

The Effect of Using Online Automated Feedback on Iraqi EFL Learners' Writings at University Level

Khaldoon Waleed Husam Al-Mofti

Department of English, College of Arts,
University of Anbar
khaldoonalmofti@gmail.com

Received: May 14, 2020 **Accepted:** June 14, 2020 **Online Published:** September 29, 2020

DOI: <http://doi.org/10.36231/coeduw/vol31no3.12>

Abstract

Feedback on students' assignments can be done in many different ways. Nowadays, the growing number of students at universities has increased the burden on the instructors to give feedback on students' writings quickly and efficiently. As such, new methods of modern online automated feedback tools, such as *Hemingway app* and *ecree*, are used to assist and help instructors. Hence, this research is an explanatory study to examine the effect of using the online automated feedback on some Iraqi EFL learners' writings at the university level. The study comprised 60 students enrolled in an English language course at the University of Anbar. They were divided randomly into two groups, experimental, and control with 30 students in each. Data were gathered through using pre and post-tests and a questionnaire. The statistical analysis of students' responses on tests showed that there was a significant improvement of the experimental group students' writings who received the intervention over the control group students who received the feedback traditionally. Besides, the questionnaire's data were analyzed both qualitatively and quantitatively. The obtained results also supported the findings reached from the tests and the students who were convinced and very satisfied with the online automated feedback. Therefore, in light of these results, the study highly supports the use and the integration of the online automated feedback tools to teach writing in EFL classrooms.

Keywords: online automated feedback, Iraqi EFL learners, students' perception, writing performance

أثر استخدام التغذية الراجعة التلقائية عبر الإنترنت على كتابات متعلمي اللغة الإنجليزية بوصفها لغة أجنبية للطلبة العراقيين في المستوى الجامعي

خلدون وليد حسام

قسم اللغة الانكليزية/ كلية الاداب
جامعة الانبار

khaldoonalmofti@gmail.com

المستخلص

يمكن إجراء التغذية الراجعة لواجبات الطلاب بعدة طرق مختلفة. حالياً، أدى العدد المتزايد من الطلاب في الجامعات إلى زيادة العبء على المدرسين لتقديم التغذية الراجعة حول كتابات الطلاب بسرعة وكفاءة. لذلك، يمكن استخدام طرق جديدة للتغذية الراجعة التلقائية الحديثة عبر الإنترنت، مثل تطبيق *Hemingway* و *ecree*، لمساعدة المدرسين ومساعدتهم. وبالتالي، فإن هذا البحث هو دراسة بيانية درست أثر استخدام التغذية الراجعة التلقائية عبر الإنترنت على بعض كتابات متعلمي اللغة الإنجليزية باعتبارها لغة أجنبية للطلبة العراقيين في المستوى الجامعي. تضمنت الدراسة 60 طالباً مسجلين في قسم اللغة الانكليزية في جامعة الانبار. تم تقسيمهم عشوائياً إلى مجموعتين، تجريبية، وضابطة. وبواقع 30 طالباً في كل

مجموعة. كما تم جمع البيانات من خلال اختبارين وأستبانة. أظهر التحليل الإحصائي لاختبار الطلاب أن هناك تحسناً كبيراً في كتابات طلاب المجموعة التجريبية الذين تلقوا التغذية الراجعة التلقائية عبر الإنترنت على حساب طلاب المجموعة الضابطة الذين تلقوا التغذية الراجعة بشكل تقليدي. إلى جانب ذلك ، تم تحليل بيانات الاستبانة بشكل نوعي وكمي. حيث دعمت نتائج الاستبانة النتائج التي تم الحصول عليها من خلال الاختبارات وكان الطلاب إيجابيين وكانوا راضين للغاية عن التغذية الراجعة التلقائية عبر الإنترنت. لذلك ، في ضوء هذه النتائج ، توصي الدراسة استخدام وأدراج أدوات التغذية الراجعة التلقائية عبر الإنترنت في تدريس مهارة الكتابة في الفصول الدراسية في اللغة الإنجليزية باعتبارها لغة أجنبية.

الكلمات المفتاحية: التغذية الراجعة التلقائية، متعلمي اللغة الإنجليزية باعتبارها لغة أجنبية للعراقيين، أدراك الطلاب، أداء كتابة الطلبة

1. Introduction

Online automated feedback (OAF), known as automated feedback, 'automated essay evaluation' (Hoang & Kunnan, 2016, p. 14) (AEE), or 'automated writing evaluation' (Huang & Renandya, 2018, p.6) (AWE) are applications and tools used in different areas to score or evaluate students' writings either in EFL/ ESL or L1 (First language) education (Hyland & Hyland, 2006; Lai, 2010; Ferris & Hedgcock, 2005; Truscott, 1996). These OAFs include different types and kinds of software and tools according to their purpose and application: such as, IntelliMetric, Criterion, ETS', e-rater, MY Access, Pearson's Intelligent Essay, Grammarly, Ginger, hemingway app, paper rater, and Write To Learn. These OAF applications are known for their wide applicability and significance as powerful learning tools (Cheng, 2017, p.24) that assist instructors to grade or evaluate learners' writings as proved by many studies (Warschauer & Ware, 2006). Moreover, they are known for their ability to reduce the burden over the instructors' shoulders, especially in a large number of classes (Yubing, 2016).

At Iraqi universities, the increasing number of students enrolled each year has enlarged the amount of pressure on instructors and university lecturers, providing educational guidance, advice, and feedback as a part of an assessment process that is crucial for learning (Van der Kleij et al, 2015). Teachers work laboriously and exhaustively to provide their students with written feedback on their assignments. As such, it becomes a daunting process for instructors, especially when it is done in a traditional manner. Consequently, it requires considerable time and effort to score or evaluate students' writings either in EFL/ ESL or L1 (First language) education (Choi, 2010). Thus, modern teaching systems voices represented by educationalists and researchers are calling for using modern technology tools into today's classrooms and especially in second language teaching (Huang & Renandya, 2018; Bai & Hu, 2017; Lawley, 2016; Cheung, 2016). It is known that in teaching EFL/ ESL writing, 'giving feedback is one of the most appropriate ways of instruction' (Samiei et al, 2017, p.108). Therefore, most recently, these OAFs have gained a lot of attention with a 'bigger role in writing assessment' (Hoang & Kunnan, 2016, p.1) because, in reality, they assist instructors in evaluating and giving feedback to the students' assignments. However, it still seems that their effectiveness and use in EFL contexts has not yet been proved on a large scale of research investigation in the Iraqi context.

Generally, some researchers would argue that automated feedback use and validity cannot be authorized without taking the perspectives of both instructors and learners and specifically by measuring its effect on students' writing development (Briody et al, 2013). It is known that students who receive immediate feedback may learn from their mistakes, and consequently, their writing improves. Contrarily, they may feel disappointed, demotivated, and dissatisfied with the feedback they receive (Swain, 2006). Therefore, including students' perception on this learning phenomenon through online technology is necessary for pedagogical improvements (Huang & Renandya, 2018), especially when teachers' feedback is considered demotivating as in the case of the Arab learners who also reported that it restrains their ability to edit and revise their writings (Ouahidi & Lamkhanter, 2020).

More specifically then, the use of OAF is something quite unfamiliar to the Iraqi EFL teachers and learners. Although, it is now widely reinforced and encouraged by many educationalists around the world (Yubing, 2016; Shermis & Burstein, 2003; Ware, 2005). Recently, the educational system in Iraq has emphasized the role of using modern technology in education, especially in teaching EFL, to keep pace with the global development taking place using technology in modern teaching (Salim & Ismail, 2018). However, it is not easy for the instructors to use these tools without any solid background regarding their usefulness in the Iraqi EFL teaching settings. As such, there is a need for valid studies in the field of using OAFs in learning and teaching writing in Iraqi setting as well as the Iraqi students' perception and reaction to their applicability in teaching.

1.1 Statement of the Problem

According to EFL teachers, the effect of using OAFs is subject to a debate regarding their usefulness to assist them to improve or assess students' writings by giving corrective or evaluative feedback. Therefore, there is a need to study their applicability, usefulness, and effectiveness in the Iraqi EFL classrooms, especially that using technology is something recently introduced and requires time to be fully integrated into modern teaching. As such, the present study aims to conduct a clear, concise, and informative study on an experimental group of Iraqi EFL learners' writings in an experimental study.

1.2 Objective of the Study

As stated, the current research investigates the effect of using automated feedback on Iraqi EFL learners' writings via an online platform; specifically, measuring the amount of progress on EFL students' writings after receiving automated feedback on their writing assignments. Moreover, the study also seeks to identify students' perceptions of this new kind of feedback. This, in return, either helps to support or negate the validity of using these applications in the Iraqi EFL classrooms. Also, the study aims to pave the way into integrating technological applications in teaching English a second or a foreign language.

1.3 Research Questions

In order to achieve the objective of this study, the following questions are raised:

1. What are the effects of using computer-generated feedback on Iraqi EFL learners' writings?
2. What are the perceptions of the Iraqi EFL learners of automated feedback during learning?
3. How can using technology, such as computer-generated feedback, be integrated into the teaching of a second or foreign language?

2. Review of Related Literature

2.1 Online Automated Feedback Tools

Technology for language learning or Language technology according to Silva (2010, p. 284) is 'a potential tool for the learning of a foreign language and has been integrated into classroom instruction for some time'. Language technology is the use of applications, software and other online tools to facilitate language learning. They are kinds of software that work by analyzing features in a text-based on comparing it to an already stored database of writings, having the same field or the genre (Hockly, 2018), and specifically, they 'evaluate an essay and promptly returns the results using artificial intelligence and several language analysis mechanisms' (Choi, 2010, p.38). So, they can easily analyze a large amount of data (corpus) in a very short time (Al-Mofti, 2015). Similarly, and more interestingly, they can be used in both summative and formative assessments (Al-Mofti, 2020) as well as in placement tests in language learning (Elliot et al, 2013; Hockly, 2019).

These OAF tools are generally divided into two main types based on their methods of analysis and purpose (Hoang & Kunnan, 2016). Some of them are specific to primary functions purely in the linguistic paradigms such as checking grammar, syntax, spelling, writing conventions of short texts, and essays. An example of these includes e-rater, and IntelliMetric. These kinds of tools' results rely on statistical estimates showing features of a text using calculations rather than general of good writing indicators. Therefore, they are called statistics-based programs.

Other groups of software are called the Rule-based or knowledge-based systems. These groups of software work on larger data and

features related to what is called natural language process methods. These include but are not limited to semantic association or analysis, characterizing word categories in sentences using machine-learning method, and also, showing genres of texts using special methods.

2.2 Previous Studies on OAF

There is no doubt that studies on automated online feedback tools are vast and have come up with many different results, conclusions, and interpretations about the applicability, usefulness, effects, and generalizability of conducting these tools in language learning (Ware, 2011). However, 'insufficient studies have been conducted to clarify to what extent this additional integration can benefit the learners (Huang & Renandya, 2018, p.2). For instance, the debate is still among scholars about their effectiveness in assisting instructors in grading, evaluating, and storing a huge number of written texts in a short amount of time.

Previous studies varied in their objectives, implementation, methods, and results. Some studies on the OAF effects and integration have emphasized that these tools and software are not like a human being in their grading and scoring and could not possibly detect mistakes or incorrect sentences (Tsuda, 2014). They can only be a kind of extra assistance to the instructors but not rather a substitute. So, assessment and grading should be done by instructors rather than by software because writing is a human act, and there is no correlation between how human react and software (Carr, 2014; Koskey & Shermis, 2013; Landauer, 2003). Moreover, other rejections about OAF are that the feedback itself is vague and sometimes incomprehensible to the students (Hegelheimer, 2015; Lai, 2010; Tsuda, 2014).

On the other hand, some other researchers proved the opposite. They claimed that these tools are accurate and precisely agreeing with human intervention and grading. They validated the work of these tools for assisting instructors, scoring, evaluating, and grading in a short time. The conclusions of these studies are based on practical studies with accurate results referring to the good correlation between the scores given by the software and the instructors. Also, these studies attributed their validities to the recent development in the natural language processing and technology, and also found good agreement and correlation between the OAF and the human ratings (Xi 2010; Jones 2006; Ben-Simon & Bennett, 2007).

Furthermore, other studies focused their attention on students' perception of OAF in L1 learning (Grimes and Warschauer, 2010; Lai, 2010). A study conducted by Chen and Cheng (2008) on 68 students investigated students' reflection towards using OAF software in grading and scoring written assignments over some time.

The study concluded that the students were concerned about the authenticity of OAF in using grading, and they expressed their dissatisfaction with using it in their written assignment assessment. Contrarily, another study by Fang (2010) showed different results on how students reacted towards the use of OAF in automated feedback. The sample of that study included 45 university students. They found that OAF software such as My Access was a very useful and

beneficial writing tool with its proofreading techniques. Similar to the findings of Fang's (2010), a study of Dikli and Bleyle (2014) stated that students were in favor of using OAF because they received feedback about their writings in terms of grammar, vocabulary choice, and mechanics. It is quite apparent that these studies recorded positive and negative aspects of using OAF in L1 writings (Dikli & Bleyle, 2014; Fang, 2010). However, those studies were mainly about L1 students (Chen & Cheng, 2006), and no studies combined investigating the effect and the students' perception as well as the integration of OAF in L2 writings in one study according to the previous studies reviewed in this study.

Studies on EFL learners' writings have received limited or no attention among scholars (Coniam, 2009). The limited knowledge about the effect of OAF and its integration in the process of giving feedback on EFL Iraq students' writings needs further research to approve its use and application in teaching writing or disapprove its effectiveness in evaluating and assessing writings. Also, it is hoped that the present study motivates EFL learners in writing practices and IL2 learning.

3. Methodology

3.1 Context

The present study was conducted in the context of a 16-week long course at the Department of English, College of Arts, at the University of Anbar. An academic writing course is mandatory for the first and second year students. The course is mainly examination-oriented with summative assessment only where students have 2-3 tests during the course and one exam at the end of the course. As such, the instructors are under pressure of grading, and scoring and giving summative assessment results because the learning process is test-driven. Usually, the assessment of the students' writings is based on the textbook's essay evaluation form and the error log chart in Zemach & Rumisek (2005, p.126-127). With the context of an academic writing course at universities, the workload over the teachers' shoulders makes the grading process mostly summative.

3.2 Participants

The participants of the current study comprised 60 sophomores students enrolled at the College of Arts, Department of English, at the University of Anbar. Their ages ranged between 19-21 years old. They were taught by the same instructor during the course time with the same kind of instructions and materials about writing. They were divided randomly into Experimental group (N=30) and control group (N30). The participants wrote 120 essays during the research intervention from two prompts (writing tasks) taken from Zemach & Rumisek (2005).

3.3 Research Instruments

To answer the questions of the current study, two instruments were used. The first one was writing task prompts and the second was a questionnaire to identify students' perception of the online automated feedback. Two writing prompts were conducted; one as a pretest and one as a posttest.

These writing tasks were modeled according to Zemach & Rumisek (2005) writing task and the students were familiar to the given topics. The questionnaire was adapted and inspired by Huang & Renandya (2018). Huang & Renandya (2018) were inspired by previous studies as in ARE (2011) and Xi (2010), and developed their questionnaire. The searcher adapted this questionnaire for its reliability, which was over 80, according to the Alpha coefficient, effectiveness, usefulness, and practicality.

It was used in this study to identify the perceptions of the students towards using automated feedback tools, like Hemingway and ecree, in their writing process. It had two

sections with a six-point Likert scale as was adapted from Huang & Renandya (2018) with some minor modifications to suit the current study. The first section seeks to identify:

- a. students' comprehensibility of the automated feedback,
- b. value, and usefulness of the feedback for revision,
- c. value, and effects of the feedback, and
- d. value of the peer review activity.

The second part of the questionnaire is an open-ended question that seeks to identify the students' feedback on the above-mentioned content areas. On the other hand, the two writing prompts were about two different topics; a problem /solution topic entitled "Air Pollution" and a comparison/contras topic entitled "Reading a story and seeing a film" taken from Zemach & Rumisek (2005, p.45).

3.4 Procedure

The experiment lasted 4 weeks. The students of the experiential and control groups were asked to write an essay in the pre-test prompt in the first phase of the study. In the second phase, the students in the experiential group were introduced to *Hemingway app* and *ecree* tools, and their writing assignments will be assessed using automated feedback, while the students in the control group were told that that their assignments are going to be evaluated according to essay evaluation form and error log chart mentioned in Zemach & Rumisek (2005, p.126-7). In the lab, the experiential group, during the third phase, were asked to write their essays in 45 minutes followed by 15 minutes copying what they wrote in *Hemingway app* and *ecree* to receive feedback. They were then asked to revise their essays according to the automated feedback they received from the two tools. During the other 15 minutes, the experimental group students were asked to do peer review activity. At the same time, students in the control group were asked to write their essays and conduct an anonymous peer-review process to review their peers' writing assignments traditionally, i.e. without using these two applications. In the last phase, students of both groups had to hand in their final draft. The experiential group also had to complete the questionnaire about their perceptions towards the online automated feedback. The following table summarizes the whole procedural process:

Table (1): Summary of the Research Procedural Process

	Experiential group	Control group
Phase 1	Pre-test writing Prompt (for writing performance).	Pre-test writing Prompt.
Phase 2	Students were introduced to Hemingway app and ecree tools.	-----
Phase 3	a. The 1st essay draft is written by the students. b. Anonymous peer review is conducted by students.	a. The 1st essay draft is written by the students b. Anonymous peer review is conducted by students.
Phase 4	a. Students revise their essays based on the feedback they receive from the online tools. b. Students submit the final draft of the essays. c. Students complete the questionnaire on their perception towards OAF.	a. Students revise the essays based on peer feedback. b. Students submit the final draft of the essays.

4. Data Analysis

In this section, the questions raised in this study were addressed. As for the first question, the essays collected from the students in phase 4 were analyzed and scored using 100-point grading scheme by following its descriptors as suggested by Jacobs et al. (1981). Then, both inferential and descriptive statistics were used to analyze the scores generated by the grading scheme. Next, independent-sample t-tests were used to identify (1) that both groups in the current study were of the same level in the pre-test and (2) to compare the two groups in the post-test scores. The purpose of the pre-test writing prompt was to ensure that the students of

both groups were at the same level of English proficiency. As for the second question, data obtained from the closed- ended items of the questionnaire were analyzed using SPSS in descriptive statistics. While the open questions of the questionnaire were coded into categories and subcategories and labeled using the notions of Jacobs et al. (1981). Then, they were analyzed using the “grounded theory” method of analysis.

4.1 Data Analysis and Discussion of Results

4.1.1 The Effects of using Online Automated Feedback on Iraqi EFL learners’ Writings

In fact, the first independent-samples t-test value of the pre-test has shown that both the experiential and the control groups are similar in terms of English proficiency, while the second independent-samples t-test value of the post-test showed statistically significant differences between the two groups as shown in table (3). The t-test was applied to measure each component of Jacobs’ et al. (1981) and identify any significance improvement in students’ writings during the automated feedback intervention. The following table shows the differences between the two groups in the pre-test writing prompts:

Table (2): The statistical values of the two groups in the pre-test writing prompt

The compared groups	The experiential N=30		The control N= 30		Statistical Values df
	M	SD	M	SD	
The scores results	5.78	1.34	5.44	1.26	55

Table 2 above illustrates the similarity of results between the two groups under investigation of the current study. The obtained results of (M) and (SD) values of the two groups shown were very close. Therefore, it means that they both have the same degree of proficiency in English writing to some extent. Consequently, this would support the results obtained from the post-test with the intervention to the experiential group, as shown in Table 3 below:

Table (3): The statistical values of the two groups in the post-test writing prompt

The compared groups	The experiential N=30		The control N= 30		Statistical Values df
	M	SD	M	SD	
Jacobs et al (1981) grading scheme					
Content	23.63	.94	16.98	1.30	55
Organization	18.5	1.76	13.32	1.11	55
Vocabulary	16.66	.43	14.63	.86	55
Language use	18.87	1.77	17.03	1.23	55
Mechanics	4.22	1.14	3.79	.92	55
Total	81.88	6.04	65.75	5.42	55

The results presented in table (3) above seemed to support the intervention of the online automated tools. The descriptive statistics indicate that the *Hemingway app* and *ecree* tools have an effective pedagogical outcome, and for the benefit of the experiential group. The total mean (M) value of the variables in the table (3) of each group shows statistically significant differences between the experiential group, who received the automated feedback as compared to that of the control group who studied in the traditional way. The significance is supported by the fact that the pre-test proved that the students of both groups had the same level (see table 2) before conducting the automated feedback intervention.

In term of the grading scheme of essay components, the mean (M) value of ‘content’ is the highest and this illustrates the obvious improvement that the experiential group students had, then followed by the ‘organization’ variable which its mean value is greater than ‘language use’. Thus, all these components had the highest significant statistical improvement

than those of the control group essay's components. This significance can be attributed to the following reasons:

- Both the *Hemingway app* and *ecree* tools provided the learners with feedback on the 'content' by highlighting the sentences that require improvement or change in the essay.
- *Hemingway app* tools presented information to the learners related to the language use and readability of the essay submitted by them.
- *Ecree* is more detailed in terms of feedback according to the sections of the essay and to Jacobs' et al (1981) scoring scheme. Therefore, it seemed that having feedback on the essay parts is very helpful in revising and improving the essay.
- *Ecree*, moreover, provided the writers with feedback on the components of each section of the essay. For example, for the introduction part of the essay, *ecree* gives detailed information about the components of this section. For instance, the presence of a thesis statement, supporting details and ideas, and even examples if they are found. This comprehensive feedback certainly draws the attention of the writers to what is missing in their essays and subsequently to have better revised and upgraded essays.

It is quite apparent that the online automated feedback has positive consequences on the students' performance in essay writing. Therefore, it can be said that providing more feedback on the length of an academic course may result in writing better essays. Guided automated feedback is a new way for EFL instructors to support and encourage their students to write and receive quick, reliable, and constructive feedback.

4.1.2 Students' Perception of Online Automated Feedback

Table (4) below is used to present the descriptive statistics obtained from the second part of the questionnaire (Likert-scale questionnaire), which contained (13) items. The questionnaire was distributed to the experiential group students at the end of the intervention to identify their perception and attitude about the OAF. The table is divided into four sections: the comprehensibility, usefulness, affects, and peer review value, based on the four involved criteria.

Table (4): The Experimental Group' Perception of Online Automated Feedback

Comprehensibility	No. of Item	Scale of agreement						Average response	Average
		Strongly disagree	Disagree	Somehow disagree	Somehow agree	Agree	Strongly agree		
	1	0	0	1	11	8	10	4.9	4.6
	2	0	0	5	13	10	2	4.3	
	3	0	0	2	11	9	8	4.7	
Usefulness	4	0	0	0	4	18	8	5.1	5.0
	5	0	0	0	14	7	9	4.8	
	6	0	0	0	7	10	13	5.2	
Effects	7	0	1	3	16	6	4	4.2	4.6
	8	0	0	0	7	15	8	5.3	
	9	0	2	0	6	13	9	4.7	
	10	0	0	0	19	6	5	4.5	
Peer review Value	11	0	4	2	9	10	5	4.3	3.9
	12	0	6	8	6	8	2	3.7	
	13	2	6	4	7	6	5	3.8	

The comprehensibility factor, the first part of table (4), presents the statistical results that are in favor of using online automated tools. The high frequency with 4.6 average in comparison with the students who disagreed with the understandability of the online feedback tools, supports the claim that these online tools are understandable. As for this factor, the first item has got the highest percentage as 10 students responded 'strongly agree' on that item. While the number of students who responded as "somehow disagree" was the highest on the second item as 5 students chose it. At the same time, "somehow agree" has got the highest number of responses as 13 students chose it when they responded to the second item. These descriptive results are highly supported by the students' answers in the part of the open-ended questions in the questionnaire where most of the students have a positive perception about the comprehensibility of the online automated feedback that they had received.

For the usefulness factor, the second part of table 4, the students' responses are high and in favor of the online automated feedback with the highest average score (5.0) among the other factors. There are no disagree responses by the students to the three items of this factor. As for item no.4, the number of responses to the option 'agree' was the highest as 18 students selected it. This item reads whether the students have any knowledge about these online tools. Thus, this means that a good number of the students already knew the value of the revision and feedback received from the online automated feedback. While (13) students strongly agreed on item 6 that reads whether the feedback was clear to the students. Certainly, the students would have more benefit if they have common sense about the value of the OAF as well as the comments and notes on their writings are clear and understandable to all the students.

The effect factor in the third part of the questionnaire shows similar results to those of the comprehensibility factor, with a 4.6 % average score. Very few students (3 students) have responded with 'disagree' to the usefulness and effect of the online automated tool. While 19 students have expressed their agreement to item 10 and to show that they are convinced with the effect and long term beneficial consequences of using online automated feedback on their writing performance. Moreover, 15 students have agreed that online automated feedback affects improving their grammar by having revision on the grammatical mistakes they have in their writings. Also, 13 students agreed on the statement in item (9) about increasing their vocabulary through the feedback they have received with multiple vocabulary options in the revision of the OAF.

Finally, the responses of the students on the fourth-factor 'peer review value' has got the lowest average score (3.9) among the other factors. . This could be attributed to the students' knowledge and exposure of peer review in their learning process in EFL writing classes where the students can only have feedback from their teachers and not from their peers. Therefore, and because of this lack of exposure to this kind of peer feedback, it could then most probably create conflict among the students themselves if they receive negative feedback from their classmates. The highest number of the students' disagreement is on the "Peer Review" part as 32 of them chose the last three items (11, 12, 13) of the questionnaire, among them 2 students have responded with strongly disagree to item 13.

According to the students' responses, it is concluded that almost 93 % of them are convinced because of the high value and usefulness of the online automated feedback. It is proved statistically with a high percentage on the first three factors. Therefore, they seemed to have a high perception of the effectiveness of using automated feedback on their writing performance.

4.2 The second part of the questionnaire – the open questions

The answers on the open questions of the questionnaire by the experiential group were analyzed using ‘grounded theory’ and coded into themes. They support the results found in both the writing prompts as well as the inferential results of the closed item questions of the questionnaire. The responses state clearly that the feedback received from the online automated tools is clear and understandable. As for the question related to whether the OAF is clear to them, 26 students out of 30 in the experiential group have answered positively on the clarity and comprehensibility of the OAF. In fact, only 4 students have responded negatively. One of the students’ answers on the first question states: “I had problems with the way the revision appeared on my screen because I am used to having the comments and revisions annotated to my paper and given directly by the teachers, this is something new to me”. The answer seemed against the use of OAF because the student needed to have more knowledge and exposure to this new form of revision in advance and before its implementation.

As for the second question, it seems that 25 students had a positive attitude to ‘what extent has the OAF improved their writings and the quality of the essays’. The students’ answers to this question were in favor of the OAF except for very few students. One of the students answered by saying: “That is our focus on the language form of the essay I was not paying attention to the meaning so in the feedback that we received only trying to organize the content of the essay”. Thus, with the progress of the students’ level, it is believed that they will have more focus on the meaning and structure of the whole essay. Having revisions from OAF could support their aim and objective towards writing an essay that is correct in terms of form and meaning.

In the context of the current study, the students are used to receive feedback from their teachers based on Zemach & Rumisek (2005) essay evaluation form, so having them received online machine feedback is something new to them. Therefore, 6 students out of 30 have answered the third question based on their earlier impression of the revision process to their writings. The question which reads whether they think that the OAF will improve their writing performance has got 6 responses that show they are unconvinced with the positive effect of the new intervention.

One of the students has reported that these OAFs should be introduced first to the students and they should be trained for a period of time because they lack knowledge of technology. On the other hand, other students’ answers are positive saying that it had a positive impact and improved their writing performance. Generally, almost 90 % percent of the students are enthusiastic and responded positively to the effect of the OAF on their writing performance and the good revisions they received on the final draft of their essays.

5. Conclusion

The current study has examined the effect of using OAF on Iraqi EFL learners’ writing performance at the university level. The study has also explored the EFL students’ perception of OAF tools, namely the *Hemingway app* and *ecree*. Findings show that the writing performance and the revised drafts of the experiential group have significantly improved after the intervention with better inferential and descriptive results. Moreover, over 93 % of the students under investigation have expressed their positive perception about the use of OAF to improve their writing performance.

6. Recommendations

The current study results have been obtained from only one intervention in the learning process of EFL learners. So, any generalizability cannot be done unless a longitudinal study is implemented with more OAF tools in similar contexts. In addition to that, the intervention

could be conducted for a complete course to allow the students to practice and learn more about the new methods of OAF. For example, a study may be conducted for three months including a large number of participants with different methods and strategies.

The current study recommends that instructors can use multiple methods to give feedback to their students by incorporating both immediate and online tools in order to improve students' writing performance. Moreover, instructors can have better results by using OAF to conduct revision, assessment, evaluation, and scoring through a complete course planning on when and how the intervention is being implemented. On the other hand, instructors can investigate their students' perceptions and feedback about the OAF to amend their methods of revising and giving feedback, especially the pedagogical implications that can be recorded with close observation and repeated evaluation.

References

- Al-Mofti, K. (2015). Lexical Semantic Representations of Psychological States of Positive and Negative Emotions in Modern American War Novels: A Corpus Linguistics based Approach. *International Journal of English Language, Literature and Humanities*, 3(3), 123-143
- Al-Mofti, K. (2020). "Challenges of Implementing Formative Assessment: by Iraqi EFL Instructors at University Level", *Koya University Journal of Humanities and Social Sciences*, 3(1), 181-189. doi: 10.14500/kujhss.v3n1y2020.
- Bai, L., & Hu, G. (2017). In the Face of Fallible AWE Feedback: How Do Students Respond? *Educational Psychology*. 37(1), 67–81.
- Briody, P., Wang, Y. & Shang, H. (2013). Exploring the Impact of Using Automated Writing Evaluation in English as a Foreign Language University Students' Writing. *Computer Assisted Language Learning* 26(3), 234–257.
- Carr, N. (2014). 'Computer-automated scoring of writing responses'. In A. J. Kunnan (Ed.), *The companion to language assessment*, (pp. 1063–1078). Malden, MA: Wiley.
- Chen, C., & Cheng, W. (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.
- Cheng, G. (2017). The impact of online automated feedback on students' reflective journal writing in an EFL course. *The Internet and Higher Education*, 34, 18-27.
- Cheung, Y. (2016). Feedback from Automated Essay Evaluation Systems: A Review of Selected Research. *TESL Reporter*, 48(2), 1–15.
- Choi, J. (2010). The Impact of Automated Essay Scoring (AES) For Improving English Language Learner's Essay Writing. (Ph.D. Dissertation), University of Virginia.
- Coniam, D. (2009). Experimenting with a computer essay-scoring program based on ESL student writing scripts. *ReCALL*, 21, 259–279.
- Dikli, S., & Bleyle, S. (2014). Automated Essay Scoring Feedback for Second Language Writers: How Does It Compare to Instructor Feedback? *Assessing Writing*, 22,1–17.
- Elliot, N., Gere, A. G., Gibson, C., Toth, C., & Presswood, A. (2013). Use and limitations of automated writing evaluation software. *WPA-CompPile Research Bibliographies*, 23, 1-26.
- Fang, Y. (2010). Perceptions of the Computer-Assisted Writing Program among EFL College Learners. *Educational Technology & Society*, 13(3), 246–256.
- Ferris, D. & Hedgcock, J. (2005). Teaching ESL Composition Purpose, Process, and Practice. New Jersey: Routledge.
- Hegelheimer, V., Dursun, A. & Li, Z. (2016). Automated Writing Evaluation in Language Teaching: Theory, Development, and Application. *CALICO JOURNAL*, 33(1), i–v.
- Hoang, G. & Kunnan, A. (2016). Automated Essay Evaluation for English Language Learners: A Case Study of MY Access. *Language Assessment Quarterly*, 13(4), 359-376.
- Hockly, N. (2018). *Technology for the Language Teacher: Automated writing evaluation*. Oxford: Oxford University Press.
- Huang, S. & Renandya, W. (2018). Exploring the integration of automated feedback among lower-proficiency EFL learners, *Innovation in Language Learning and Teaching*. Taylor Francis, 14(1), 15-26. DOI: 10.1080/17501229.2018.1471083
- Hyland, K., & Hyland, F. (2006). Feedback on Second Language Students' Writing. *Language Teaching*, 39(02), 83.
- Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). Testing ESL Composition: A Practical Approach. Rowley, MA: Newbury House.

- Jones, E. (2006). ACCUPLACER's essay-scoring technology: When reliability does not equal validity. In P. F. Ericsson, & R. H. Haswell (Eds.), *Machine scoring of student essays: Truth and consequences* (pp. 79–92). Logan, UT: Utah State University Press.
- Koskey, K., & Shermis, M. (2013). Scaling and norming for automated essay scoring. In: M. D. Shermis, & J. Burstein (Eds.). *Handbook of automated essay evaluation: Current applications and new directions* (pp. 199–220). New York, NY: Routledge
- Lai, Y. (2010). Which Do Students Prefer to Evaluate their Essays: Peers or Computer Program. *British Journal of Educational Technology*, 41 (3), 432–454.
- Landauer, T., Foltz, P., & Laham, D. (1998). The intelligent essay assessor: Applications to educational technology. *Interactive Multimedia Educational Journal of Computer-Enhanced Learning*, 1(2), 939-944.
- Lawley, J. (2016). Spelling: Computerised Feedback for Self-Correction. *Computer Assisted Language Learning*, 29(5), 868–880.
- Ouahidi, M. & Lamkhanter, F. (2020). Students' Perceptions about Teachers' Written Feedback on Writing in a Moroccan University Context. In: A. M. Ahmed & S. Troudi & Susan, R. (eds.). *Feedback in L2 English Writing in the Arab World*. Palgrave Macmillan.
- Salim, M. & T. Ismail (2018). An Insight on the Education Technology in Universities of Kurdistan Region in Iraq. Proceeding of a joint Scientific Conference between Soran University & PTI of Rwandz, Aynda & Akre.
- Samiei, M. & Rouhi, A. (2017). Efficacy of Written Corrective Feedback in the Short and Long Term. *Journal of Modern Languages*, 27, 107.
- Shermis, M., & Burstein, J. (Eds.). (2013). *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. New York: Routledge.
- Silva, I. (2010). Designing Computer-Generated Pedagogical Feedback for Spanish Students. *RESLA*, 23, 281-295.
- Swain, M. (2006). Linguaging, Agency and Collaboration in Advanced Language Proficiency. In: *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. (edited by H. Byrnes), 95–108. New York: Continuum.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369.
- Tsuda, N. (2014). Implementing Criterion (Automated Writing Evaluation) in Japanese College EFL Classes. *Language and Culture. The Journal of the Institute for Language and Culture*, 18, 25-45.
- Van der Kleij, M., Feskens, R. & Eggen, T. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, Month 201X, XX, (X), 1–37.
- Ware, P. (2011). Computer-Generated Feedback on Student Writing. *TESOL Quarterly*, 45(4), 769–774.
- Ware, P., & Warschauer, M. (2005). Electronic feedback and second language writing. In: K. Hyland & F. Hyland (Eds.). *Feedback and second language writing*. Cambridge: Cambridge University Press.
- Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, 10, 157–180.
- Xi, X. (2010). Automated Scoring and Feedback Systems: Where Are We and Where Are We Heading?. *Language Testing* 27(3), 291–300.
- Yubing, Q. (2016). Pigai Smart Essay Scoring System and Its Implications for Teaching English Writing. *Journal of Applied Science and Engineering Innovation*, 3(6), 217-219.

Zemach, D. & Rumisek, L. (2005). Academic Writing: From Paragraph to Essay. Oxford: Macmillan.