5. The rating level of excellent to very good for content has a maximum score of 30 to a minimum of 27 representing a piece of writing which is well-informed, practical, (has) thorough growth of thesis, related to given topic, whereas very poor content of an essay had a maximum score of 16 and a minimum score of 13 showing that the essay "does not show knowledge of subject, (is) not substantive, not pertinent, or not enough to evaluate" (Jacobs et al., 1981).

In the group receiving computer-generated feedback, the students' writings received feedback from two electronic systems, *Grammarly* and *Write&Improve*. The default version of *Grammarly is* freely obtainable from its official online website and it can easily be installed on other devices. Grammarly claims to be the best grammar checker software in the world. Its free default version checks for grammar, contextual spelling, punctuation, sentence structure and style, whereas the premium version provides a check for plagiarism, vocabulary enhancement, as well as professional proofreading. The second system, *Write&Improve*, uses technology developed at the University of Cambridge. Learners submitted their work and *Write & Improve* scored participants' writings from A1 (lowest) to C2 (highest) on the Common European Framework of Reference (CEFR) scale. It also indicates parts of text that need improvement. It points out the errors within seconds in terms of content, grammatical structures, use of lexical items, style, and organization. It mostly highlights or pinpoints the errors and does not provide writers with the immediate correct answer. Therefore, learners can work on areas that require their attention and keep improving.

To examine the role of DOP in the incorporation of corrective feedback on the writings of EFL language learners, each participant's level of processing for lexical and grammatical items was determined by following Leow's operationalization of DOP (2015) (See Appendix E).

3.3. Procedures

The study was conducted in two stages; in the first stage, participants wrote an argumentative essay in 250 words. They were given 40 minutes to write the essay and were not allowed to use dictionary or any other sources. They were asked to proofread their essays before handing them to the teacher or submitting them to the software to make sure that the possible existing errors were not mistakes or slips of pen or typing errors, which could have been possibly caused by their lack of concentration and could be self-corrected by the writer (James, 1998; Qi & Lapkin, 2001; Poulisse, 1999). The participants were instructed by the teacher on how to think aloud to ensure that they could let their thoughts flow as they attempted to write the second version of their essays. The researcher and two other teachers also modeled how to think aloud while reflecting on the feedback on a piece of writing. The instruction aimed to help students to explain what they were thinking while correcting their essays and hence helped the researchers in gaining insights into their minds and thoughts.

After two weeks, students were required to think aloud concurrently while revising their essays based on the feedback they had received. For this stage, they were given the same amount of time to proofread their work. The two writings of each participant were scored by the researcher and two raters, using the ESL composition profile. The inter-rater reliability was calculated 80%. The collected essays were corrected either by the teacher, or the software in terms of their form (grammatical accuracy), content, semantics or word choice. The participants in the Teacher feedback group received indirect coded feedback on their essays. The teacher feedback included comments on grammatical errors (grammar, wrong tense, wrong format), and punctuation and lexical (wrong word, spelling, extra word and extra or repeated words). In the computer-generated feedback group, participants received feedback from the software on grammar, spelling, punctuation, word choice and passive voice. An example of each type of feedback is provided in the Appendix A, B and C.

4. Results

Table 1 displays what aspects of students' essays received feedback in each group. More errors were found by the teacher than by the computer. The computer-generated feedback mainly focused on grammar, spelling, punctuation and partially on "suspicious word" which generally focused on any lexical problem, extra or even the wrong format. No further explanation was given by the system in regard to how the word is suspicious or why it was an error.

Table 1

Aspects of writing receiving teacher written and computer-generated feedback

	Teacher Written Feedback	Computer-generated Feedback
Grammar	78	55
Spelling	42	32
Wrong Format	55	-
Wrong Word	37	-
Punctuation	26	30
Extra word	11	-
Suspicious word	10	50

To compare the effects of teacher- and computer-generated feedback on writing quality, participant's writing was assessed according to ESL Composition Profile proposed by Jacobs et al. (1981).As seen in Table 2 the mean score of writings in both groups increased suggesting that that feedback improves the quality of writing.

Table 2

Descriptive Statistics of two groups receiving two types of feedback

	Group	N	Mean	Std. Deviation	Std. Error Mean
writing quality v1	Teacher Feedback+TA	15	80.26	8.97	2.31
	Computer Feedback+TA	15	82.80	5.32	1.37
writing quality v2	Teacher Feedback+TA	15	85.00	6.81	1.75
	Computer Feedback+TA	15	86.06	4.41	1.14

Table 3 displays the results of the independent sample t-test comparing the difference between the quality of writing in the groups receiving teacher feedback and computer-generated feedback.

Table 3

Independent Sample t-test of two groups of teacher and computer feedback

	Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
writing quality v1	4.226	.049	-.940	28	.355	-2.53333	2.69509	-8.05397	2.98730
writing quality v2	4.413	.045	-.509	28	.615	-1.06667	2.09641	-5.36096	3.22763

The results reveal that there is no significant difference between the writing quality of students who received teacher or computer-generated feedback (sig=.355, sig=.615). However, the comparison of the mean scores (Table 2) indicate that both types of feedback helped students

revise their writing accurately and hence as a result of receiving feedback the quality of writings improved.

To examine the role of DOP in the incorporation of corrective feedback, each participant's level of processing for lexical and grammatical items was determined following Leow's operationalization of DOP (2015) (See Appendix E). Table 4 displays the frequency analysis for each level of processing (low, intermediate and high) regarding lexical or grammatical items after receiving teacher- and computer-generated feedback.

Table 4

Descriptive Statistics of lexical and grammatical items in the Teacher and Computer-generated feedback

Participants (TFB/TA)	Lexical items			Grammatical items			Total Low	Total Medium	Total High	Accuracy		Quality	
	Low	Medium	High	Low	Medium	High				1st version	2nd version	1st version	2nd version
1.	6	10	0	23	9	2	29	19	2	0	9	60	71
2.	10	0	1	15	3	0	25	3	1	10	17	88	90
3.	6	1	0	3	4	0	9	5	0	7	10	90	92
4.	10	9	10	7	8	11	17	17	21	5	24	78	90
5.	0	2	3	11	2	1	11	4	4	2	7	70	85
6.	5	4	2	9	0	0	14	4	2	2	4	70	75
7.	3	3	1	0	2	7	3	5	8	7	10	85	87
8.	4	5	3	7	13	5	11	18	8	7	16	86	87
9.	4	6	2	5	11	6	9	17	8	8	13	76	79
10.	2	4	1	7	8	4	9	12	5	5	10	77	79
11.	6	3	0	8	10	7	14	13	7	4	12	88	89
12.	4	7	4	5	12	9	9	19	13	10	12	89	91
13.	5	8	3	4	7	3	9	15	6	14	23	91	92
14.	2	9	0	9	8	0	11	17	0	10	16	77	78
15.	6	11	2	10	5	6	16	16	8	7	17	79	90
Grand Total	73	82	32	123	102	61							
Participants (CFB/TA)													
16.	1	0	3	4	5	1	5	5	4	14	17	92	95

17.	4	4	3	3	1	1	7	5	4	12	17	84	86
18.	3	2	1	4	2	5	7	4	6	10	14	87	89
19.	14	7	6	4	9	1	18	16	7	5	18	90	92
20.	5	3	1	4	7	2	9	10	3	11	18	80	86
21.	1	0	4	7	4	0	8	4	4	12	18	75	83
22.	4	8	5	5	5	1	9	13	6	13	19	83	86
23.	2	3	1	4	3	1	6	6	2	10	15	78	82
24.	2	2	2	3	4	4	5	6	6	10	12	80	83
25.	4	3	1	2	5	3	6	8	4	11	19	81	84
26.	5	1	3	5	2	5	10	3	8	7	16	76	79
27.	2	0	4	4	3	6	6	3	10	8	18	90	92
28.	1	9	2	5	10	7	6	19	9	13	19	87	88
29.	7	4	3	1	6	4	8	10	7	12	20	80	84
30.	4	2	5	8	5	9	12	7	14	14	22	79	82
Grand Total	59	48	44	63	71	50							

The results in Table 4 indicate that participants processed lexical items more at a medium level (82) than at a low level (73) or than at a high level (32). Participants in computer-generated feedback group processed the lexical items more at a low level (59) than at a medium (48) or at a high level (44). These DOP results suggest that learners spent more time processing lexical items according to the collected think-aloud data, which means that participants employed a greater level of cognitive effort to focus on the meaning of the target item. In teacher feedback, most participants processed the grammatical items at a low level (123) rather than medium (102) or a high level (61), while in computer-generated feedback group, participants processed the grammatical items more at a medium level (71), than at a low (63) or a high level (50). This would suggest that after receiving teacher feedback, participants had lower level of cognitive effort for grammatical items than for lexical items. However, in the computer group they had a medium level of processing for grammatical items while writing the second version of their essay.

The quality of the first and revised writings were compared to find out how DOP is related to changes in the written product. Table 5 presents the descriptive statistics for the first and the second writings.

Table 5

Descriptive Statistics of the pre and posttest in teacher and compute-generated feedback

	Group	N	Mean	Std. Deviation	Std. Error Mean
writing quality v1	Teacher Feedback+TA	15	80.2667	8.97987	2.31859
	Teacher Feedback+Silent	15	73.9333	5.54806	1.43250
writing quality v2	Teacher Feedback+TA	15	85.0000	6.81385	1.75933
	Teacher Feedback+Silent	15	80.1333	5.13902	1.32689
writing quality v1	Computer Feedback+TA	15	82.8000	5.32112	1.37391
	Computer Feedback+Silent	15	74.7333	5.43095	1.40227
writing quality v2	Computer Feedback+TA	15	86.0667	4.41534	1.14004
	Computer Feedback+Silent	15	77.6667	4.99524	1.28976

Table 5 indicate that the mean scores of students' writings after receiving both types of feedback improved. Both types of feedback resulted in quality improvement. Comparing the mean scores for the second version of students' writings who thought aloud and those who remained silent after receiving the same type of feedback shows higher scores for both groups and hence more improvement for those who thought aloud in both feedback groups.

A paired sample t-test was run to examine how DOP is related to the changes in the writings. As it can be seen in Table 6, there is a significant difference between the writing quality in the pre and posttest after receiving feedback.

Table 6

Paired sample T- test comparing DOP in the first and revised writing

		Levene's Test		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Teacher FB	writing quality v1	3.903	.058	2.324	28	.028	6.33333	2.72542	.75056	11.91611
	writing quality v2	2.317	.139	2.209	28	.036	4.86667	2.20360	.35279	9.38055
Computer FB	writing quality v1	.000	1.000	4.109	28	.000	8.06667	1.96315	4.04533	12.08800
	writing quality v2	.649	.427	4.880	28	.000	8.40000	1.72139	4.87390	11.92610

Table 6 show that higher processing led to higher quality. Both groups had improvements in their second writing. However, when comparing the DOP results, it can be seen that teacher feedback resulted in more instances of deeper processing for both lexical and grammatical items, therefore it proved to be more effective.

To find out whether the improvements from the first to the second writing were significant, the gain scores were compared using paired t-test.

Table 7

Gain scores in two groups of teacher and computer-generated feedback

Teacher FB	Pre test	Post test	Gain Scores	Computer FB	Pre test	Post test	Gain Scores
1	60	71	**11**	1	92	95	**3**
2	88	90	**2**	2	84	86	**2**
3	90	92	**2**	3	87	89	**2**
4	78	90	**12**	4	90	92	**2**
5	70	85	**15**	5	80	86	**6**
6	70	75	**5**	6	75	83	**8**
7	85	87	**2**	7	83	86	**3**
8	86	87	**1**	8	78	82	**4**
9	76	79	**3**	9	80	83	**3**
10	77	79	**2**	10	81	84	**3**
11	88	89	**1**	11	76	79	**3**
12	89	91	**2**	12	90	92	**2**
13	91	92	**1**	13	87	88	**1**
14	77	78	**1**	14	80	84	**4**
15	79	90	**1**	15	79	82	**3**
16	70	76	**6**	16	70	73	**3**
17	71	79	**8**	17	67	71	**4**
18	75	83	**8**	18	65	70	**5**
19	80	88	**8**	19	72	74	**2**
20	83	89	**6**	20	73	75	**2**
21	86	90	**4**	21	75	77	**2**
22	77	80	**3**	22	79	81	**2**
23	69	76	**7**	23	77	83	**6**
24	70	79	**9**	24	76	77	**1**
25	73	78	**5**	25	80	82	**2**
26	76	79	**3**	26	79	83	**4**
27	68	74	**6**	27	84	87	**3**
28	70	78	**8**	28	79	80	**1**
29	68	74	**6**	29	68	73	**5**
30	73	79	**6**	30	77	79	**2**

Table 7 presents the results of each participant's gain score after receiving teacher or computer feedback in order to make a comparison between pretest and posttest results. Table 8 presents the results of independent sample t-test comparing gain scores in both feedback groups.

Table 8

Independent Sample t-test comparing gain scores in two groups of teacher and computer-generated feedback

	Mean teacher	Mean computer	Levene's Test		t-test for Equality of Means						
										95% Confidence Interval of the Difference	
							Sig. (2-tailed)	Mean Difference	Std. Error Difference		
			F	Sig.	t	df				Lower	Upper
	5.13	3.10	16.714	.000	2.796	58	.007	2.033	.727	.578	3.489

A significant difference was found between the gain scores in teacher and computer-generated feedback groups indicating that teacher feedback is more effective than computer feedback in improving the quality of writings.

5. Discussion

The present study focused on the aspects which received teacher and computer feedback. The findings indicated that computer-generated feedback focused on the basic aspects of writing such as grammar, spelling, punctuation and a vague, general category, called suspicious word category. Teachers also commented on grammar, spelling, punctuation; however, rather than providing "suspicious word" feedback type, teacher feedback was detailed in expressing the lexical problem and dividing errors into wrong format and wrong word and extra word. Although computer-generated feedback has a number of advantages such as providing the opportunity for the students to submit various versions of their writings and receive immediate feedback, offering holistic and analytic scores, presenting graphs and charts to

display learner improvement, it is mechanical and does not take into account the implied meaning in complex sentences; hence, for complex structures computer feedback is unable to detect more complicated errors.

The effect of teacher- and computer-generated feedback on the quality of EFL students' writings were compared. The results indicated a significant improvement in the quality of writings in both feedback groups. Students in the teacher feedback group had deeper processing of lexical and grammatical items and higher writing quality improvement than those receiving computer-generated feedback (Ware, 2011; Leung, 2017). Students who received teacher feedback had to interpret the error codes, discover the nature of it, and then produce a revised version of it. Contrary to students in the computer-generated feedback group who directly received the corrected version of their errors and did not process their errors, students in the teacher feedback group cognitively analyzed the received feedback in order to revise their paper; their involvement enhanced the chances of long-term learning and ultimately enabled them internalize the correct forms. This is highlighted by Leow and Mercer (2015) who noted that higher levels of cognitive processing are more likely to result in learning and leaving memory traces for future retrieval.

The findings of the present study related to the second research question focusing on improvement in the writing quality are in line with Ebyary and Windeatt (2010) who reported a significant improvement for most participants and a fairly dramatic improvement from their first to second essay after receiving teacher feedback. The findings regarding the higher quality of writings after receiving teacher feedback compared to computer-generated feedback are in line with those of Ferris (2002), Ferris and Roberts (2001), Shizuka (2000), Kahraman (2013) and Ware (2011) who reported that teacher-feedback in a coded or indirect style, is most effective in improving the participants' writing skill and accuracy, developing their metalinguistic awareness, reducing anxiety and fostering language learners' cognitive

and affective, as well as contributing to long-term learning of specific linguistic forms.

Regarding the differences in DOP of the participants receiving teacher and computer-generated feedback, it was found that participants had low processing of lexical items and medium processing of grammatical items in the computer-generated feedback group and in teacher feedback medium processing of lexical items and low processing of grammatical items was reported. For computer-generated group, this shows that participants employed more cognitive effort while processing the grammatical items, as they spent more time comparing their mistakes against the universal grammatical rules. This might be due to the way they received feedback; the correct answer was directly given to them by the software, while the teacher did not directly provide the correct answer and only pointed out that an error has occurred which demanded some levels of cognitive activity and processing on the part of students. In the computer-generated feedback group only a few more instances of medium (71) than low (63) processing were reported when processing the grammatical items. Students who received teacher feedback processed the lexical items at medium level (82) and the grammatical items at low level (123). As Sachs and Polio (2007) noted direct error correction highlights the errors more clearly, like the computer software, whereas in teacher feedback the errors were not directly corrected, but only located and coded based on the type of error. On the other hand, as in teacher feedback the correct answer was not given to the participants, they had to employ higher cognitive effort to discover the meaning. As Leow and Mercer (2015) discussed learners apply a greater cognitive effort to fully understand the meaning of the words and their prior knowledge. Deep processing takes place when learners pay attention to features beyond the appearance of lexical items, when learners process and analyze a word for its meaning or consider its relation with other words (Leow & Mercer, 2015). Software such as *Grammarly* or *Write & Improve* which offer computer-generated feedback have yet to be strong enough to replace a teacher's feedback (Ware, 2011). In these programs

writing analysis occurs in terms of a defined collection of rules and formulas; therefore, they are unable to provide feedback on deeper, more complicated concepts conveyed through language (Ware, 2011; Leung, 2017). Based on the findings it can be argued that computer-generated feedback is effective in improving the quality of students' writings. Although it has the potential to be used as for writing instruction courses due to its distinctive features such as the speed in providing feedback and scoring writings as well as the ability to provide charts and statistics on students' improvement or proficiency level, it is recommended to be employed as a supplement for writing courses and not as a replacement (Chen & Cheng, 2008; Shermis & Burstein, 2003; Ware, 2005; Warschauer & Ware, 2006).

Hence the students are prompted to process the teacher feedback deeply compared to computer feedback. The higher level of processing for lexical items in the teacher feedback group than in the computer group might be due to the fact that inferring the meaning and trying to find the problem in the detailed coded feedback offered by the teacher demanded more attention and cognitive effort on the part of the learners. As Leow and Mercer (2015) describe deeper processing of lexical items happens when learners apply a greater cognitive effort to the meaning of the words and use prior knowledge. One more concern about computer-generated feedback can be due to its mechanic and formulaic nature, as it could be seen in the results participants had low processing of lexical term when receiving computer-generated feedback. Processing the lexical items relates to meaning and demand deeper understanding of the target language. As it has been seen in the present study, computer-generated feedback can sometimes miss the imbedded meaning in some complex sentences and hence cause the sense that it is detached from the real world contexts, which may cause some shortcomings when computer is the main source of correction (Warschauer & Grimes, 2008; Lai, 2010; Lee et al, 2009; Tuzi, 2004).

6. Conclusion
Teacher feedback can be more time consuming and in some occasions it can even become subjective; however, when it comes to more

complex language structures it seems more fruitful than computer-generated feedback due to the reasons discussed at the results section which relate to giving indirect, coded feedback as well as being able to attend to more errors. The computer-generated feedback, on the other hand, can be delivered in seconds as soon as learners submit their essays. It has various advantages such as providing student privacy, model essays and a scoring system or level identification. However, at times it is unable to decipher complicated notions. Hence, teacher feedback is perceived to be more effective in providing comprehensive type of feedback and resulting in higher quality scores.

The findings regarding DOP pointed to differences in processing of lexical and grammatical items in teacher and computer-generated feedback. The reason for this difference might be due to students' lack of experience working with a computer software that provided them with feedback on their writings.

More research is required to test the situations in which the computer-generated feedback software is used, the content of the essay, and the purpose of the students are taking the writing course to further explore the results of the present study regarding the higher quality of students' essays after receiving teacher feedback Still, to increase knowledge about the use of technology in providing feedback, teachers need to learn how to benefit from computer-generated feedback in improving quality and accuracy of students' writing. Overall, there might not be a one-size-fits-all kind of feedback for improving students' writing quality, hence computer-generated feedback can be employed as a supplement alongside with teacher feedback to deliver more effective results.

References

Alavi, S.M., Kaivanpanah, S. Danesh. F. (2019). A Comparative Study of Writing Assessment Using Activity Theory-Based Assessment Model (ATBAM) and a Traditional Approach. *Journal of English Language Teaching and Learning*. 11(23). 1-25.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9, 227–258.

Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the Paper presented at the National Council on Measurement in Education, San Diego, CA.

Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater v.2. *Journal of Technology*, *Learning, and Assessment*, 4(3), 1–30.

Beach, R. (1976). Self-evaluation strategies of extensive and non-revisers. *College Composition and Communication, 27*(5), 160-164.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 18, 102–118.

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

Bitchener, John & Storch, Neomy. (2016). Written Corrective Feedback for L2 Development.

Bitchener, J. and Young, S. & Cameron, D. (2005). The effect of different types of feedback on ESL student writing. *Journal of Second Language Writing*. 14. 191-205.

Carroll, S. and Swain, M. (1993). Explicit and implicit negative feedback: an empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition* 15: 357-86.

Chandler, J. (2003). The Efficacy of Various Kinds of Error Feedback for Improvement in the Accuracy and Fluency of L2 Student Writing. *Journal of Second Language Writing*, 12, 267-296.

Chen, C-F., & Cheng, W-Y. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112.

Coniam, D. (2009). Experimenting with a computer essay-scoring program based on ESL student writing scripts. *ReCALL,* 21, 259-279.

Ebyary, Kh., Windeatt, S,. (2010). The Impact of Computer-Based Feedback on Students Written Work. *IJES, 10* (2), 121-142.

Ellis, R. (2008). A typology of written corrective feedback types. *Elt Journal*. 63. 97-107.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281–318.

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.

Enginarlar, H. (1993). Student response to teacher feedback in EFL writing. *System*, 21, 193-204.

Ericsson, P. F., & Haswell, R. (Eds.). (2006). *Machine scoring of student essays: Truth and consequences*. Logan, UT: Utah State University Press.

Fathman, A., & 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. R. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL Quarterly*, 29, 3353.

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

Ferris, D. R. (2002). Treatment of error in second language student writing. Ann Arbor: The University of Michigan Press.

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

Ferris, D. R. (2004). The "Grammar Correction" Debate in L2 Writing: Where are we, and where do we go from here? (and what do we do in the meantime …?). *Journal of Second Language Writing*. 13. 49-62.

Ferris, D. (2006). 'Does error feedback help student writers? New evidence on short- and long-term effects of written error correction' in K. Hyland and F. Hyland (eds.). *Feedback in Second Language Writing: Contexts and Issues.* Cambridge: Cambridge University Press.

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

Ferris, D. R. and Roberts. B. (2001). 'Error feedback in L2 writing classes: How explicit does it need to be?'. *Journal of Second Language Writing* 10: 161–84.

Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *Modern Language Journal*, 79, 329–344.

Guénette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*. 16. 40-53.

Grimes, D., and Warschauer, M. (2010). Utility in a Fallible Tool: A Multi-Site Case Study of Automated Writing Evaluation. *JTLA*, 8(6), 1-43.

Han, Z-H. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly* 36: 543-72.

Hedgcock, J. and Lefkowitz, N. (1996). Some input: Two analyses of student response to expert feedback in L2 writing. *The Modern Language Journal*, 80, 287-308.

Hutchison, D. (2007). An evaluation of computerised essay marking for national curriculum assessment in the UK for 11-year-olds. *British Journal of Educational Technology*, 38(6), 977-989.

Hyland, F. (2010). Future directions in feedback on second language writing: overview and research agenda. *International Journal of English Studies*, 10(2), p. 171-182.

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

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). Testing ESL composition; A practical approach. Rowley, MA: Newbury House.

James, C. (1998). Errors in language learning and use. Essex, UK: Addison Wesley Longman.

Kahraman, A. (2013). Affective and cognitive effects of coded teacher feedback on FL writing students. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education], 28*(1), 189-201.

Kim, H. R., Bowles, M. (2019). How deeply do second language learners process written corrective feedback? Insights gained from think alouds. *TESOL Quarterly. 0*(0), 1-26.

Krashen, S. D. (1984). Writing: Research, theory, and application. Oxford: Pergamon Press.

Lai, Y.-h. (2010). Which do students prefer to evaluate their essays: Peers or computer program. *British Journal of Educational Technology*, 41(3), 432-454.

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

Lee, I. (2004). Error correction in L2 secondary writing classrooms. *Journal of Second Language Writing,* 13, 285-312.

Lee, Y-W., Gentile, C., & Kantor, R. (2010). Toward automated multi-trait scoring of essays: Investigating links among holistic, analytic, and text feature scores. *Applied Linguistics*, 31(3), 391–417.

Lee, C., Wong, K., Cheung, W., & Lee, F. (2009). Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation. *Computer Assisted Language Learning,* 22, 57-72.

Leung, B. W. (2017) The Study of Teacher Written Feedback: The Effectiveness of Electronic Feedback on Student Writing Revisions, Durham theses, Durham University.
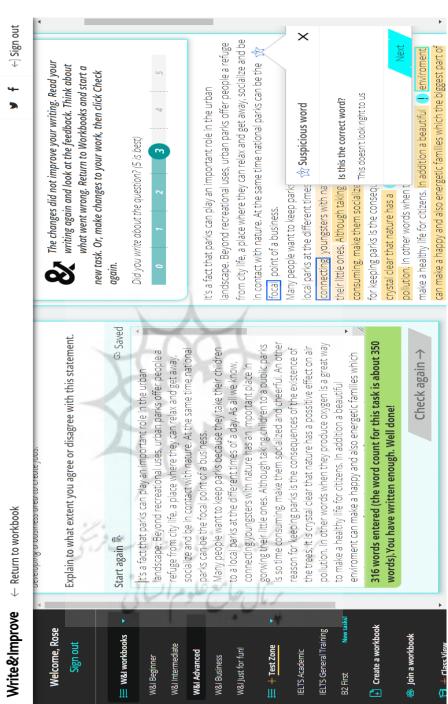
Leow, R. P. (2015). Explicit learning in the L2 classroom. New York. Routledge.

Leow, R. P., Hsieh, H. C., & Moreno, N. (2008). Attention to Form and Meaning Revisited. *Language Learning*, *58*(3), 665–695.

Leow, R. P. & Mercer, J. D. (2015) Depth of processing in L2 learning: theory, research, and pedagogy, *Journal of Spanish Language Teaching*, 2(1), 69-82.

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, 54, 153–188.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497.

Miao, Y., Badger,R. and Zhen Y. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179-200.

Morgan-Short, K., Heil, J., Botero-Moriarty, A., & Ebert, S. (2012). Allocation of attention to second language form and meaning: Issues of think alouds and depth of processing. *Studies in Second Language Acquisition, 34*, 659–685.

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

Poulisse, N. (1999). Slips of the tongue: Speech errors in first and second language production. Amsterdam: John Benjamins.

Qi, D. S., and 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.

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29(1), 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.

Shermis, M. D., & Burstein, J. (2003). ''Introduction.'' Automated essay scoring: *A cross disciplinary perspective* (pp. xiii–xvi). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shizuka, T. (2000). Effects of different editing methods on EFL writing quality: What do the students think? *Kanto Koshinetu Eigo Kyoiku Gakkai, 14*, 43-54.

Sommers, N. (1980). Revision strategies of student writers and inexperienced adult writers. *College Composition and Communication*, 31, 378-388.

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 against grammar correction in L2 writing classes": a response to Ferris. *Journal of Second Language Writing*, 8, 111-122.

Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, 21(2), 217-235.

Ware, P. (2005). Automated writing evaluation as a pedagogical tool for writing assessment. In A. Pandian, G. Chakravarthy, P. Kell, & S. Kaur (Eds.*), Strategies and practices for improving learning and literacy* (pp. 174–184). Selangor, Malaysia: Universiti Putra Malaysia Press.

Ware, P. (2011). Computer-Generated Feedback on Student Writing. *TESOL Quarterly.* 45(4), 769-774.

Warschauer, M. (2004). Technology and writing. In J. Cummins and C. Davison (Eds.), The Kluwer international handbook of English language education. Kluwer: Dordrecht, Netherlands.

Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies*, 3(1), 22–36.

Warschauer, M., and Ware, P. (2006). Automated writing evaluation: defining the classroom research agenda. *Language Teaching Research*, 10(2), 1-24.

Yoshida, R. (2010). How Do Teachers and Learners Perceive Corrective Feedback in the Japanese Language Classroom?. *The Modern Language Journal*. 94. 293 - 314.

Appendix
Teacher Feedback Sample

S = Spelling (L)
E = Extra (L)
G = grammar (G)
WW = wrong word (L)
WF = wrong format (G)
P = punctuation (G)

An argumentative essay: Public parks

It is true that whilst many people use their public parks, this space could be used for other purposes, such as building houses for young people or developing a business area to create jobs.

Explain to what extent you agree or disagree with this statement in 350 words.

① The most important effect of green space in cities is

translation

E
on their environmental functions that make cities

M

more convenient). ② Green spaces confront with destructive

effects of enPanding industry and transPortation).

ww — M.
It Promotes quality of citizen's life. ③ Parks are

green spaces   WF   WF

G.H   G.H
the best Place to relax and relieve fatigue.)

G.H
WF
Green space helps, to have concentration. (Entertainment
L   us   ④
and games help training and breeding the children's

body and mind). Although, Parks have little Playing equipment

⑤   G
(they help children to gain skills and creativity).   M

Appendix B
Computer-generated Feedback (Write&Improve)

Write&Improve

Welcome, Rose
Sign out

≡ W&I workbooks
W&I Beginner
W&I Intermediate
W&I Advanced
W&I Business
W&I Just for fun!

≡ + Test Zone
IELTS Academic
IELTS General Training
B2 First   New tasks!

Create a workbook
Join a workbook

← Return to workbook

Sign out

↩ Sign out

Explain to what extent you agree or disagree with this statement.

Start again

⟳ Saved

It's a fact that parks can play an important role in the urban landscape. Beyond recreational uses, urban parks offer people a refuge from city life, a place where they can relax and get away, socialize and be in contact with nature. At the same time national parks can be the focal point of a business.

Many people want to keep parks because they take their children to a local parks at the different times of a day. As all we know, connecting youngsters with nature has an important place in gorwing their little ones. Although taking children to a public parks is so time consuming, make them socialized and cheerful. An other reason for keeping parks is the consequences of the existence of the trees. It is crystal clear that nature has a positive effect on air pollution. In other words when they produce oxygen is a great way to make a healthy life for citizens. In addition a beautiful enviroment can make a happy and also energetic families which

316 words entered (the word count for this task is about 350 words). You have written enough. Well done!

Check again →

The changes did not improve your writing. Read your writing again and look at the feedback. Think about what went wrong. Return to Workbooks and start a new task. Or, make changes to your work, then click Check again.

Did you write about the question? (5 is best)

0   1   2   **3**   4   5

It's a fact that parks can play an important role in the urban landscape. Beyond recreational uses, urban parks offer people a refuge from city life, a place where they can relax and get away socialize and be in contact with nature. At the same time national parks can be the focal point of a business.

Many people want to keep parks at the different times local parks at the different times

connecting youngsters with na their little ones. Although taking consuming, make them socialize for keeping parks is the consequ crystal clear that nature has a pollution. In other words when t make a healthy life for citizens. In addition a beautiful enviroment can make a happy and also energetic families which the biggest part of

×

Suspicious word

Is this the correct word?
This doesn't look right to us

Next

Appendix C
Computer-generated Feedback (Write&Improve)

Appendix D
ESL Composition Profile (Jacobs et al., 1981)

ESL COMPOSITION PROFILE			
STUDENT DATE TOPIC			
SCORE	LEVEL	CRITERIA	COMMENTS
CONTENT	30- 27	**EXCELLENT TO VERY GOOD:** knowledgeable • substantive • thorough development of thesis • relevant to assigned topic	
	26-22	**GOOD TO AVERAGE:** some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail	
	21-17	**FAIR TO POOR:** limited knowledge of subject • little substance • inadequate development of topic	
	16-13	**VERY POOR:** does not show knowledge of subject • non-substantive • not pertinent • OR not enough to evaluate	

ORGANIZATION	20-18 17-14 13-10 9-7	**EXCELLENT TO VERY GOOD:** fluent expression • ideas clearly stated/supported • succinct • well-organized • logical sequencing • cohesive **GOOD TO AVERAGE:** somewhat choppy • loosely organized but main ideas stand out • limited support • logical but incomplete sequencing **FAIR TO POOR:** non-fluent • ideas confused or disconnected • lacks logical sequencing and development **VERY POOR:** does not communicate • no organization • OR not enough to evaluate	
VOCABULARY	20-18 17-14 13-10 9-7	**EXCELLENT TO VERY GOOD:** sophisticated range • effective word/idiom choice and usage • word form mastery • appropriate register **GOOD TO AVERAGE:** adequate range • occasional errors of word/idiom form, choice, usage but meaning not obscured **FAIR TO POOR:** limited range • frequent errors of word/idiom form, choice, usage • meaning confused or obscured **VERY POOR:** essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate	

LANGUAGE USE	25-22	**EXCELLENT TO VERY GOOD:** effective complex construction • few errors of agreement, tense, number, word order/ function, articles, pronouns, prepositions	
	21-18	**GOOD TO AVERAGE:** effective but simple constructions • minor problems in complex construction • several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions but meaning seldom obscured	
	17-11	**FAIR TO POOR:** major problems in simple/complex constructions • frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, deletions • meaning confused or obscured	
	10-5	**VERY POOR:** virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate	

MECHANICS	5	**EXCELLENT TO VERY GOOD:** demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paragraphing
	4 3	**GOOD TO AVERAGE:** occasional errors of spelling, punctuation, capitalization, paragraphing but meaning not obscured
	2	**FAIR TO POOR:** frequent errors of spelling, punctuation, capitalization, paragraphing • poor handwriting • meaning confused or obscured
		VERY POOR: no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paragraphing • handwriting illegible • OR not enough to evaluate

Appendix E
Operationalization of Depth of Processing (DOP), Leow (2015)
Operationalization of Depth of Processing (DOP): Lexical Items

	LEVEL 1	LEVEL 2	LEVEL 3
	Low depth of processing	Medium depth of processing	High depth of processing
Description	Shows no potential for emerging form-meaning	Provides some evidence of processing target item	Provides evidence of making accurate form - meaning connection
Descriptors	reads target quickly translates the phrase to English but leaves the target in Spanish says s/he isn't sure what it is says s/he will click something repeats the target item carefully pronounces target word does not spend much time processing target item low level of cognitive effort to get meaning	spends a bit more time processing target item makes a comment that indicates some processing of target item some level of cognitive effort to get meaning of target item	spends time processing target item provides an accurate translation of target item or finds a different way to say almost the same thing high level of cognitive effort to get meaning of target item

Operationalization of Depth of Processing (DOP): Grammatical Items

	Low depth of processing	Medium depth of processing	High depth of processing
Level of awareness	*Noticing*	*Reporting*	*± Understanding (based on accuracy of underlying rule or form- meaning connection)*
Description	Shows no potential for processing target form grammatically	Comments on target item in relation to grammatical features	Arrives at an inaccurate, partially or fully accurate target underlying grammatical rule
Descriptors	reads target quickly translates the phrase to English but leaves the target in Spanish carefully pronounces target item repeats target item	spends a bit more time processing target item makes comments that indicate some processing of target item some level of cognitive effort to process target	makes hypotheses regarding target item provides an inaccurate, accurate and/or partially accurate rule corrects previous translatio n spends much time processing target item high level of

Says s/he isn't sure what it is	item grammatically	cognitive effort to process target item grammatically
does not spend much time processing target		