

# The Effect of Task Type and Word Type on Vocabulary Learning: A Comparison Based on Involvement Load Hypothesis and Technique Feature Analysis

Maryam Ehsani<sup>1</sup> , Hossein Karami<sup>2</sup> , Omid Mallahi<sup>3</sup> 

<sup>1</sup>Ph.D. Candidate, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, Email: [Maryam.Ehsani1996@ut.ac.ir](mailto:Maryam.Ehsani1996@ut.ac.ir)

<sup>2</sup>*Corresponding author*, Associate Professor, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, Email: [Hkarami@ut.ac.ir](mailto:Hkarami@ut.ac.ir)

<sup>3</sup>Assistant Professor, Department of English Language Teaching, University of Hormozgan, Bandar Abbas, Iran, Email: [o.mallahi@hormozgan.ac.ir](mailto:o.mallahi@hormozgan.ac.ir)

## Abstract

This study aimed to investigate the effect of task type (i.e., sentence fill-in/sentence writing) and word type (i.e., real/pseudo) on initial learning and retention of 10 word meanings, taking the predictions of Involvement Load Hypothesis (ILH) and Technique Feature Analysis (TFA) into account. Participants were 59 intermediate-level EFL learners from eight intact classes. Each intact class was randomly assigned to one of the four learning conditions: 1) sentence fill-in with real words, 2) sentence writing with real words, 3) sentence fill-in with pseudowords, and 4) sentence writing with pseudowords. Initial learning was measured by administering a meaning recall test immediately after the tasks and medium-term retention was measured by administering the same test with rearranged items one week after the tasks phase. The results of a mixed between-within subjects ANOVA did not show any significant interaction effect between test time and learning condition. Furthermore, the main effect for learning condition was not statistically significant but there was a significant main effect for test time, suggesting that participants' scores dropped significantly from the immediate posttest to the delayed posttest. The findings of two independent-samples t-tests failed to show any significant difference between the immediate and delayed posttest scores of the participants who received either sentence fill-in or sentence writing tasks. However, some tentative findings demonstrated that those participants who were assigned to the sentence writing task achieved higher scores on the posttests. This finding indicates that TFA has probably more predictive power than ILH and it also provides some evidence in favor of the heavier weight of the evaluation component of the ILH when compared to its search component.

**Keywords:** involvement load hypothesis, technique feature analysis, L2 vocabulary learning, vocabulary learning task

Received: May 15, 2022

Revised: August 28, 2022

Accepted: January 11, 2023

Article type: Research Article

DOI: 10.22111/IJALS.2023.45695.2355

Publisher: University of Sistan and Baluchestan

© The Author(s).



How to cite: Ehsani, M., Karami, H., & Mallahi, O. (2023). The effect of task type and word type on vocabulary learning: A comparison based on involvement load hypothesis and technique feature analysis. *Iranian Journal of Applied Language Studies*, 15(1), 169-190. <https://doi.org/10.22111/IJALS.2023.45695.2355>

## 1. Introduction

Vocabulary has been demonstrated to play a major part in language proficiency (Stæhr, 2008; Renandya & Richards, 2012). Extensive knowledge of vocabulary is the essential component of successful communication and mastery of grammar is unlikely to compensate for lack of word knowledge. Wilkins (1972) claimed that “without grammar very little can be conveyed, without vocabulary nothing at all can be conveyed” (p. 111). Vocabulary knowledge, as a rather complex and multidimensional construct involving various aspects such as form, meaning, register and collocation, is considered the foundation for the efficient learning and improvement of receptive and productive language skills (Nation, 2001). There has been a growing body of research on effective ways to enhance students' vocabulary development (e.g., Ellis & Beaton, 1993; Iravi & Malmir, 2022; Nation, 2006; Pigada & Schmitt, 2006). However, the issue of how to design and use vocabulary learning tasks in classroom settings to improve learners' vocabulary development is not well-researched (Hazrat & Read, 2022; Yanagisawa & Webb, 2021).

Different tasks have been suggested to help students learn words more effectively, among which are word cards, original sentence writing, choosing definitions, fill-in-the-blanks and so on. Two models that have been proposed to assess the effectiveness of different vocabulary learning tasks and the elaborate processing needed for L2 vocabulary acquisition are the Involvement Load Hypothesis (ILH), which conceptualizes depth of processing in terms of need, search and evaluation (Laufer & Hulstijn, 2001) and Technique Feature Analysis (TFA), which necessitates motivation, noticing, retrieval, generation, and retention components for vocabulary processing (Nation & Webb, 2011). Little empirical studies have been carried out to compare the predictive power of these two models in terms of vocabulary learning gains. Inspired by the study conducted by Hu and Nassaji (2016) and Jafari Gohar et al. (2018), the current study intends to compare the predictability of these two models using two different task types.

In order to control for the effect of pre-knowledge of the target items, many vocabulary learning studies replace real words with pseudowords (e.g., Arndt & Woore, 2018; Batterick & Neville, 2011; Keating, 2008; Montero Perez, 2020; Pellicer- Sanchez, 2016; Pulido, 2009; Webb, 2007). However, the effects of using pseudowords instead of real words in vocabulary learning studies have not been investigated yet. Some studies suggest that the use of pseudowords may underestimate the learning gains (e.g., Papagno et al., 1991, as cited in Arndt & Woore, 2018) and others suggest that using pseudowords may overestimate the learning gains (e.g., Arndt & Woore, 2018). These contradictory assumptions regarding the effect of using pseudowords in vocabulary learning research seem to call for further investigation of the issue. This is what the present study aims to accomplish.

## 2. Review of the Literature

### 2.1. *Involvement Load Hypothesis (ILH)*

Craik and Lockhart (1972, p. 675) first introduced the concept of deep processing in their “depth of processing” model. In this model, they related the retention of new information to the depth with which it is processed. In other words, according to the model, the deeper the information is processed, the better it can be retained and stored in memory. They proposed that processing has several levels, some deeper and some shallower. For example, they contended that processing the meaning of words occurs deeply but processing the phonological form of them occurs shallowly. Craik and Tulving (1975), however, argued that it is the richness of material encoding or elaboration, rather than the presence of semantic encoding, that determines the retention of lexical items. Craik and colleagues’ (1972, 1975) model of the depth of processing has been criticized for the vagueness of the definition of the level of processing and lack of explicit criteria to distinguish a deeper level of processing from a shallower one (Laufer & Hulstijn, 2001).

Recognizing the lack of a clear-cut operationalization for the concepts proposed by Craik and colleagues’ (1972, 1975) (e.g., elaboration) at the level of vocabulary learning tasks, Laufer and Hulstijn (2001) put forward the Involvement Load Hypothesis (ILH). According to them, the construct of involvement is made of three components of need, search, and evaluation with their associated degrees of prominence (i.e., 0, 1, 2). ILH assumes that the retention of words in memory is dependent on the amount of involvement in processing those words. For example, a vocabulary learning task with a higher involvement load is hypothesized to yield better retention of words than a task with a lower involvement load, irrespective of task mode (i.e., aural, oral, visual). Table 1 shows the three components of involvement load with their degrees of prominence.

**Table 1**

*Components of Involvement Load with Their Associated Degrees of Prominence*

	Search	Need	Evaluation
Absent (0)	✓	✓	✓
Moderate (1)	✓	✓	✓
Strong (2)	(*) (✓)	✓	✓

The original IL proposed by Laufer and Hulstijn (2001) comprises the sum of different degrees (i.e., 0, 1, 2) of need, search, and evaluation. However, some researchers suggest that the three components of IL, and even the different degrees of each of them, might not equally contribute to vocabulary learning (Laufer & Hulstijn, 2001; Kim, 2008; Keating, 2008). In an attempt to find an answer to these theoretical speculations, Zou (2016) carried out a study which examined the influence of the evaluation component of the ILH on word learning. The results of the posttests revealed that the two writing tasks (IL: 3) led to significantly better learning outcomes than the fill-in-the-blanks task (IL: 2). This finding was in line with ILH. The finding that was inconsistent with the predictions of ILH was that, despite having the same IL indices, the

composition-writing task (evaluation score: 2) resulted in significantly better word learning than the sentence-writing task (evaluation score: 1). In order to account for the differences among the three tasks in terms of word learning effectiveness, “an augmented evaluation framework” was proposed in this study according to which evaluation can be moderate (e.g., fill-in-the-blanks), strong (e.g., sentence writing), and very strong (e.g., composition writing). Karami and Esrafil’s (2021) study supported Zou’s finding about the greater weight of the evaluation component of the IL in comparison with the other two components.

## 2.2. Technique Feature Analysis (TFA)

Nation and Webb (2011) criticized ILH on the grounds that although the model is simple to use, it neglects many factors that research has shown to be important for vocabulary learning. They proposed Technique Feature Analysis (TFA) as a model that can compensate for the shortcomings of the ILH. Being made of five main components and 18 criteria, TFA is evidently more comprehensive than ILH. It seems noteworthy to point out that TFA does not disregard ILH components, but views them from a different perspective. Table 2 displays TFA components along with their associated criteria.

**Table 2**

*Technique Feature Analysis Checklist (from Nation & Webb, 2011, p. 7)*

Criteria	Scores	
<b>Motivation</b>		
Is there a clear vocabulary learning goal?	0	1
Does the activity motivate learning?	0	1
Do the learners select the words?	0	1
<b>Noticing</b>		
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
<b>Retrieval</b>		
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
<b>Generation</b>		
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
<b>Retention</b>		
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
<b>Maximum score</b>	18	

### ***2.3. Studies on the Role of ILH and TFA in Vocabulary Learning***

In order to compare the effectiveness of the two vocabulary learning frameworks, ILH and TFA, Hu and Nassaji (2016) conducted a study with 96 EFL learners using four different vocabulary learning tasks. Three of those tasks had consistent rankings based on the two models and one (i.e., reading and fill-in-the-blanks) had an inconsistent ranking based on the two models. Compared with the other three tasks, the reading and fill-in-the-blanks task was ranked higher on the TFA but lower on the ILH. Therefore, one cannot understand whether, based on the TFA, this task will result in better learning gains than the other tasks or based on the ILH, it will lead to lower learning gains than the other tasks. Hence, more research seems warranted to investigate the reliability of the models and to discover which one can better predict the learning gains resulting from different vocabulary learning tasks. The results of the hierarchical multiple-regressions in Hu and Nassaji's study indicated that the TFA framework was significantly more powerful in predicting lexical gains than the ILH model.

Khoshsima and Eskandari (2017) also tried to compare the predictive power of ILH and TFA and conducted their study with 76 EFL university students. Students who were from four intact classes were randomly assigned to one of the four vocabulary learning tasks suggested by Nation and Webb (2011) including reading and fill-in-the-blanks (IL index: 2, TF score:7), reading and rewording sentences (IL index:3, TF score: 6), reading and choosing definitions (IL index:3, TF score:6), and reading and multiple-choice questions on the text content (IL index:3, TF score:6). In order to confirm that the participants had similar levels of English language proficiency, an Oxford Placement Test was given to 100 students and those who scored one standard deviation above or below the mean were selected as participants of the main study; 76 participants in total. A modified version of the Vocabulary Knowledge Scale (Paribakht & Wesche, 1997) developed by Folse (2006) was administered as both a pretest and an immediate posttest. Students' performance during each task was also examined. During-task performance results showed no significant difference between the tasks and did not provide support for the assumptions made by either the TFA or the ILH. Posttest results demonstrated a significant difference between the fill-in-the-blanks task and the other three tasks. This result lent support to TFA rather than ILH because the fill-in-the-blanks task, which is the task with the highest TFA score, resulted in significantly better vocabulary posttest gains. The results of the hierarchical multiple regression, however, did not confirm the superiority of the TFA framework over the ILH framework in predicting vocabulary learning gains. Time-on-task was not controlled in this study and no delayed posttests were used.

In another attempt to compare the explanatory power of ILH and TFA in terms of vocabulary learning gains, Jafari Gohar et al. (2018) randomly assigned advanced EFL learners to three tasks with different IL indices and TF scores. Participants in the first group were assigned to the sentence-writing task (IL index: 3, TF score: 7), those in the second group were assigned to the composition-writing task (IL index: 3, TF score: 9), and those in the control group were assigned to

the reading comprehension task (IL index: 1, TF score: 3). The first two groups, which were the experimental groups, were given five minutes to read the target words along with their meanings and example sentences prior to the tasks. In order to verify students' unfamiliarity with the target vocabulary, the researchers consulted the comments of two of the students' teachers and also used a vocabulary pretest one week before the main experiment. The during-task performance scores and pretest-posttest gains were submitted to analysis. The results of during-task performance scores were shown to be significantly different, the difference lying between the first two tasks and the reading comprehension task. The findings of pretest-posttest scores revealed that the composition-writing task brought about the highest vocabulary gains; but the difference between the scores that resulted from this task and those of sentence-writing task did not differ from one another in a statistically significant way. Contrary to findings of Hu and Nassaji's (2016) study, in this study the results of the hierarchical multiple regression failed to find any considerable evidence that showed a more powerful predictability of TFA over ILH, at least in terms of significant differences. Therefore, more research seems to be required to investigate the yet unresolved problem of whether ILH or TFA offers a more reliable framework for predicting vocabulary learning gains.

Adopting ILH theory, Kaivanpanah, Alavi and Ravandpour (2020) investigated the effect of input-based (focusing on comprehension) and output-based (focusing on both comprehension and production) tasks with identical and different degrees of involvement on 120 pre-intermediate Iranian EFL learners' incidental vocabulary learning. These tasks and conditions were configured and administered in three different phases. The results indicated the superiority of students receiving output-based tasks in vocabulary acquisition in almost all phases, while the students receiving input-based tasks outperformed their counterparts in terms of vocabulary retention in the delayed posttest.

Moreover, using ILH and TFA theories, Mateo-Valdehi and de Diego (2021) explored the receptive and productive vocabulary acquisition of 308 French students of Spanish as a second language by comparing the effectiveness of three task conditions, namely definition-choosing, gap-filling and sentence writing with a given word. Among these tasks, sentence writing, which demands a higher cognitive effort and mental processing, provided a better opportunity for the students to learn receptively and productively and to incorporate a higher number of vocabularies into their mental lexicon.

Furthermore, Karami and Esrafil (2021) empirically investigated the impact of ILH and task types on Iranian EFL learners' incidental vocabulary acquisition and retention. The participants were randomly assigned to four groups to complete one of the four task types having similar load indices: 1) reading comprehension plus dictionary use, 2) Cloze exercise plus dictionary use, 3) inferencing, and 4) Sentence writing plus the meaning of target words in the first and second language. A multiple-choice word recognition test and a vocabulary knowledge scale were administered once immediately after completing the tasks (immediate posttest) and three weeks

later (delayed posttest) to measure the learners' retention of words. Contrary to what is postulated by ILH theory, the results indicated significant differences in the learners' levels of performance on the four different tasks and a memory loss from immediate to delayed posttest, which signifies the significance of factors such as multiple exposures, depth of processing and involvement in vocabulary learning.

In a recent attempt, Ehsani and Karami (2022) compared the predictive power of ILH and TFA by applying one control task (i.e., Reading and comprehension questions) to examine the learners' ability to infer and recall the meaning of words and four experimental vocabulary learning tasks including 1) Reading and true/false statements, 2) Reading and choosing definitions, 3) Reading and fill-in-the-blanks, and 4) Reading and rewording the sentences (suggested by Nation and Webb, 2011). A total of 114 young EFL students from five intact classes were assigned to one of these conditions. Besides assessing the learners' during-task performance, the initial learning and the medium-term retention were measured by administering an English-to-Persian translation once immediately after the completion of the test (as an immediate posttest) and with rearranged items one week later (as a delayed posttest). The results of the study confirmed the higher predictive power of TFA in accounting for the vocabulary learning gains of the learners. In addition, corroborating the findings of previous studies, the researchers highlighted "the efficiency of word-focused activities and the so-called 'reading plus' condition as opposed to 'reading only' condition for vocabulary learning" (p. 182).

#### ***2.4. Using Pseudowords as Target Items in Vocabulary Research***

Vocabulary learning studies are found to use both real words and pseudowords as their target items. Some researchers have argued that the use of pseudowords may underestimate vocabulary learning gains. For example, Papagno et al. (1991, as cited in Ardnt & Woore 2018) stated that if learners are aware of the use of pseudowords, they may be less motivated to learn those words because they will recognize that those words will be of no use to them in the real world. On the other hand, there are other researchers who have argued that the use of pseudowords in lieu of real words may overestimate vocabulary learning gains. Ardnt and Woore (2018), for instance, argue that if the pseudowords were used instead of high-frequency words, the participant's task to infer their meanings might be made easier and thus may lead to an overestimation of learning gains. Considering the fact that some vocabulary learning studies have chosen pseudowords as the target items of their research and others have used real words, the need to examine whether the use of pseudowords may underestimate or overestimate the learning gains can be felt.

The study conducted by Hu and Nassaji (2016) showed that the TFA was more powerful in predicting vocabulary learning gains than the ILH. However, Jafari Gohar et al. (2018) failed to provide convincing evidence regarding the better predictability of the TFA. Additionally, Hu and Nassaji (2016) found some evidence that indicated differential weights of different ILH and TFA



components. The study conducted by Zou (2016) supported the heavier weight of the evaluation component of ILH, but TFA components have not been examined in terms of their possible differential weights. Comparing the posttest scores resulting from two vocabulary learning tasks, the results of this study shed more light on whether either ILH or TFA is a more reliable predictor of vocabulary learning. It also explores the possible differences in the weights of ILH and TFA components. Also, it examines the effects of two different task types on learning words. Furthermore, the present study aims to examine whether using pseudowords in lieu of real words may underestimate, overestimate, or neither underestimate nor overestimate the vocabulary learning gains.

To the knowledge of the researchers, although vocabulary learning studies have used both real words and pseudowords as the target items, there seems to be an underlying assumption behind the majority of them that the use of pseudowords does not distort the vocabulary learning results. However, as explained in the previous section, there are some researchers who call this assumption into question by arguing that the use of pseudowords may either underestimate or overestimate the learning gains. This study intends to discover whether learners' vocabulary learning pattern differs when pseudowords or real words are used as the target items. The results of this study are of pedagogical interest to language teachers, materials developers, and researchers who are interested in vocabulary learning studies.

The present study was guided by the following research questions:

- RQ1. Which of the four learning conditions (real word/fill-in, real word/sentence writing, pseudoword/fill-in, pseudoword/sentence writing) leads to better meaning recall scores measured at two time periods (i.e., immediate/delayed posttests)?
- RQ2. Are there any significant differences between the posttests scores of those participants who receive real words as their target vocabulary and those who receive pseudowords as their target vocabulary?
- RQ3. Are there any significant differences between the posttest scores of those participants who do the sentence fill-in task and those who do the sentence writing task?

## 3. Method

### *3.1. Research Design*

In this study, each of the eight intact classes was randomly assigned to one of the four learning conditions (see section 4.1). This study was a quasi-experimental one because although the groups were randomly assigned to one of the four learning conditions, all the participants in each group were required to do the same task and were given the same word type and the assignment of participants to learning conditions was not totally on a random basis.



### 3.2. Participants and Sampling

Fifty-nine intermediate-level EFL learners from eight intact classes of a language institute participated in this study (male: 29, female: 30). The participants were all native speakers of Farsi and their ages ranged from 13 to 38 (mean age: 21). Each intact class was randomly assigned to one of the four learning conditions: 1) real word/fill-in; 2) real word/sentence writing; 3) pseudoword/fill-in; and 4) pseudoword/sentence writing. Hence, each condition was randomly given to two of the intact classes.

### 3.3. Target Words

First, 30 real English words were given to the four instructors of the participants and they were asked to choose 15 words which they believed students most probably did not know their meanings. Based on the teachers' responses and suggestions, 15 words were selected. Next, two English language learners whose proficiency level was higher than the target group of learners selected 12 words the meanings of which they did not know. From these 12 words, the researchers chose 10 words from three major parts of speech for the main study: four verbs, four nouns, and two adjectives. The target words of learning conditions 1 and 2 were real English words and the target words of learning conditions 3 and 4 were pseudowords which were used in lieu of the real English word forms. The pseudowords had the same meaning as the real English words. All the pseudowords conformed to the phonological and morphological properties of the English language. Table 3 shows both real and pseudo words that were used in this study.

**Table 3**

*Target words (real, pseudo)*

Real word	Pseudo word equivalent
Accessory (noun)	Proffendo
Assiduous (adjective)	Asmious
Caprice (noun):	Masco
Diagnose (verb):	Muliate
Disseminate (verb):	Tasperate
Fade (verb):	Rade
Hilarious (adjective)	Londrous
Impress (verb)	Starify
Infection (noun)	Pacon
Itch (noun)	Sanch

### 3.4. Instruments

The instruments used in this study are discussed below and all of them are available in the appendices.

### ***3.4.1. Tasks***

Two task types that were suggested in Nation and Webb (2011) were employed in this study: 1) sentence fill-in (IL: 3 [need: 1, search: 1, evaluation: 1], TFA score: 7): in this task, the learners read the 10 sentences and filled in the gaps with the appropriate words from the word list. The participants were not given the word meanings but were instructed to ask their teacher to provide them with the word meanings. Therefore, this task induced a search of one. This was in line with Laufer and Hulstijn's (2001) contention that if learners consult their teacher to find the word meaning the search induced by the task is one; 2) original sentence writing (IL: 3 [need:1, search: 0, evaluation: 2], TFA score: 8): the participants were asked to write one original sentence with each of the 10 words. For each word, some information including L1 translation, L2 explanation (or synonym), and a sample sentence containing the target word was provided.

### ***3.4.2. Tests***

To measure participants' initial learning and medium-term retention of the target word meanings, one meaning recall test with a different order of items was administered once immediately after task completion (immediate posttest) and once one week later (delayed posttest). In this test, the participants had to provide either Persian translations or English explanations/synonyms of the 10 target words. The rearrangement of task items in the two posttests was done in accordance with prior research (e.g., Hu & Nassaji, 2016; Keating, 2008), so that any potential ordering effects could be precluded. Both immediate and delayed posttests were scored dichotomously, a correct Persian translation or English explanation/synonym was given a score of 1 and an incorrect response received a score of 0.

### ***3.5. Data Collection Procedure***

Each of the eight intact classes was randomly assigned to one of the following learning conditions: 1) sentence fill-in with real words, 2) sentence writing with real words, 3) sentence fill-in with pseudowords, and 4) sentence writing with pseudowords. Upon finishing the tasks, the participants were unexpectedly tested on their initial learning of the target words. One week later, the same recall test whose items were rearranged was given to them as the delayed posttest to evaluate their medium-term retention of the newly acquired target words. All the tasks and tests were presented during students' regularly scheduled online classes which were held through the mobile application "WhatsApp". In order to ensure that the learners did not have access to the target words during the two posttests, they were required to delete the data from the previous stages and send a screenshot of the messages they deleted to their instructors. All the stages of the study were observed by a trained supervisor and the researchers. Following Hulstijn and Laufer (2001), time on task was regarded as inherent to the task and it was not controlled. Participants who did the

sentence fill-in task (groups 1 and 3) spent about 20 minutes on this task and those who did the sentence writing task (groups 2 and 4) completed their task in about 25 minutes.

## 4. Results

A mixed between-within subjects ANOVA and a number of independent-samples t-tests were conducted to analyze the data of the study. In order to answer RQ1, a mixed between-within subjects ANOVA was performed, with learning condition (i.e., sentence fill-in with real words, sentence writing with real words, sentence fill-in with pseudowords, and sentence writing with pseudowords) as the between-subjects independent variable and test time (i.e., immediate, delayed) as the within-subjects independent variable. The descriptive statistics for the immediate and delayed posttest scores of the 59 participants who had been present at all the phases of the study are displayed in Table 4.

**Table 4**

*Descriptive Statistics for Immediate and Delayed Posttest Scores Per Condition*

	Condition	Mean	Std. Deviation	N
Immediate	Real/Fill-in	5.937	2.264	16
	Real/Sentence-writing	5.733	2.344	15
	Pseudo/Fill-in	4.500	2.278	14
	Pseudo/Sentence-writing	7.000	1.961	14
	Total	5.796	2.332	59
Delayed	Real/Fill-in	4.375	2.526	16
	Real/Sentence-writing	4.000	2.267	15
	Pseudo/Fill-in	2.785	2.259	14
	Pseudo/Sentence-writing	5.071	2.555	14
	Total	4.067	2.483	59

The figures in Table 4 show that in general, students scored higher on the immediate posttest than they did on the delayed posttest ( $5.79 > 4.06$ ). Additionally, it shows that Group 4 members who wrote original sentences with pseudowords had the highest mean score on both the immediate posttest ( $M=7$ ) and the delayed posttest ( $M=5.07$ ) and Group 3 members who filled in the blanks with pseudowords received the lowest mean score on both immediate ( $M=4.5$ ) and delayed posttests ( $M=2.78$ ). Participants' performance on both immediate and delayed posttests can be succinctly shown as Condition 4 > Condition 1 > Condition 2 > Condition 3. Table 5 shows the multivariate tests results of the mixed ANOVA.

**Table 5***Multivariate Tests Results of the Mixed ANOVA*

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	.687	120.690 <sup>b</sup>	1.000	55.000	.000	.687
	Wilks' Lambda	.313	120.690 <sup>b</sup>	1.000	55.000	.000	.687
	Hotelling's Trace	2.194	120.690 <sup>b</sup>	1.000	55.000	.000	.687
	Roy's Largest Root	2.194	120.690 <sup>b</sup>	1.000	55.000	.000	.687
Time * Condition	Pillai's Trace	.012	.228 <sup>b</sup>	3.000	55.000	.876	.012
	Wilks' Lambda	.988	.228 <sup>b</sup>	3.000	55.000	.876	.012
	Hotelling's Trace	.012	.228 <sup>b</sup>	3.000	55.000	.876	.012
	Roy's Largest Root	.012	.228 <sup>b</sup>	3.000	55.000	.876	.012

As can be seen in Table 5, Wilks' Lambda value of Time and Condition interaction was not statistically significant,  $F(3, 55) = .228, p = .87$ . Looking at the value of Wilks' Lambda for the main effect of Time, one can see that this factor was statistically significant,  $F(1, 55) = 120.69, p < .05$ , partial eta squared = .687. This suggests that participants' scores on the immediate and delayed posttests differed significantly. It was demonstrated in Table 4 that participants' scores on the immediate posttest were higher than their scores on the delayed posttest. The significant difference between them that was detected in Table 5 reveals that the participants scored significantly higher on the immediate posttest than on the delayed posttest. This significant main effect had a partial eta squared of .687 which is considered a very large effect size according to Cohen's commonly used guidelines (as cited in Pallant, 2016). In order to obtain information about the between-subjects factor of Task Type, Tests of Between-Subjects Effects table needs to be consulted (see Table 6).

**Table 6***Tests of Between-Subjects Effects*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	2853.784	1	2853.784	285.338	.000	.838
Condition	82.379	3	27.460	2.746	.052	.130
Error	550.078	55	10.001			

Table 6 does not show any significant main effect for learning condition,  $F(3, 55) = 2.746, p > .05$ . This indicates that participants assigned to different learning conditions did not gain significantly differential scores on the posttests.

In order to answer RQ2, first an independent-samples t-test was conducted to compare the immediate posttest scores for participants who received real words and those who received pseudowords. As Table 7 shows, there was no significant difference in immediate posttest scores for participants who received real words as their target vocabulary ( $N = 31, M = 5.83, SD = 2.26$ ) and

those who received pseudowords as their target vocabulary ( $N=28$ ,  $M=5.75$ ,  $SD=2.44$ ;  $t(57)=-.145$ ,  $p=.89$ , two-tailed).

**Table 7**

*Independent Samples T-Test for Immediate Posttest Scores (IV: word type)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Immediate	Equal variances assumed	.181	.672	.145	57	.886	.08871	.61332	-1.13945	1.31687
	Equal variances not assumed			.144	55.246	.886	.08871	.61571	-1.14507	1.32249

Next, an independent-samples t-test was conducted to compare the delayed posttest scores for participants who received real words and those who received pseudowords. According to Table 8, there was no significant difference in the delayed posttest scores for participants who received real words as their target vocabulary ( $N=31$ ,  $M=4.19$ ,  $SD=2.37$ ) and those who received pseudowords as their target vocabulary ( $N=28$ ,  $M=3.92$ ,  $SD=2.63$ ;  $t(57)=.406$ ,  $p=.69$ , two-tailed).

**Table 8**

*Independent Samples t-test for Delayed Posttest Scores (IV: word type)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2- tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Delayed	Equal variances assumed	.524	.472	.406	57	.686	.26498	.65220	-1.04102	1.57098
	Equal variances not assumed			.404	54.623	.688	.26498	.65577	-1.04942	1.57937

In order to answer RQ3, first an independent-samples t-test was conducted to compare the immediate posttest scores for participants who did the sentence fill-in task and those who did the sentence writing task. As Table 9 reveals, there was no significant difference in the immediate posttest scores for participants who filled in the blanks ( $N=30$ ,  $M=5.26$ ,  $SD=2.34$ ) and those who wrote sentences ( $N=29$ ,  $M=6.44$ ,  $SD=2.22$ ;  $t(57)=1.8$ ,  $p=.076$ , two-tailed).

**Table 9***Independent Samples t-test for Immediate Posttest Scores (IV: task type)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Immediate	Equal variances assumed	.004	.950	-1.809	57	.076	-1.07816	.59586	-2.27134	.11502
	Equal variances not assumed			-1.811	56.978	.075	-1.07816	.59530	-2.27024	.11392

Next, another independent-samples t-test was conducted to compare the delayed posttest scores for participants who did the sentence fill-in task and those who did the sentence writing task (see Table 10). There was no significant difference in immediate posttest scores for participants who filled in the blanks (N=30, M=3.63, SD=2.49) and those who wrote sentences (N=29, M=4.51, SD=2.42;  $t(57)=1.37$ ,  $p=.174$ , two-tailed).

**Table 10***Independent Samples t-test for Delayed Posttest Scores (IV: task type)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Delayed	Equal variances assumed	.332	.567	-1.377	57	.174	-.88391	.64180	-2.16908	.40127
	Equal variances not assumed			-1.378	56.998	.174	-.88391	.64149	-2.16847	.40065

## 5. Discussion

The first research question asked about the effectiveness of the four learning conditions (real word/fill-in-real word/sentence writing- pseudoword/fill-in- pseudoword/sentence writing) and aimed to explore if any of them would lead to higher meaning recall scores measured once immediately after the tasks and once one week later. The results of the mixed between-within subjects ANOVA failed to provide any evidence in favour of any of the four learning conditions. The results also showed that participants' scores decreased significantly from the immediate posttest to the delayed posttest, a result which can be attributed to lack of practice or what Hulstijn

and Laufer (2001) call “rehearsal” or additional encounter with the target vocabulary. In fact, upon each retrieval, the link between the form and meaning of the words gets stronger in the learners’ mind (Nation, 2001). In addition, the tasks might not have actively involved and engaged the learners and, hence, adequate processing for acquiring the meaning of words might not have occurred. In addition, other variables such as the aspect of vocabulary knowledge, learners’ language proficiency levels, the quality and frequency of their exposure to the target words and time on task can influence their final attainment (Hazrat & Read, 2022; Liu & Reynolds, 2022; Yanagisawa & Webb, 2021).

The second research question asked whether there were any significant differences between posttests scores of those participants who received real words as their target vocabulary and those who received pseudowords as their target vocabulary. The findings of the two independent-samples t-test on the immediate and delayed posttest scores demonstrated that the answer to this research question is no. This reveals that in this study no firm evidence showed that using pseudowords in lieu of real words may underestimate or overestimate learning of word meanings. Nevertheless, since knowing a lexical unit has various aspects that gradually and continuously can be incorporated into the learners’ mind, when selecting vocabulary types and learning tasks, teachers must consider the usability of the lexical items and the degree of knowledge of these target words learners are required to achieve (Mateo-Valdehi & de Diego, 2021).

The third research question asked whether there were any significant differences between posttest scores of those participants who did the sentence fill-in task and those who did the sentence writing task. The results of the independent-samples t-tests showed that there was no significant difference between the immediate and delayed posttest scores of the participants who did either of the tasks. In line with the hypothesis of the study, participants who wrote sentences outperformed those who filled in the blanks on the posttests. This finding confirms Mateo-Valdehi and de Diego’s (2021) assertion that since sentence writing demands higher levels of cognitive processing and effort, students can acquire and retain a higher number of words in their mental lexicon as a result of using this task. In fact, according to Meara et al. (2005), writing “is a good way to consolidate your knowledge of words”, and it “does not put you under time pressure, so it lets you access and rehearse vocabulary that you can then use later in speech” (pp. 4-5). However, some other researchers believe that writing tasks expend too many resources and might not effectively enrich learners’ vocabulary knowledge (Barcroft, 2004; Folse, 2006).

In addition, the better performance of the sentence writing group was consistent with the predictions of the TFA model because according to this model, the sentence writing task has a higher TFA score in comparison to the sentence fill-in task ( $8 > 7$ ) and this model predicts that the sentence writing task will lead to better meaning recall scores. Accordingly, this finding provides some hints that the TFA is probably a more reliable predictor of vocabulary learning gains, which corroborates the findings of previous studies in terms of the superiority of TFA in predicting “initial and medium-term retention of the meaning of the target words” since it “covers more factors that



have been previously found to be effective for vocabulary learning than does ILH” (Ehsani & Karami, 2022, pp. 181-182). In fact, based on TFA principles, for learning new vocabularies and putting them in their repertoire of linguistics competence, learners need to have clear goals for learning (i.e., motivation), must attend to and negotiate the meaning of new words (i.e., noticing), be able to retrieve the new words (i.e., retrieval) and creatively use them (i.e., generation), and keep the target words in their mental lexicon for subsequent usage (i.e., retention) in receptive and productive activities.

In terms of ILH, although both tasks possess the same overall ILs, the distribution of their IL components differs. The sentence writing task has an evaluation score of 2 and a search score of zero and the sentence fill-in task has an evaluation score of 1 and a search score of 1. Viewing the results from an ILH perspective, one can state that the better performance of the sentence writing group may be due to the higher score of the sentence writing task on the evaluation component, a component that has been suggested to have a heavier weight than the search component (e.g., Zou, 2016). This interpretation should be taken with a grain of salt due to the lack of significant differences in this study. In fact, despite being a useful framework, ILH could only account for and accurately predict a small proportion of variance in incidental vocabulary learning and retention due to the fact that the involvement components (evaluation, need and search) in this framework are given equal importance while the analyses confirmed their different weights and effectiveness (Laufer, 2019; Yanagisawa & Webb, 2021). In fact, besides these elements, the role of other factors such as contextual variables, nature of tasks and modes of input/output must be considered to account for the vocabulary knowledge of learners (Hazrat & Read, 2022).

## 6. Conclusion and Implications

This study aimed to compare the use of two different vocabulary learning tasks (i.e., sentence fill-in and sentence writing) and two word types (i.e., real and pseudo) in terms of their resulting scores on the immediate and one-week delayed meaning recall posttests. The results of the mixed ANOVA did not provide any support for the superiority of any of the four learning conditions (i.e., sentence fill-in with real words, sentence writing with real words, sentence fill-in with pseudowords, and sentence writing with pseudowords) over the other. The findings also revealed that participants’ scores dropped significantly from the immediate posttest to the delayed posttest. A comparison of the two-word types failed to show any overestimation or underestimation of learning due to the use of pseudowords. By comparing the overall results of the posttest scores of the participants who carried out the sentence fill-in and sentence writing tasks some tentative evidence was found in support of the TFA predictions and heavier weight of the evaluation component of the ILH. However, because the results did not reach significant levels, this interpretation should be treated with caution. Nevertheless, educational practitioners need to take into account the significant role of motivation, noticing, retention, productive retrieval and generation components

while designing and giving word tasks to the learners and encouraging their active involvement and engagement in the learning process.

A number of implications for L2 pedagogy and research can be drawn based on the findings of the current study. As for L2 research on vocabulary learning, this study provided some clues that the findings of vocabulary learning studies which use pseudowords can be regarded as comparable with the findings of studies which use real words. This implication is of significance for research on ILH and TFA because previous research on these two framework have used both pseudowords (e.g., Ehsani & Karami, 2022; Keating, 2008) and real words (e.g., Hu & Nassaji, 2016; Kim, 2008) and researchers may wonder how the use of different word types (i.e., pseudowords or real words) may have affected the results of such studies.

Like any other study, the current study had a number of limitations and delimitations which need to be taken into consideration when interpreting the results. One of the limitations of this study can be said to be the uncontrolled time on task. Time on task is related to the amount of time the individual spends on the task and it suggests that if the person spends more time on something, he becomes better at it (Nation & Webb, 2011). This factor was considered important in some studies of vocabulary learning (Knight, 1994; Tu, 2004; Folse, 2006; Keating, 2008; Kim, 2008; Eckerth & Tavakoli, 2012). But some other studies regard time on task as an inherent feature of different tasks (e.g., Laufer & Hulstijn, 2001). According to Eckerth and Tavakoli (2012), different times on tasks may lead to different interpretations of the results of the study because one is uncertain whether the results are due to the task or due to different times spent on the task. The other limitation of this study is that the findings of this study cannot be generalized to other settings because of the relatively small sample which was selected from intact classes. One delimitation of this study was that only intermediate learners were chosen for this study. Further research will be needed to investigate the topic on learners with other characteristics and from other contexts.

## References

- Arndt, H. L., & Woore, R. (2018). Vocabulary learning from watching YouTube videos and reading blog posts. *Language Learning & Technology, 22*(1), 124–142.
- Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research, 20*(4), 303-334.
- Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience, 23*, 3181–3196.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.
- Craik, F. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology, 104*, 268-294.
- Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*(2), 227-252.
- Ehsani, M., & Karami, H. (2022). Comparing the predictive power of involvement load hypothesis and technique feature analysis. *International Journal of Language Studies, 16*(2), 165-188.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, N., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559–617. <https://doi.org/10.1111/j.1467-1770.1993.tb00627.x>
- Folse, K. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly, 40*(2), 273-293.
- Hazrat, M., & Read, J. (2022). Enhancing the involvement load hypothesis as a tool for classroom vocabulary research. *TESOL Quarterly, 56*(1), 387-400. <https://doi.org/10.1002/tesq.3051>
- Hu, H., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. *System, 56*, 28-39.
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning 51*(3). 539-558.
- Iravi, Y., & Malmir, A. (2022). The effect of lexical tools and applications on L2 vocabulary learning: A case of English academic core words. *Innovation in Language Learning and Teaching, 17*(3), 636-649.
- Jafari Gohar, M., Rahmanian, M., & Soleimani, H. (2018). Technique feature analysis or involvement load hypothesis: Estimating their predictive power in vocabulary learning. *Journal of Psycholinguistic Research, 47*(4), 859-869.
- Kaivanpanah, S., Alavi, S. M., & Ravandpour, A. (2020). The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners' incidental vocabulary learning. *Cogent Psychology, 7*(1), 1731223. <https://doi.org/10.1080/23311908.2020.1731223>
- Karami, H., Esrafil, M. (2021). The impact of task type and involvement load index on Iranian EFL learners' incidental vocabulary learning and retention. *Journal of Language Horizons, 5*(1), 251-266. <https://dx.doi.org/10.22051/lghor.2020.31501.1311>
- Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365-386.
- Khoshsima, H., & Eskandari, Z. (2017). Task effectiveness predictors: Technique feature analysis versus involvement load hypothesis. *Iranian Journal of English for Academic Purposes (IJEAP), 6*(2), 50-69.

- Kim, Y. J. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 58*(2), 285-325.
- Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal, 78*, 285-299.
- Laufer, B. (2019). Evaluating exercises for learning vocabulary. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 351-368). Taylor & Francis Ltd.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.
- Meara, P., Batia, L. & Nation, P. (2005). Ten best ideas for teaching vocabulary. *The Language Teacher, 29* (7), 3-6.
- Montero Perez, M. (2020). Incidental vocabulary learning through viewing video. *Studies in Second Language Acquisition, 1-25*.
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Liu, S., & Reynolds, B. L. (2022). Empirical Support for the Involvement Load Hypothesis (ILH): A Systematic Review. *Behavioral Sciences, 12*(10), 354-377. <https://doi.org/10.3390/bs1210035>.
- Nation, I. (2006). How large a vocabulary is needed for reading and listening?. *Canadian Modern Language Review, 63*(1), 59-82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nation, I., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle, Cengage Learning.
- Pallant, J. (2016). *SPSS survival manual* (6th ed.). Open University Press.
- Paribakht, T., S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 174-200). Cambridge University Press.
- Pellicer-S´anchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading. *Studies in Second Language Acquisition, 38*, 97-130.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*(1), 1-28.
- Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading?. *Studies in Second Language Acquisition, 31*(01), 31-58.
- Renandya, W., & Richards, J. (2012). *Methodology in language teaching*. Cambridge University Press.
- San Mateo-Valdehí ta, A., & Criado de Diego, C. (2021). Receptive and productive vocabulary acquisition: Effectiveness of three types of tasks. *Onomázein: Journal of Linguistics, Philology and Translation, 51*, 36-56. <https://dx.doi.org/10.7764/onomazein.51.05>
- Stæhr, L. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal, 36*(2), 139-152.
- Tu, H. (2004). *Effects of task-induced involvement on incidental vocabulary learning in a second language* [Unpublished master's thesis]. National Tsing Hua University, Hsinchu, Taiwan.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*, 46-65.
- 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*(1), 121-145. <https://doi.org/10.1111/lang.12444>.
- Zou, D. (2016). 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.

## Appendices

### Appendix A

#### Learning Condition 1: Sentence fill-in with real words

##### Directions:

Please fill in the blanks with one of the given words. Each word should be used just once. Ask your teacher to give you the word meanings. Four words are extra.

**Disseminate- accessories- impasses- hilarious- miserly- impress- tremble- itch- fades- caprice- diagnosed- assiduous- belittled- infection**

1. The \$300 million tower was built to satisfy the ..... of one man.
2. One of the organization's aims is to ..... information about the COVID-19.
3. I had a painful ear ..... last week.
4. The book offers advice on choosing fabrics, furniture, and .....
5. She didn't like the film at all but I thought it was .....
6. My memory of childhood ..... as time goes by.
7. In 1894 her illness was ..... as flu.
8. Albert Einstein was one of the most ..... researchers in all history.
9. He tried to ..... his teachers by using big words in all his essays.
10. I had a slight ..... on my back.

### Appendix B

#### Learning Condition 2: Sentence writing with real words

Please write one sentence for each of the words below. The meanings and example sentences are provided in the "vocabulary list".

1. Disseminate:
2. Accessory:
3. Hilarious:
4. Impress:
5. Itch:
6. Fade:
7. Caprice:
8. Diagnose:
9. Assiduous:
10. Infection:

##### "Vocabulary list"

**a) Accessory (noun):** Something added to something else to make it more useful, attractive, or effective.

(وسيله فرعى/جانبى)

Exp. She bought some fashion *accessories* such as bracelets and rings.

**b) Assiduous (adjective):** Showing hard work, care, and attention to detail. (سخت كوش، با پشتكار)

Exp. Mary was an *assiduous* student.

**c) Belittle (verb):** Describe as little or unimportant (كوچك شمردن، تحقير كردن)

Exp. The critic belittled the writer's work.

**d) Caprice (noun):** A sudden and unexpected change in someone's opinion or behavior without any good reason.

(هوى و هوس)

Exp. Workers have complained of being at the service of the manager's *caprice*.

- d) Diagnose (verb):** To say exactly what an illness or the cause of a problem is. (تشخیص دادن)  
Exp. The test is used to *diagnose* several diseases.
- e) Disseminate (verb):** To cause something (e.g., information) to go to many people. (انتشار دادن)  
Exp. The internet allows us to *disseminate* information faster.
- f) Fade (verb):** To cause to lose strength gradually; to become weaker. (به تدریج ضعیف شدن)  
Exp. My grandmother's hearing *faded* as she grew older.
- g) Hilarious (adjective):** very funny (بسیار خنده دار)  
Exp. Some people don't like his comedy, but I think he's *hilarious*.
- h) Impasse (noun):** Situation in which no progress seems possible (ین بست، تنگنا)  
Exp. She had reached an *impasse* in her career.
- i) Impress (verb):** if a person or thing impresses you, you admire them or it. (تحت تاثیر قرار دادن)  
Exp. I am *impressed* that you can play the piano so well.
- j) Infection (noun):** The act or process of causing or getting a disease. (عفونت)  
Exp. They tested the patient's blood for signs of the *infection*.
- k) Itch (noun):** An uncomfortable feeling on the skin that makes you want to rub it with your nails. (خارش)  
Exp. I had an *itch* on the back of my neck.
- l) Miserly (adjective):** Stingy, hating to spend money (خسیس)  
Exp. The team's *miserly* owner refused to pay for new equipment.
- m) Tremble (verb):** To shake slightly (لرزیدن)  
Exp. My voice *trembled* as I began to speak.

## Appendix C

### Learning Condition 3: Sentence fill-in with pseudowords

Please fill in the blanks with one of the given words. Each word should be used just once. Ask your teacher to give you the word meanings. Four words are extra.

tasperate - proffendos - impasses- londrous - miserly- starify - tremble- sanch - rades - masco - muliated - asmious - belittled- pacon

- The \$300 million tower was built to satisfy the ..... of one man.
- One of the organization's aims is to ..... information about the COVID-19.
- I had a painful ear ..... last week.
- The book offers advice on choosing fabrics, furniture, and .....
- She didn't like the film at all but I thought it was .....
- My memory of childhood ..... as time goes by.
- In 1894 her illness was ..... as flu.
- Albert Einstein was one of the most ..... researchers in all history.
- He tried to ..... his teachers by using big words in all his essays.
- I had a slight ..... on my back.

## Appendix D

### Learning Condition 4: Sentence writing with pseudowords

Please write one sentence for each of the words below. The meanings and example sentences are provided in the "vocabulary list".

- Tasperate:
- Proffendo:

3. Londrous:
4. Starify:
5. Sanch:
6. Rade:
7. Masco:
8. Muliare:
9. Asmious:
10. Pacon:

#### “Vocabulary list”

- a) Asmious (adjective):** Showing hard work, care, and attention to detail. (سخت کوش، با پشتکار)  
Exp. Mary was an *asmious* student.
- b) Belittle (verb):** Describe as little or unimportant (کوچک شمردن، تحقیر کردن)  
Exp. The critic belittled the writer’s work.
- c) Impasse (noun):** Situation in which no progress seems possible (بن بست، تنگنا)  
Exp. She had reached an impasse in her career.
- d) Londrous (adjective):** very funny (بسیار خنده دار)  
Exp. Some people don’t like his comedy, but I think he’s *londrous*.
- e) Masco (noun):** A sudden and unexpected change in someone’s opinion or behavior without any good reason.  
(هوی و هوس)  
Exp. Workers have complained of being at the service of the manager’s *masco*.
- f) Miserly (adjective):** Stingy, hating to spend money (خسیس)  
Exp. The team’s miserly owner refused to pay for new equipment.
- g) Muliare (verb):** To say exactly what an illness or the cause of a problem is. (تشخیص دادن)  
Exp. The test is used to *muliare* several diseases.
- h) Pacon (noun):** The act or process of causing or getting a disease. (عفونت)  
Exp. They tested the patient’s blood for signs of the *pacon*.
- i) Proffendo (noun):** Something added to something else to make it more useful, attractive, or effective.  
(وسیله فرعی اجنبی)  
Exp. She bought some fashion *proffendos* such as bracelets and rings.
- j) Rade (verb):** To cause to lose strength gradually; to become weaker. (به تدریج ضعیف شدن)  
Exp. My grandmother’s hearing *raded* as she grew older.
- k) Sanch (noun):** An uncomfortable feeling on the skin that makes you want to rub it with your nails. (خارش)  
Exp. I had a sanch on the back of my neck.
- l) Starify (verb):** if a person or thing impresses you, you admire them or it. (تحت تاثیر قرار دادن)  
Exp. I am *starified* that you can play the piano so well.
- m) Tasperate (verb):** To cause something (e.g., information) to go to many people. (انتشار دادن)  
Exp. The internet allows us to *tasperate* information faster.
- n) Tremble(verb):** To shake slightly (لرزیدن)  
Exp. My voice trembled as I began to speak.