Examining the effects of oral reproduction and summary writing vocabulary tasks on L2 word learning: Technique feature analysis on trial

Mojtaba Kamali¹, Fatemeh Behjat²* and Mohammad Sadegh Bagheri¹

Abstract: The current study investigated the influence of oral reproduction and summary writing vocabulary tasks on English as a Foreign Language (EFL) learners’ learning and retention of new foreign language (L2) words. To this end, 66 advanced EFL learners were randomly selected and assigned to two experimental and one control conditions. The participants in all three groups were required to read eight texts which included a total of 40 target words during eight weeks of treatment and to perform designated tasks which were: reading and writing a summary incorporating the target words, and reading and reproducing the content of the passage orally using the target words. The results of immediate and delayed recognition and recall vocabulary post-tests indicated that while both tasks resulted in significant vocabulary learning, the “oral reproduction” condition was more effective than “summary writing” in terms of long-term retention of target words. The findings are explained and justified in light of the predictions of Technique Feature Analysis.

ABOUT THE AUTHORS
Mojtaba Kamali is a PhD candidate of TEFL at Islamic Azad University, Shiraz, Iran. He has been teaching English at several universities and language institutes. His main areas of interest include L2 vocabulary acquisition, dynamic assessment, and learning issues.

Fatemeh Behjat is an assistant professor in TEFL and a faculty member at Islamic Azad University, Iran. Her areas of interest include applied linguistics, language teaching and learning issues in particular.

Mohammad Sadegh Bagheri is an assistant professor in TEFL. Currently, he is a faculty member at the English Department, Islamic Azad University, Shiraz, Iran. His main fields of interest are international exams (IELTS, TOEFL), teacher training, teacher development, assessment and research.

PUBLIC INTEREST STATEMENT
This study investigated the effects of two reading-based L2 vocabulary tasks on EFL learners’ learning and retention of 40 target words. The two tasks were evaluated and scored according to the theoretical framework of the Technique Feature Analysis proposed by Nation and Webb (2011). To this end, 66 advanced EFL learners were randomly assigned to two experimental conditions and one control group. The Participants in the summary writing (SW) condition were required to write a summary and in the oral reproduction (OR) condition to orally reproduce the content of the passages incorporating the target words only once. The participants in the control group only read the same passages and answered the true/false questions that followed the passages. Two types of vocabulary measures were used to assess learners’ learning and retention of the target words. The results indicated that both SW and OR incorporating the target words were effective tasks. However, OR was more effective than SW due to its higher task-induced generation.
Subjects: Language Teaching & Learning; Languages of the Middle East; Persian

Keywords: Involvement load hypothesis; oral reproduction; summary writing; task-induced generation; technique feature analysis; vocabulary tasks

1. Introduction

Vocabulary knowledge has always been an integral component of second or foreign language knowledge. The importance of lexical knowledge and second language vocabulary acquisition is widely acknowledged by language teachers, learners, and researchers. Studies have indicated that a learner’s word knowledge is a basic indicator and determiner of their language proficiency levels (Nation, 2001; Zou, 2016). It is generally believed that a large vocabulary is needed to function in a language. However, learning a sufficient amount of vocabulary is one of the biggest challenges facing language learners.

1.1. Task-based language teaching

Task-Based Language Teaching (TBLT) refers to the use of tasks as central units of planning and instruction in language teaching. TBLT is a logical development of Communicative Language Teaching since it advocates the inclusion of such principles in language teaching as promoting activities that involve “real communication” for language learning, using language for carrying out meaningful tasks to promote learning, and using language that is meaningful to the learner.

The key assumptions of TBLT are summarized by Feez (1998) as follows:

The focus is on process rather than product, basic elements are purposeful activities and tasks that emphasize communication and meaning, learners learn language by interacting communicatively and purposefully while engaged in the activities and tasks, activities and tasks can be either pedagogical or real-life tasks, activities and tasks of a task-based syllabus are sequenced according to difficulty, the difficulty of a task depends on a range of factors including the previous experience of the learner, the complexity of the task, the language required to undertake the task, and the degree of support available. (p. 17)

In spite of the fact that TBLT was motivated primarily by a theory of learning rather than a theory of language, one of the major assumptions concerning its theory of language asserts that lexical units are central in language used and language learning. Thus, vocabulary is considered to play a pivotal role in language learning than was traditionally assumed since L2 learners need to know a large vocabulary to carry out communicative tasks. Therefore, learning L2 vocabulary gains prominence in TBLT, and this perspective should be incorporated in designing, evaluating, and employing effective vocabulary learning tasks.

1.2. L2 vocabulary learning

While L1 learners acquire the majority of L1 words from massive exposure to spoken and written input and not by vocabulary instruction, L2 learners, especially in the classroom learning context, are not provided with the same amount of input and consequently need to perform effective vocabulary learning tasks to boost their L2 lexical knowledge. These tasks can take a variety of forms and are expected to lead to differential vocabulary gains based on the degree of involvement they induce. A number of L2 vocabulary researchers have indicated that L2 learners should not only pay deliberate attention to target words but also deeply process their different aspects in order to learn them effectively (Hu & Nassaji, 2012; Hulstijn & Laufer, 2001; Laufer, 2005; Nassaji, 2003, 2004; Nassaji & Hu, 2012; Schmidt, 2000). This refers to the concept of “elaborate processing” which has been considered as an essential factor in L2 vocabulary learning.

The concept of “elaborate processing” was originally introduced by Craik and Lockhart (1972) in their “depth of processing” model. The depth of processing model suggests that the degree to which new information is retained and stored in long-term memory depends on how the
information is processed. In this model, elaboration is the key to learning and retention of vocabulary. In their revised version, Lockhart and Craik (1990) further expanded those ideas by highlighting at least two stages for effective learning: an input analysis stage whereby sensory features, such as orthographic and phonological features of word forms, are analyzed and a retrieval stage in which semantic and conceptual features are retrieved with deeper analysis (Eckerth & Tavakoli, 2012). In this model, not only are initial attention, noticing and processing of words essential but also their subsequent retrieval and consolidation of the semantic encoding of the word features in memory are critical for learning. In the context of L2 vocabulary learning, there are two theoretical frameworks to operationalize the concept of “depth of processing”.

The first framework is referred to as the Involvement Load Hypothesis (ILH) (Lafer & Hulstijn, 2001) and the second one as the Technique Feature Analysis (TFA) (Nation & Webb, 2011). These two frameworks differ in the way they conceptualize the depth of processing and in the parameters they propose for elaboration learning. The ILH stimulated a body of research on designing and comparing L2 vocabulary tasks which tried to test the hypothesis and provide empirical evidence for its validity (Eckerth & Tavakoli, 2012; Hulstijn & Lafer, 2001; Keating, 2008; Kim, 2008; Lee & Pulido, 2017; Zou, 2017).

In spite of the fact that a body of research has been conducted on the most effective L2 vocabulary tasks within the theoretical framework of the ILH, very little research is conducted to evaluate, refine, and improve the predictive power of the TFA in terms of designing the most effective L2 vocabulary learning tasks. In addition, very little is known about the potential effects of the summary writing task in comparison with the oral reproduction task incorporating the target words on L2 vocabulary development within the theoretical framework of the TFA. Consequently, the present study aims at filling this gap in the literature through investigating the effects of two reading-based L2 vocabulary tasks, namely reading plus summary writing and reading plus oral reproduction incorporating the target words, on advanced EFL learners’ vocabulary learning and retention.

2. Literature review
As mentioned earlier, there are currently two frameworks to operationalize the concept of “depth of processing” in the context of L2 vocabulary learning: the Involvement Load Hypothesis and the Technique Feature Analysis. These frameworks differ in the way they operationalize the “depth of processing” and also the components and the criteria they propose for lexical elaboration. As a result, they allocate different weights to each attentional component leading to variations in predicting the effectiveness of L2 vocabulary tasks. Each framework will be described in the following sections.

2.1. The involvement load hypothesis
The concept of depth of processing in the realm of L2 vocabulary acquisition was first operationalized by Lafer and Hulstijn (2001) as the construct of task-induced involvement or the Involvement Load Hypothesis (ILH). The ILH comprises the three components of need, search, and evaluation. The need component is the motivational, non-cognitive dimension of the hypothesis, which concerns the need to achieve. Need can be moderate or strong. Need is moderate if it is externally imposed by the teacher (e.g., the teacher requires the learner to find the meaning of a new word) and is strong when it is intrinsically motivated or self-imposed by the learner (e.g., the learner decides to look up the meaning of a word in a dictionary). Search and evaluation are the two cognitively-oriented (information processing) components of the ILH. Search is the attempt to find the meaning of a new word and can take two forms: moderate or strong. Search is moderate when it is receptive retrieval, as when a learner has to look for or retrieve the meaning of a word, and is strong when it is productive retrieval, as when a learner needs to find the L2 word form for expressing a concept (e.g., trying to find the L2 translation of an L1 word) by consulting a dictionary or his/her teacher. Evaluation entails comparison of a given word with other words, or the specific meaning of a word with its other meanings, or combining the word with other words to assess whether a word fits or does not fit its original context. Evaluation can be moderate or strong. It is moderate if the learner has to compare the meaning of a specific word with its other
meanings, and is strong if the learner needs to assess whether a word’s meaning fits a specific linguistic context or not. The major contention and assumption of the ILH is that vocabulary learning tasks which are designed in such a way that induce higher degrees of involvement are expected to result in higher degrees of retention of new words. Table 1 illustrates some vocabulary learning activities with their task-induced involvement load.

As Table 1 presents, “writing a composition” task, either when concepts are selected by the teacher or when they are selected and looked up by L2 learner-writers, is supposed to induce the highest degrees of involvement load and will result in better vocabulary learning.

To examine the predictions of the ILH in terms of the efficacy of L2 vocabulary tasks, some researchers investigated and compared the effectiveness of several L2 vocabulary learning tasks which induced differential degrees of task-induced involvement load (Eckerth & Tavakoli, 2012; Folse, 2006; Hulstijn & Laufer, 2001; Keating, 2008; Kim, 2008; Lee & Pulido, 2017; Min, 2008; Sun, 2017; Zou, 2017).

In an experimental study conducted to provide empirical evidence for the usefulness of the ILH in designing effective L2 vocabulary learning tasks, Hulstijn and Laufer (2001) explored the effects of three vocabulary learning tasks on vocabulary learning by advanced Dutch- and Hebrew-speaking learners of English. Their tasks included reading comprehension with marginal glosses, reading comprehension plus fill-in, and composition writing while incorporating target words. Need and search components of the ILH were held constant in the three tasks. Since need was induced by the task, it was moderate (1), and as glosses were provided, search was absent (0). Therefore, the ILH index of tasks varied only in terms of the degree of evaluation that they induced. The ILH indices of Tasks 1, 2, and 3 were 1, 2, and 3, respectively. They found that Task 3, that is, “composition writing” was superior to Tasks 1 (reading with marginal glosses) and 2 (reading plus fill-in). In addition, Task 3 was only superior to Task 1 in the Hebrew-English experiment.

Contradictory results were found by Folse (2006) who investigated the effects of three types of written exercises on L2 vocabulary retention: (a) one fill-in-the-blank exercise (b) three fill-in-the-blank exercises, and (c) one original sentence-writing exercise. He found that the number of

| Tasks | Status of target words | Need | Search | Evaluation |
|-------|------------------------|------|--------|------------|
| 1. Reading and comprehension questions | Glossed in text but irrelevant to task | – | – | – |
| 2. Reading and comprehension questions | Glossed in text and relevant to task | + | – | – |
| 3. Reading and comprehension questions | Not glossed but relevant to task | + | + | –/+ (depending on word and context) |
| 4. Reading and comprehension questions and filling gaps | Relevant to reading comprehension. Listed with glosses at the end of text. | + | – | + |
| 5. Writing original sentences | Listed with glosses | + | – | + + |
| 6. Writing a composition | Concepts selected by the teacher (and provided in L1). The L2 learner-writer must look up the L2 form | + | + | + |
| 7. Writing a composition | Concepts selected (and looked up) by L2 learner-writer | + | + | + |

Table 1. Task-induced involvement load (adapted from Laufer & Hulstijn, 2001, p. 18)
retrievals, induced by the particular task for each word when carrying out an L2 vocabulary exercise, overrode the depth of processing for that word. His findings also suggested the value of an L2 vocabulary exercise that required multiple retrievals of or encounters with the target words. These findings were in line with the psycholinguistic and educational psychology studies on *rehearsal* (Baddeley, 1990) and *distributed practice* (Atkins & Baddeley, 1998). In fact, Fols’s study showed that doing multiple retrievals of a target word, regardless of the superficiality of that activity, was more effective than the depth of processing and the task-involvement load in L2 vocabulary learning.

In a similar study, Keating (2008) investigated the potential effects of L2 vocabulary learning tasks, which induced differential degrees of involvement load indices, on learners’ vocabulary uptake. The tasks included reading comprehension with marginal glosses, reading comprehension plus fill-in, and writing original sentences using target words. The ILH index Task 1 was 1 (moderate need). The task induced no search or evaluation. The ILH index of Task 2 was 2 since need was moderate, search was absent, and evaluation was moderate. Task 3 induced moderate need (1), no search (0), and strong evaluation (2) because the learners had to evaluate the appropriateness of the target words in writing original sentences. Therefore, Task 3 induced the highest ILH index of 3 among the three tasks. The results of Keating’s study indicated that learners who completed Tasks 2 and 3 outperformed learners who completed Task 1. In terms of learners’ passive knowledge of the target words, Tasks 2 and 3 resulted in higher retention scores than Task 1. Regarding learners’ active word knowledge, the results fully supported the ILH in learning the target words.

Kim (2008) obtained similar results from comparing reading comprehension with marginal glosses, with the ILH index of 1 (moderate need), reading comprehension with marginal glosses plus gap-fill, with the ILH index of 2 (moderate need, moderate evaluation), and writing a composition and incorporating the target words, with the ILH index of 3 (moderate need, strong evaluation). The findings of the study provided partial support for the ILH in that learners acquired more vocabulary through tasks which induced a higher level of involvement. Furthermore, the results indicated that the two tasks of “composition writing” and “writing original sentences”, which had identical involvement load indices (moderate need, no search, and strong evaluation), were equally effective in promoting both the initial learning and long-term retention of target lexical items. Learners’ level of language proficiency in Kim’s study seemed to have no interaction effect on task type effectiveness.

The potential of three types of reading-based written output activities was investigated by Rassaei (2015) to see whether different types of written output tasks would lead to differential L2 vocabulary gains. Among the three output activities of “reading plus generating comprehension questions and answering them incorporating target words”, “reading plus summarizing the text incorporating target words”, and “reading plus making predictions about what would come next in the story incorporating target words”, “reading plus prediction” condition outperformed the other conditions in terms of both immediate learning and long-term retention of target lexical items. Rassaei concluded that written output activities that involve creative writing were more beneficial for L2 vocabulary learning than tasks which only require reconstructive production.

Other L2 vocabulary researchers focused on examining the contribution of each component of the ILH and evaluated the degree to which each component contributed to L2 vocabulary learning. In this regard, Lee and Pulido (2017) examined the motivational factor of the ILH, that is, the need factor, by investigating the effects of topic interest, L2 proficiency, and gender on L2 vocabulary acquisition through reading. They found that, in line with the predictions of the ILH, topic interest significantly affected incidental vocabulary acquisition. In other words, learners’ level of interest in reading topics positively influenced the involvement load and facilitated word learning. Furthermore, it was shown that learners’ L2 reading proficiency was a strong predictor of vocabulary learning through reading. In addition, gender difference was found to have no significant effect on lexical development through reading in an L2. However, a significant interaction was found between gender and topic interest in word form recognition.
Investigating the allocation of involvement load to the evaluation component of the ILH, Zou (2017) examined how three approaches to evaluation, namely “cloze exercises”, “sentence writing”, and “composition writing” promoted word learning. The results of Zou’s study partially supported the predictions of the ILH in that the two reading-based written output activities, that is, “sentence writing” and “composition writing”, were found to be more effective than “cloze exercises” in promoting L2 word knowledge. However, these two more effective activities which had been hypothesized to induce the same degree of involvement load, and hence to result in equal vocabulary uptake, yielded differential gains. More specifically, “composition writing” task led to significantly better word learning than “sentence writing”. These findings were justified in terms of different organization methods and degrees of pre-task planning of the three tasks.

Nguyen and Boers (2018) investigated the potential benefits of a sequence “input-output-input” activity on L2 vocabulary uptake. They compared two groups of EFL learners who watched a TED Talk video twice. However, only participants in the experimental group were asked to sum up the content of the video, that is, to produce output, based on what they had watched immediately after the first exposure to the input. Immediate and delayed post-tests showed significantly better word-meaning recall in the “input-output-input” condition. It was also revealed that words that learners had attempted to use in their oral summaries stood a good chance of being recalled later. These findings are line with the predictions of Swain’s (1995) Output Hypothesis, Laufer and Hulstijn (2001) Involvement Load Hypothesis, and Nation and Webb (2011) Technique Feature Analysis.

Since it was revealed that vocabulary tasks with the same ILH index resulted in differential vocabulary gains by L2 learners and that the ILH did not include many other factors that other studies had demonstrated to be important when designing vocabulary learning techniques, Nation and Webb (2011) built on the ILH and proposed a more comprehensive framework to operationalize the depth of processing model. The second framework is referred to as the Technique Feature Analysis (TFA).

2.2. The technique feature analysis

Technique Feature Analysis (TFA) is a theoretical framework proposed by Nation and Webb (2011) to more precisely operationalize the concept of depth of processing in L2 vocabulary learning. TFA consists of 5 components and 18 criteria to assess each component. The five components, which are considered as essential to any vocabulary learning technique, are Motivation, Noticing, Retrieval, Generation, and Retention. Each component has several criteria to render the five components quantifiable and to make different vocabulary learning techniques comparable according to their TFA total scores. Motivation refers to whether the vocabulary learning activity is motivating enough for learners to do. Noticing refers to learners’ attention to and awareness of new vocabulary items to be acquired in addition to negotiation of target words. Retrieval refers to whether the vocabulary technique requires learners to recognize or recall the target lexical items as well as requiring multiple retrievals and spacing between each retrieval of the target word. Generation refers to the fact that words that are met in new ways (receptive generative use) or used in ways that learners have not met before (productive generative use) strengthens memory for the word (Joe, 1998). According to Nation and Webb (2011), productive generative use is more demanding than receptive generative use. If a vocabulary learning technique requires generative productive use, the degree of generation can be evaluated against a scale of generativity. Joe (1998) devised a scale of generativity for techniques that induce various degrees of generation. These include no generation (repeating what is in the text), low generation (small grammatical or inflectional changes), reasonable generation (used with new collocations or remarkable grammatical changes), and high generation (elaborating or stretching the meaning, using collocations that stretch meaning, and manipulating derivational affixes). The last component of the TFA is Retention, which refers to whether the vocabulary technique ensures successful linking of form and meaning, the involvement of seeing an instance of the new word in context, the provision of a mental or visual image of the new word, and avoiding interference through teaching words that are not members of the same lexical set. Table 2 presents the components and related criteria of the TFA.
Table 2. A checklist for technique feature analysis (adapted from Nation & Webb, 2011, p. 7)

| Criteria                                              | Scores |
|-------------------------------------------------------|--------|
| Motivation                                            |        |
| Is there a clear vocabulary learning goal?            | 0 1    |
| Does the activity motivate learning?                  | 0 1    |
| Do the learners select the words?                     | 0 1    |
| Noticing                                              |        |
| Does the activity focus attention on the target words?| 0 1    |
| Does the activity raise awareness of new vocabulary learning? | 0 1 |
| Does the activity involve negotiation?                | 0 1    |
| Retrieval                                             |        |
| Does the activity involve retrieval of the word?      | 0 1    |
| Is it productive retrieval?                          | 0 1    |
| Is it recall?                                         | 0 1    |
| Are there multiple retrievals of each word?           | 0 1    |
| Is there spacing between retrievals?                  | 0 1    |
| Generation                                            |        |
| Does the activity involve generative use?             | 0 1    |
| Is it productive?                                     | 0 1    |
| Is there a marked change that involves the use of other words? | 0 1 |
| Retention                                             |        |
| Does the activity ensure successful linking of form and meaning? | 0 1 |
| Does the activity involve instantiation?              | 0 1    |
| Does the activity involve imaging?                    | 0 1    |
| Does the activity avoid interference?                 | 0 1    |
| Maximum score                                         | 18     |

Hu and Nassaji (2016) compared the two frameworks of ILH and TFA to see which one had a stronger predictive power in determining more effective vocabulary learning tasks. They employed four vocabulary learning tasks which had been ranked differently by the two frameworks. The tasks included reading a text with multiple-choice items, reading a text and choosing definitions, reading plus fill in the blanks, and reading and rewording the sentences. Task 1 had an ILH index of 3 and a TFA score of 6; task 2 had an ILH index of 3 and a TFA score of 6; task 3 had an ILH index of 2 and a TFA score of 7, and task 4 had an ILH index of 3 and a TFA score of 6. Four groups of EFL learners were required to read a text containing 14 unknown target words, and then carry out their designated tasks. The findings suggested that all four tasks led to L2 vocabulary uptake by learners. In addition, “reading plus fill in the blank” task, which had a higher TFA score led to better task performance than the other tasks. In terms of long-term retention of new L2 words, task 4 which entailed productive generation by requiring the participants to rewrite the target sentences yielded the best results. Overall, the results indicated that TFA had a stronger predictability power than the ILH in vocabulary learning. In particular, TFA appeared to suggest more sensitive factors measuring the potential effectiveness of vocabulary learning tasks than the ILH.

Zou et al. (2018) examined the effectiveness of vocabulary learning tasks from the perspective of Nation and Webb (2011) Technique Feature Analysis. Three word learning tasks (reading comprehension, cloze-exercises, and sentence writing) and two types of annotations for target words (pictorial annotations, and textual annotations) were investigated. Participants were required to complete four tasks: reading comprehension with pictorial annotations, cloze-exercises with textual annotations, cloze-exercises with pictorial annotations, and sentence-writing with textual annotations.
annotations. The post-test scores indicated that the tasks of reading comprehension with pictorial annotations and doing cloze-exercises with textual annotations were similarly effective and cloze exercises with pictorial annotations were similarly effective as sentence-writing with textual annotations. The results were consistent with the predictions of the TFA, indicating that this framework was reliable in evaluating and predicting task effectiveness. It was also found that the inclusion of imaging in a vocabulary learning task was conducive to word learning, and pictorial annotations promoted effective learning.

As the above literature suggests, vocabulary tasks which require learners to read a text with new lexical items and carry out written output activities such as writing original sentences using target words, composition writing incorporating the target words, predicting what would come next in the story, and fill in the blank exercises promoted L2 vocabulary learning the most. This can be explained in terms of the higher amount of involvement loads they induce and the deeper cognitive processing and elaboration they require. However, the need for more research on effective L2 vocabulary tasks is felt since previous studies on L2 vocabulary tasks mainly investigated the effects of written activities on vocabulary learning. Consequently, the potential of oral output incorporating target words in promoting L2 vocabulary development is under-researched. Furthermore, it is not clear which of the two techniques of vocabulary learning, that is, oral reproduction and summary writing incorporating target words, is more conducive to L2 vocabulary development. Investigating the potential effects of these two vocabulary learning techniques which are evaluated and scored based on the TFA can also contribute to the development and modification of the TFA and increase the predictive power of the framework in terms of designing the most effective L2 vocabulary learning tasks. To this end, the following research questions were formulated:

1. Does reading a passage plus writing a summary incorporating target words promote advanced EFL learners’ learning and retention of target lexical items?
2. Does reading a passage plus oral reproduction using target words promote advanced EFL learners’ learning and retention of target lexical items?
3. Which of the two reading-based L2 vocabulary learning tasks, that is, summary writing or oral reproduction, is more effective for promoting advanced EFL learners’ L2 vocabulary development?

3. Method

3.1. Participants

A total of 66 advanced Iranian male EFL learners participated in this study who were all native speakers of Farsi and had 5 years of English language education. They all came from the same sociocultural background whose age ranged from 16 to 18. They were selected from a population of 90 EFL learners, who were studying general English in a major English language institute at the advanced level. During their English classes at the institute, the participants studied a new unit’s lexical items, read the following conversation or the reading passage while their language teacher explained the meaning of new words and paraphrased the conversation or the reading sentences for them. In addition, they performed listening, writing, and grammar tasks following the instruction of grammar and writing by their teacher.

Based on their scores on Oxford Quick Placement Test (2001). Based on the results of the OPT, 24 learners whose scores were under 48 (2 SDs below the Mean) were discarded from the study. The results of the Oxford Placement Test (OPT) are presented in Table 3.

| Table 3. Descriptive statistics for learners' scores on OPT |
|------------------------------------------------------------|
| N         | Mean | Std. deviation |
| Statistic | 66   | 52.06          | 1.508  |
With reference to the Common European Framework of Reference (CEFR), the participants of this study fell into C1 level who had Effective Operational Proficiency. According to CEFR guidelines, C1 users can understand a wide range of demanding, longer texts, and recognize implicit meaning, express him/herself fluently and spontaneously without much obvious searching for expressions, use language flexibly and effectively for social, academic and professional purposes, produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors and cohesive devices.

CEFR is a guideline for describing levels of achievement in language learning for foreign languages across Europe, including English, developed by the European organization the Council of Europe. It includes six levels of achievement divided into three broad divisions, which describe what a learner should be able to do in reading, listening, speaking and writing at each level. It is intended to provide a common basis for describing communicative performance and to serve as a basis for developing language syllabuses, curriculum guidelines, examinations, and textbooks, regardless of the target language. (Richards & Schmidt, 2010). Table 4 displays levels of CEFR.

Following participants’ consent to participate in the study, they were randomly assigned to two experimental groups and one control group. Therefore, there were 22 participants in each group. They attended their English classes twice a week for an hour and forty-five minutes.

3.2. Design
This study employed a pre-test, post-test, and delayed post-test experimental design. Following the administration of the Oxford Placement Test, a vocabulary pre-test of 40 multiple-choice items was given to all participants to ensure that all the target words are unknown to them and that they are homogeneous in terms of their L2 vocabulary knowledge. Following the vocabulary pre-test, participants were randomly assigned to two experimental conditions and one control group. Participants in the two experimental conditions and the control group read one reading passage each session twice a week during 4 weeks of treatment and performed the designated post-reading task in the class. Two immediate post-tests were administered one day after the completion of the treatment sessions and two delayed-post-tests were administered two weeks later. Figure 1 displays the design of the study.

3.3. Identification of target words
A total of 40 target words were selected from eight reading passages based on the following criteria. First, the knowledge of selected target words was essential for understanding the content of passages. In other words, target lexical items were selected in such a way that the learners could not fully comprehend the content of passages by skipping the unknown target words. Second, learners were given a list of 50 target words and were asked to provide a written L1 translation or an L2 synonym or definition for each word. Learners’ answers revealed that from the initial 50 target words, 40 words were completely unfamiliar to the learners. Third, a multiple-choice recognition pre-test of these 40 words was administered to all participants to make sure that all the selected target words were unknown to the participants. Table 5 presents the 40 target words.

| Table 4. Levels of CEFR |
|-------------------------|
| **Division**           | **Level** | **Level name** |
| Proficient user        | C2        | Mastery        |
|                        | C1        | Effective Operational Proficiency |
| Independent user       | B2        | Vantage        |
|                        | B1        | Threshold      |
| Basic user             | A2        | Breakthrough   |
|                        | A1        | Waystage       |
3.4. Reading materials

The reading materials were eight passages taken from *The IL1 English Series: Advanced 1 student’s book* (Iran Language Institute, 2006), which was the advanced EFL learners’ coursebook. The reading passages included 730–952 words and were the same for all the three groups. They were selected based on the degree to which they facilitated learners’ tasks of writing a summary and reproducing the content orally. In each reading passage, five target words the knowledge of which was essential to the comprehension of the main ideas of the passage were carefully selected and typed in boldface followed by their L1 translations in parentheses. Therefore, both experimental conditions, namely, reading plus oral reproduction and reading

| Table 5. Target lexical items |
|-------------------------------|
| **Verbs** | **Nouns** | **Adjectives** |
| Proliferate acquiesce | Cardiologist exodus | Oblivious viable |
| Outperform usurp | Moelstrom crisis | Unprecedented |
| Advocate stereotype | Funeral struggle | Equitable alienated |
| Feminize deprive | Manifesto literacy | Inevitable anecdotal |
| Segregate retreat | Demographer sojourn | Simultaneous |
| Embark on gauge | Adjustment sabbatical | Tangible vital |
| Capsize suffocate | Futility | |
| Subside regret teem | | |
Two readability indices were calculated and reported for the passages used in this study: Flesch Reading Ease Readability and Flesch–Kincaid Grade Level Readability. Flesch Reading Ease Readability is a formula proposed by Flesch (1948), which rates the readability of passages on a 100-point scale. The higher the score, the easier the passage and the lower the score, the more difficult the passage to read and understand. Flesch–Kincaid Grade Level Readability rates texts on a U.S. school grade level. For example, a score of 9 means that a 9 grader can understand the document. Both readability indices were calculated using Microsoft Office word-processing software. The Flesch Reading Ease Readability indices of these passages were between 37.9 and 67.2, and Flesch–Kincaid Grade Level Readability indices of these passages were between 8.2 and 12.0. It can be concluded that the passages that were used in this study were fairly long and mostly difficult or very difficult which suited advanced learners’ level of language proficiency.

3.5. Procedures of the two instructional conditions

As mentioned earlier, the purpose of the current study was to investigate the effects of two reading-based L2 vocabulary learning tasks, namely, oral reproduction and summary writing incorporating target words, on learning and retention of 40 target words by Iranian advanced male EFL learners. To this end, participants in each experimental condition were provided with eight reading passages each of which included five target words and were then required to perform their designated tasks using the target words.

The two vocabulary tasks in this study were analyzed and scored according to the criteria and elaboration parameters proposed by the TFA framework and, based on the total TFA score of each task, were hypothesized to induce the same amount of involvement load and consequently to result in the same degree of L2 vocabulary learning and retention. Table 7 illustrates the feature analysis of the two vocabulary tasks used in this study.

3.5.1. Task 1. Oral reproduction (OR)

During the eight treatment sessions, participants of the OR condition read one passage of between 700 and 800 words each session and were asked to summarize the content of the passage orally while incorporating the 5 target words in each reading passage. Therefore, they read eight different passages on a variety of topics during the treatment sessions. In the first treatment session, participants were instructed by their language teacher how to present the oral summary of a reading passage through understanding the main ideas of the passage. To model the required performance, a passage similar to the reading materials was distributed among the learners in which five words had been typed in boldface. The learners were asked to read the passage but were not required to orally reproduce the content of the passage at this stage. However, their language teacher presented the oral summary of the passage to the learners and incorporated the boldfaced words. Following the language teacher’s model performance, the first reading passage was distributed among the learners to read. They were not allowed to consult their dictionaries to look up the target words during reading since the L1 translations of target words were provided in parentheses. Upon finishing reading the passages, reading papers were collected by the teacher and learners were asked to present their oral summaries. It was emphasized that it was essential that they use the stipulated words only once in their oral productions. To control for the confounding effect of hearing a target word several times from other participants during their oral productions, each participant was asked to present his oral summary privately only to the teacher while other learners could not hear him summarize. Consequently, participants encountered each target word once in the reading passage and used it once in their oral reproduction activity. Since Hulstijn and Laufer (2001) recommended that the time needed to complete a vocabulary task be considered to be an inherent property of the task by, participants were given enough time to present their oral summaries. Participants’ oral reproductions
| Passages | Passage 1 | Passage 2 | Passage 3 | Passage 4 | Passage 5 | Passage 6 | Passage 7 | Passage 8 |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|
| Word Count | 730 | 859 | 946 | 780 | 851 | 952 | 880 | 888 |
| Flesch Reading Ease Readability | 47.8 | 40.2 | 37.9 | 67.2 | 58.4 | 47.7 | 64.3 | 42.2 |
| Flesch-Kincaid Grade Level Readability | 11.0 | 12.0 | 12.0 | 8.2 | 9.4 | 10.3 | 9.1 | 12.0 |

Table 6. Word count, Flesch Reading Ease Readability, and Flesch-Kincaid Grade Level Readability of the passages
were recorded using a voice recorder for further analysis which revealed that the mean length of participants’ oral reproductions was 340 words.

3.5.2. Task 2. Summary writing (SW)
Participants in the SW condition read the same reading passages as participants in the OR group did. Similar to the first treatment session for the OR condition in which learners were instructed about how to do the required performance, the language teacher instructed the learners with regard to how they can write a summary of a passage they read. Again, a sample reading passage similar in length and readability to treatment passages in which five target words had been typed in boldface was distributed among the learners to read. When the learners finished reading the passage, the language teacher provided the learners with instructions as to how they can write a summary of a passage using certain summarizing strategies including reading the passage for the main ideas, identifying the main ideas of the passage, skipping trivial or redundant information, combining the main ideas to reconstruct the main points of the passage, and finishing their writing with a concluding sentence. To further clarify the required performance, the language teacher put the summary outline of the distributed passage on the board and recommended that learners follow the same procedure in writing their summaries while incorporating the target words. Following the distribution of the first reading passage in which five target words were typed in boldface along with their L1 translations in parentheses, the language teacher asked the learners to write a summary between 150 and 250 words on their worksheets and include the target words only once. Thus, participants encountered the target words once in the reading passage and used them once in their written productions. No time limit was set for writing the summary and learners were given enough time to complete their

| Criteria                                      | OR | SW |
|-----------------------------------------------|----|----|
| Is there a clear vocabulary learning goal?    | 0  | 0  |
| Does the activity motivate learning?          | 1  | 1  |
| Do the learners select the words?             | 1  | 1  |
| Does the activity focus attention on the target words? | 1  | 1  |
| Does the activity raise awareness of new vocabulary learning? | 1  | 1  |
| Does the activity involve negotiation?        | 0  | 0  |
| Does the activity involve retrieval of the word? | 1  | 1  |
| Is it productive retrieval?                   | 1  | 1  |
| Is it recall?                                 | 1  | 1  |
| Are there multiple retrievals of each word?   | 0  | 0  |
| Is there spacing between retrievals?          | 0  | 0  |
| Does the activity involve generative use?     | 1  | 1  |
| Is it productive?                             | 1  | 1  |
| Is there a marked change that involves the use of other words? | 0  | 0  |
| Does the activity ensure successful linking of form and meaning? | 1  | 1  |
| Does the activity involve instantiation?      | 0  | 0  |
| Does the activity involve imaging?            | 0  | 0  |
| Does the activity avoid interference?         | 1  | 1  |
| Total score                                   | 11 | 11 |
writing task. As suggested by Hulstijñ and Laufer (2001) and also Webb (2005), the time needed to perform a task is considered as an inherent property of the task itself and therefore the learners were given enough time that they needed to complete the summary writing task. The analysis of learners’ writings showed that they had used the target vocabulary only once and the mean length of their productions was 190 words.

3.5.3. Reading only (RO)
Participants in the RO condition also read the same eight passages as did the two experimental conditions during their normal class sessions. The five target words in each reading passage had been typed in boldface along with their L1 translations in parentheses. Having read each passage, participants in the RO condition answered 10 true/false items in which the target words appeared only once. As a result, participants in the RO group were exposed to the 40 target vocabulary only twice; once in the reading passage and once in the post-reading exercise they did.

To measure learners’ immediate learning and long-term retention of the 40 target vocabulary in both OR and SW conditions as well as in the RO condition, two unannounced vocabulary immediate post-tests were administered one day after the last treatment session and two unannounced vocabulary delayed post-tests were given two weeks after the last treatment session to ensure that learners’ vocabulary gains as demonstrated on vocabulary measures would be the result of incidental vocabulary learning.

3.6. Vocabulary measures
A 40-item multiple-choice recognition pre-test was administered to all participants to make sure that the 40 target words were unfamiliar to all learners and that participants of the three conditions were homogeneous in terms of target vocabulary knowledge. In addition, two types of vocabulary tests were developed and administered as immediate and delayed post-tests to measure advanced L2 learners’ learning and retention of 40 target vocabulary including a 40-item multiple-choice active recognition (MCAR) test (Appendix 1), and a 40-item cued response active recall (CRAR) test (Appendix 2). Following Laufer and Razovski-Roitblat (2015), in the MCAR test, participants were presented with an English sentence in which the target word was missing from the sentence and the L1 translation of the target word was provided as a prompt in parentheses. Learners were required to select the correct item among four options in L2. For example:

Computer courses continue to ______ (کورس کامپیوتر کنات) in many different places.

a) procrastinate  b) inundate  c) acknowledge  d) proliferate

In the CRAR test, participants were required to supply the L2 target word form for its L1 translation equivalent while the first letter of the L2 target word form was provided as a cue (Laufer & Razovski-Roitblat, 2015). For example:

Members of the security council of the UN passed a resolution to prevent nuclear weapons from ______ (کنترل اتمی)

The same two types of vocabulary tests were used as immediate and delayed post-tests. However, the order of items was re-arranged and re-ordered to control for the testing effect and to prevent participants from learning from the previous tests.

3.7. Data analysis
Four scores were obtained using the two types of immediate and delayed MCAR and CRAR vocabulary tests for each learner. For the MCAR test, 1 point was given if the learner selected the correct choice among the four options and wrong or no answers were given a zero. With regard to the CRAR test, the correct L2 target word form supplied was given 1 point. In addition, close
approximations to the target words which were semantically understandable or slightly misspelled but still recognizable were also given 1 point. For example, the provision of “proliferation” instead of the correct word “proliferation” was considered as a correct answer and received 1 point. The maximum score each learner could obtain on each of the four tests was 40. Learners’ answer sheets on the CRAR tests were rated by two experienced language teachers and the correlation coefficient between the two ratings indicated 0.95 interrater reliability.

The collected data were submitted to SPSS software (Version 24) for statistical analysis. A one-way between-groups analysis of variance (ANOVA) was conducted to explore the effects of post-reading SW and OR incorporating target words on advanced EFL learners’ learning and retention of 40 target lexical items as measured by MCAR and CRAR vocabulary tests.

4. Results

4.1. Descriptive statistics
Tables 8 and 9 present the descriptive statistics for learners' scores on MCAR and CRAR immediate and delayed vocabulary post-tests.

Since we observed differences in mean scores among the three groups in the MCAR and CRAR vocabulary tests, we conducted ANOVAs to further investigate if such differences were statistically significant or not followed by Scheffe’s post hoc tests to see where the differences among the groups existed.

4.2. Inferential statistics
Tables 10 and 11 display the inferential statistics for learners' scores on MCAR and CRAR immediate and delayed vocabulary post-tests.

As Table 10 displays, regarding learners' active recognition vocabulary knowledge as measured by MCAR on the immediate post-test, inferential statistics indicated that there was a statistically significant difference at the p < .05 level for the three conditions: F (2,63) = 378.90, p = .00. The effect size, calculated using eta squared, was .92, which according to Cohen’s (1988) guidelines, is a large effect size. In addition, Scheffe’s post hoc comparisons indicated that the mean scores for

| Summary writing | Immediate post-test | Delayed post-test |
|-----------------|---------------------|------------------|
|                 | N  | Mean | Std. deviation | Mean | Std. deviation |
| Reading only    | 22 | 34   | 2.1            | 30  | 2.2             |
| Oral reproduction | 22 | 36   | 1.6            | 34  | 1.8             |

| Summary writing | Immediate post-test | Delayed post-test |
|-----------------|---------------------|------------------|
|                 | N  | Mean | Std. deviation | Mean | Std. deviation |

| Reading only    | 22 | 15   | 2.2            | 10  | 1.8             |

For CRAR tests, the ANOVA results showed statistically significant differences at the p < .01 level for the three conditions: F (2,63) = 188.80, p = .00. The effect size, calculated using eta squared, was .89, which is a very large effect size. Scheffe’s post hoc comparisons indicated that the mean scores for
both SW and OR conditions were significantly different from RO condition and both experimental conditions outperformed the control group in terms of learning the target vocabulary. Furthermore, the OR condition significantly outperformed the SW condition.

Regarding learners’ retention of the target words as measured by the MCAR delayed post-test, there was a statistically significant difference at the $p < .05$ level for the three groups: $F(2,63) = 551.83$, $p = .00$. The effect size, calculated using eta squared, was .94, which is a large effect size.

In addition, Scheffe’s post hoc comparisons revealed that both SW and OR experimental conditions significantly outperformed the RO condition. Similar to the results for the MCAR immediate post-test, OR condition significantly outperformed the SW condition.

With regard to learners’ active recall vocabulary knowledge as measured by CRAR on the immediate post-test, inferential statistics indicated a statistically significant difference among the three groups at the $p < .05$ level for the three conditions: $F(2,63) = 517.37$, $p = .00$. The effect size, calculated using eta squared, was .94, which is a large effect size. Furthermore, Scheffe’s post hoc comparisons indicated that both SW and OR conditions outperformed the RO condition in terms of active recall vocabulary knowledge. In addition, OR condition was found to be more conducive to vocabulary learning than the SW condition.

In terms of learners’ retention of target vocabulary as measured by CRAR delayed post-test, the analysis indicated a statistically significant difference at the $p < .05$ level for the three instructional conditions: $F(2,63) = 646.80$, $p = .00$. The effect size, calculated using eta squared, was .95, which is considered a large effect size. Further analysis using Scheffe’s post hoc comparisons indicated that, similar to the results for CRAR immediate post-test, both SW and OR instructional conditions outperformed the RO condition while the OR condition outperformed the SW condition. Table 11 summarizes the inferential statistics for learners’ scores on CRAR immediate and delayed vocabulary post-tests.
5. Discussion
The current study aimed at investigating the effects of two reading-based L2 vocabulary learning tasks, namely, reading plus summary writing incorporating target vocabulary (SW) and reading plus oral reproduction of the content of reading passages incorporating target words (OR), on advanced EFL learners’ learning and retention of 40 target words. The two vocabulary tasks were analyzed and scored according to the components and criteria of the TFA framework and were supposed to result in equal vocabulary gains by advanced EFL learners since they both had the same TFA score of 11.

The first research question of the current study asked if reading a passage followed by writing a summary incorporating target words promotes advanced EFL learners’ immediate learning and long-term retention of target lexical items. The results of one-way ANOVA along with Scheffe’s post hoc comparisons indicated that the SW condition significantly outperformed the RO (control group) condition on both MCAR and CRAR immediate and delayed vocabulary post-tests. In other words, it was revealed that writing a summary while including the target vocabulary following reading a passage is an effective vocabulary learning technique. The superiority of the SW task over RO condition can be justified in terms of the amount of “involvement” which is induced by the SW task. In particular, since the SW task triggered generative retrieval of target words and established form–meaning relationship, it seems to have required a deeper cognitive processing than merely reading a passage. In addition, learners had to evaluate the appropriateness of using the target words in a new context while writing a summary using new lexical items. Furthermore, the SW task pushed learners to produce comprehensible output, which is, according to Swain’s (2005) Output Hypothesis, promotes L2 acquisition. This finding is in line with the findings of previous studies on the positive effect of written output tasks using target words on learning new L2 vocabulary (Hulstijn & Lafer, 2001; Joe, 1998; Keating, 2008; Lafer, 2003; Pichette et al., 2012; Rassaei, 2015; Webb, 2005). However, it is in contrast with the findings of Folse (2006) who found that the frequency of word retrieval overrides the depth of word processing.

The second research question asked whether reading a passage followed by oral reproduction of the content of the passage including target words promotes advanced EFL learners’ learning and retention of target lexical items or not. The results of one-way ANOVA and Scheffe’s post hoc comparisons indicated that the OR condition significantly outperformed the RO condition in terms of immediate learning and long-term retention of target words. In other words, the OR task was found to be a more effective L2 vocabulary task than merely reading a passage. This finding is in line with the findings of Nguyen and Boers (2018) who found that words that learners had attempted to use in their oral summaries stood a good chance of being recalled later. Similarly, the superiority of OR task over the RO condition can be attributed to the higher amount of involvement load induced by orally reproducing the content of the passages and the depth of cognitive processing required by this output activity. Furthermore, learners in the OR condition were required to produce pushed output and were engaged in bottom-up processing which is an important component of the Output Hypothesis.

The last research question posed in this study asked which of the two vocabulary tasks, namely SW and OR, which had the same TFA score of 11, was more effective for L2 vocabulary learning and retention. Although both tasks had the same TFA score and were expected to lead to equal vocabulary gains by the advanced learners, the results of one-way ANOVA indicated that OR condition significantly outperformed the SW condition on both MCAR and CRAR immediate and delayed post-tests. In other words, the OR condition not only resulted in a higher amount of vocabulary learning on both aspects of recognition and recall knowledge than the SW task but was also more effective in terms of long-term retention of target words and the vocabulary gains by OR condition did not fade over time.

A point which merits further discussion concerns the possible reason(s) why the OR task was found to be a more effective vocabulary task than the SW task while both tasks had the same TFA score. According to the Nation and Webb (2011) theoretical framework of TFA, which was suggested to include more elaboration components and criteria in operationalizing the depth of
processing model in designing and evaluating vocabulary tasks, tasks with higher TFA scores are expected to result in more effective vocabulary learning. However, as we observed in the current study, the two supposedly as effective tasks of OR and SW incorporating target words yielded differential vocabulary uptake. The analysis of learners’ productions indicated that learners in the OR condition produced more words (340) than learners in the SW group (190). Meanwhile, learners in the OR condition spent more time presenting their oral summaries than learners in the SW condition writing their summaries. Consequently, this longer time-on-task needed by OR condition and their longer productions may suggest that this task entailed a deeper cognitive processing and rehearsal and that learners attempted to produce their oral output in such a way that it would be both meaningful and comprehensible. Furthermore, vocal production of target vocabulary while embedding them in a meaningful context might have positively affected learners’ memory for the new words and have significantly contributed to their superior word learning than the SW group. The positive effect of vocalization of target words on L2 vocabulary learning was also confirmed by Icht and Mama (2019) whose findings indicated that vocally produced words were more durable and showed less memory decay over time.

A further possible factor concerning the superiority of OR task over SW condition and the higher durability of new word learning and retention over time in the former condition in this study is that the two vocabulary tasks of OR and SW turned out to be different regarding the degrees of generation they stimulated. As mentioned earlier, the fourth elaboration component of the TFA concerns generation which can earn a maximum score of 3. According to the Generation criteria of the TFA, both tasks of OR and SW involve generative use of target vocabulary and are productive activities. Theoretically, both tasks earn a TFA generation score of 2 out of the maximum score of 3. However, as the analysis of learners’ productions revealed, since learners in the OR condition incorporated the target words in their output tasks along with some other words and phrases, which were not included in the reading passages they had read, in novel sentences whereas learners in the SW condition merely reconstructed previously read passages without the inclusion of some other words, the two tasks induced different degrees of generation. As a result, although both tasks involved generative use of language, the OR task triggered more productive use of target language through using other words and phrases and was more conducive to vocabulary development. As mentioned earlier, Joe (1998) devised a scale of generativity which holds that tasks can be put on a continuum according to their generativity including no generation, low generation, reasonable generation, and high generation (p. 364). As the results of this study suggest, OR task induced reasonable generation, that is, learners used the target words with some new collocations and made substantial grammatical changes in the original sentences they had read in the passages. In contrast, the SW task stimulated no generation as learners merely summarized what was already in the passages without including new words or making significant grammatical changes.

Based on the results of this study, we suggest an augmented TFA framework by taking into consideration the differential degrees of generation which is induced by a given L2 vocabulary task. As this study showed, vocabulary tasks which were theoretically expected to result in equal degrees of word learning appeared to differ in terms of learning outcomes and task effectiveness. Therefore, examining and including the degrees of task-induced generation into the generation component of the TFA would add to its predictive power with regard to designing and evaluating L2 vocabulary learning tasks. This can be achieved by adding a fourth criteria to the current Generation component of the TFA which would precisely specify the degree of task-induced generation, that is, no generation, low generation, reasonable generation, and high generation. Therefore, vocabulary tasks which trigger more generation and push learners to use the new words along with other words, collocations, and grammatical structures such as oral reproduction can be expected to lead to more effective word learning than tasks which induce lower degrees of generativity through merely requiring learners to reconstruct previously read information.

6. Conclusions, implications, limitations, and future directions
Overall, the results of the current study confirmed the positive effect of reading-based L2 vocabulary tasks of oral reproduction and summary writing incorporating the target words on advanced
EFL learners’ learning and retention of L2 vocabulary. More specifically, while both experimental conditions, that is, SW and OR tasks, outperformed the RO (control) group, OR task was found to be a more effective L2 vocabulary task than SW regarding both immediate learning and long-term retention of target words. The results can be justified in light of the positive effect of higher task-induced generation of OR than SW. In other words, since oral reproduction pushed learners to produce more comprehensible and extended output in which target words were used along with other words, phrases, collocations, and grammatical structures in novel sentences, target words stood a higher chance of learning and retention compared to summary writing task. As a result, we propose a refined and expanded TFA framework in order to include degrees of task-induced generation under the Generation component of the framework through assigning a point for tasks with a reasonable or a high degree of generation. Consequently, a task which induces a reasonable or a high degree of generation gets 1 point and the maximum score for a vocabulary task designed and evaluated based on the TFA would be 19.

The results of this study have important pedagogical implications for language teachers and textbook developers. As far as language teachers are concerned, one implication of this study is that language teachers can assign both post-reading activities of summary writing and oral reproduction including target words in order to help their students boost their L2 vocabulary knowledge. Preferably, they may require L2 learners to provide oral summaries of reading passages including target words as an effective vocabulary learning task. Furthermore, textbook developers can diversify and enrich their language coursebook vocabulary activities by including vocabulary tasks that encourage learners to write summaries or compositions or prepare their own oral summaries of the reading passages with target lexical items.

A few points should be taken into account while interpreting the results of this study. Like other experimental studies, this study also suffers from some limitations. First, this study was conducted on advanced EFL learners who were proficient users of the language and the results cannot be generalized to learners of different levels of proficiency. Due to their high command of the language, advanced learners benefited from the tasks in this study. It remains unclear whether similar results will be obtained from learners of various proficiency levels. Second, since vocabulary knowledge is a multidimensional construct, more sensitive and various vocabulary measures could have been used in the study.

Additionally, it should be made clear that EFL learners’ subsequent learning and long-term retention of target words following the treatment sessions in this study does not mean that their overall language proficiency improved since L2 vocabulary knowledge is just one of the components of the construct of language proficiency.

Therefore, future research can investigate whether similar results will be obtained with learners across other proficiency levels and use more communicative measures of vocabulary knowledge to arrive at a clearer picture of learners’ word knowledge following the treatment sessions. A further suggestion for future research can be including retrospective measures such as interviews following the treatment sessions to examine learners’ behaviors and processes they undergo during task performance.

Funding
The authors received no funding for this research.

Author details
Mojtaba Kamali1
E-mail: mojvkamali@yahoo.com
ORCID ID: http://orcid.org/0000-0003-2966-3433
Fatemeh Behjat2

E-mail: fb_30k@yahoo.com
ORCID ID: http://orcid.org/0000-0002-9060-3125
Mohammad Sadegh Bagheri1
E-mail: Bagheres@gmail.com
1 Department of English, Shiraz Branch, Islamic Azad University, Pardis Campus, Shiraz, Iran.
2 Department of English, Abadeh Branch, Islamic Azad University, Abadeh, Iran.
correction
This article has been republished with minor changes. These changes do not impact the academic content of the article.

Citation information
Cite this article as: Examining the effects of oral reproduction and summary writing vocabulary tasks on L2 word learning: Technique feature analysis on trial, Mojtaba Kamali, Fatemeh Behjat & Mohammad Sadegh Bagheri, Cogent Education (2020), 7: 1795966.

References
Atkins, P., & Baddeley, A. (1998). Working memory and distributed vocabulary learning. Applied Psycholinguistics, 19(4), 537–552. https://doi.org/10.1017/S0104217460001035
Baddeley, A. D. (1990). Human memory: Theory and practice. Lawrence Erlbaum Associates.
Cohen, J. W. (1988). Statistical power analysis for the behavi- oral sciences (2nd ed.). Lawrence Erlbaum Associates.
Croft, F. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11(6), 571–588. https://doi. org/10.1016/0022-5371(72)90001-X
Eckerth, J., & Tavokoli, P. (2012). The effects of word exposure frequency and elaboration of word proces- sing on incidental L2 vocabulary acquisition through reading. Language Teaching Research, 16(2), 227–252. https://doi.org/10.1177/1362168814313777
Fezz, S. (1998). Text-based syllabus design. National Center for English Teaching and Research.
Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32(3), 211–223. https://doi.org/10.1037/h0057532
Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. TESOL Quarterly, 40(2), 273–293. https://doi.org/10.2307/40264523
Hu, Hsu-hao Marcella., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement Load Hypothesis versus Technique Feature Analysis. System, 56, 28–39. doi:10.1016/j.system.2015.11.001
Hu, M., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with retention of word meanings inferred from context. Canadian Modern Language Review, 68(1), 54–77. https://doi.org/10.3138/cmlr.811.056
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
Icht, M., & Moma, Y. (2019). The effect of vocal production on vocabulary learning in a second language. Language Teaching Research, 1362168819883894. https://doi.org/10.1177/1362168819883894
Iran Language Institute. (2006). The ILI English series: Advanced I student’s book. Kanoonchop.
Joe, A. (1990). What effects do text-based tasks promot- ing generation have on incidental vocabulary acquisition? Applied Linguistics, 19(3), 357–377. https://doi.org/10.1093/applin.19.3.357
Keating, G. D. (2008). Task effectiveness and word learn- ing in a second language: The involvement load hypothesis on trial. Language Teaching Research, 12 (3), 365–386. https://doi.org/10.1177/1362168808089922
Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. Language Learning, 58(2), 285–325. https://doi.org/10.1111/j.1467-9922.2008.00442.x
Laufer, B. (2003). Vocabulary acquisition in a second lan- guage: Do learners really acquire most vocabulary by reading? Some empirical evidence. Canadian Modern Language Review, 59(4), 565–585. https://doi.org/10.3138/cmlr.59.4.567
Laufer, B. (2005). Focus on form in second language vocabulary learning. In S. H. Foster-Cohen, M. Garcia-Mayo, & J. Cenoz (Eds.), Eurosla yearbook (Vol. 5, pp. 223–250). John Benjamins.
Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. Applied Linguistics, 22(1), 1–26. https://doi.org/10.1093/applin/22.1.1
Laufer, B., & Rozovski-Rotblat, B. (2015). Retention of new words: Quantity of encounters, quality of task, and degree of knowledge. Language Teaching Research, 19(6), 687–711. https://doi.org/10.1177/1362168815597977
Lee, S., & Pulido, D. (2017). The impact of topic interest, L2 proficiency, and gender on EFL incidental vocabulary acquisition through reading. Language Teaching Research, 21(1), 118–138. https://doi.org/10.1177/1362168816637381
Lockhart, R. S., & Croft, F. I. M. (1990). Levels of proces- sing: A retrospective commentary on a framework for memory research. Canadian Journal of Psychology, 44(1), 87–112. https://doi.org/10.1037/h0084237
Min, H.-T. (2008). EFL Vocabulary Acquisition and Retention: Reading Plus Vocabulary Enhancement Activities and Narrow Reading. Language Learning, 58, 73–115 1 doi:10.1111/j.1467-9922.2008.58.1
Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. TESOL Quarterly, 37(4), 645–670. https://doi.org/10.2307/3588216
Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2 learners’ lexical inferencing strategy use and success. The Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 61(1), 107–134. https://doi.org/10.3138/cmlr.61.1.107
Nassaji, H., & Hu, M. (2012). The relationship between task-induced involvement load and learning words from context. International Review of Applied Linguistics in Language Teaching (IRAL), 51(1), 69–86.
Nassaji, H., & Croft, F. I. M. (2001). Learning vocabulary in another language. Cambridge University Press.
Nations, P., & Webb, S. (2011). Researching and analyzing vocabulary. Heinle.
Nguyen, C. D., & Boers, F. (2018). The effect of content retelling on vocabulary uptake from a TED talk. TESOL Quarterly, 53(1), 5–29. https://doi.org/10.1002/tesq.441
Oxford Quick Placement Test. (2001). Quick Placement Test. Oxford University Press.
Pichette, F., De Serres, L., & Lofantoina, M. (2012). Sentence reading and writing for second language vocabulary acquisition. Applied Linguistics, 33(1), 66–82. https://doi.org/10.1093/applin/amr037
Rasouei, E. (2015). Effects of three forms of reading-based output activity on L2 vocabulary learning. Language Teaching Research, 21(1), 76–95. https://doi.org/10.1177/13621688150606160
Richards, J. C., & Schmidt, R. W. (2010). Longman dic- tionary of language teaching and applied linguistics (4th ed.). Longman.
Schmidt, R. (2000). Attention. In P. Robinson (Ed.), Cognition and second language instruction (pp. 3–32). Cambridge University Press.
Sun, C.-H. (2017). The value of picture-book reading-based collaborative output activities for vocabulary retention. Language Teaching Research, 21(1), 96–117. doi:10.1177/1362168816655364

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), Principle and Practice in Applied Linguistics (pp. 125–144). Oxford: Oxford University Press

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), Handbook of research in second language teaching and learning (pp. 471–483). Lawrence.

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. https://doi.org/10.1017/S0272263105050023

Zou, D. (2016). Comparing dictionary-induced vocabulary learning and inferencing in the context of reading. Lexikos, 26(1), 372–390. https://doi.org/10.5788/26-1-1345

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. Language Teaching Research, 21(1), 54–75. https://doi.org/10.1177/1362168816652418

Zou, D., Wang, F. L., Kwan, R., & Xie, H. (2018). Investigating the effectiveness of vocabulary learning tasks from the perspective of the technique feature analysis: The effects of pictorial annotations. In S. K. S. Cheung, J. Lam, K. C. Li, O. Au, W. W. K. Ma, & W. S. Ho (Eds.), Technology in education: Innovative solutions and practices (Vol. 843, pp. 3–15). Springer Singapore.
Appendix 1. MCAR vocabulary test

Read the sentences below and choose the correct answer a, b, c, or d.

1. Computer courses continue to _______ (کورس پردازی). In fact, they are increasing quickly in many different places.
   a) procrastinate  b) inundate  c) acknowledge  d) proliferate

2. He specializes in heart problems. In other words, he is a(n) _______ (کاردیولوژی).
   a) oculist  b) psychologist  c) cardiologist  d) optometrist

3. _______ (به توجه) to danger, he was driving too fast.
   a) oblivious  b) chagrined  c) obese  d) optimistic

4. During the war, many young people were sucked into the _______ (کردیتی) of war.
   a) allergy  b) marsupial  c) maturity  d) maestral

5. The _______ (مال خیمه) inflation created many economic-related problems in the country.
   a) sabbatical  b) culprit  c) unprecedented  d) counterproductive

6. Canadian Ministry of Education found that girls are _______ (منکن پرتو داشتن) boys in literacy skills by a wider margin than previously thought.
   a) muttering  b) pilfering  c) manufacturing  d) outperforming

7. During protests, extremists were openly _______ (حمله کرد) violence and inviting people to destroy state buildings.
   a) confronting  b) reflecting  c) advocating  d) exploring

8. Women are increasingly resisting attempts to _______ (عملی زنان لبیش ایجاد) in society.
   a) materialize  b) feminize  c) implement  d) assign

9. It is essential that the government follow a policy which ensures a(n) _______ (علیه اعیان) distribution of food supplies among people.
   a) equitable  b) contemptible  c) reversible  d) inevitable

10. It was quite common in the United States that blacks were _______ (چاپ کرد) from whites in schools.
    a) integrated  b) surrendered  c) stimulated  d) segregated

11. Analysts believe that at this stage, war between the two countries is _______ (اجتناب ناپذیر) and may break out at any time.
    a) improbable  b) intriguing  c) inevitable  d) indispensable

12. According to its developers, up to twenty users can have _______ (رسیدنی) access to the system.
    a) courageous  b) simultaneous  c) illiterate  d) infallible

13. The police have announced that they are looking for _______ (شکنجه) evidence of his guilt.
    a) irresponsible  b) tangible  c) recurrent  d) ambiguous

14. Knowledge about cultural differences is _______ (دروری و جایزه) for the success of a businessman in a foreign country.
    a) negligent  b) ignorant  c) minor  d) vital

15. It can be argued that global standardization of products to fit all markets is unlikely to be the most _______ (علمی) option.
    a) viable  b) latent  c) external  d) extrinsic

16. He _______ (آغاز کرد) a new project even though he had failed many projects before.
    a) put off  b) embarked on  c) gave up  d) look after

17. They were caught off guard by a sudden storm that _______ (وازکن کرد) their boat.
    a) pronounced  b) screamed  c) capsized  d) sighted

18. When the sea was calmer and the waves had _______ (فریکش کرد) he looked around the wreckage of their boat for his friend.
    a) qualified  b) subsided  c) floated  d) sought

19. We all _______ (حفر دفتر حدود) the chances we have missed in our lives.
    a) regret  b) nudge  c) dehydrate  d) authenticate

20. When he telephoned his family, their grief turned to joy, for they had already held a _______ (مراسم ترک) service for him.
    a) sabbatical  b) liable  c) sophisticated  d) funeral

21. We live in an overcrowded world _______ (پر بودن از) with billions of humans who are struggling for food and clean water.
    a) envisaging  b) teeming  c) declining  d) inhibiting

22. The animal seizes its prey by the throat and _______ (دفن کرد) it to death.
    a) erodes  b) revives  c) suffocates  d) supports
23. A massive ________ (مهاجرت دسته جمعی) of doctors is forcing the government to recruit from abroad.
   a) drawback     b) enterprise     c) exodus     d) exposure

24. The impact of some drugs on our body cannot be completely ________ (منجیدن).
   a) gauged     b) staggered     c) preceded     d) prefabricated

25. In Europe, ________ (جمعیت شناسان) forecast a major drop in the numbers who will work and earn money, while the population of the older people will soar.
   a) refugees     b) victims     c) photographers     d) demographers

26. The first stage of ________ (منطق شدن) begins before the travelers even leave home.
   During this stage, they form ideas and images of what life in a new country will be like.
   a) surroundings     b) adjustment     c) foundations     d) identity

27. The language of the foreign country is the real ________ (مشکل) for many immigrants who lack academic backgrounds.
   a) struggle     b) symptom     c) withdrawal     d) assumption

28. Minority students begin to feel more and more ________ (منزوی) if they are not appropriately integrated into the community of students.
   a) overwhelming     b) extroverted     c) alienated     d) mild

29. The rebels were attacked and forced to ________ (عقب نشینی کردند) to mountains.
   a) prioritize     b) unwrap     c) violate     d) retreat

30. Travelers’ friends and family are only mildly interested in the exciting things that happened to them during their ________ (اقامت) abroad.
   a) accomplishment     b) sojourn     c) tribute     d) debut

31. A lot of these children have been ________ (محروم) of a normal home life.
   a) pondered     b) escalated     c) acknowledged     d) deprived

32. Today some universities offer ________ (فرستد مطالعاتی) to their lecturers to travel and study abroad for a period of time.
   a) homogeneity     b) indifferences     c) sabbaticals     d) potentials

33. The main point of many modern novels is the ________ (پیهودگی) of war, that is, the uselessness of violence in our time.
   a) futility     b) analogy     c) phenomenon     d) pedagogy

34. The new president is going to issue his ________ (بيانیه تبییغاتی) for the next four years this week.
   a) manifesto     b) anthropologist     c) imperative     d) eruption

35. Unfortunately, black people have been ________ (برداشت کلیه ای داشتند) throughout history as uncivilized and violent.
   a) impeded     b) sedated     c) exorted     d) stereotyped

36. The government is planning to implement educational programs to improve adult ________ (مواد).
   a) illiteracy     b) extinction     c) literacy     d) ambiguity

37. There were a couple of attempts to ________ (از قدرت کنار زدن) the young king and appoint his brother as his successor.
   a) interrogate     b) depict     c) fluctuate     d) usurp

38. His findings are based mainly on ________ (حکایت گونه) evidence rather than serious and scholarly research.
   a) empirical     b) anecdotal     c) scientific     d) laboratory

39. Oil companies have been accused of ________ (به اجازر راهیت دادن) in the pollution of the ocean.
   a) embezzling     b) degrading     c) acquiescing     d) stagnating

40. The president was criticize for the way in which he handled the ________ (بحران).
   a) crisis     b) audacity     c) sketch     d) rhetoric
Appendix 2. CRAR vocabulary test

Read the sentences below and choose the correct answer a, b, c, or d.

1. Members of the security council of the UN passed a resolution to prevent nuclear weapons from p_______ (گسترش کردن).
2. We need more e_______ (Nguồn cung cấp) in this hospital to help patients with heart problems.
3. He drove off, o_______ (بکارگیری) to the damage he had caused.
4. The m______ (گردآوری داده) of war affected the lives of many people.
5. The u______ (استقلال) economic emergency forced many companies to lay off their workers.
6. Experts believe that stocks generally o_______ (ملکرد پهنی داشتن) other investments.
7. Those who a_______ (خسارت دیدن) for doctor-assisted suicide say the terminally ill should not have to suffer.
8. Some governments have tried to f_______ (عجله زنان تبعیض ایجاد کردن) women in both public and private sectors.
9. Protesters urged the government to pay a(n) e_______ (عدل آری) salary to all workers.
10. It is a tradition in some countries to s_______ (چهار کردن) boys from girls in schools.
11. Inflation and unemployment were the t_______ (اثربخشی ناپایدار) consequences of the war.
12. The speeches will be broadcast live, with s_______ (ترجمه معنی) (translation).
13. The manager of the company asked for t_______ (علائم) evidence of my skills and experience.
14. Doctors advise that regular exercise is v_______ (برآور و حیاتی) for your health.
15. The committee discussed the issue but could not come forward with a v_______ (عملی) (عملی) solution.
16. He is willing to e_______ (اخذ کردن) a new career as a teacher.
17. A rescue team was sent to save the young men whose boat had been c_______ (واژگون شدن) by a sudden storm.
18. The pains in his head had s_______ (درد کردن), but he still felt dizzy and sick.
19. I r_______ (درست خوردن) leaving school when I was so young.
20. Hundreds of mourners attended the f_______ (پردازش تکمیل) of the two boys.
21. During summer holidays, the island was t_______ (پر بودن از) tourists from around the world.
22. There was evidence that he had put a pillow over her face and s_______ (قتل کردن) her.
23. The mass e_______ (مهاجرت دسته جمعی) of skilled workers to foreign countries left the country with lack of labor force.
24. The thermostat will g_______ (ازنجیره) the temperature and control the heat.
25. D_______ (بمعنی شناسان) believe that the current birth rate of the country will cause the population to drop dramatically in future.
26. The a_______ (سطح قرنی) period for each immigrant varies depending on many factors.
27. For many people who decide to lose weight, it is a real s_______ (مشکل) to stop eating sweets.
28. She had become a_______ (منوز) from her family because her parents were drug addicts.
29. Our troops were attacked and forced to r_______ (غیر محروم کردن).
30. During my a_______ (افراد) abroad, I learned many cultural differences.
31. D_______ (محموم) children tend to do less well at school than children from wealthy families.
32. During his s_______ (پردازش مطالعاتی) in Australia, he gave lectures, attended conferences, and published two research papers.
33. After five years of devastating war, even its supporters realized the f_______ (بی‌هویتی) of war.
34. The presidential candidates are expected to publish their m_______ (به نسبی) for the upcoming election tomorrow.
35. Human rights activists urged that racial s_______ (بردیدن کلیه یک داشتن) in the media should be stopped.
36. Digital l_______ (مورد) has become one of the most important aspects of modern life.
37. The older brothers made a couple of attempts to u_______ (به چنین آوردی) the ownership of the company.
38. We cannot make our judgement based on a_______ (هکیمیت گونه) evidence. We need empirical evidence.
39. Some government officials were accused of a_______ (به اجبار رضا دادن) in the cover-up of the incident.
40. The true power of president was seen in the way he handled the c_______ (پرداخت).
