LALEj

LANGUAGE AND LITERACY EDUCATION JOURNAL

2(1) (2026)

61 - 80

ENHANCING LEXICAL COMPETENCE: COMPARATIVE ANALYSIS OF PRODUCTIVE LEXICON ACQUISITION TASKS IN IRANIAN EFL LEARNERS

Elham Basirat* & Zuwati Hasim

Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia

elham.basiratesfahani@gmail.com

ABSTRACT

The development of productive vocabulary knowledge represents a persistent challenge for second-language learners, particularly when learners possess receptive knowledge of lexical items but have difficulty retrieving and using them productively. This study compared the effectiveness of three productive lexical activities—open-cloze, sentence translation, and sentence writing—in developing productive lexical competence among Iranian pre-university EFL learners. It also examined whether gender influenced learners' productive lexical performance across the three task conditions. A quasi-experimental pretest–posttest design was employed with 384 intermediate-level EFL learners, who were assigned by intact class to one of three task conditions. Participants' receptive and productive vocabulary knowledge was assessed before the intervention, and 20 lexical items known receptively but not productively were selected as the target items. Posttest productive vocabulary performance was analyzed using analysis of covariance (ANCOVA), controlling for initial productive vocabulary knowledge. In addition, participants responded to three open-ended questions concerning their experiences, perceived effectiveness, challenges, and future use of the lexical activities, and their responses were analyzed using thematic content analysis. The quantitative results revealed a significant overall effect of task type, $F(2, 380) = 6.063$, $p = .003$, $\eta p^2 = .031$. Sentence writing produced the highest adjusted posttest performance and significantly outperformed both open-cloze and sentence translation, whereas the difference between open-cloze and sentence translation was not statistically significant. Gender did not significantly influence productive lexical performance, $F(1, 377) = 0.238$, $p = .626$, $\eta p^2 = .001$, and the Gender × Task interaction was also nonsignificant, $F(2, 377) = 2.829$, $p = .060$, $\eta p^2 = .015$. The qualitative findings indicated that learners generally perceived the productive activities as useful for vocabulary development, confidence, and active lexical use, while also reporting differences in cognitive demands and individual learning experiences. Taken together, the findings suggest that productive task type is associated with differences in productive lexical performance, with sentence writing providing the strongest performance outcomes among the three activities under the conditions examined. The findings also indicate that these task effects did not vary significantly by gender.

Keywords: *Productive lexical competence, vocabulary learning tasks, sentence writing, sentence translation, open-cloze tasks, gender differences, EFL learners.*

INTRODUCTION

Lexical competence is widely acknowledged as a central component of second language (L2) proficiency because vocabulary knowledge provides the foundation for meaningful communication and successful language use. Although grammatical knowledge contributes to structural accuracy, insufficient lexical knowledge can considerably restrict learners' ability to express ideas, comprehend messages, and participate effectively in communicative contexts. Vocabulary knowledge is fundamental to learners' ability to express meanings and participate effectively in communication (Wilkins, 1972). Therefore, developing adequate lexical competence is essential for learners seeking communicative and academic success in a second or foreign language.

However, acquiring an L2 lexicon involves more than learning individual words' meanings. Lexical knowledge is a complex and multidimensional construct that includes knowledge of word form, meaning, and use (Nation, 2001). Knowledge of form involves recognizing and producing a word's spoken and written representations, whereas knowledge of meaning includes understanding its semantic associations and conceptual relationships. Knowledge of use further involves the ability to employ lexical items appropriately according to grammatical, collocational, and contextual constraints. Consequently, lexical competence requires not only recognition of vocabulary items but also the ability to access and use them accurately in meaningful communication.

A major distinction within lexical competence concerns receptive and productive vocabulary knowledge. Receptive lexical competence refers to the ability to recognize and understand lexical items during listening and reading, whereas productive lexical competence involves retrieving and producing lexical items in speaking and writing (Henriksen, 1999; Nation, 2001). Although receptive knowledge typically develops earlier and tends to be larger than productive knowledge, the transformation of receptive vocabulary knowledge into productive ability remains a persistent challenge for many L2 learners (Henriksen, 1999; Schmitt, 2019; Webb & Nation, 2017). Learners may recognize a word and understand its meaning when encountering it in a text, yet experience difficulty retrieving and using the same word when required to produce language independently.

This receptive–productive gap is particularly evident in foreign language learning contexts, where learners often encounter vocabulary through reading comprehension activities and explicit vocabulary instruction but have fewer opportunities to actively use lexical items in productive contexts. Traditional vocabulary instruction has frequently emphasized receptive knowledge because recognition-based activities are relatively easier to implement and assess. Nevertheless, knowing a word receptively does not necessarily guarantee the ability to retrieve and use it productively. Therefore, instructional approaches that require active lexical production are important for supporting productive lexical competence.

The characteristics of learning tasks may also be crucial for vocabulary development. Productive lexical tasks differ in the degree of contextual support and cognitive involvement they provide, which may influence learners' ability to retrieve and use newly acquired vocabulary. According to the Involvement Load Hypothesis (Laufer & Hulstijn, 2001), tasks involving greater levels of need, search, and evaluation encourage deeper processing and may result in stronger lexical learning. From a cognitive perspective, vocabulary tasks also differ in the amount of mental effort, contextual processing, and independent lexical generation they require. Therefore, comparing different productive vocabulary tasks can provide valuable insights into how instructional activities contribute to the development of productive lexical competence.

Three commonly used productive vocabulary activities are open-cloze tasks, sentence translation tasks, and sentence writing tasks. Open-cloze tasks require learners to retrieve appropriate lexical items based on contextual, semantic, and grammatical cues. Because these tasks provide partial linguistic support, they may facilitate vocabulary retrieval while still requiring active production. Sentence translation tasks require learners to transfer meaning from their first language into the target language and select appropriate L2 lexical items within a given context. Although such tasks provide

conceptual support through the source language, they require learners to establish connections between meaning and L2 lexical forms. In contrast, sentence writing tasks require learners to independently retrieve and appropriately use the target lexical items within self-generated sentences. This process places greater demands on lexical retrieval because learners must select appropriate words, construct grammatically acceptable sentences, and communicate intended meanings without direct linguistic prompts.

Despite the theoretical importance of productive vocabulary tasks, previous research has provided limited empirical evidence regarding the comparative effectiveness of different productive lexical activities. Many previous studies have investigated vocabulary size, receptive vocabulary development, or individual instructional techniques; however, fewer studies have directly compared multiple productive task types within the same experimental framework. Consequently, it remains unclear which productive tasks effectively facilitate the development of productive lexical competence.

Furthermore, learner-related variables may influence vocabulary learning outcomes. Among these variables, gender has received considerable attention in second language research; however, findings regarding gender differences in vocabulary learning remain inconsistent. While some studies have reported advantages for female learners, others have identified male advantages or no significant differences between genders. These conflicting findings suggest that further investigation is needed to determine whether gender influences the effectiveness of productive vocabulary learning tasks.

The present study addresses these gaps by examining the effectiveness of three productive lexical tasks—open-cloze, sentence translation, and sentence writing—in developing productive lexical competence among Iranian EFL learners. Additionally, the study investigates whether the effectiveness of these tasks differs according to learners' gender. By comparing different productive task types within the same instructional context, this study aims to provide empirical evidence regarding productive vocabulary development and offer pedagogical implications for EFL vocabulary instruction. The current study aims to answer the following research questions:

1. **RQ1.** Do open-cloze, sentence translation, and sentence writing tasks differ significantly in their effectiveness in developing productive lexical competence among Iranian intermediate EFL learners?
2. **RQ2.** Does gender significantly influence productive lexical competence among Iranian intermediate EFL learners?

LITERATURE REVIEW

Lexical Competence and the Receptive–Productive Distinction

Lexical competence is a multidimensional component of second language (L2) proficiency that extends beyond recognizing individual word meanings. Nation (2001) conceptualized vocabulary knowledge in terms of knowledge of word form, meaning, and use, while Henriksen (1999) proposed that lexical development can be viewed along several dimensions, including the movement from receptive knowledge toward productive use. This perspective highlights that knowing a word is not a single, dichotomous state; rather, learners may possess different degrees and dimensions of knowledge of a lexical item.

An important distinction within lexical competence is therefore that between receptive and productive vocabulary knowledge. Receptive knowledge concerns the ability to recognize and understand lexical items encountered in listening or reading, whereas productive knowledge involves retrieving and using lexical items in language production. Productive use consequently requires a greater degree of accessibility because learners must retrieve an appropriate lexical form rather than simply recognize it when it is encountered. Henriksen (1999) explicitly identifies the movement from receptive to productive use as a dimension of lexical development.

The distinction has important implications for vocabulary instruction. A learner may recognize a word

and understand its meaning but still struggle to retrieve it during production. Productive vocabulary development therefore cannot be assumed to follow automatically from receptive learning. Instead, learners need opportunities to retrieve lexical items and use them in contexts that require active production. This makes the transition from receptive knowledge to productive mastery an important area of investigation in L2 vocabulary research.

Recent research has specifically examined this transition. Teng and Xu (2025), for example, investigated how different task types and repetition frequencies could help learners move vocabulary knowledge from receptive knowledge toward productive mastery. Their study compared several vocabulary exercises, including gap-filling, sentence translation, and sentence writing, and explicitly focused on the development of productive mastery of words that learners already knew receptively.

This line of research provides a particularly relevant foundation for the present study because it shifts attention from vocabulary acquisition in general to the specific problem of making receptively known vocabulary available for productive use.

Productive Vocabulary Development in L2 Learning

Productive vocabulary development presents a particular challenge because learners must retrieve lexical forms accurately and integrate them into meaningful linguistic contexts. Whereas receptive vocabulary tasks can often be completed through recognition or comprehension, productive tasks require learners to access lexical knowledge and generate an appropriate response.

Vocabulary knowledge itself is multidimensional. Schmitt (2014) distinguishes between vocabulary breadth, referring broadly to the number of lexical items known, and vocabulary depth, referring to the richness of knowledge associated with individual words. Productive competence is particularly relevant to this broader conception of lexical knowledge because successful production requires learners to access not only a word's meaning but also its form and appropriate use.

The movement toward productive mastery may therefore depend on the amount and type of retrieval that learners experience. Research on vocabulary learning has repeatedly demonstrated the importance of active engagement with lexical items. In a meta-analysis of output tasks, Huang et al. (2012) found that learners who completed output tasks outperformed learners who only read texts, and the findings generally supported the importance of task involvement in vocabulary learning.

More specifically, Teng and Xu (2025) examined the effects of task type and repetition frequency on the development of productive mastery among EFL learners. Their study is particularly relevant to the present investigation because it selected words that learners already knew receptively and then examined how different vocabulary exercises supported their movement toward productive knowledge.

These findings suggest that productive vocabulary development should be considered not simply in terms of vocabulary size, but also in terms of learners' ability to retrieve previously encountered lexical items when required to produce language. Consequently, the characteristics of the task through which retrieval occurs may be an important determinant of productive vocabulary development.

Productive Lexical Tasks and Task Involvement

One influential framework for explaining differences among vocabulary learning tasks is the Involvement Load Hypothesis (ILH) proposed by Laufer and Hulstijn (2001). The hypothesis argues that the effectiveness of a vocabulary task is related to the degree of involvement it generates through three components: need, search, and evaluation. Tasks that require learners to engage more deeply with lexical items are expected to produce stronger vocabulary learning and retention. Hulstijn and Laufer's (2001) empirical work found greater vocabulary retention following a composition task than following fewer demanding activities, providing initial support for the role of task-induced involvement in vocabulary learning.

Subsequent research has provided broader empirical support for examining vocabulary learning through task characteristics. Huang et al. (2012), in a meta-analysis of output-based vocabulary learning studies, reported that learners completing output tasks generally demonstrated greater vocabulary gains than learners who only encountered the target words through reading. More recently, Yanagisawa and Webb (2021) synthesized 398 effect sizes from 42 studies and found that involvement load significantly predicted vocabulary learning, although its predictive power varied across learning conditions and test types.

Importantly, the ILH does not imply that all productive tasks are equivalent. Tasks may differ in the amount of contextual support they provide and in the extent to which learners must independently retrieve and use lexical items. A systematic review of research on the ILH identified a wide range of vocabulary task types, including sentence writing, fill-in-the-blank activities, translation, and composition writing, demonstrating the diversity of task formats used to promote vocabulary learning.

Within this framework, open-cloze tasks provide learners with contextual and linguistic information that can guide lexical selection while still requiring retrieval of the target word. Sentence translation tasks require learners to establish a connection between an intended meaning expressed in the first language and an appropriate lexical form in the target language. Sentence writing tasks, in contrast, require learners to generate their own sentence contexts and integrate the target lexical item into an independently constructed linguistic structure.

These differences are theoretically relevant to productive vocabulary development because the tasks vary in the degree of contextual support and independent lexical generation they require. Sentence writing, in particular, has been investigated as a potentially high-involvement vocabulary task. Hulstijn and Laufer (2001) found greater vocabulary retention following a composition-writing task than following less demanding vocabulary-learning activities.

However, task involvement should not be treated as a simple linear scale in which a supposedly more demanding task must always produce greater learning. The meta-analysis by Yanagisawa and Webb (2021) showed that the predictive contribution of involvement load is influenced by factors such as test format and frequency of encounters or use. Therefore, empirical comparison of specific productive tasks remains important rather than assuming their relative effectiveness solely from theoretical task rankings.

Comparative Research on Productive Vocabulary Tasks

Although vocabulary research has extensively examined individual task types, comparatively fewer studies have directly investigated how different productive tasks contribute to the development of productive lexical competence from an existing base of receptive knowledge. This distinction is important because a task that promotes general vocabulary learning does not necessarily have the same effect on learners' ability to retrieve words productively.

A particularly relevant study is Teng and Xu (2025), who investigated the effects of task type and repetition frequency on moving receptive word knowledge toward productive mastery. Their study included several vocabulary exercises, including gap-filling, sentence translation, and sentence writing, and focused specifically on words that learners already knew receptively. This provides a relevant empirical basis for examining how different task types may facilitate the transition from receptive knowledge to productive use.

The present study extends this line of research by directly comparing three productive lexical tasks—open-cloze, sentence translation, and sentence writing—within the same instructional framework. The three tasks differ in the amount of contextual support and independent lexical generation they provide. Open-cloze tasks provide sentence-level contextual and grammatical information while requiring learners to retrieve an appropriate lexical item. Sentence translation provides a first-language meaning-based prompt while requiring learners to retrieve and produce the corresponding L2 lexical form. Sentence writing provides less direct linguistic support and requires learners to generate a

sentence while retrieving and appropriately using the target lexical item.

These differences make the three tasks particularly relevant for investigating productive lexical development. Rather than assuming that one task must be more effective on the basis of its presumed cognitive demands, the present study empirically compares their effectiveness in developing productive lexical competence.

However, existing research does not provide a definitive hierarchy among open-cloze, sentence translation, and sentence writing specifically for transforming receptively known lexical items into productive vocabulary knowledge. Differences in participants, target vocabulary, task design, learning duration, repetition frequency, and assessment format make direct conclusions difficult.

The present study therefore addresses this gap by comparing the three productive lexical tasks using the same target lexical items and learner population. By focusing specifically on lexical items that learners knew receptively but had not yet mastered productively, the study examines whether the type of productive task influences subsequent productive lexical performance.

Gender Differences in L2 Vocabulary Learning

In addition to task characteristics, learner-related variables may contribute to variation in vocabulary learning outcomes. Gender has been examined in research on vocabulary learning, particularly in relation to vocabulary learning strategies and vocabulary knowledge. However, findings have not consistently demonstrated an advantage for either male or female learners.

For example, Lee (2019), in a study of 1,124 EFL high school students, found no significant gender difference in vocabulary knowledge, although significant differences were observed in the vocabulary learning strategies reported by male and female learners. The findings led the author to argue that assumptions about female superiority in vocabulary learning should be reconsidered across different learning contexts.

Other research has reported gender differences in vocabulary learning strategies usage. For example, Okyar (2021) found that female Turkish EFL learners reported significantly greater use of several vocabulary learning strategies than male learners, although the pattern varied across strategy categories. Such findings suggest that gender differences in vocabulary learning may be related more closely to particular learning behaviors or strategies than to vocabulary achievement itself.

The inconsistency across studies indicates that gender should not be assumed to have a uniform effect on vocabulary learning. Rather, its possible influence may depend on the learning context, learner characteristics, instructional approach, and type of vocabulary outcome being measured. This issue is particularly relevant to the present study because it examines productive vocabulary performance under controlled instructional conditions. Comparing male and female learners within the same task conditions provides an opportunity to determine whether gender is associated with differences in productive lexical outcomes when initial productive vocabulary knowledge is statistically controlled.

Research Gap and the Present Study

The literature reviewed above points to three related issues. First, lexical development involves a progression from receptive knowledge toward productive use, yet possessing receptive knowledge of a lexical item does not necessarily mean that the item can be retrieved successfully during language production.

Second, research on vocabulary learning tasks indicates that task characteristics and the amount of learner involvement can influence vocabulary development. However, empirical findings do not justify assuming that one task type will always be more effective than another.

Third, although gender has been investigated in vocabulary learning research, findings concerning

gender differences are inconsistent. Some studies have reported differences in vocabulary learning strategies, whereas others have found no significant gender differences in vocabulary knowledge.

Against this background, the present study addresses a more specific question: whether different productive lexical tasks differ in their effectiveness in transforming receptive lexical knowledge into productive lexical competence, and whether gender influences productive vocabulary outcomes under these instructional conditions.

The study therefore compares three productive lexical tasks—open-cloze, sentence translation, and sentence writing—using the same target lexical items and the same learner population. By focusing specifically on receptively known but not yet productively mastered vocabulary, the study extends previous task-based vocabulary research toward the receptive-to-productive dimension of lexical development. It also examines gender as a learner-related variable while controlling statistically for initial productive vocabulary knowledge.

METHODOLOGY

The study employed a quasi-experimental pre-test-post-test design to examine whether different productive lexical tasks were associated with differences in productive lexical competence among Iranian intermediate EFL learners.

Participants were assigned to one of three task conditions at the intact-class level: open-cloze, sentence translation, or sentence writing. Productive lexical competence was assessed using the Productive Vocabulary Levels Test (PVLT) before and after the instructional period. For RQ1, a one-way analysis of covariance (ANCOVA) was conducted to compare the three task conditions while controlling for pre-test productive vocabulary scores. For RQ2, a factorial ANCOVA was conducted to examine the main effect of gender, the main effect of task condition, and the Gender × Task interaction, with PVLT pre-test scores entered as the covariate.

In addition to the quantitative analyses, an open-ended questionnaire was administered immediately after the PVLT post-test to obtain complementary qualitative evidence concerning learners' perceptions of the instructional tasks. The qualitative data were used primarily to provide learner-centered context for interpreting the findings related to RQ1. Because the questionnaire responses were not analyzed separately by gender, they were not used to test or interpret RQ2.

Participants

The participants were 384 Iranian pre-university EFL learners enrolled in two private language institutes in Tehran. The sample comprised 192 female and 192 male learners aged between 17 and 22 years. Their English proficiency was at the intermediate (B1) level, as determined through the Oxford Placement Test (OPT). Only learners classified at the B1 level according to the OPT scoring criteria were included.

The participants were drawn from 63 intact classes, with approximately six to seven learners per class. The classes were randomly assigned to one of three instructional task conditions: open-cloze, sentence translation, or sentence writing. Each condition comprised 128 learners distributed across 21 intact classes. Assignment was therefore conducted at the intact-class level rather than at the individual learner level, which helped minimize disruption to the existing instructional arrangements.

The instructional intervention was conducted over a 26-week period, with classes meeting twice a week as part of their regular instructional schedule. During the intervention, learners completed productive lexical activities based on the same 20 target lexical items in their assigned task condition. The target vocabulary, instructional objectives, and general instructional content were kept consistent across the three conditions, while the format through which learners practiced and produced the target items differed according to task condition. Participants completed the instructional activities within their regular classes, and the same feedback procedures were applied across all three conditions.

Ethical principles were observed throughout the study. Before participation, learners were informed in Persian about the general purpose and procedures of the study, the voluntary nature of participation, and the use of their data for research purposes. Participants were assured that their personal information and research data would remain confidential and anonymous and would be used solely for research purposes. They were also informed that participation and performance on the research assessments would not affect their course grades or academic standing. Participants were informed of their right to withdraw from the study without negative consequences. Because the participants included learners aged 17 years and above, consent procedures were implemented in accordance with the applicable institutional and research requirements governing participation in educational research. Testing and instructional procedures were standardized across the three experimental conditions to ensure that participation did not create unnecessary academic disadvantage for any group.

Instruments

Four instruments were used to establish participants' proficiency levels, measure receptive and productive vocabulary knowledge, and obtain learners' perceptions of the instructional tasks.

First, the Oxford Placement Test (OPT) was administered to determine learners' general English proficiency and to identify those whose proficiency corresponded to the B1 level. The test consisted of 60 multiple-choice items and required approximately 60 minutes to complete. Each correct response was awarded one point, whereas incorrect or blank responses received zero points. The total score was used to categorize learners according to the Common European Framework of Reference for Languages (CEFR) proficiency levels. Only learners classified at the intermediate (B1) level according to the OPT scoring criteria were included in the study.

Second, the New Vocabulary Levels Test (NVLT; McLean & Kramer, 2015) was administered to measure learners' receptive vocabulary knowledge. The test consisted of 40 items and required approximately 20 minutes to complete.

Third, the Productive Vocabulary Levels Test (PVLT; Laufer & Nation, 1999) measured productive lexical competence. The test was administered as part of the initial vocabulary screening and subsequently as the pre-test and post-test measure of productive lexical competence. The PVLT required approximately 35 minutes to complete.

The NVLT and PVLT were first used together as a screening procedure to identify lexical items that were known receptively but not productively. Specifically, items that were correctly recognized by every participant on the NVLT but incorrectly produced by every participant on the PVLT were retained as candidate target items. From this candidate pool, 20 common lexical items were selected for the instructional intervention. The target set consisted of five nouns, five verbs, five adjectives, and five adverbs. The items were drawn from the 3,000–5,000 frequency bands using frequency information from the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). The same 20 target lexical items were then used throughout the instructional intervention and assessed in the PVLT pre-test and post-test.

Fourth, an open-ended questionnaire entitled Questionnaire on Lexical Acquisition Activities was administered immediately after the PVLT post-test to obtain learners' perceptions of the instructional tasks. The questionnaire consisted of three open-ended questions addressing (a) how the task supported remembering and using new vocabulary, including its use in writing, speaking, and outside the classroom; (b) problems or negative aspects encountered during task completion; and (c) learners' willingness to use the task for future vocabulary learning and their reasons for doing so. The questionnaire also included a participant identification code and date. Participation was voluntary, and responses were treated confidentially and anonymously.

Target Vocabulary and Instructional Tasks

The same 20 target lexical items were used across all three instructional conditions. The task

materials were developed by the researcher in consultation with a native English-speaking instructor. Example sentences were adapted from the Oxford Advanced Learner's Dictionary (2022) and the Cambridge Advanced Learner's Dictionary (2022), with modifications where necessary to ensure that the language was appropriate for B1-level learners.

In the open-cloze condition, learners encountered the target vocabulary in contextualized sentences or short passages in which the target lexical item was removed. Learners were required to retrieve and supply the appropriate word based on the surrounding contextual, semantic, and grammatical cues.

In the sentence translation condition, learners were provided with Persian sentences that conveyed the intended meaning of the target vocabulary in context. They were required to translate the sentences into English and produce the appropriate target lexical item.

In the sentence-writing condition, learners were asked to produce their own English sentences using the target lexical items. This condition provided less linguistic support than the other two tasks and required learners to retrieve the target item and integrate it into an independently generated sentence.

Although the three tasks differed in the amount and type of contextual support they provided, the target vocabulary, instructional objectives, and general instructional content were kept consistent across conditions. Learners in all three groups therefore worked with the same lexical items and received comparable opportunities to encounter and produce the target vocabulary, while the format of productive practice differed according to the assigned task condition.

Procedure

Data collection was conducted in several stages. First, the OPT was administered to establish participants' English proficiency. Each correct response received one point, while incorrect or blank responses received zero points. The total score was used to categorize learners according to CEFR proficiency levels, and learners classified at the B1 level according to the study's inclusion criterion were retained for participation.

The NVLT was then administered to assess receptive vocabulary knowledge, followed by the PVLT at a separate testing session to measure productive vocabulary knowledge. The NVLT and PVLT results were compared to identify lexical items that learners knew receptively but had not yet mastered productively. Based on these results, 20 target lexical items were selected. These 20 items were subsequently used both in the instructional activities and in the PVLT pre-test and post-test.

Participants then completed the instructional activities associated with their assigned task condition over the 26-week instructional period. Classes met twice a week as part of their regular instructional schedule. The three groups worked with the same 20 target lexical items and received comparable opportunities to encounter and produce the vocabulary. The instructional activities were implemented within the learners' regular classes, with the principal difference among the conditions being the format of productive lexical practice.

To maintain consistency across conditions, written feedback was provided individually on the same day. Feedback focused specifically on the target lexical items, including lexical form, spelling or morphology where relevant, meaning, and appropriate lexical choice. Grammar, sentence organization, and stylistic features were not corrected unless they directly affected the target lexical item. The same feedback criteria and procedures were applied across the three conditions.

Following the instructional phase, the PVLT was administered as the post-test. The PVLT post-test scores served as the primary quantitative outcome measure of productive lexical competence. Because learners differed somewhat in their initial productive vocabulary knowledge, PVLT pre-test scores were subsequently used as a covariate in the statistical analyses.

Immediately after completing the PVLT post-test, all 384 participants completed the open-ended questionnaire. The questionnaire was administered before any subsequent activities so that learners could report their perceptions of the instructional tasks without reference to later learning activities. Responses were collected using participant identification codes rather than names.

Data Analysis

The quantitative data were analyzed using the Statistical Package for the Social Sciences (SPSS). Descriptive statistics were first calculated for the main quantitative variables, followed by checks of the assumptions required for the inferential analyses.

For RQ1, a one-way analysis of covariance (ANCOVA) was conducted to determine whether the three instructional task conditions differed significantly in productive lexical post-test performance after controlling for initial productive vocabulary knowledge. Instructional task (open-cloze, sentence translation, and sentence writing) was entered as the independent variable, PVLT post-test score as the dependent variable, and PVLT pre-test score as the covariate. When a significant main effect of task was obtained, pairwise comparisons using the least significant difference (LSD) procedure were conducted to determine which instructional conditions differed from one another. Effect size was reported using partial eta squared (ηp^2).

For RQ2, a factorial ANCOVA was conducted to examine the main effect of gender while controlling for initial productive vocabulary knowledge and accounting for instructional task. Gender and instructional task were included as factors, PVLT post-test score served as the dependent variable, and PVLT pre-test score was entered as the covariate. The analysis examined the main effect of gender, the main effect of instructional task, and the Gender × Task interaction. Partial eta squared (ηp^2) was reported as the measure of effect size.

Before conducting the ANCOVAs, the assumptions of normality, homogeneity of variance, and homogeneity of regression slopes were examined. Statistical significance was evaluated at the .05 level.

To complement the quantitative findings, responses to the three open-ended questionnaire questions were analyzed using conventional content analysis with an inductive coding approach. All 384 participants responded. The analysis focused on learners' perceptions of the usefulness of the instructional activities, difficulties encountered during task completion, and their willingness to use the activities for future vocabulary learning. Because the qualitative data were analyzed across the full sample rather than separately by task condition or gender, the findings present overall patterns of learner perception and are not used to establish differences between experimental groups or genders.

For the first questionnaire question, which asked learners how the task helped them remember and use new vocabulary, including in writing, speaking, and outside the classroom, three broad themes emerged: perceived utility for productive vocabulary development, increased confidence in vocabulary retrieval and language production, and variability in vocabulary retention and transfer. Overall, 65% of responses were classified as predominantly positive, 19% as neutral, 7% as negative, and 9% as mixed. Positive responses commonly indicated that the activities helped learners remember target words, retrieve them when needed, and use them more actively in language production. Some learners also reported greater confidence in using newly learned vocabulary, while other responses indicated that retention and transfer varied across learners.

For the second questionnaire question, which addressed problems or negative aspects of completing the task, three themes were identified: productive tasks were challenging but manageable, cognitive load and task engagement, and practical difficulties rather than instructional problems. Across the full sample, 54% of responses were classified as predominantly positive, 24% as neutral, 13% as negative, and 9% as mixed. Although many participants did not identify substantial problems, some described the productive activities as demanding, particularly when they had difficulty retrieving an appropriate lexical item or producing language independently. The responses therefore indicated that

productive vocabulary activities could involve noticeable cognitive demands while remaining manageable for many learners.

The third questionnaire question examined whether learners would use the task to learn new lexical items in the future and the reasons for their responses. Three themes emerged: intention to continue using productive lexical activities, learner autonomy and strategy development, and preference for flexible or combined learning approaches. Overall, 39% of responses were classified as predominantly positive, 28% as neutral, 12% as negative, and 21% as mixed. Positive responses commonly reflected learners' willingness to continue using productive activities because they perceived them as useful for remembering and actively using vocabulary. Some learners emphasized the value of developing independent vocabulary-learning strategies, whereas others prefer combining productive activities with other vocabulary-learning techniques.

Taken together, the qualitative findings showed that learners generally perceived productive vocabulary activities as useful, while also reporting variation in task difficulty, vocabulary retention, and preferences for future use. The findings provide complementary information about learners' experiences of the instructional activities and contextualize the quantitative results concerning productive lexical performance. However, because the qualitative responses were not analyzed separately by task condition or gender, they do not provide evidence of differences in perceptions across the three instructional groups or between male and female learners.

RESULTS

Preliminary Analysis

Descriptive statistics were calculated for the three task groups. Each group consisted of 128 learners. The mean PVLT pre-test scores were 14.48 (*SD* = 2.28) for the open-cloze group, 15.19 (*SD* = 2.08) for the sentence-translation group, and 15.90 (*SD* = 1.98) for the sentence-writing group. Because some differences were observed in learners' initial productive vocabulary knowledge, PVLT pre-test scores were controlled for in the subsequent ANCOVA analyses. Levene's test indicated that the assumption of homogeneity of variance was satisfied, $F(2, 381) = 2.080$, $p = .126$.

RQ1: Effect of Task Type. To address RQ1, an analysis of covariance (ANCOVA) was conducted with PVLT post-test score as the dependent variable, task type as the independent variable, and PVLT pre-test score as the covariate.

The analysis revealed a statistically significant main effect of task type, $F(2, 380) = 6.063$, $p = .003$, partial $\eta^2 = .031$. The PVLT pre-test was also a significant covariate, $F(1, 380) = 493.326$, $p < .001$, partial $\eta^2 = .565$.

The adjusted post-test means were 15.82 (*SE* = .18) for the open-cloze group, 16.06 (*SE* = .17) for the sentence-translation group, and 16.68 (*SE* = .18) for the sentence-writing group.

Table 1

Adjusted PVLT Post-test Means by Task Type

Task	Adjusted M	SE	95% CI
Open-cloze	15.82	.18	15.47–16.17
Sentence translation	16.06	.17	15.72–16.40
Sentence writing	16.68	.18	16.33–17.03

Pairwise comparisons indicated that the sentence-writing group significantly outperformed the open-cloze group (*MD* = 0.859, *SE* = .255, *p* = .001) and the sentence-translation group (*MD* = 0.617, *SE* = .248, *p* = .013). However, the difference between the open-cloze and sentence-translation groups was not statistically significant (*MD* = -0.242, *SE* = .248, *p* = .330).

Table 2
Pairwise Comparisons Among Task Conditions

Comparison	Mean Differ	SE	*p*	95% CI
Open-cloze – Sen translation	-0.242	.248	.330	-0.730–0.246
Open-cloze – Sen writing	-0.859	.255	.001	-1.361—0.358
Sentence translati Sentence writing	-0.617	.248	.013	-1.105—0.129

Overall, the results indicate that sentence writing produced significantly higher adjusted productive lexical performance than both open-cloze and sentence translation. Although the sentence-translation group obtained a slightly higher adjusted mean than the open-cloze group, the difference between these two conditions was not statistically significant.

RQ2: Effect of Gender. To address RQ2, a second ANCOVA was conducted with gender and task type as between-subjects factors, PVLT post-test score as the dependent variable, and PVLT pre-test score as the covariate. The Gender × Task interaction was also examined to determine whether task effectiveness differed according to gender. Levene's test indicated that the assumption of homogeneity of variance was satisfied, $F(5, 378) = 1.095$, $p = .363$.

Table 3
PVLT Post-test Scores by Task Type and Gender

Task	Gender	N	M	SD
Open-cloze	Male	64	15.39	2.75
Open-cloze	Female	64	14.75	3.55
Sentence translation	Male	64	15.56	3.10
Sentence translation	Female	64	16.56	2.94
Sentence writing	Male	64	17.34	2.74
Sentence writing	Female	64	17.52	2.65

The ANCOVA results are presented in Table 4.

Table 4
ANCOVA Results for Gender, Task Type, and Gender × Task Interaction

Source	df	F	*p*	Partial η^2
Gender	1	.238	.626	.001
Task	2	6.117	.002	.031
PVLT pre-test	1	490.923	<.001	.566
Gender × Task	2	2.829	.060	.015
Error	377			

The results showed no statistically significant main effect of gender, $F(1, 377) = .238$, $p = .626$, partial $\eta^2 = .001$. Thus, after controlling for initial productive vocabulary knowledge, male and female learners did not differ significantly in their productive lexical performance.

The Gender × Task interaction was also not statistically significant, $F(2, 377) = 2.829$, $p = .060$, partial

$\eta^2 = .015$. This indicates that there was insufficient evidence that the relative effectiveness of the three task conditions differed according to gender. The main effect of task type remained statistically significant, $F(2, 377) = 6.117$, $p = .002$, partial $\eta^2 = .031$.

Overall, the findings provide no evidence that gender significantly influenced productive lexical competence or moderated the effectiveness of the three productive vocabulary tasks.

Qualitative Findings

To complement the quantitative findings, responses to the three open-ended questionnaire questions were analyzed using conventional content analysis with an inductive coding approach. All 384 participants responded. The analysis focused on learners' perceptions of the usefulness of the instructional activities, difficulties encountered during task completion, and their willingness to use the activities for future vocabulary learning. Because the qualitative data were analyzed across the full sample rather than separately by task condition or gender, the findings are presented as overall patterns of learner perception and are not used to establish differences between experimental groups or genders.

For the first questionnaire question, which asked learners how the task helped them remember and use new vocabulary, including in writing, speaking, and outside the classroom, three broad themes emerged: perceived utility for productive vocabulary development, increased confidence in vocabulary retrieval and language production, and variability in vocabulary retention and transfer. Overall, 65% of responses were classified as predominantly positive, 19% as neutral, 7% as negative, and 9% as mixed. Positive responses commonly indicated that the activities helped learners remember target words, retrieve them when needed, and use them more actively in language production. Some learners also reported greater confidence in using newly learned vocabulary, while other responses indicated that retention and transfer varied across learners.

For the second questionnaire question, which addressed problems or negative aspects of completing the task, three themes were identified: productive tasks were challenging but manageable, cognitive load and task engagement, and practical difficulties rather than instructional problems. Across the full sample, 54% of responses were classified as predominantly positive, 24% as neutral, 13% as negative, and 9% as mixed. Although many participants did not identify substantial problems, some described the productive activities as demanding, particularly when they had difficulty retrieving an appropriate lexical item or producing language independently. The responses therefore indicated that productive vocabulary activities could involve noticeable cognitive demands while remaining manageable for many learners.

The third questionnaire question examined whether learners would use the task to learn new lexical items in the future and the reasons for their responses. Three themes emerged: intention to continue using productive lexical activities, learner autonomy and strategy development, and preference for flexible or combined learning approaches. Overall, 39% of responses were classified as predominantly positive, 28% as neutral, 12% as negative, and 21% as mixed. Positive responses commonly reflected learners' willingness to continue using productive activities because they perceived them as useful for remembering and actively using vocabulary. Some learners emphasized the value of developing independent vocabulary-learning strategies, whereas others prefer combining productive activities with other vocabulary-learning techniques.

Taken together, the qualitative findings showed that learners generally perceived productive vocabulary activities as useful, while also reporting variation in task difficulty, vocabulary retention, and preferences for future use. The findings provide complementary information about learners' experiences of the instructional activities and contextualize the quantitative results concerning productive lexical performance. However, because the qualitative responses were not analyzed separately by task condition or gender, they do not provide evidence of differences in perceptions across the three instructional groups or between male and female learners.

DISCUSSION

Discussion of Research Question 1

The first research question examined whether open-cloze, sentence translation, and sentence writing differed in their effectiveness in developing productive lexical competence. The ANCOVA revealed a statistically significant effect of instructional task after controlling for learners' initial productive vocabulary knowledge, $F(2, 380) = 6.063$, $p = .003$, partial $\eta^2 = .031$. The adjusted post-test means were 15.82 for open-cloze, 16.06 for sentence translation, and 16.68 for sentence writing. Pairwise comparisons showed that sentence writing significantly outperformed both open-cloze ($p = .001$) and sentence translation ($p = .013$), whereas the difference between open-cloze and sentence translation was not statistically significant ($p = .330$).

These findings indicate that sentence writing had a relative advantage over the two more constrained productive tasks under the conditions of the present study. Nevertheless, the relatively small effect size suggests that task type accounted for a modest proportion of the variance in adjusted productive lexical performance. The finding should therefore be interpreted as evidence of a relative advantage for sentence writing in this particular instructional context rather than as evidence that sentence writing is universally superior to other vocabulary-learning activities.

One possible explanation for the advantage of sentence writing is the greater degree of generative lexical engagement required by the task. Open-cloze activities provide learners with substantial contextual support, while sentence translation provides semantic content and an L1 linguistic frame. In contrast, sentence writing requires learners to independently retrieve the target lexical item and construct an original English sentence around it. Learners therefore need to attend not only to the form and meaning of the target item but also to its appropriate use within a self-generated context.

This interpretation is compatible with the Depth of Processing perspective (Craik & Lockhart, 1972), according to which more elaborate processing may facilitate subsequent access to learned information. Sentence writing may encourage such elaboration because learners must actively make decisions about how a lexical item can be used rather than simply reproduce or supply it within a predetermined context.

The finding is also consistent with the Involvement Load Hypothesis (Laufer & Hulstijn, 2001), which proposes that greater involvement in lexical processing may contribute to vocabulary learning. Sentence writing arguably requires greater independent retrieval and evaluation of lexical knowledge than the other two tasks. However, the present study did not independently manipulate or measure the specific components of involvement load. Thus, the results should be interpreted as being consistent with the hypothesis rather than as direct empirical confirmation of it.

The findings also correspond with previous research emphasizing productive engagement and active lexical processing. Allal-Sumoto et al. (2023) highlighted the contribution of productive vocabulary use to lexical processing and comprehension, while Hopp (2025) showed that L2 lexical processing can influence the learning of grammatical structures, highlighting the importance of lexical processing in L2 language development. The present findings extend these observations to Iranian intermediate EFL learners by indicating that independently generating sentences with target vocabulary was associated with higher productive lexical performance than the two more constrained production formats.

The absence of a significant difference between open-cloze and sentence translation is also noteworthy. Although sentence translation produced a slightly higher adjusted mean than open-cloze, the difference was not statistically significant. Thus, the results do not support a clear hierarchy in which sentence translation is more effective than open-cloze. Rather, the most consistent distinction was between sentence writing and the two more supported task formats.

The qualitative findings enrich these quantitative results. Across the 384 participants, 65% of

responses to the question concerning the usefulness of the activities were classified as predominantly positive, while smaller proportions were neutral, negative, or mixed. Learners commonly described the activities as useful for remembering target vocabulary, retrieving words when needed, and using new lexical items more actively in language production. Some participants also associated the activities with increased confidence in vocabulary retrieval and language use. These perceptions provide a complementary learner-centered context for the quantitative finding that the instructional task was significantly associated with differences in productive lexical performance.

At the same time, the qualitative findings revealed variability in learners' experiences. Participants differed in their reported ability to retain and transfer vocabulary, as well as in their perceptions of task difficulty. Responses to the question concerning difficulties suggested that productive activities were often viewed as challenging but manageable. This is particularly relevant to sentence writing because its greater independence may increase cognitive demands while simultaneously providing opportunities for active lexical engagement. Importantly, however, these qualitative responses cannot establish that perceived cognitive effort caused the superior performance of sentence writing. Rather, they provide learner-centered contextual evidence that productive vocabulary activities were generally viewed as useful despite their demands.

Learners' responses concerning future use were also somewhat varied. Although a substantial proportion indicated an intention to continue using productive lexical activities, some participants expressed neutral, negative, or mixed preferences, and some appeared to favor combining such activities with other forms of vocabulary practice. This finding suggests that effectiveness and learner preference should not be treated as equivalent constructs. A task can contribute to productive vocabulary development without necessarily being preferred by every learner as a stand-alone learning activity.

Overall, the quantitative and qualitative findings together suggest that the nature of the productive task matters, while also highlighting variation in learners' experiences of task demands. Sentence writing produced the strongest quantitative performance, and learners generally recognized the value of active productive vocabulary work. The evidence therefore supports a relative advantage of generative sentence writing while also suggesting that more supported activities can remain valuable components of a broader vocabulary-learning sequence.

Discussion of Research Question 2

The second research question examined whether gender significantly influenced productive lexical competence after controlling for learners' initial PVLT performance. The ANCOVA showed that the main effect of gender was not statistically significant, $F(1, 377) = 0.238$, $p = .626$, partial $\eta^2 = .001$. The extremely small effect size indicates that gender accounted for virtually none of the variance in adjusted productive vocabulary performance.

The analysis also revealed a significant main effect of instructional task, $F(2, 377) = 6.117$, $p = .002$, partial $\eta^2 = .031$, which is consistent with the finding for RQ1. However, the Gender × Task interaction was not statistically significant, $F(2, 377) = 2.829$, $p = .060$, partial $\eta^2 = .015$. Thus, there was insufficient evidence that the relative effectiveness of the three productive tasks differed systematically according to gender.

The descriptive results showed some variation across task conditions. In the open-cloze condition, males obtained a mean post-test score of 15.39 ($SD = 2.75$), compared with 14.75 ($SD = 3.55$) for females. In sentence translation, females obtained a higher mean ($M = 16.56$, $SD = 2.94$) than males ($M = 15.56$, $SD = 3.10$). In sentence writing, the means were very similar for males ($M = 17.34$, $SD = 2.74$) and females ($M = 17.52$, $SD = 2.65$). However, these descriptive differences were not statistically significant after controlling for initial productive vocabulary knowledge and therefore should not be interpreted as evidence of meaningful gender-based differences.

The absence of a significant gender effect is consistent with research suggesting that differences in

vocabulary-learning behavior or strategy preferences do not necessarily result in differences in vocabulary-learning outcomes. Bhatti and Mukhtar (2020), for example, reported gender-related differences in vocabulary-learning strategy preferences without establishing corresponding differences in vocabulary achievement. Similarly, Jiménez Catalán (2003) found significant gender differences in the number and types of vocabulary-learning strategies used by male and female learners.

Other studies have reported mixed findings concerning gender and language learning strategies (Oxford & Nyikos, 1989; Rao, 2016). Such findings highlight the importance of distinguishing between learning preferences or strategy use and actual learning outcomes. The present study contributes to this distinction by showing that gender was not significantly associated with productive lexical performance after initial vocabulary knowledge was controlled.

The nonsignificant Gender × Task interaction is particularly important. Although descriptive means varied across the task conditions, there was insufficient evidence that one task systematically benefited one gender more than another. Thus, the data did not provide sufficient evidence that the relative performance advantage observed for sentence writing varied systematically by gender.

The qualitative findings should not constitute a corresponding claim about gender because the questionnaire responses were analyzed across the full sample rather than separately by gender. Nevertheless, the qualitative analysis revealed individual variation in perceived task difficulty, vocabulary retention, and willingness to continue using productive activities. Such variation suggests that learner characteristics other than gender may contribute to differences in how productive vocabulary activities are experienced.

Accordingly, the absence of a significant gender effect should not be interpreted as demonstrating that gender can never influence vocabulary learning. Rather, it indicates that under the instructional and learner conditions examined in this study, there was insufficient evidence of a gender-related difference in productive lexical development. Future research could investigate whether gender interacts with other learner characteristics, such as motivation, learning strategies, proficiency, or task familiarity.

Integrative Interpretation

Taken together, the findings provide evidence that instructional task type was associated with differences in productive lexical performance, whereas gender was not. Sentence writing produced significantly higher adjusted post-test performance than both open-cloze and sentence translation, while no significant difference was found between the latter two tasks. At the same time, neither the main effect of gender nor the Gender × Task interaction reached statistical significance.

The relative advantage of sentence writing may reflect its greater generative demands. Whereas open-cloze and sentence translation provide learners with relatively strong contextual or semantic support, sentence writing requires more independent retrieval and contextual generation. The task therefore expands opportunities for learners to process lexical form, meaning, and use simultaneously.

The qualitative evidence contextualizes this interpretation but does not independently establish differences among the task conditions. Learners generally reported that productive activities helped them remember and use new vocabulary, and many viewed the activities positively despite acknowledging challenges. The qualitative responses also revealed variation in perceived difficulty, retention, and future preferences. Thus, the qualitative evidence supports the broader interpretation that productive vocabulary activities can be useful and engaging while also demonstrating that their demands are not experienced uniformly by all learners.

Importantly, the qualitative evidence should not be interpreted as independent proof of task effectiveness. The quantitative PVLT results ground conclusions about productive lexical

performance, whereas the questionnaire provides information about learners' perceptions and experiences. The convergence of these two forms of evidence is therefore best understood as complementary rather than causal.

The relatively small effect size for instructional task also warrants attention. Although the task effect was statistically significant, task type explained only a modest proportion of the variance in adjusted post-test performance. This suggests that productive lexical development is influenced by multiple factors and cannot be attributed to task design alone. Initial productive vocabulary knowledge was strongly associated with post-test performance, emphasizing the importance of learners' starting levels.

Theoretical Implications

The findings contribute to theoretical discussions of productive vocabulary learning in several ways. First, the significant task effect provides support for perspectives emphasizing active productive engagement in lexical development. Sentence writing, which required the high independent lexical generation among the three conditions, produced the highest adjusted productive vocabulary performance.

Second, the results provide qualified support for the Involvement Load Hypothesis. The superior performance of sentence writing is compatible with the assumption that greater involvement in lexical processing may facilitate learning. However, because need, search, and evaluation were not independently manipulated or measured, the present findings cannot establish which specific component or combination of components accounted for the observed difference.

Third, the findings are consistent with the Depth of Processing perspective. The generative nature of sentence writing may have encouraged learners to process target lexical items more elaborately by considering their meanings and appropriate use in self-generated contexts. Nevertheless, depth of processing was not directly measured and should therefore be regarded as a plausible theoretical explanation rather than a directly tested mechanism.

The qualitative findings provide indirect contextual support for these interpretations. Learners frequently described the activities as requiring them to recall and use vocabulary actively, and some associated such engagement with increased confidence in vocabulary retrieval and language production. These perceptions are compatible with theoretical accounts emphasizing active lexical processing, although they cannot be taken as direct evidence of cognitive depth or involvement load. Finally, the absence of a significant gender effect or Gender × Task interaction suggests that the task-related differences observed in productive lexical performance were not systematically moderated by gender in this sample. This finding contributes to a more task-sensitive interpretation of productive vocabulary learning in which the design of the learning activity may be more consequential than gender-based assumptions about learners.

Pedagogical Implications

The findings have several implications for vocabulary instruction. First, teachers seeking to develop productive lexical competence may benefit from incorporating sentence-writing activities alongside more constrained vocabulary exercises. Requiring learners to independently use target vocabulary in original sentences may provide opportunities for active retrieval and meaningful language production. Second, the findings do not suggest that open-cloze or sentence translation activities should be abandoned. The lack of a significant difference between these two tasks indicates that both can provide useful forms of supported productive practice. They may be particularly appropriate when learners are developing familiarity with new lexical items before progressing toward more independent production.

Third, the qualitative findings suggest that teachers should consider both the potential benefits and demands of productive activities. Although many learners perceived the activities positively and reported benefits for vocabulary recall and use, some experienced cognitive or practical difficulties.

Teachers may therefore need to provide appropriate scaffolding and gradually increase task demands according to learners' proficiency and familiarity with productive vocabulary activities.

Fourth, learners' varied preferences concerning future use suggest that a flexible combination of task types may be preferable to exclusive reliance on one activity. A sequence that moves from contextualized or supported retrieval toward increasingly independent sentence generation may allow learners to benefit from different forms of productive practice.

Finally, the absence of significant gender differences suggests that teachers should avoid assigning particular productive vocabulary tasks on the basis of assumptions about male or female learners. Instead, task selection may be better informed by learners' proficiency, prior vocabulary knowledge, task familiarity, and individual learning needs.

Limitations and Directions for Future Research

Several limitations should be acknowledged. First, the participants were Iranian intermediate EFL learners aged 17–22 who were recruited from private language institutes in Tehran. The findings may therefore not generalize directly to learners at different proficiency levels, age groups, educational settings, or sociocultural contexts.

Second, the study examined only three productive vocabulary tasks and a defined set of 20 target lexical items. Other task types, lexical categories, vocabulary difficulty levels, or target-word characteristics may produce different outcomes. Future research could therefore examine a broader range of productive activities and lexical variables.

Third, although initial productive vocabulary knowledge was statistically controlled, other individual-difference variables were not examined. Motivation, language aptitude, working memory, learning strategies, and task familiarity may influence how learners respond to productive vocabulary activities. Future studies could investigate whether such variables moderate task effectiveness.

Fourth, the qualitative component relied on self-reported responses to an open-ended questionnaire. Although these responses provided useful information about perceived usefulness, difficulty, confidence, and future intentions, they represent subjective experiences rather than objective measures of vocabulary development. Furthermore, the qualitative analysis was conducted across the full sample and was not designed to establish statistical relationships between learner perceptions and post-test performance. Future research could combine questionnaires with interviews, classroom observations, learner diaries, or other qualitative methods and examine how learner perceptions relate to measurable vocabulary outcomes.

Finally, the present study relied on the PVLT post-test as the primary measure of productive lexical performance. Although the PVLT provided the study's primary measure of productive lexical performance, future research could complement it with delayed retention measures and more open-ended assessments of spontaneous lexical use to determine whether task-related advantages extend beyond controlled productive vocabulary assessment.

CONCLUSION

This study examined whether open-cloze, sentence translation, and sentence writing differed in their effectiveness in developing productive lexical competence among Iranian intermediate EFL learners and whether gender influenced productive lexical performance.

The findings demonstrated a statistically significant effect of instructional task after controlling for initial productive vocabulary knowledge. Sentence writing produced the highest adjusted post-test performance and significantly outperformed both open-cloze and sentence translation, whereas the difference between open-cloze and sentence translation was not significant. The relatively small effect size indicates that this advantage should be interpreted as a task-specific relative benefit rather than

evidence of universal superiority.

The study also found no statistically significant main effect of gender and no significant Gender × Task interaction. Thus, there was insufficient evidence that gender influenced productive lexical performance or systematically moderated the relative effectiveness of the three tasks.

The qualitative findings provided complementary learner-centered evidence. Participants generally perceived the productive activities as useful for remembering, retrieving, and using new lexical items, and many reported increased confidence in vocabulary use. At the same time, learners differed in their experiences of task difficulty, retention, and willingness to continue using the activities. These findings suggest that productive vocabulary activities can be both demanding and beneficial and that learner experiences may vary even when the instructional conditions are broadly comparable.

Overall, the findings support a task-sensitive approach to productive vocabulary instruction. Sentence writing showed a relative advantage under the conditions examined in this study when the goal is to encourage independent and generative lexical use, while open-cloze and sentence translation may serve as useful forms of more supported productive practice. The results also suggest that task-related differences in productive lexical performance should not be interpreted through gender-based assumptions. Instead, productive vocabulary instruction may benefit from combining different levels of contextual support and productive demand while taking learners' individual needs and experiences into account.

REFERENCES

Allal-Sumoto, T. K., Miyoshi, K., & Mizuhara, H. (2023). The effect of productive vocabulary knowledge on second language comprehension. *Frontiers in Psychology, 14*, 1049885. https://doi.org/10.3389/fpsyg.2023.1049885

Bhatti, M. S., & Mukhtar, R. (2020). Impact of vocabulary learning strategies on gender based ESL learners. *The Eurasia Proceedings of Educational and Social Sciences, 18*, 100–108.

Cambridge University Press & Assessment. (2022). *Cambridge advanced learner's dictionary* (4th ed.). Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303–317. https://doi.org/10.1017/S0272263199002089

Hopp, H., Reifegerste, J., & Ullman, M. T. (2025). Lexical effects on second language grammar acquisition: Testing psycholinguistic and neurocognitive predictions. *Language Learning, 75*(2), 424–457. https://doi.org/10.1111/lang.12672

Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal, 96*(4), 544–557. https://doi.org/10.1111/j.1540-4781.2012.01394.x

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539–558. https://doi.org/10.1111/0023-8333.00164

Jiménez Catalán, R. M. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics, 13*(1), 54–77. https://doi.org/10.1111/1473-4192.00037

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*(1), 33–51. https://doi.org/10.1191/026553299672614616

Lee, J.-W. (2019). EFL learners' vocabulary learning in terms of gender and vocabulary knowledge. *English Language & Literature Teaching, 25*(2), 1–19.

McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken, 19*(2), 1–11.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
Okyar, H. (2021). Vocabulary learning strategies of Turkish EFL learners: A focus on gender. *Journal of English Teaching, 7*(1), 43–54. https://doi.org/10.33541/jet.v7i1.2289
Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal, 73*(3), 291–300. https://doi.org/10.1111/j.1540-4781.1989.tb06367.x
Oxford University Press. (2022). *Oxford advanced learner's dictionary* (10th ed.). Oxford University Press.
Rao, Z. (2016). Language learning strategies and English proficiency: Interpretations from information-processing theory. *The Language Learning Journal, 44*(1), 90–106.
Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951. https://doi.org/10.1111/lang.12077
Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching, 52*(2), 261–274. https://doi.org/10.1017/S0261444819000053
Teng, M., & Xu, J. (2025). Pushing vocabulary knowledge from receptive to productive mastery: Effects of task type and repetition frequency. *Language Teaching Research, 29*(2). https://doi.org/10.1177/13621688221077028
Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.
Yanagisawa, A., & Webb, S. (2021). To what extent does the involvement load hypothesis predict incidental L2 vocabulary learning? A meta-analysis. *Language Learning, 71*(2), 487–536. https://doi.org/10.1111/lang.12444