Research Paper
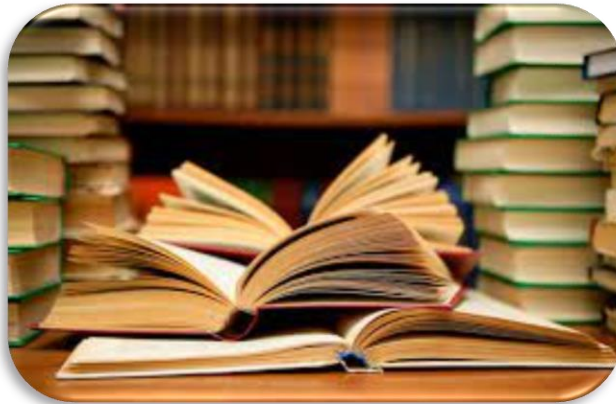
Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms: The Case of Receptive and Productive Types of Vocabulary Testing

Dariush Aliakbari[1], Morteza Aslrasouli[2*], Davud Kuhi[3]

[1]Ph.D. Candidate, English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, dariush.aliakbari22@gmail.com

[2]Ph.D., English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, Iran. mortazaaslrasuli@gmail.com

[3]Ph.D., English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, Iran. davudkuhi@iau-maragheh.ac.ir

ABSTRACT

The research on both receptive and productive vocabulary knowledge has attracted the attention of researchers, however, there is paucity of studies in these strategies on productive skills' enhancement. Hence, this study aimed at finding the effects of receptive and productive types of vocabulary testing on overall and lexical writing performance of Iranian EFL learners. Multiple-choice format of cloze tests and C-tests were selected as receptive and productive forms of vocabulary testing, respectively. For this purpose, 60 Iranian EFL learners with the age range of 20 to 27 at the intermediate level took part in this study. The data gathering tools were Oxford placement test, a writing pretest, a writing post-test, and a rubric. The students were randomly assigned to one control group and two experimental groups whose vocabulary knowledge was assessed using Cloze tests and C-tests. The results of one-way ANOVA showed that using both C-tests and multiple-choice format of Cloze tests was effective in enhancing learners' overall and lexical aspect of writing performance. The quantitative results approved by the semi-structured interviews from the C-test group that represented more positive attitudes towards the strategy. Analysis of the students' responses revealed that they found C-tests more effective in improving their active vocabulary knowledge that assisted them in writing better essays. The findings of this study might provide new insights for language teachers and curriculum designers to apply receptive and productive measure of vocabulary to improve different aspects of the writing ability.

Keywords: C-test; Cloze test; Receptive testing; Productive testing; Writing performance

تأثیر انواع آزمون‌های دریافتی و سازنده واژگان بر عملکرد نوشتاری کلی و واژگانی زبان‌آموزان ایرانی

تحقیق در مورد دانش واژگان دریافتی و مولد توجه محققین را به خود جلب کرده است، با این حال مطالعات کمی در این راهبردها در مورد افزایش مهارت های تولیدی وجود دارد. از این رو، این پژوهش با هدف یافتن تأثیر انواع آزمون‌های دریافتی و سازنده واژگان بر عملکرد نوشتاری کلی و واژگانی زبان‌آموزان ایرانی انجام شد. قالب‌های چندگزینه‌ای کلوز و آزمون سی به‌ترتیب به‌عنوان فرم‌های دریافتی و سازنده آزمون واژگان انتخاب شدند. برای این منظور ۶۰ زبان آموز ایرانی زبان انگلیسی بین ۲۰ تا ۲۷ سال در سطح متوسط در این پژوهش شرکت کردند. ابزار جمع‌آوری داده‌ها، آزمون تعیین جایگاه آکسفورد، پیش‌آزمون نوشتاری، پس‌آزمون نوشتاری و روبریک بود. دانش‌آموزان به‌طور تصادفی در یک گروه کنترل و دو گروه آزمایشی قرار گرفتند که دانش واژگان آنها با استفاده از آزمون‌های کلوز و آزمون C مورد سنجش قرار گرفت. نتایج تحلیل واریانس یک طرفه نشان داد که استفاده از آزمون‌های C و قالب چندگزینه‌ای آزمون‌های کلوز در تقویت جنبه کلی و واژگانی عملکرد نوشتاری زبان‌آموزان مؤثر است. نتایج کمی توسط مصاحبه های نیمه ساختاریافته از گروه C-test تایید شد که نشان دهنده نگرش مثبت تری نسبت به استراتژی بود. تجزیه و تحلیل پاسخ‌های دانش‌آموزان نشان داد که آنها تست‌های C را در بهبود دانش واژگان فعال آن‌ها مؤثرتر می‌دانستند که به آنها در نوشتن مقاله‌های بهتر کمک می‌کرد. یافته‌های این مطالعه ممکن است بینش‌های جدیدی را برای معلمان زبان و طراحان برنامه درسی فراهم کند تا معیارهای دریافتی و سازنده واژگان را برای بهبود جنبه‌های مختلف توانایی نوشتن به کار گیرند.

کلمات کلیدی : تست C، تست کلوز، تست دریافتی، تست تولیدی، عملکرد نوشتاری

INTRODUCTION

With more international interactions, transactions, and study abroad chances, the value of writing abilities is rising. The increased focus on writing skills in English as a Foreign Language courses and the present tendency for standardized exams like TOEFL and IELTS to include this skill in mandatory sections of tests (Karakoç & Köse, 2017) are proof of the increasing importance paid to this skill (Ebadijalal & Moradkhani, 2023). Some academics have distributed a variety of writing skill viewpoints, such as the idea of "Free composition" (Silva, 1993), which places a greater value on writing quality than on quantity or fluency. According to Yi (2009), this perspective is a descendant of structuralism and the bottom-up theory of processing. Writing thus appears as a secondary issue that serves to strengthen oral habits based on the fundamental tenet that the main channel of language is oral: speech is language... speech has an upper hand in language teaching (Richards & Rodgers, 1986). The method has behaviorist characteristics psychologically because it educates students to mimic, duplicate, and modify examples found in literature and taught by instructors (Hyland, 2019). Writing ability is described as the capacity to react to a stimulus in accordance with what an authority deems to be the appropriate reaction for instructors and scholars who hold the "Texts-as autonomous objects" viewpoint (Nunan, 1999). Discourse analysts, who have a quite different perspective, have since questioned this viewpoint. They contend that the placement of information in a sentence and the choice of grammatical forms depend on the discourse context or the setting in which the phrase is formed. There was a consciousness that there was more to the act of writing than constructing grammatical statements (Silva, 1993), which led to the introduction of the idea of texts-as-discourse (Hyland, 2002). In light of this, the purpose of this method is to impart the knowledge that writing is not merely a collection of unrelated sentences but rather includes linking them to create an integrated discourse. Consequently, the logical structure of writing is the main focus of this viewpoint. Students are instructed to look to uncover how authors employ structures of language choices to produce logical, meaningful writing (Hyland, 2002). The majority of current writing textbooks and programs still use this methodology.

In addition to methods to writing competence, studies acknowledged that teaching and evaluation methodologies play important roles in writing classrooms (Fathi & Rahimi, 2022; Lam, 2017). Employing various test formats, such as receptive and productive vocabulary testing kinds, appears to be one strategy that helps EFL learners improve their language proficiency. In reality, the majority of eminent academics concurred that vocabulary knowledge should be segmented according to how it is used in writing, reading, listening, and speaking. In order to measure the amount of knowledge of comprehension and production of lexis (Read, 2000), they separate the understanding of vocabulary into productive and receptive vocabulary (Schmitt, 2014). In contrast to productive knowledge, which is the linguistic proficiency needed to produce words while speaking and writing, receptive understanding is sometimes characterized as the capacity to receive words while listening and reading (Schmitt, 2010) .

Receptive and productive forms for test items are another option. For instance, while C-test and open-ended items are productive, matching and multiple-choice items are receptive (productive article). Receptive forms may support passive teaching and learning while productive formats promote more active instruction, which may result in a more engaged learning process (Read & Chapelle, 2001). These testing methodologies may have various washback effects.

35

Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms ...

Additionally, according to Ebadi and Rahimi (2019), aligning learning activities and evaluation procedures is a crucial aspect of language instruction that must change. Additional study can surely shed light on this alignment's murky areas. There is a dearth of research examining the impact of receptive and productive vocabulary testing on writing performance of EFL learners, despite the fact that a majority of research has been carried out on receptive and productive types of vocabulary testing in numerous settings and on different abilities and sub-skills (Gogolin et al., 2021; Hartono & Prima, 2021; Zhong, 2018). In order to address this, this study set out to compare the effects of receptive (multiple-choice items) and productive vocabulary testing (C-test questions) on writing performance in an EFL context similar to the one being studied. Moreover, the study was an attempt to explore the attitudes of EFL learners in a group that outperformed the other groups on the pros and cons of the strategy under the study to validate the results. So, based on the objectives of the study, the following research questions were formulated:

RQ1: Do receptive and productive forms of vocabulary testing have different effects on Iranian EFL students' overall writing performance?

RQ2: Do receptive and productive forms of vocabulary testing have different effects on Iranian EFL students' lexical aspect of writing performance?

RQ3: What are the students' perceptions about the use of cloze tests and C-tests as receptive and productive forms of vocabulary testing to improve their writing performance?

LITERATURE REVIEW

Receptive versus Productive Knowledge

Different facets of vocabulary knowledge are involved in recognizing a word (Nation, 2001). The receptive-productive scale, which can aid in estimating the degree of mastery of understanding and production of lexis, is one method of differentiating different forms of vocabulary knowledge (Schmitt, 2014). In contrast to productive knowledge, which is the linguistic proficiency needed to produce words while speaking and writing, receptive knowledge is sometimes characterized as the capacity to comprehend words when listening and reading (Fitzpatrick & Clenton, 2017). The degree of comprehension and lexis production proficiency can be assessed using the receptive-productive dimension, another subtype of vocabulary knowledge (Garnier & Schmitt, 2016; Schmitt, 2014). Three perspectives are suggested by Henriksen (1999, p. 304) to analyze vocabulary learning: a "partial-precise knowledge" perspective, a "depth of knowledge" perspective, and a "receptive-productive" perspective. He adds that the majority of academics concur that receptive and productive vocabulary are distinct. Receptive knowledge is sometimes defined as the ability to perceive words when listening to and reading, whereas productive knowledge is the linguistic competency required while producing words in speaking and writing (Schmitt, 2014). According to Melks Teichrow (1982, as cited in Pignot-Shahov, 2012), students pick up more vocabulary once receptive knowledge evolves into productive knowledge. A number of scholars agree on the existence of a receptive-productive spectrum, which typically suggests that vocabulary knowledge is organized in a hierarchy from passive detection to active recall.

Cloze Test versus C-test

Both in our daily life and the process of learning a language, testing is a common occurrence (Shirzadi & Amerian, 2020). Language examinations are given for a variety of reasons and have a big impact on people's social and professional lives, as McNamara (2000) demonstrates. Since they must assess their students, language teachers participate with language testing, which is also done for research objectives (Jin, 2023). The cloze procedure is one of the most widely used methods for evaluating test takers' proficiency in reading side of language. Its popularity has grown due to its reliability, substantial validity, and exceptional usability. Cloze tests may come in a variety of styles, but they all essentially consist of a reading passage with every nth word (often the seventh) blank, leaving the learner to either fill in the precise word or any suitable replacement in order to finish the passage (Sattar, 2022). Afterwards, cloze test is sometimes described as a text or paragraph with every nth word eliminated, which is of acceptable length and complexity. The definition appears to be pretty simple. Yet, specific security precautions should be implemented in relation to the ideas of suitable length, complexity, value, and so forth. The C-test method and cloze tests are theoretically connected. It makes use of several deletion procedures as well as additional criteria, in opposition to cloze test approaches. A universal language competence test with less repetition is called the C-test (Zhong, 2011). It is made up of five to six clear, accurate, and concise texts. The first and last sentences in these compositions are preserved unaltered. The "rule-of-two," which states that every other word beginning with the second word in the second sentence should be cut in half, causes damage to some parts of the text when used with the "C-test." This exam format is productive and context-dependent, requiring an understanding of the context and lexical signals, such as word relationships (Wang-Taylor & Milton, 2019).

The two primary cloze test types are the C-test and the Cloze-Elide test. These tests come in many different variations. The former consists of four to six brief texts that students may be familiar with, but in which the first and last sentences have been preserved while every second word—starting with the second word of the second sentence—has its second half deleted in place of a whole word (Jensen & Elbro, 2022). It is asserted that the C-test has some advantages over the cloze test since the range of texts allows for a better representation of subject matter, literary genres, and vocabularies. Additionally, since every other word is eliminated, there is a more representative sample of all linguistic components; scoring is simple and objective because there is typically just one right answer. Contrary to cloze tests, learners appear to enjoy doing C-tests, whether as a classroom assignment or as a more formal evaluation exercise. This positive feedback loop benefits students (Drackert & Timukova, 2020).

The cloze-elide test is the name of the later kind. This type adds a specific number of words to the text rather than eliminating a particular group of words from the passage (Esfandiari, 2019). The exam takers must read the text and mark any excess or unnecessary words by crossing them out. In reality, the processing of the text differs slightly in the cloze-elide test from the traditional cloze test (Chae & Shin, 2015). Despite the cloze procedure's popularity as a test, many language educators are unaware of its benefits and drawbacks across a wide range of subjects. The influence of various test kinds is one area where cloze process has, at least in part, been disregarded. Therefore, this paper is an attempt to provide its readers, especially language teachers, with some useful information about the same issue.

37

Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms ...

Empirical Studies

As it is clear crystal, in educational settings, tests are used for the aim of assessment, however, sometimes they are used as a tool for research. In the current study, receptive and productive types of vocabulary testing (cloze test and C-test) were employed as assessment strategy tools to investigate whether using different test types resulted in writing performance of EFL learners. The C-test is one of the most effective language tests for assessing vocabulary knowledge. The impact of C-tests on L2 achievement has been the subject of numerous research. For instance, Harsch and Hartig (2016) compared the embedded contextualized C-test format and the discrete decontextualized Yes/No vocabulary exam in order to establish whether format can account for the majority of variation in measures of reading and listening comprehension. The results showed that in both methodological methods, the C-test performed better than the Yes/No format. The contextualized aspect of the C-test, according to the researchers, appears to be able to clarify a significant percentage of the discrepancy in indicators of receptive language skills. Gogolin et al., (2021) recently concentrated on the growth of language proficiency in the foreign languages of French and English in Germany as determined by a C-Test. The use of both scores in their study allowed them to discriminate between the receptive and productive components of students' literacy development as adolescents and revealed that secondary school pupils are still working on mastering spelling in both of their foreign languages. The findings also suggest that at multiple phases of learning a foreign language, the receptive and productive elements of language ability exhibit different patterns of development.

The cloze test is regarded as an integrative test since candidates are required to utilize their total language skills to reconstruct the passage by looking at the relationships between phrases or within a sentence and then filling in the blanks with the correct responses. It is widely utilized by professionals and scholars equally as an indicator of overall proficiency for second language (L2) learners and has been shown in multiple studies to be an accurate and trustworthy instrument (Brown, 1983). According to the findings, learners who scored higher on the cloze test were more likely to compose lengthy sentences, t-units, and clauses, as well as an increased number of complicated nominalizations and verb phrases, a broader range of vocabulary, and a higher percentage of function words than those who scored lower on the test. The study's findings showed that, contrary to the commonly held opinion that the c-test is more effective than the cloze test, participants did better on the cloze test as a gauge of their reading comprehension.

Considering the significance of the problem, there are few empirical research that explore how various exam kinds affect learning. Additionally, there are inconsistencies in the findings of the earlier studies that spur further investigation. The current study, which attempted to investigate how multiple-choice cloze tests and C-tests, two types of vocabulary assessments, could enhance EFL learners' active understanding of vocabulary and, in turn, might impact their writing efficiency, sought to shed some light on the present problem. Because productive testing may improve the use of vocabulary in writing, which impacts the general level of the writing, this research made the assumption that receptive and productive forms of testing, determined via multiple-choice formats of cloze test and C-tests, accordingly, could be associated with the productive skill of writing. This also relates to the concept of washback, which was discussed above in relation to how testing might affect the standard of instruction. In order for learners

to create words that others can comprehend to express their thoughts and emotions; productive vocabulary should be viewed as an active word system (Webb, 2005).

METHOD

Participants

The purpose of this study was to investigate the effects of receptive and productive forms of testing measured through multiple-choice format of cloze test and C-tests on overall and lexical aspect of writing performance of EFL learners. Besides, the study explored the attitudes of the participants towards the treatment. To this end quasi-experimental research with the pretest-posttest control group was used. "A typical experimental study usually uses comparisons or control groups to investigate research questions" (Mackey & Gass, 2011, p.146). Moreover, this study used a sequential mixed method since it gathered the data both quantitatively (via pre-test, treatment, and post-test) and qualitatively (via interview). The study population consisted of all of the intermediate EFL students at one of the private language institutes of Maragheh, Iran. The sample of this study was a total of 75 intermediate-level Iranian EFL students (male and female) in the age range of 20-27. An Oxford placement Test and a writing test in the form of writing a 250-word essay were used to test and homogenize the participants' general English language proficiency and writing proficiency at the beginning of the course. Finally, 60 students were chosen based on their scores from the Oxford placement test and writing test and they were randomly divided into three groups of the receptive vocabulary group, productive vocabulary group, and traditional or control group. Each class had an equal number of students (N=20). The participants are homogeneous in age and language proficiency at the time of data collection. In addition, they share the same mother tongue, Turkish, but they are also fluent in Persian. The treatments were performed on these three groups to see whether multiple-choice cloze test or C-tests were more effective in the development of EFL learners' written performance by recalling their active vocabulary knowledge. In the control group, no use of testing as a learning form of instruction was used.

Instruments

Oxford Placement Test

A standardized test called the Oxford Placement Test (OPT) was selected to gauge language proficiency. The average administration duration for this 60-item test is 70 minutes. Those scores falling between X and 1 SD were chosen since the acquired scores did not significantly deviate from the normal distribution. The test's reliability was estimated and reported to be .81, showing the test items' high level of internal consistency.

Pretest and Posttest in Writing

Two descriptive writing tests were administered as pretest and posttest for all the three groups before and after the treatment. Both tests had similar topics (a favorite place in your city) which asked the students to describe something or someone. The students had to write at least 250 words for each test. The students were supposed to write a 250-words essay about the topic and to support their points of view by the reasons. The allotted time was 45 minutes. Upon completion of the treatment, a writing topic (a favorite place in your home) that was different from the pre-test was administered to all students in the groups.

Writing Rubric

The analytic writing scale developed by Jacobs et al. (1981) and revised by Polio (2013) was used to evaluate the students' writings on both pretest and posttest. It is worth noticing that five aspects of writing that are focused on in the rubric are content, organization, grammar, language use, and mechanics. Content section focuses on the logical development of ideas as well as categories for evaluating the main ideas, the supporting ideas, and examples. The organization section of the analytic scoring rubric evaluates the sequence of introduction, body and conclusion as well as the use of cohesive devices. The language section evaluates the choice of vocabulary, register, and grammar that is mainly at the sentential level. The mechanics including spelling, capitalization, indentation, and punctuation are also evaluated based on the mentioned rubric. Each component in this scale has separate scores. This scale was used in the study because it separated scores which provide the researcher with more useful diagnostic information and a more accurate picture of the individuals' writing skill. The essays were rated by two expert examiners who had been trained to use the rubrics and gave scores out of 100 for the five dimensions. The scores given by the two raters were compared to obtain inter-rater reliability. The coefficients are provided in Table 1 which indicates that reliability values were high.

Table 1

Inter-rater Reliability Coefficients for the Writing Test

	N	Cont.2	Org.2	Voc.2	L.U.2	Mech.2	O.W.2
Content1	60	.802**					
Organization1	60		.784**				
Vocabulary1	60			.786**			
Language use1	60				.583**		
Mechanics1	60					.911**	
Overall writing1	60						.941**

Note: 1 = rater one; 2 = rater two; **$p < .001$.

Cloze Test

Based on the words unfamiliar to the participants, a series of 6 multiple-choice cloze tests consisting of 20 items were designed to see EFL learners' active vocabulary development. The test was designed to see if the multiple-choice cloze test affects the first experimental group and how receptive testing could affect active vocabulary knowledge over time.

C-test

Based on the words unfamiliar to the participants, a series of 6 C-tests consisting of 20 items was designed for each productive vocabulary learning group session. To make the C-test, the rule-of-two was applied to construct items using passages based on the target words taught to both groups. The test was used to measure the development of EFL learners' active vocabulary knowledge in the second experimental group. The C-test has completion items formed by mutilating half of every other word

Islamic Republic of
Iran-Kerman

beginning from the second word of the second sentence. The first and last sentences are kept intact. The tests consisted of passages with five gaps that required active vocabulary knowledge of the students.

Semi-structured Interview

A semi-structured interview was the last tool utilized in the present study to examine the participants' perspectives toward the employment of the test type that outperformed the other groups in writing performance. Hence, to gain insights, the researcher encouraged students in the group to reflect on their overall perception of the instruction. This encompassed their preferences for the class and whether they would recommend it to others. Additionally, they were expected to provide rationales for their responses. Furthermore, participants were tasked with sharing their thoughts on the strengths and weaknesses of the instruction and offering suggestions for enhancements. Each student was interviewed individually, with their voices recorded for subsequent analysis and presentation. Notably, the students provided their responses in Farsi, which were then translated into English and included as part of the study's qualitative data. To estimate the credibility of the interview questions, the researcher presented the information gathered from the interviewees and asked them to pay attention to each question, and the answers were given themselves to see whether there was any problem with or differences in the answers. The participants confirmed the accuracy of the data, so the interview's credibility was approved. For dependability, 30% of the interview results were re-checked by two of the researcher's colleagues who were familiar with the data analysis section. The inter-rater reliability results were reported to be .83.

Procedures

At the initiation of the study and in adherence to a general research ethics code, the research received approval from the Academic Affairs department of the university's language department. This ensured that students were well-informed about the research's objectives and the confidentiality of their personal information. It was crucial to ensure that both the research institution and the participants were aware of their right to remain anonymous. In the initial phase, two weeks before commencing the treatment, an Oxford Placement Test (OPT) was administered to 75 EFL students to ascertain their homogeneity in terms of language proficiency. This test was used to determine whether the participants exhibited uniform language proficiency. Following the analysis of the proficiency test results, 15 students were excluded from the study due to their exceptionally high or low scores, leaving a total of 60 learners who were then divided into three groups: receptive vocabulary (N=20), productive vocabulary (N=20), and a control group (N=20). Subsequently, in this phase, the learners in the groups completed a descriptive writing pretest one week before the study began, with a designated time limit of 45 minutes. Following this pretest, the treatment phase commenced.

The treatment lasted for three months. This course was completed within 6 weeks, including two weekly sessions. This instructor as the researcher of the study instructed three classes. The instruction of writing in three groups was the same, however, the key distinction between the control and experimental groups was that the control group received no testing treatment and only followed the conventional teaching methodology of their textbook, whereas in the first experimental group, multiple-choice cloze tests were used throughout the treatment. In the second experimental group, C-tests were used as productive measure of vocabulary that could activate new words. In addition to a weekly vocabulary

41

Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms ...

exam, students were also required to compose a brief essay as a writing assignment. In the groups, the learners receive writing instruction in such a way that the students were required to write descriptively after that the teacher gave some information on the manner of writing including the main idea, supporting sentences, and concluding remarks. They received explicit written corrective feedback in their writing samples delivered to the instructor. In the groups, the instructor underlined the learners' writing errors and provided the correct form for them; and then, the corrected writings were given back to the learners in order to make them aware of their errors and the corrected form and to improve them in the following writing sessions. It is worth noting that in the groups, the teacher attracted the attention of the learners to the writing phases that are pre-writing, while writing, and post-writing phase using different strategies that are specific to these phases. For example, brainstorming technique was used for pre writing, and during-writing activities, the teacher tried to engage learners in recursive writing, self-editing and revisions. As the students were guided through writing and re-writing, post-writing activities helped learners reflect on and revise their writing based on feedback from an audience, such as peers and/or an instructor.

After completion of the treatment, both groups were given a 45-minute essay writing exam. The topic of the task was similar to writing pretest and the students had to write at least 250 words for that. The writing performances of both pretest and posttest were rated by two professors who used the analytic writing scale developed by Jacobs et al. (1981) and revised by Polio (2013). The framework is based on five components: content, organization, vocabulary, language use and mechanics. The essays were rated by two expert examiners who had been trained to use the rubrics and give scores out of 100 for each of the five dimensions. The scores given by the two raters were compared to obtain inter-rater reliability. The data were statistically analyzed by SPPS software (Version 22) to compare the scores obtained from the experimental and control groups to assess the progress in EFL learners' active vocabulary knowledge. To this end, independent samples ANOVA was run to probe the differences among the groups and check how receptive and productive testing types could affect the writing performance of the students.

To give more credit to the findings of this study, a semi-structured interview was also administered to know how students evaluate the effectiveness of the most effective method in improving their writing performance compared with the other groups. The researcher individually interviewed each student, capturing their voices for subsequent analysis and inclusion in the presentation. These interviews were conducted in Farsi, and the responses were translated into English and incorporated into the study's qualitative data. The central constructs were subjected to thematic analysis, and the findings were presented as selected excerpts.

RESULTS
Answer to the First Research Question
The descriptive statistics for the scores of the participants on OPT, writing pretest, overall writing score of posttests and its vocabulary component are provided in Table 2 which includes the maximum possible score, mean, standard deviation and range of their scores.

Table 2

Descriptive statistics for students' scores on each test and sub-test

Test	MPS	Range	Mean	SD	N
OPT	60	11	31.55	2.84	60
Writing Pretest	100	16	55.28	3.17	60
Overall Writing	100	39	69.01	10.50	60
Vocabulary	20	8.5	13.92	2.00	60

Note: MPS = Maximum possible score; SD = Standard deviation

To answer the first research question, a one-way ANOVA was run to compare the means of the three groups (See Table 3).

Table 3

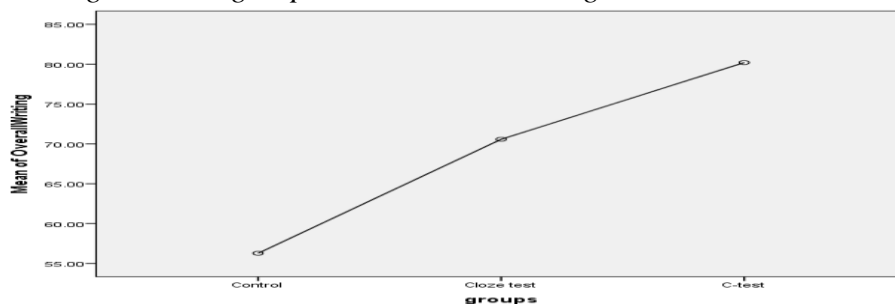*Results of One-way ANOVA for the Overall Writing Performance*

(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
Control	Cloze	-14.300	1.116	.000***
	C-test	-23.925	1.116	.000***
Cloze test (receptive)	Control	14.300	1.116	.000***
	C-test	-9.625	1.116	.000***
C-test (productive)	Control	23.925	1.116	.000***
	Cloze	9.625	1.116	.000***

*** $p < .001$

The results, as shown in Table 3, revealed that there were significant differences between the three groups. More specifically, the difference between the descriptive writing of the students in the control group and the receptive group using multiple-choice cloze tests was statistically significant ($p < .001$). Similarly, the differences between the descriptive writing of the students instructed with productive vocabulary testing and the students in the control group was statistically significant ($p < .001$). It is also evident that the effects of productive vocabulary testing were higher than receptive vocabulary testing ($p < .001$). The overall results, therefore, showed that using cloze tests and C-tests could positively affect the descriptive writing of Iranian EFL students. However, the effect of productive testing was more than the receptive type ($p < .001$). The means of the three groups and their differences are also illustrated in Figure 1.

Figure 1

Mean differences among the three groups in the overall writing score

43

Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms ...

Answer to the Second Research Question

To answer the second research question, another one-way ANOVA was run to compare the means of the three groups to see the difference among the groups on the lexical aspect of writing test (see Table 4).

Table 4

*Results of One-way ANOVA for the Lexical Aspect of Writing Performance*

(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.
Control	Cloze test	-1.900	.4464	.000***
	C-test	-3.500	.4464	.000***
Cloze test (receptive)	Control	1.900	.4464	.000***
	C-test	-1.600	.4464	.003***
C-test (productive)	Control	3.500	.4464	.000***
	Cloze test	1.600	.4464	.003***

*** $p < .001$

   In particular, the difference between the vocabulary aspect of descriptive writing of the students in the control group and the receptive group was statistically significant ($p < .001$). Additionally, the differences between the lexical component of descriptive writing of the students in the productive vocabulary testing group and the students in the control group was statistically significant ($p < .001$). It is also evident that the effects of productive vocabulary testing were higher than the receptive vocabulary testing ($p < .001$). The overall results revealed that using cloze tests and C-tests could positively affect the vocabulary use in descriptive writing of Iranian EFL students, but the effect of productive testing was more than the receptive type ($p < .001$. The means of the three groups and their differences are also illustrated in Figure 2.

Figure 2

Mean differences among the three groups in the vocabulary aspect of writing score
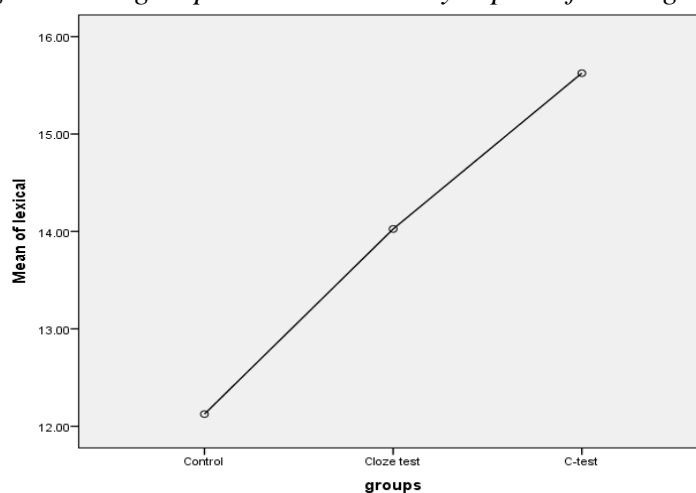
Answer to the Third Research Question

Based on the quantitative data analysis, the learners in productive vocabulary knowledge group outperformed the other groups in both lexical and overall writing performance. Hence, the learners in the C-test group interviewed to explore the merits and demerits of the strategy under study. For this reason, the researcher decided to conduct one by one interview from the participants in C-test group to investigate the themes and sub-themes of the used strategy in terms of both positive and negative points. The first question of the semi-structured interview was on the positive attitudes of the C-tests as productive vocabulary knowledge. The answers of the respondents along with the themes in the form of extracts were presented. Furthermore, the same themes (without extracts) are shown in Tables 5.

(Triggered background knowledge) Respondents 1, 2, 3, 4, 5,6,7, 8, 9, 10, 11, 13, 14, 16, 17 and 20, respectively.

Extract 1: The use of C-tests was influential in my writing as while writing, the vocabularies that I already learned came to my mind and increased the level of lexicon as well as the level of writing.

Extract 2: The use of C-tests triggered my past knowledge of learned vocabularies, so it increased my speed and consciousness in writing by removing my stress.

Extract 3: I was not afraid of my shortcomings or misunderstanding since the group ignored my problems.
(High inspiration) Respondents 2, 4, 5, 6, 7, 10, 13, 14, 15, 16, 18, and 19, respectively.

Extract 4: The instructor created a welcoming environment for learning via C-test that inspired the learners to discover the answers.

Extract 5: The teacher was very collaborative, although I depended on the strategy used in the classroom and it promoted high inspiration to write the text with high awareness.
(Class context) Respondents 2, 4, 7, 9, 11, 12, 13, 14, 17, and 20, respectively.

Extract 6: Active participation and happiness in the class context increased my motivation and decreased my anxiety level.

Extract 7: The fun nature of the class with novel strategy that was the use of C-test, which increased my vocabulary knowledge, all and all, were attractive for us.
(Self-assessment) Respondents 1, 2, 4, 7, 8, 9, 10, 11, 14, 15, 17, and 18.

Extract 8: Most of the reading strategies were practiced. Also, self-assessment and feedback should not be ignored while writing descriptive essays.

Extract 9: Self-evaluation approach was much more effective in this class may be due to the role of C-tests that repeatedly practiced. I had enough time to write carefully as it already had lots of new words

that I could retain them during writing.

It is worth noting that the mentioned positive points on the strategy was repeated by most of the students until the saturation happened. Table 5 shows summary of the positive attitudes on the strategy under the study that is productive vocabulary knowledge via C-tests.

Table 5
Summary of the Positive Attitudes on the Strategy

Items	Frequency	Percent
Background knowledge	16	80%
High inspiration	12	60%
Class context	10	50%
Self-assessment	12	60%

As the data of the Table 5 shows, 80% pointed out to the role of background knowledge that triggered via C-tests while writing essays. Furthermore, around half of the students mentioned to the other positive point that is class context that created fun atmosphere. Two plus points of high inspiration and self-assessment devoted 60% of the answers.

The learners moreover criticized some of the teaching processes and the strategy under the study, and they thought they had to be altered for better. These criticisms fall within the following categories. (Time wasting on small points) Respondents 5, 6, 7, 8, 9, 11, 13, 14, 15, and 20.

Extract 10: Prior to taking this course, I had some familiarity with C-tests and essay writing. However, one drawback of this form of instruction was that there were occasions when excessive time was devoted to addressing minor details.

Extract 11: We acquired novel and insightful strategies during the course. Nevertheless, the methodology and procedures used seemed inefficient and consumed more of my time. (High fatigue) Respondents 1, 2, 3, 4, 5, 6, 10, 12, 16, 17, 18, and 19.

Extract 12: Sometimes the class was boring for me as taking C-tests most of the times annoyed me.

Extract 13: The class context and repeated practice of C-tests bothered me as they resulted in high fatigue.

The mentioned negative points (similar to the positive ones) on the strategy were repeated by most of the students until the saturation happened. Table 6 shows summary of the negative attitudes on the productive vocabulary knowledge via C-tests.

Table 6

Summary of the Negative Attitudes on the Strategy

Items	Frequency	Percent
Time wasting on small points	10	50%
High fatigue	12	60%

As the above table shows, 60% of the participants mentioned high fatigue that rooted in repeated practice of C-tests. Also, half of the interviewees pointed out to the time wasting on method when they were exposed to productive vocabulary knowledge strategy.

As evident from the tables above, the interviewees shared their thoughts on the advantages and disadvantages of the course, including concerns about time being inefficiently utilized and a dull class environment. Nonetheless, the course served as a source of motivation due to its appealing nature. The strategy had novelty to the learners as C-tests were helpful in improving their active vocabulary knowledge and writing skills in terms of recalling vocabulary knowledge. They said that after performing C-tests throughout the treatment phase, memorizing words for writing was easier. The students also said that although answering the C-test was more difficult than the typical tests they take in classes, challenging, and time-consuming, they liked to be tested by C-tests because it is more dynamic, they feel more active while doing the C-tests, and the quality of their vocabulary knowledge was higher.

DISCUSSION

The present study aimed at investigating the effects of cloze tests and C-tests as receptive and productive forms of vocabulary assessment on Iranian EFL learners' descriptive writing ability as well as its lexical component. The results of quantitative data analysis showed significant differences among the three groups, indicating in particular that both receptive and productive testing of vocabulary could significantly affect the overall writing ability as well as its lexical component, but the effects of C-tests as productive measure of vocabulary was more significant. The results were confirmed by the learners who experienced productive vocabulary knowledge in writing courses as they have extra positive attitudes and less negative point of views towards the employment of C-tests in writing performance.

The findings are consistent with Harsch and Hartig's (2016) study, which revealed that the C-test is an excellent choice for placement and screening because it is a dependable, affordable, and rigorous assessment. They are capable of providing account for guessing and can be seen as measurements of vocabulary breadth and guessing variables, accordingly. Babaii and Jalali Moghadam's (2006) findings that the C-test necessitates macro-level processing and requires test takers to search for contextual indicators like lexical chains have been previously proven. As a result, the C-test's beneficial benefits might be explained by this kind of processing, which results in a greater degree of processing, active recall of vocabulary knowledge, and improved overall understanding (Chae & Shin 2015). To put it another way, productive testing could result in improved active recall of vocabulary knowledge. This was supported by Janebi Enayat and Derakhshan (2021), who demonstrated that productive rather than receptive vocabulary knowledge measures were more strongly correlated with L2 speaking ability.

The findings of this study are consistent with earlier research, which indicates that the development of productive skills is more influenced by productive vocabulary measurements. Janebi Enayat and

47

Mediating EFL Learners' Overall and Lexical Writing Skills in English Classrooms ...

Derakhshan (2021), for instance, looked at the relationship between L2 speaking proficiency and receptive and productive measures of vocabulary size. They discovered that both the lexical and general speaking abilities were improved by the productive evaluations of vocabulary knowledge. Additionally, semantic sets, which are found in C-tests (Janebi Enayat & Babaii, 2018), according to Atai and Dabbagh (2010), may aid EFL students in producing higher-quality compositions. In other words, the more fruitful and contextual nature of C-tests (Read & Chapelle, 2001) may aid EFL students in producing essays that are richer in vocabulary and more cohesive and coherent.

The findings of this study demonstrated that C-tests were successful in improving students' writing abilities on both an overall and lexical level. The findings differ from those of Chai et al.'s study from 2020, which found that cloze-based passages aid in measuring and extending the vocabulary growth of EFL learners. Their findings may suggest that providing some passages from cloze exams can be helpful in boosting EFL learners' active vocabulary knowledge. This result is consistent with Klçkaya's (2019) study, which found that participants used more receptive expertise than productive knowledge when answering multiple-choice questions.

In order to support the findings, Coxhead (2018) contends that participants' receptive vocabulary extent is greater than their productive vocabulary extent and that this explains why learners do better on the multiple-choice exam than on the c-test. The current study's findings are somewhat similar to those of Ajideh and Mozaffarzadeh (2012), who came to the conclusion that the multiple-choice cloze test is less effective than the C-test in measuring reading comprehension. However, the general finding was that the participants did more effectively on the cloze test. Additionally, the data analysis findings showed that, when compared to the other two experimental groups, the mean writing posttest scores for EFL learners in the control group were the lowest. One explanation for the control group's poor performance is rote learning, which involves memorizing a list of unfamiliar words. Rote learning makes it possible for students to memorize new words rather than learning them to understand them (Tayebi & Marefat, 2019). Because repetition without understanding led to short-term rather than long-term learning, this memory technique might not be appropriate for all students. Therefore, it is the role of the teachers to encourage, support, and direct the students to apply various vocabulary acquisition methodologies in order to acquire greater improvement and precise outcomes in the broad arena of vocabulary acquisition (Li & Cutting, 2011).

CONCLUSION

This study was an attempt to trigger EFL learners' receptive and productive vocabulary knowledge using C-tests and cloze tests on writing overall and lexical performance. Although the results of some studies mentioned above showed that C-tests are not effective in L2 achievement, actually the results of semi-structured interviews from the learners in C-test group approved the results of ANOVAs since the students in the C-test group expressed more positive perceptions towards the strategy and found productive testing as very effective in the development of their L2 writing ability due to the reasons such as enhancing background knowledge and fun nature of the class. In addition, the negative attitudes of the students can be a guide for those teachers who try to activate EFL learners' productive vocabulary awareness and help them to remove the barriers existed in using C-tests in writing courses. In sum, it can

be concluded that the using C-tests as a vocabulary improvement tool can have positive effects on the English writing development of EFL learners compared to the multiple-choice cloze tests.

The results of this study have repercussions for educators, students, test producers, and content creators. It is recommended that teachers employ a variety of vocabulary assessment procedures, such as the C-test and multiple-choice cloze tests, to assist students improve their active vocabulary and writing ability. However, C-tests are more strongly advised because they are more effective vocabulary exams. The results of this study can be used by teachers to decide which of these two strategies is more advantageous. Such exams might also be taken by students as a way to gradually advance their writing skills. The potential of such tests as alternate forms of assessment as well as useful resources and exercises to be utilized in textbooks and supplemental materials should be taken into consideration by test producers and materials writers.

This study had several limitations, like all the others, which may be fixed by additional research. First off, since just one level of language proficiency was examined in this study, it's likely that this research may be repeated with students who are more or less proficient in English as a foreign language. The second restriction is the sampling size, and results from bigger samples may be more broadly applicable. Thirdly, while this study only looked at descriptive writing, future research might examine other writing genres, such as argumentative or letter writing, as well as the productive skill of speaking, to see how different types of testing, such as receptive and productive forms of testing, might impact learning over time. Lastly, the duration of the treatment can be replicated by longitudinal studies that might reveal more reliable results.

REFERENCES

Ajideh, P., & Mozaffarzadeh, S. (2012). C-Test vs. Multiple-Choice Cloze Test as Tests of Reading Comprehension in Iranian EFL Context: Learners' Perspective. *English Language Teaching*, *5*(11), 143-150.

Atai, M. R., & Dabbagh, A. (2010). Exploring the role of vocabulary depth and semantic set in EFL learners' vocabulary use in writing. *Teaching English Language*, *4*(2), 27-49.

Babaii, E., & Moghaddam, M. J. (2006). On the interplay between test task difficulty and macro-level processing in the C-test. *System*, *34*(4), 586-600.

Brown, J. D. (1983). A closer look at cloze: Validity and reliability. In J. W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 237-250). Rowley, MA: Newbury House.

Chae, E. Y., & Shin, J. A. (2015). A study of a timed cloze test for evaluating L2 proficiency. *English Teaching*, *70*(3), 33-49.

Chai, Z. F., Swanto, S., & Din, W. A. (2020). Variants of Cloze-Test Based Tasks and Vocabulary Achievement. *Universal Journal of Educational Research*, *8*(7), 2980-2989.

Coxhead, A. (2018). Vocabulary assessment. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1-6). John Wiley & Sons, Inc.

Drackert, A., & Timukova, A. (2020). What does the analysis of C-test gaps tell us about the construct of a C-test? A comparison of foreign and heritage language learners' performance. *Language Testing*, *37*(1), 107-132.

Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, *32*(5–6), 527–555.

Ebadijalal, M., & Moradkhani, S. (2023). Impacts of computer-assisted collaborative writing, collaborative prewriting, and individual writing on EFL learners' performance and motivation. *Computer Assisted Language Learning*, 1-25.

Esfandiari, R. (2019). Cloze-Elide Test as an Alternative Test Method: Linking Personality Types to Test Method Performance. *Applied Research on English Language*, *8*(1), 1-26.

Fathi, J., & Rahimi, M. (2022). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: a case of EFL students. *Computer Assisted Language Learning*, *35*(7), 1668-1706.

Fitzpatrick, T., & Clenton, J. (2017). Making sense of learner performance on tests of productive vocabulary knowledge. *Tesol Quarterly, 51*(4), 844 - 867.

Garnier, M., & Schmitt, N. (2016). Picking up polysemous phrasal verbs: How many do learners know and what facilitates this knowledge? *System, 59*, 29-44.

Gogolin, I., Schnoor, B., & Usanova, I. (2021). Crossing the bridge to literacy in foreign languages: C-test as a measure of language development. *Multilingua*, *40*(6), 771-790.

Harsch, C., & Hartig, J. (2016). Comparing C-tests and Yes/No vocabulary size tests as predictors of receptive language skills. *Language testing*, *33*(4), 555-575.

Hartono, D. A., & Prima, S. A. B. (2021). The correlation between Indonesian university students' receptive vocabulary knowledge and their reading comprehension level. *Indonesian Journal of Applied Linguistics*, *11*(1), 21-29.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in second language acquisition*, *21*(2), 303-317.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics*, *34*(8), 1091-1112.

Hyland, K. (2019). *Second language writing*. Cambridge university press.

Janebi Enayat, M., & Haghighatpasand, M. (2019). Exploiting adventure video games for second language vocabulary recall: A mixed-methods study. *Innovation in Language Learning and Teaching*, *13*(1), 61-75.

Jensen, K. L., & Elbro, C. (2022). Clozing in on reading comprehension: a deep cloze test of global inference making. *Reading and Writing*, *35*(5), 1221-1237.

Jin, Y. (2023). Test-taker insights for language assessment policies and practices. *Language Testing*, *40*(1), 193-203.

Karakoç, D., & Köse, G. D. (2017). The impact of vocabulary knowledge on reading, writing and proficiency scores of EFL learners. *Journal of language and linguistic studies*, *13*(1), 352-378.

Kilickaya, F. (2019). Assessing L2 vocabulary through multiple-choice, matching, gap-fill, and word formation items. *Lublin Studies in Modern Languages and Literature*, *43*(3), 155-166.

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test1. *Language Testing*, *1*(2), 134-146.

Lam, R. (2017). Taking stock of portfolio assessment scholarship: From research to practice. *Assessing Writing*, *31*, 84–97.

Mackey, A., & Gass, S. M. (Eds.). (2011). *Research methods in second language acquisition: A practical guide* (Vol. 7). John Wiley & Sons.

McNamara, T. (2000). *Language testing*. Oxford University Press.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nunan, D. (1999). *Second Language Teaching & Learning*. Heinle & Heinle Publishers.

Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working Papers*, *4*(1), 37-45.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language testing*, *18*(1), 1-32.

Richards, J.C. & Rodgers, T.S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.

Sattar, A. (2022). Validation of the Cloze Test as an Overall Measure of English Language Proficiency among Iraqi EFL Learners. *North American Journal of Psychology*, *24*(1), 147-147.

Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In *Insights into non-native vocabulary teaching and learning* (pp. 28-40). Multilingual Matters.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language learning*, *64*(4), 913-951.

Shirzadi, D., & Amerian, M. (2020). Washback effects of multiple-choice, cloze and metalinguistic tests on EFL students writing. *Indonesian Journal of Applied Linguistics*, *9*(3), 536-544.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL quarterly*, *27*(4), 657-677.

Stoeckel, T., McLean, S., & Nation, P. (2021). Limitations of size and levels tests of written receptive vocabulary knowledge. *Studies in Second Language Acquisition*, *43*(1), 181-203.

Tayebi, G., & Marefat, S. (2019). The impact of rote learning on vocabulary learning: The case of Iranian EFL learners with visual and auditory learning styles. *Journal of Studies in Learning and Teaching English*, *8*(1), 129-145.

Teng, M., & Xu, J. (2022). Pushing vocabulary knowledge from receptive to productive mastery: Effects of task type and repetition frequency. *Language Teaching Research*, 13621688221077028.

Wang-Taylor, Y., & Milton, J. (2019). Predicting international students' academic success with vocabulary and intercultural communicative competence. *Journal of Applied Linguistics*, (32), 83-101.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in second language acquisition*, *27*(1), 33-52.

Yaseen, M. A., & Rasheed, W. A. (2022). Investigating EFL university students' written performance on cloze test. *Journal of Positive School Psychology*, *6*(10), 1273-1281.

Yi, J. Y. (2009). Defining Writing Ability for Classroom Writing Assessment in High Schools. *Journal of Pan-Pacific Association of Applied Linguistics*, *13*(1), 53-69.

Zhong, H. (2011). Learning a word: From receptive to productive vocabulary use. In *The Asian conference on language learning: Official conference proceedings*, 116-126.

Zhong, H. F. (2018). The relationship between receptive and productive vocabulary knowledge: A perspective from vocabulary use in sentence writing. *The Language Learning Journal*, *46*(4), 357-370.

Biodata

Dariush Aliakbari, Ph.D. Candidate, English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, Iran**.** Dariush Aliakbari received his B.A. and M.A. degrees in English Language Teaching from Islamic Azad University, Maragheh branch respectively. In 2003, he joined to the Ministry of Education in Maragheh. He has been teaching in high schools of Maragheh and Marefat Languages Institute for 20 years. His major area of research and interest is testing. He has had several publications in domestic and international journals.

Email:*dariush.aliakbari22@gmail.com*

Morteza Aslrasouli, Ph.D., Assistant professor, English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, Iran. Mortaza Aslrasouli received his B.A. and M.A. degrees in English Language Teaching from Islamic Azad University, Maragheh and Tabriz branches respectively. In 2011, he received his Ph.D. degree in Applied Linguistics from University of Hyderabad, India. He has been teaching in Islamic Azad University, Maragheh Branch, for 22 years. His major area of research and interest are ESP (English for Medical Purposes), Material Development, and testing. He has authored 'Bridging the Gap', an English text book for students of medical sciences. In addition to supervising over 50 MA and 3 Ph.D. students, he has had several publications in domestic and international journals.

Email: *mortazaaslrasuli@gmail.com*

Davud Kuhi, Ph.D., Assistant professor, English Language Department, Maragheh Branch, Islamic Azad University, Maragheh, Iran. Davud Kuhi is a full-time academic member of the department of English language at Islamic Azad University, Tabriz Branch. He has been researching on different properties of academic discourses for approximately two decades and has published a large number of articles in prestigious national and international academic journals. His major research interests are English for Academic Purposes, Academic Discourse Analysis, Academic Writing, and Academic Genre Studies.

Email: *davudkuhi@iau-maragheh.ac.ir*