

The Effect of Oral vs. Written Feedback in EFL Writing

Emel Küçükali *

Marmara University, Istanbul, Turkey

Abstract

This mixed-method study aimed to investigate the effect of Oral feedback (OF) vs. Written feedback (WF) on overall L2 writing performance and on the relevant sub-scores: organization, lexis, grammar, and coherence & cohesion. EFL (English as a Foreign Language) university students (N=74) from English Preparatory Program at a state Turkish University were selected and then divided into comparison (N=35) and experimental group (N=39) through convenient sampling. While completing three writing tasks on an argumentative essay for four weeks, the comparison group was exposed to traditional WF and the experimental group was exposed to OF. The quantitative data was collected from the scores of the first and last writing papers and was analyzed through independent and paired samples t-tests. The qualitative data was collected from semi-structured interviews with three participants from the experimental group and was analyzed through open and structural coding. The empirical findings indicated a significant effect of OF on both the overall writing performance and on the sub-scores. The qualitative results showed that the opportunity for meaningful interaction, clarification, and negotiation between the students and the teacher during OF helped the students from the experimental group to receive more detailed and constructive feedback, and to significantly improve their writing skills.

Keywords: EFL writing, negotiation, oral feedback, teacher-student conferences, written feedback

INTRODUCTION

Statement of the Problem

The manifestations of feedback in L2 writing could be seen in several forms such as, 'teacher correction (with comments), error identification, commentary, teacher-student conference, peer correction, and self-correction' (Saito, 1994). From the versions above, the present study will focus on teacher-student conferences and, specifically, it aims to investigate both empirically and in-depth, the effect of Oral feedback vs. Written feedback on the overall writing performance of EFL university students and also on their writing performance in term of organization, lexis, grammar, and coherence & cohesion.

Theoretically, the dependent variable (DV) is defined by CEFR as the writing ability of a B2 level learner who “can write an essay or report that develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail. Can evaluate different ideas or solutions to a problem (Council of Europe, 2001). As for the writing performance on an argumentative essay (this essay type is participants’ writing task), it requires macro-skills, academic writing and genre knowledge such as taking a position on an issue, proposing a thesis, supporting it with facts and logic by considering both sides, describing alternative positions effectively and providing persuasive explanation for one’s position (Douglas Brown, 2004; Grabe & Kaplan, 2014). In the context of the present study DV or student’s performance on argumentative essay writing, will be defined as the score of their graded writing tasks based on the analytic writing rubric (see Appendix A) developed by the testing office department of the university.

The independent variable (IV) or feedback is defined as ‘specific descriptions and suggestions with a particular student’s work’ (Brookhart, 2008, p.1). The audience determines whether the feedback is individual or group, and modality classifies it into oral, written or visual. Good oral and written feedback need some decisions on what to say or write to the student such as, whether to focus on the work, process or student; whether to use criterion-referenced, norm-referenced or self-referenced feedback; whether to describe or evaluate; whether to give positive, negative or integrated comments; how much to clarify and specific to be; and how much to involve the student (Brookhart, 2008; Duke, 1975). Written feedback itself, is described as responding to student writing through written comments on content and organization, and strengths and weaknesses of essay (Grabe & Kaplan, 2014). On the other hand, oral feedback is defined as in-class conferences (5-10 minutes) with individual students, while the rest of the class is engaged in other activities; or out-of-class longer (15-30 minutes) conferences with individual students or groups (Brookhart, 2008; Grabe & Kaplan, 2014). Also, the oral feedback or conferences should follow planned procedures and stages of opening, student-initiated comments, teacher-initiated comments reading of the paper and closings (Reid, 1993). As for the feedback on an argumentative essay, it should include the requirements of rhetorical conventions and argument - evidence organization of the genre (Douglas Brown, 2004; Grabe & Kaplan, 2014). Because analytic scoring is suggested to have a positive washback effect on L2 academic writing (Douglas Brown, 2004), corrective feedback on grammar and vocabulary also should be integrated. Corrective feedback itself is defined as the form of responses to students’ errors in their performance, and these responses can vary in just identifying the error, providing the correct form in L2, or suppling with metalinguistic information about the error. (Ellis, Loewen, & Erlam, 2006). From the variations above, the teachers participating in the present study use direct corrective feedback (providing the correct form in L2) (Van Beuningen, De Jong, & Kuiken, 2012), that is why in this paper the oral and written feedback will be integrated with direct corrective feedback.

In the context of the present study the independent variable (IV), or the feedback on L2 writing of argumentative essay, will be defined as individual feedback given in class by

the teacher to a student. Written feedback (WF hereafter) consists of direct corrective feedback (DCF hereafter) on lexical and grammatical errors, analytic score and overall written commentary on the content and organization of the argumentative essay task. Oral feedback (OF hereafter) will be implemented as mini-individual conferences in-class, consisting of DCF, analytic score and oral comments given in the form of discussion by asking clarification questions and giving suggestions to the student.

Justification

According to Grabe and Kaplan (2014), the research on OF on writing papers consists mainly of case studies which evaluate teachers and students after conferences (Fassler, 1978; Patthey-Chavez & Ferris, 1997; Rose, 1982) or explore the nature of teacher-student talk and interaction during teacher-students conferences (Ewert, 2009; Freedman, 1981; Goldstain & Conrad 1990). The findings are in favor of OF as they indicate that students receive more effective feedback during conferences than they do through written feedback, especially when students actively participate and negotiate with their teacher. Several advantages of conferences for both students and teachers, have been reported in these case studies such as, the chance to discover each other through non-verbal language in a speech situation and thus see and solve the problems on the spot (Rose, 1982). Also, Fassler (1978) observed that personal contacts at the conferences made the grading more interesting for both the students and teachers, as the writer and reader can interact face-to face, and the student could witness how his/her paper is read, reacted to and evaluated. Other studies have displayed that the nature of the interaction may influence the effect of OF, that is, higher negotiation on single topics and focus on rhetoric organizations during conferences may result in students' participation and more intensive revisions (Ewert, 2009). On the other hand, when students did not negotiate meaning with their teacher during the conference, they did not tend to improve their writing skills qualitatively (Goldstain & Conrad 1990). Even in a study on L1 writing (Carnicelli, 1980) students viewed writing conferences as more beneficial than written feedback due to the interaction with the teacher, and the given opportunity for clarification and expressing themselves in conferences. However, the valuable research above tends to investigate the effect of OF on L2 writing mainly from qualitative perspective. Along with qualitative methods, this study will be an attempt to provide some empirical evidence on the issue, as well.

Significance

The present paper may help researchers and teachers to improve L2 learners' academic writing. This study also may add to the field by author's attempt to examine the effect of OF vs. WF on L2 writing through both qualitative and quantitative methods. The probable statistical value of the findings might give more support to OF as a part of formative assessment and a facilitator of cognition and motivation in learning, specifically, when there are interaction and negotiation with the teacher (Brookhart, 2008). This also leads to the association of OF with social constructivist approach of learning within social context (Roberts, 2016), with cooperative learning by teacher-

student interaction (Slavin, 1990) and with Vygotskian view of teachers' scaffolding of students during conference (Newman, Griffin & Cole, 1989).

Research Questions

The following research questions and hypothesis were posed:

1. What is the effect of OF vs. WF on the overall writing performance on an argumentative essay of B2 level EFL university students?
2. Is there a significant difference between the pre-test and post-test results of Oral (experimental) group in terms of organization, lexis, grammar, and coherence & cohesion?"
3. What are the reflections of students from Oral group on OF and WF after experiencing OF?

Research Hypotheses

H1: There is not a significant difference between the effect of OF and WF on the overall writing performance on an argumentative essay of B2 level EFL university students.

H2: There is not a significant difference between the pre-test and post-test results of the Oral group in terms of organization, lexis, grammar, and coherence & cohesion.

METHOD

Participants

The participants consist of B2 level university students (N=74) from English Preparatory Program at a state Turkish University. They were not randomly sampled but four classes were selected from 14 B2 level classes through convenient sampling (Cohen, Manion, & Morrison, 2013) considering students' age (18-24), attendance (regular), gender (46% female in experimental; 57% female in comparison group) and background (L1 Turkish). The students had been distributed in B2 level classes according to a placement test administered by the institution at the beginning of the semester. Two classes were randomly assigned to be the Written (comparison) (N=35) and two classes to be the Oral (experimental) group (N =39). For the interview session three students were selected on a voluntary basis from the Oral group.

Instrumentation

Three writing tasks on argumentative essay were administered to collect the quantitative data. Task 1 and 3 represented the pre- and post-test, respectively, while Task 2 was used in the intervention session and its grades were not used for the data collection. The tasks were graded according to the writing rubric developed by the testing office of the institution (See Appendix A). The tasks are parts of B2 level syllabus and their topics were decided on by the Curriculum Office of the institution. The topic of writing Task 1 was on sport, of Task 2 was on job descriptions, and of Task 3 was on

technology (See Appendix B for Task 1 samples). Qualitative data were collected from semi-structured individual interviews (Frankel & Wallen, 2009) with three students.

A checklist of strategies and procedures of giving OF and WF was used by the teachers while giving feedback to the participants (See Table 1 in 2.3. Procedure). The checklist in Table 1 is adapted from Brookhart's (2008, p.4-5) list of strategies of good feedback. Brookhart (2008) has built her strategies on the so-called double-barreled approach, which activates both cognitive and motivational factors simultaneously. While the cognitive factor gives diagnostic information to the students about where they are, the motivational factor fosters the feeling of autonomous learning. Also, Brookhart supports her strategies with constructivist learning or so-called self-regulation, which is learning by controlling one's own cognitive processes. From this constructivist perspective, the student is a self-agent, who filters teacher's feedback through his/her perception, prior knowledge, experiences, and motivation, and then decides where he/she is and sets her learning goals and needs accordingly. Both OF and WF strategies in Table 1 are based on the theoretical background above (constructivism) but OF has some unique and additional features because it is context dependent, the feedback is more student-specific and it is in private (Brookhart, 2008). Because there is a face to face interaction and negotiation in OF, it could be additionally supported by social constructivist approach to learning within a social context (Roberts, 2016). In the context of the study and in Table 1 OF is defined as in-class conferences (5-10 minutes) with individual students (Brookhart, 2008; Grabe & Kaplan, 2014).

Procedure

Data were collected in March and April 2017 in the classrooms of the participants. Mixed-method design with a combination of pre- post-test quasi-experimental and qualitative design was followed within procedures which lasted four weeks and encompassed four sessions.

Pre-Test Session: Pre-test was administered in week 1 to both groups and consisted of in-class writing task (Task 1) on an argumentative essay for 45 minutes. The writing papers were rated through double-check procedure according to a rubric developed by the institution. Overall five teachers from the institution were involved in the grading of the papers.

Treatment Session: The treatment session lasted two weeks. In week 2, individual feedback was given to both groups by their writing teachers on Task 1. However, while the Written (comparison) group followed the routine procedure of corrective plus written feedback, the Oral (experimental) group was given corrective plus oral feedback (See Table 1 below for OF and WF checklist, and Appendix B1/2 for sample papers with two types of feedback). In week 3, both groups completed a second, in-class writing task (Task 2) on an argumentative essay with a different topic, followed by the same feedback procedure of Task 1. So, in this session Written and Oral groups received written and oral feedback respectively, both on Task 1 and Task 2. Table 1 below

displays the procedures and strategies which the teachers followed in the feedback process:

Table 1. Checklist of strategies and procedures of giving written and oral feedback.

Checklist	Written Feedback	Oral Feedback
Timing	If possible daily or within a two-day time after the task completion	If possible immediately or within a two-day time after the task completion
Amount	Prioritize the most important points and relate to major learning goals	Prioritize the most important points and relate to major learning goals
Mode	Written/ One-paragraph written commentary at the end of the writing paper	Oral/ Mini-conferences in-class, lasting from 5 to 10 minutes
Audience	Individual	Individual, teacher-student individual talk ,oral discussion, negotiations
Content& Procedures	Give direct corrective feedback on lexical and grammatical errors + analytic score + written comment on content and organization requirements of the argumentative essay task	Give direct corrective feedback on lexical and grammatical errors & analytic score in the presence of the student + oral comment on content and organization requirements of the argumentative essay task
Focus	Describe both the work and the process	Describe both the work and the process
Comparison	Use norm- and criterion- referenced feedback	Use norm- and criterion- referenced feedback
Function	Describe. Don't judge	Describe. Don't judge
Valence	Integrate positive and negative comments	Integrate positive and negative comments
Clarity & Specificity	Be clear to the student & identify and correct errors	Be clear to the student & identify and correct errors
Tone	Choose words that cause students to think or wonder	Choose words that cause students to think or wonder

Note: Adapted from Brookhart, 2008, pp. 5-7.

Post-Test Session: In week 4, both groups took a post-test in the form of an in-class writing task (Task 3) on an argumentative essay with a different topic from Task 1 and 2. The writing papers of Task 3 were rated through double-check procedure according to a rubric developed by the institution. Overall five teachers from the institution were involved in the grading of the papers.

Interview Session: Several days after Task 3, an individual interview was conducted once, with each of three voluntary participants from the Oral group. The interviews lasted approximately 10 minutes each, were audio recorded and then transcribed. Interview questions were developed by the author and aimed to understand students' perceptions of OF and WF, the differences between them, and which one they had benefitted more and why (See Appendix C for interview questions).

Data Analysis

Inferential statistics were used to analyze the quantitative data. Independent samples t-test and paired samples t-tests (Urdu, 2016) were used to estimate the significance of OF vs. WF effect on writing performance. In addition, SPSS program version 23 was used for the calculations and statistical analysis of the study.

Qualitative data were analyzed through structural (Saldaña, 2013) and open coding (De Cuir-Gunby, Marshall & McCulloch, 2011) by two independent coders. While the first reading was assisted by open coding to elicit the key concept from the raw data, in the second reading structural coding was used to match the key concepts with the RQ and then to reframe them as codes. After two coders compared each other's codes to check consistency, the inconsistent and overlapping codes were negotiated and modified or added as sub-codes to larger codes, which after a unanimous decision resulted in the formation of six codes in the final codebook (see Appendix D for the codebook). Then, the codes were classified into three themes which respond to RQ3 of the present study.

Reliability

The trustworthiness of the current study was supported by inter-rater reliability of the writing scores and inter-coder reliability analysis. Spearman's rho correlation analysis (Urdu, 2016) was used to estimate the inter-rater reliability of the writing scores rated by two different instructors. The correlation coefficient was found to be statistically significant between the raters of the pre-test Written ($\rho=0.934$) and Oral group ($\rho=0.909$), and post-test Written ($\rho=0.856$) and Oral group ($\rho=0.820$) at a highly significant level of 0.01. For the reliability of the codes used for the interview analysis, inter-coder reliability analysis using the Kappa statistic was performed to determine consistency between two independent coders of the data (Landis & Koch, 1977). The inter-coder reliability for the initial 10 codes of interview data was found to be Kappa = 0.89 (Sig= 0.000; $p < 0.001$) which is a significant result and considered to be a substantial agreement between two coders (Viera, & Garrett, 2005).

RESULTS

The Findings of the Data Analysis Related to RQ1:

What is the effect of Oral Feedback (OF) versus traditional Written Feedback (WF) on the overall writing performance on an argumentative essay of B2 level EFL university students?

Paired samples t-test was conducted to compare the means between pre-test and post-test of the Written (comparison) and Oral (experimental) group. Paired samples t-test analysis in Table 2 indicated a statistically significant difference between pre- and post-test of the Oral group ($t = -9.087$, Sig= 0.000, $p < .05$) and non-significant result for the Written group ($t = .409$, Sig= .685, $p > .05$). Therefore, it could be claimed that OF had a significant effect vs. WF on students' overall writing performance in the Oral group.

Table 2. Paired samples t-test of overall writing performance in Oral and Written groups

Group	Test	Mean	Sd	t	df	Sig.(2-tailed)
Written	Pre-test	72,14	11,33	,409	34	,685
	Post-test	71,29	9,73			
Oral	Pre-test	69,04	10,49	-9,087	38	,000
	Post-test	82,95	8,49			

The results of the Independent samples test also were in favor of the Oral group. Table 3 below shows that there is a significant difference between the means of Written and Oral group for the post-test ($t = -5.509$, Sig= .000, $p < .05$), which means that the Oral group has outperformed the Written group in terms of overall writing performance after the treatment. On the other hand, Table 3 displayed a non-significant difference between the groups for the pre-test ($t = 1.224$, Sig= .225, $p > .05$), a result which reveals the homogeneity and supports the representativeness of the sample.

Table 3. Inter-group statistics (Independent samples test) of pre-test and post-test of overall writing performance

Test	Group	Mean	Sd	t	df	Sig.(2-tailed)
Pre-test	Written	72,14	11,33	1,224	72	,225
	Oral	69,04	10,49			
Post-test	Written	71,29	9,73	-5,509	72	,000
	Oral	82,95	8,49			

All in all, the results from Paired and Independent samples T-test allowed the Null hypothesis to be rejected and RQ1 to be answered: OF had a significant effect vs. WF on overall writing performance on an argumentative essay of B2 level EFL university students.

The Findings of the Data Analysis Related to RQ2:

Is there a significant difference between the pre-test and post-test results of Oral group in terms of organization, lexis, grammar, and coherence & cohesion?

The data analysis in the previous section revealed that there is a significant difference between the pre-test and post-test results of the Oral group in terms of the holistic score of the participants. Table 4 below illustrates the comparison of the analytic scores of pre-test and post-tests for the Oral group.

Table 4. Paired samples T-test of analytic score in experimental group

Group	Test	Mean	Sd	t	df	Sig.(2-tailed)
Oral	Organization Pre-test Organization Post-test	-3,21	3,71	-5,388	38	,000
Oral	Lexis Pre-test Lexis Post-test	-2,37	3,24	-4,569	38	,000
Oral	Grammar Pre-test Grammar Post-test	-3,27	3,25	-6,278	38	,000
Oral	Coherence & Cohesion Pre-test Post-test	-5,06	3,22	-9,827	38	,000

The paired samples t-test of the analytic score in Table 4 displayed four sub-scores rated according to the four bands of the writing rubric: organization, lexis, grammar, and coherence & cohesion. The analysis resulted in significant difference between pre-test and post-test results of Oral (experimental) group in terms of all sub-scores: organization ($t=-5.39$, $\text{Sig}=.000$, $p<.05$), lexis ($t=-4.57$, $\text{Sig}=.000$, $p<.05$), grammar ($t=-6.29$, $\text{Sig}=.000$, $p<.05$), and coherence & cohesion ($t=-9.83$, $\text{Sig}=.000$, $p<.05$). Therefore, the Null hypothesis can be rejected and it may be stated that OF had significantly improved not only the overall writing performance but also organization, vocabulary, grammar, and cohesion and coherence devices used to compose the argumentative essay.

The Findings of the Data Analysis Related to RQ3:

What are the reflections of students from Oral group on OF and WF after experiencing OF?

Data analysis of the interviews resulted in six codes which were classified under three main themes (See Appendix D). These themes display three main focusing points of participants' comments: 1) WF, 2) OF vs. WF, and 3) Why OF.

WF (Written feedback)

Reflections on WF were diverse. While one of the three participants appreciated WF as a beneficial one, the other two participants were more skeptical and had some criticism on it. The first participant (P1) referred to WF as effective for his writing performance because it has problem-solving, diagnostic nature and gives the chance to examine the feedback at home:

R (Researcher): What do you think is the role of the written feedback in this change?

P1: Yes, because when I go home I examine that (written feedback)... wow this is my problem and I don't have to do it the mistake in the next essay so that is fruitful and very good [...]... both (OF and WF) are very important too.

On the other hand, the other two participants (P2 and P3) criticized WF and referred to it as not so effective because according to P2 it is not clearly understood:

R: What do you think is the role of the written feedback in this change, improvement? Did it help or not?

P2: Not much, because when teacher take notes sometimes I can't understand. She uses some special signs. They are not very effective.

R: Sometimes these teachers write some comments. Do you understand them?

P2: Not very easy

P3 adds to her classmate's comments by stating that WF is unnecessary, meaningless because it lacks interaction and negotiation:

P3: I think it has not got any effect to me. I think written feedback is not important because they just give papers, I think it is not necessary for me. [...]. Yes, sometimes (I read the comments) but I just look at my score and don't think anything. [...] in written feedback they just give the paper and give it back.

Oral vs. Written feedback

When asked to compare OF and WF, all participants were unanimous as three of them appreciated OF and classified it as more beneficial than WF:

P1: Oral feedback is good as written feedback but oral feedback is more effective.

P2: ... oral feedback is more beneficial, ... compared with written feedback.

R: If you compare oral and written feedback, which one was more beneficial?

P3: definitely oral I think.

Why OF (Oral feedback)

When asked why they prefer OF to WF, the participants listed several benefits of OF: it is negotiation involved, problem-solving and meaningful. In terms of negotiation, participants refer to OF as more beneficial because of the involvement of interaction, personal communication, discussion, clarification and negotiation between the teacher and the student during the oral conferences:

P1: ...because when you speak with someone face to face communication is also effective than normal (WF).... and when teachers love you or who wants to solve your problems it can be effective.

P2: I can see my mistakes and sometimes I critic, discuss with the teacher.

P3: But with oral feedback I ask questions and I learned a lot.

In addition to negotiation, participants refer to OF as more beneficial because it is problem-solving, and allows the teacher to deal with their personal needs and individual problems in writing.

P1: ...and when teachers love you or who wants to solve your problems it can be effective ..., in my first essays I never used conjunctions, relative pronouns before, and my sentences were very short. But when I talked to my teacher she said to me “please use them, and then, use long sentences” and when I started to use them my points increased... I think face to face can solve everything.

P3: after the task we checked our mistakes with our teachers and they say that “this is wrong or this is correct” and after this I raised my scores. This have a very big role on me... I talked what I can write here and she (the teacher) said that “you can write this or this way” and I learned it.

Finally, participants found OF more effective because it is more meaningful, clear and easier to understand. During the conference teacher can clarify unclear points on the spot and raise participants’ awareness of their mistakes and their actual level of writing skills:

P2: ... oral feedback can be remembered easily, compared with written feedback [...]. I can see my mistakes and sometimes I critique, discuss with the teacher.

P3: I made a lot of spelling rules (mistakes) at first but after that teachers said this is wrong and I wrote some words in a week and I have to know that I should write it with capital....yes. I learned a lot of words in weeks ...

To summarize, in terms of WF participants had different attitudes as one of them appreciated it but the rest two participants did not benefit from it to the same extent because WF is sometimes meaningless and lacks face-to-face communication with the teacher. However, in terms of OF, all participants were unanimous in their opinion that OF is more beneficial than WF because OF compensates the drawbacks of WF by being more meaningful, problem-solving and involving negotiation and interaction with the teacher.

DISCUSSIONS

The quantitative findings of the present study indicated that OF had a significant effect vs. WF on the performance of argumentative essay writing. Similarly, interviews revealed that participants found OF more beneficial than WF. Both, empirical and qualitative findings add to the qualitative studies in the literature which indicate that students receive more effective feedback during conferences than they do through written feedback (Fassler, 1978; Patthey-Chavez & Ferris, 1997; Rose, 1982).

According to the interview data in this study, participants view OF as more beneficial because unlike WF, it involves negotiation and interaction with the teacher. This is in line with the qualitative findings in the research that students benefit more from OF when they actively participate and negotiate with their teacher (Ewert, 2009; Freedman, 1981; Goldstain & Conrad 1990). Specifically, the findings of Goldstain and Conrad (1990), that students who negotiated meaning during the conferences revised and improved their drafts more compared to the students who did not negotiate meaning, were reinforced by the results of the present study.

Another quantitative finding in this study is that participants exposed to OF significantly improved their writing performance also in terms of organization, vocabulary, grammar, and cohesion and coherence. On the other hand, interview data revealed that OF is preferred because it is meaningful, easier to understand, and deals with individual problems and personal needs of students in writing. It seems that when grammatical errors are corrected, and when comments on Lexis, content, and organization are given in the presence of the student during an individual talk, students may focus on separate segments easier, may elicit clarification and explanations from the teacher and may receive more detailed and meaningful feedback. This is consistent with Carnicelli (1980) findings on L1 writing that conferences are more beneficial than written feedback due to the interaction with the teacher, and the opportunity for clarification and expressing themselves in conferences. Other findings also support these benefits of individual and detailed OF such as, seeing and solving the problems on the spot (Rose, 1982), students witnessing how their paper is evaluated (Fassler, 1978), and the results that negotiation on single topics and focusing on rhetoric organizations during conferences may result in students' participation and more intensive revisions (Ewert, 2009).

The overall findings about the significant effect of OF on writing performance due to negotiation, meaningful interaction, and dealing with individual problems, are consistent with social constructivist approach to learning within social context (Roberts, 2016), with cooperative learning by teacher-student interaction (Slavin, 1990) and with Vygotskian view of teachers' scaffolding of students during conference (Newman, Griffin & Cole, 1989). All these approaches make meaningful communication and negotiation important and specific aspects of OF, and involve both teacher and student in the feedback and learning process.

CONCLUSIONS

The present study aimed to investigate the effect of OF vs. WF on the overall writing performance of EFL university students and also on their writing performance in term of organization, lexis, grammar, and coherence & cohesion. It also aimed to explore this effect in depth through the reflections of the participants on OF and WF after experiencing OF. The following RQs were posed:

RQ1: What is the effect of OF versus WF on the overall writing performance on an argumentative essay of B2 level EFL university students?

RQ2: Is there a significant difference between the pre-test and post-test results of Oral group in terms of organization, lexis, grammar, and coherence & cohesion?"

RQ3: What are the reflections of students from the Oral group on OF and WF after experiencing OF?

The empirical findings indicated that OF had a significant effect not only on the overall writing performance but also on the sub-scores according to the writing rubric: organization, lexis, grammar, and coherence & cohesion. The qualitative findings showed that the opportunity for meaningful interaction, clarification, discussion and negotiation between the students and the teacher during OF conference helped the students to receive more detailed and constructive feedback than the students in written group, who could not experience the negotiation process of OF.

IMPLICATIONS

The findings are consistent with the social constructivist perspective of learning and Vygotskian view of teachers' scaffolding (Newman, Griffin & Cole, 1989; Roberts, 2016; Slavin, 1990). Also, unlike previous research, the present study adds to this perspective of learning with both qualitative, and quantitative, statistically significant findings of the positive influence of OF on the overall writing performance and on its important segments such as organization, lexis, grammar, and coherence & cohesion.

APPLICATIONS

As for applications, the results suggest that schools and teachers give more support to and make use of OF as a part of formative assessment and a facilitator of cognition and motivation in learning, specifically, by providing interaction and negotiation between students and the teacher.

LIMITATIONS AND DELIMITATIONS

The limitations of this study could be a potential inconsistency between the teachers while giving WF commentary and writing only several commenting sentences but not a completed paragraph as required in WF checklist. One of the delimitations is using convenient not random sampling method of selecting the participants. Other delimitations are that two classes in the Written (comparison) group were given written feedback by different teachers (their own teachers), and only three writing tasks were included in the data collection process.

SUGGESTION FOR FURTHER RESEARCH

For further research, replication studies in different context, with larger samples, and longitudinal and cross-sectional designs, are suggested. Also, triangulation of instruments, such as questionnaires, observations, diaries and think-aloud methods are recommended. Another suggestion is, exploring OF with different levels of proficiency, nationalities, gender or in a combination of the peer, written, group and other types of feedback to test the interaction effect between multiple variables.

REFERENCES

- Brookhart, S. M. (2008). *How to give effective feedback to your students*. ASCD.
- Carnicelli, T. A. (1980). The writing conference: A one-to-one conversation. *Eight approaches to teaching composition*, 101-131.
- Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. Routledge.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- De Cuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. *Field methods*, 23(2), 136-155.
- Douglas Brown, H. (2004). *Language Assessment Principles and Classroom Practice*.
- Duke, C. R. (1975). The student-centered conference and the writing process. *The English Journal*, 64(9), 44-47.
- 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.
- Ewert, D. E. (2009). L2 writing conferences: Investigating teacher talk. *Journal of Second Language Writing*, 18(4), 251-269.
- Fassler, B. (1978). The red pen revisited: Teaching composition through student conferences. *College English*, 40(2), 186-190.
- Frankel, J., R. & Wallen, N., E. (2009). *How to Design and Evaluate Research in Education*. NY: McGraw-Hill.
- Freedman, S. W. (1981). *Evaluation in the Writing Conference: An Interactive Process*.
- Goldstein, L. M., & Conrad, S. M. (1990). Student input and negotiation of meaning in ESL writing conferences. *Tesol Quarterly*, 443-460.
- Grabe, W., & Kaplan, R. B. (2014). *Theory and practice of writing: An applied linguistic perspective*. Routledge.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school*. Cambridge University Press.
- Patthey-Chavez, G. G., & Ferris, D. R. (1997). Writing conferences and the weaving of multi-voiced texts in college composition. *Research in the Teaching of English*, 51-90.
- Reid, J. M. (1993). *Teaching ESL writing*. Englewood Cliffs, NJ: Regents.
- Roberts, J. (2016). *Language teacher education*. Routledge.
- Rose, A. (1982). Spoken versus written criticism of student writing: Some advantages of the conference method. *College Composition and Communication*, 33(3), 326-330.
- Saito, H. (1994). Teachers' practices and students' preferences for feedback on second language writing: A case study of adult ESL learners. *TESL Canada journal*, 11(2), 46-70.
- Saldaña, J. (2013). *The coding manual for qualitative researchers*. Sage.
- Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Prentice-Hall,
- Urdan, T. C. (2016). *Statistics in plain English*. Routledge.
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1-41.
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med*, 37(5), 360-363.

APPENDICES

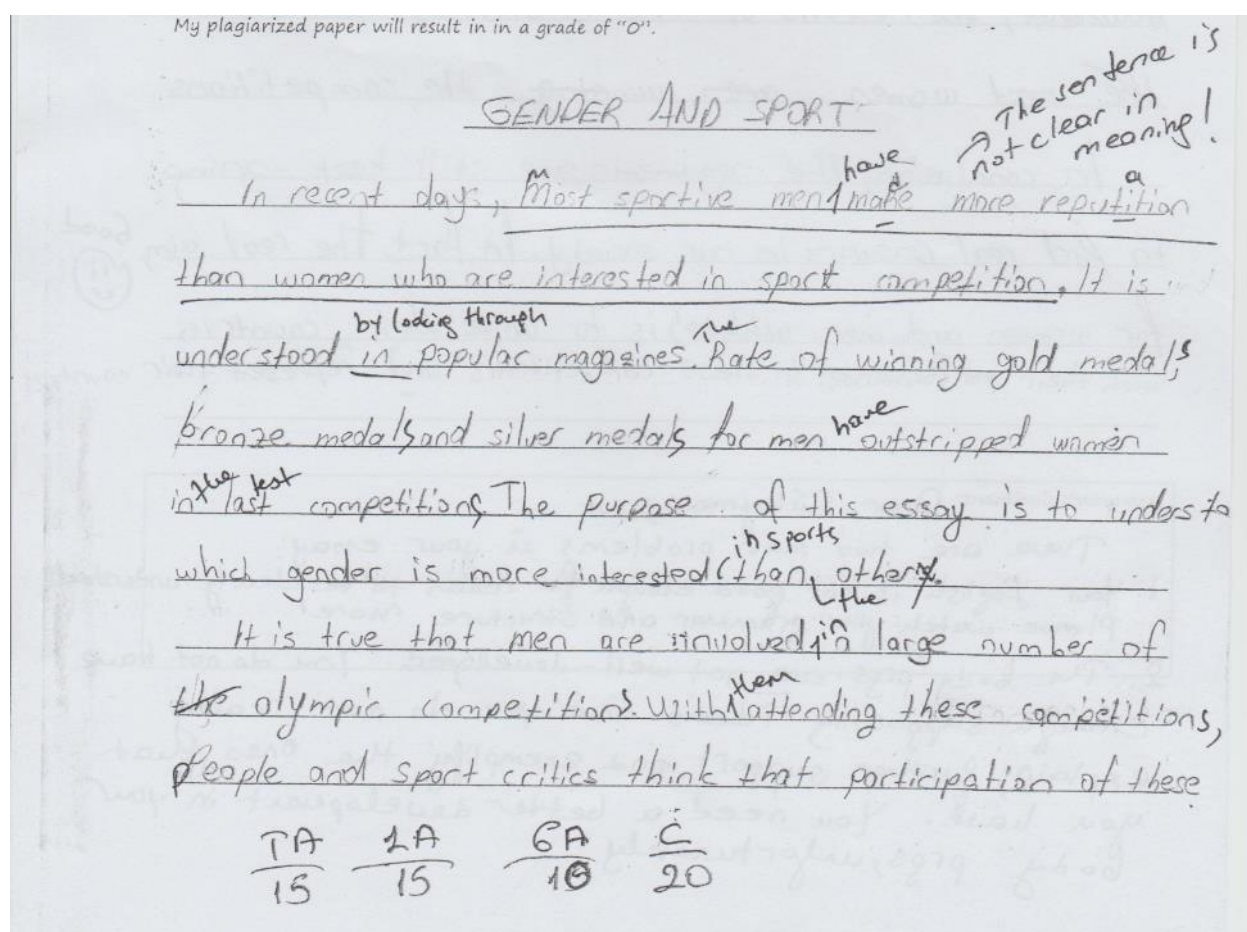
Appendix A. Writing rubric for B2 level developed by Testing Office of the institution

	TASK ACHIEVEMENT	LEXICAL RANGE & ACCURACY	GRAMMATICAL RANGE & ACCURACY	COHERENCE & COHESION
25	meets all task requirements including specified length, all content points fully dealt with; all ideas relevant to task, register and format consistently appropriate	uses a wide range of familiar vocabulary* hardly any errors in word choice and form. produces almost no errors in spelling	uses all grammatical structures* correctly few errors when complex forms are attempted. almost no errors in punctuation, capitalization	sequences information and ideas logically by using range of cohesive devices* uses referencing clearly and/or appropriately.
20	meets almost all task requirements including specified length all content points dealt with; ideas relevant to task register and format on the whole appropriate	uses a good range of familiar vocabulary* possibly some errors in word choice and form produces very few errors in spelling	uses almost all grammatical structures* correctly occasional errors when complex forms are attempted. produces very few errors in punctuation, capitalization	sequences information and ideas logically, which is incomplete at times uses cohesive devices* effectively, but cohesion within and/or between sentences may be faulty or mechanical
15	meets most of the task requirements most content points dealt with; some invalid or irrelevant ideas, and repetition reasonable, if not always successful, attempts to use appropriate register and format	uses a moderate range of vocabulary* occasional errors in word choice and form produces few errors in spelling	uses most grammatical structures* correctly some errors when complex forms are attempted. produces few errors in punctuation, capitalization	presents information with some organisation but there may be a lack of overall progression uses some cohesive devices*, but these may be inaccurate
10	meets some of the task requirements, some content points dealt with; most ideas invalid and/or repetitive, unsuccessful attempts to use appropriate register and format	uses a limited range of vocabulary* which may be used repetitively or which may be inappropriate for the task produces some errors in spelling	uses grammatical structures* with some errors, no attempt for more complex grammatical forms produces some errors in punctuation, capitalization	lacks logical sequencing ideas confused or disconnected, may use a very limited range of cohesive devices, and those used may not indicate a logical relationship between ideas
5	barely meets task requirements,	uses a very limited range of vocabulary*	uses grammatical structures* with	lacks logical sequencing

	very few content points dealt with; almost all ideas invalid and/ or repetitive, inappropriate register and format	too many errors which may impede understanding Produces frequent errors in spelling	frequent errors which may impede understanding produces frequent errors in punctuation, capitalization	ideas confused or disconnected, may use a very limited range of cohesive devices*, and those used may not indicate a logical relationship between ideas
1	does not meet task requirements incomprehensible	uses an extremely limited range of vocabulary* errors predominate and cause strain for the reader has almost no control of spelling	almost no mastery of sentence construction rules no sentence forms at all has almost no control of punctuation, capitalization	fails to communicate any message series of unrelated sentences

Appendix B. Task 1 Sample papers

Appendix B1. Task 1 Sample papers Written (Comparison) group




Bad English here!

competitions ^{are} more than women's participation. The reason why men are more active in competitions than women is that both

of these genders ~~is~~ about economical life style and where the two genders come from. Explain further please; continue ---

In contrast, Opponents of this idea trend like some women rights supporters advocated that ~~the~~ women might be ^{fewer} less in competitions than men.

However, the results of the istatistics demonstrated that the most women ~~are~~ winning the competitions

In conclusion, The arguments are still keep -going to find real answers in our society. In fact, the real aim, ^{Good}  for women and men athletes is to unite their countries with their performances in these competitions and represent their countries.

Instructor's Comment: Dear

There are two main problems in your essay:

1. Your English is not good enough for ideas to be clearly understood. Please watch your grammar and structure more!
2. The body prgs. are not well-developed. You do not have enough supporting ideas and you do not really explain, further support and exemplify the ones that you have. You need a better development in your body prgs, unfortunately.

Appendix B2.Task 1 Sample papers Oral (Experimental) group

Men And Women In Sport

Men are associated with tin of beers and chips sitting in front of the TV maybe at home or outdoors with their friend and watching a football match. Then women went to switch over to some serial or talk-show. It is generally considered that men are more interested in sport than women. In this essay we will discuss the reasons.

→ First of all, we can see that men are ^{dealing with} ~~cultivating~~ sport in daily life more than women. If we ^{are} looking back in the history we can see that most of sports are for men because they demand strength. In addition, it is believed that women can not be interested in sport.

because that is not for girls. To some people, interest in sport or war
in ^{war} our nature and it is so hard to break the rules between
men's and women's interested. ✓

Nevertheless, in the world people's interests and cultures ^{are} ~~is~~ changing.
Nowadays, women are ^{take over} having duties commonly to be doing by men
because of this the manner of thinking is also changing and women
are more interested ⁱⁿ ~~which are including~~ men's hobbies. Furthermore, some
of these hobbies are intended only for women as a women soccer league
so the rate of women interesting with sport is increasing. ✓

In conclusion, people interest in sport is growing. This growth is observed
especially among women. In spite of this, we may definitely say that men
are more interested in sport than women.

Instructor's Comment:

Appendix C: Interview Protocol

Questions for students (N=3) from experimental group (exposed to oral feedback):

1. Warm Up Questions (How are you? What is your department?)
2. Do you like writing (classes) in English? Why?
3. Do you think that your writing level/performance has changed since you started University Prep School? If, yes how?
4. What do you think is the reason for this change/improvement/no change?
5. What do you think is the role of the written feedback in this change? How do you think it has affected your writing performance? Give an example.
6. What do you think is the role of the oral feedback in this change? How do you think it has affected your writing performance? Give an example.
7. Could you compare written and oral feedback? Which one was more beneficial to you? Why?
8. Closing (This is the end of the interview. Do you have any questions? Thank you.)

Appendix D: Final codebook

Table D1

Final Codebook

Theme	Code	Definition	Example from the Text
Written Feedback	Code 1 Valuable	Participants appreciate WF and refer to it as beneficial to their writing performance (e.g., problem solving, diagnostic, gives the chance to examine the feedback at home)	<p><i>R: What do you think is the role of the written feedback in this change?</i></p> <p><i>P1: Yes, because when I go home I examine that (written feedback) wow this is my problem and I don't have to do it the mistake in the next essay so that is fruitful and very good [...]... both (OF and WF) are very important too.</i></p>
	Code 2 Problematic	Participants criticize WF and refer to it as not effective and clearly understood; as unnecessary, meaningless, lacking interaction and negotiation.	<p><i>R: What do you think is the role of the written feedback in this change, improvement? Did it help or not?</i></p> <p><i>P2: Not much, because when teacher take notes sometimes I can't understand. She uses some special signs. They are not very effective .</i></p> <p><i>R: sometimes these teachers write some comments. Do you understand them?</i></p> <p><i>P2: Not very easy</i></p> <p><i>P3: I think it has not got any effect to me. I think written feedback is not important because they just give papers, I think it is not necessary for me. [... Yes, sometimes (I read the comments) but I just look at my score and don't think anything. [...] in written feedback they just give the paper and give it back</i></p>
Oral vs.	Code 3	Participants appreciate	<i>P1: Oral feedback is good as written feedback but oral</i>

Written Feedback	More Beneficial	OF and refer to it as more beneficial than WF	<p><i>feedback is more effective.</i></p> <p><i>P2: ... oral feedback is more beneficial , ... compared with written feedback.</i></p> <p><i>R: If you compare oral and written feedback, which one was more beneficial?</i></p> <p><i>P3: definitely oral I think.</i></p>
Why Oral Feedback	Code 4 Negotiation	Participants refer to OF as more beneficial because of the involvement of interaction, personal communication, discussion, clarification and negotiation between the teacher and her students.	<p><i>P1: ...because when you speak with someone face to face communication is also effective than normal (WF).... and when teachers love you or who wants to solve your problems it can be effective.</i></p> <p><i>P2:I can see my mistakes and sometimes I critic discuss with the teacher.</i></p> <p><i>P3: But with oral feedback I ask questions and I learned a lot.</i></p>
Why Oral Feedback	Code 5 Problem Solving	Participants refer to OF as more beneficial because it is problem solving and deals with their personal needs and individual problems.	<p><i>P1: ...and when teachers love you or who wants to solve your problems it can be effective ..., in my first essays I never used conjunctions, relative pronouns before, and my sentences were very short. But when I talked to my teacher she said to me “please use them , and then, use long sentences” and when I started to use them my points increased... I think face to face can solve everything.</i></p> <p><i>P3: after the task we checked our mistakes with our teachers and they say that “this is wrong or this is correct” and after this I raised my scores. This have a very big role on me... I talked what I can write here and she (the teacher) said that “you can write this or this way” and I learned it.</i></p>
Why Oral Feedback	Code 6 Meaningful	Participants refer to OF as more beneficial because it is more meaningful, clear and easier to understand (e.g., teacher explains when not clear, participants’ awareness of their mistakes is raised)	<p><i>P2: ... oral feedback can be remembered, easily, compared with written feedback [...]. I can see my mistakes and sometimes I critic discuss with the teacher.</i></p> <p><i>P3: I made a lot of spelling rules (mistakes) at first but after that teachers said this is wrong and I wrote some words in a week and I have to know that I should write it with capital....yes. I learned a lot of words in weeks ...</i></p>
<p>OF = Oral Feedback</p> <p>P = Participant of the current study</p> <p>R = Researcher of the current study</p> <p>WF = Written Feedback</p>			