International Journal of Language Learning and Applied Linguistics World
(IJLLALW)
Volume 4 (1), September 2013; 137-151 *Abbas Ali Zarei & Maryam Afrash Ab*
EISSN: 2289-2737 & ISSN: 2289-3245 www.ijllalw.org

THE CONTRIBUTION OF WORD FORMATION, CODE-MIXING, MULTIPLE-CHOICE, AND GAP-FILLING TASKS TO L2 VOCABULARY COMPREHENSION AND PRODUCTION

Abbas Ali Zarei
Assistant professor, Imam Khomeini International University, Qazvin, Iran
Email: aazarei@ikiu.ac.ir

Maryam Afrash Ab
MA, Islamic Azad University, Takestan, Iran
Email: mafrashab@yahoo.com
(PHONE: +9183857952)

ABSTRACT
This study sought to investigate the effect of various task types including Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task on L2 vocabulary comprehension and production. A sample of 80 adult Iranian intermediate level learners participated in the study. They were in 4 groups; each group read the reading passages under one of the four conditions of Multiple-choice cloze, Gap-filling, Word formation, and Code-mixing tasks. At the end of the treatment, two posttests were administered. A 30-item vocabulary test in multiple-choice format was used to measure the participants' vocabulary comprehension. Another 30-item vocabulary test in fill-in-the-blanks format was utilized to measure the participants' vocabulary production. Two separate one-way ANOVA procedures were used to analyze the obtained data. The results showed an effect of task type, the word formation task was more effective than the other task types in vocabulary comprehension and production. It also turned out that those who were involved in the Gap-filling task were the least successful in both vocabulary comprehension and production. The present study may have theoretical as well as practical implications for the teachers, researchers, and syllabus designers.

KEYWORDS: Task-based language teaching, vocabulary learning, task type

INTRODUCTION
Vocabulary knowledge occupies an important position in second language learning. Numerous attempts have been made to solve the challenges facing learners in reference to vocabulary learning. However, doubts remain over their effectiveness in improving learners' vocabulary comprehension and production (Song, 2011).

One approach which is claimed to be beneficial in coping with the problems of vocabulary learning is task-based language teaching (TBLT). In this adaptable approach of language teaching, tasks are employed as the main pedagogical tool to teach diverse language elements. Language learning occurs through the process of completing tasks, and learners master the target

language more effectively when they are exposed to meaningful task-based activities in a natural way. Proponents of task-based language teaching believe that the uptake of vocabulary, as one of the key elements in second language learning, can be enhanced by employing proper pedagogical tasks. However, designing effective pedagogical tasks has always been a huge challenge for second language (L2) researchers and teachers.

Current second language (L2) instruction research has encouraged the use of different tasks in (L2) classrooms. Task-based language teaching provides learners with access to both explicit and implicit learning experience. It has the ability to integrate meaning-focused communication with form-focused instruction. However, the positive impact of utilizing different task types on vocabulary learning is still open to question. The fact that this debate over the effectiveness of diverse task types in promoting vocabulary uptake remains unresolved indicates a need for further examination and research in this regard. The present study attempts to investigate whether employing different task types can influence the uptake of new vocabulary items.

REVIEW OF LITERATURE

Task-based language teaching is considered as a powerful approach for maximizing language learning and teaching, which employs tasks as its main pedagogical tools to structure language teaching. In TBLT, learners are to use language to perform meaningful tasks and, thus, language which is meaningful to learners can pave the way for the learning process (Willis, 1996).

In defining task, we should draw a distinction between real-world tasks and pedagogical tasks, where the former refers to the tasks which the learner could possibly be required to carry out in the real world, and the latter refers to the tasks which would be highly unlikely for the learner to perform outside the language classroom (Rashtchi & Keyvanfar, 2007). Although pedagogic tasks seem to be of little real-world value, Widdowson (1987) argues that "tasks can be thought of as an investment to be drawn on to meet unpredictable communicative needs." (p. 68)

Nunan (1989, p.10) states that task is "a piece of classroom work which involves learners in comprehending, producing or interacting in the target language while their attention is principally focused on meaning rather than form". Willis (1996, p. 53), on the other hand, stresses that task is "a goal-oriented activity in which learners use language to achieve a real outcome". While these definitions vary somewhat, they all emphasize the fact that pedagogical tasks involve communicative language use in which the user's attention is focused on meaning rather than grammatical form (Ellis, 2000).

In TBLT literature, there are many kinds of task categories. Willis (1996) offers six main types of including listing, ordering and sorting, comparing, problem solving, sharing personal experiences, and creative tasks (Richards & Rodgers, 2001, p. 234). Furthermore, According to Rashtchi and Keyvanfar (2007, p. 115) there are different classifications of task types including the following:

- Form-focused activities (exercises)
- Meaning-focused activities (tasks)

- Focused tasks
  - Grammatical
  - Consciousness-raising
- Unfocused tasks

❖ Pedagogic tasks (Interactionally authentic)
- Information-gap
- Opinion-gap
- Reasoning-gap

❖ Target tasks (situationally authentic)

Focused tasks have two objectives: one is to encourage communicative language use and the other is to encourage learners to use some specific predetermined linguistics item(s) (Rashtchi & Keyvanfar, 2007, p. 112).

Unfocused tasks can be classified based on the degree of their similarity to real-life situation. Some tasks correspond to the everyday life of learners and, thus, they are said to have situational authenticity. On the other hand, some tasks are rather artificial in the sense that they are unlikely to happen in everyday life; but, they tend to elicit the kind of language that corresponds to the language of everyday-life interaction. These kinds of pedagogic tasks are said to have interactional authenticity (Bachman, 1991).

A large body of research has been carried out on the application of tasks in various aspects of language learning. This study is focused on the effectiveness of tasks on receptive and productive L2 vocabulary learning. Ellis and He (1999) found that the interaction between output and dialogic interaction could contribute to productive as well as receptive vocabulary knowledge. Similarly, de la Fuente (2002) found that only negotiated interaction that incorporated output appeared to have promoted both receptive and productive learning of words, and increased productive word retention.

There is a belief that learners' pushed output can contribute to vocabulary learning in a number of ways. In this regard, Swain and Lapkin (1995, p. 376) argue that output tasks can lead to the noticing of linguistic shortcomings, "pushing" learners to modify output. Actually, output has a very interactive and significant role on vocabulary acquisition. There is also a theoretical claim that output may serve as a crucial means to strengthen connections between the lemma (the particular form that is chosen by convention to represent the lexeme) and the lexeme (the set of all the forms that have the same meaning). Such strengthened connections between the lemma and the lexeme enable learners to have easy access to and efficient control of vocabulary knowledge stored in their L2 language system. As Nation (1990, p. 86) puts it, "when L2 learners are engaged in output production, they are required to actively solve problems of word form or of word meaning on their own. Through such active processing of lexical information, a learner can achieve faster, more precise and automatic use of vocabulary knowledge than when just hearing or reading L2 vocabulary."

Maftoon and Haratmeh (2012) used a pedagogical approach to investigate the relative effectiveness of tasks with different involvement loads on the vocabulary knowledge of Iranian EFL learners. The goal was to investigate the way that the construct of involvement load is related to the Input Hypothesis (Krashen, 1985) and the Output Hypothesis (Swain, 1996) to see whether the involvement load or input/output-orientation of tasks is the determining factor in task effectiveness. Contrary to the predictions of the Involvement Load Hypothesis (Laufer & Hulstijn, 2001), the results of the study indicated that involvement load is not the only determining factor in task effectiveness, but input/output-orientation of tasks is also a decisive parameter in task effectiveness. While Laufer and Hulstijn's proposal is the first valuable step towards building a theory of vocabulary learning, the results of the study indicated that involvement index may well not function independently of the task type, i.e., input or output orientation of a word-focused task.

Newton (1995) carried out a case study examining the vocabulary gains made by a group of adult learners of English as a second language as a result of performing four communicative tasks. Gains were measured on comparisons of pre- and post-tests of vocabulary from the worksheets from the four tasks. He found that the placement of a word on task worksheets and the nature of a task, whether a split information task or a shared information task, both had a strong effect on the use and learning of new vocabulary.

Joe (1998) examined the effects of text-based tasks and background knowledge (prior vocabulary knowledge and a disposition to use generative learning tactics when tackling new vocabulary) on incidental vocabulary learning. 48 adult ESL learners were randomly assigned to one of three treatments (a) reading and retelling a text with explicit generative training and without access to the text during recall, (b) reading and retelling a text without explicit generative training but with access to the text during recall, and (c) neither reading nor retelling a text. All subjects sat a pre-test (individual interviews and a read and retell task) and post-tests (individual interviews and two multiple-choice tests) designed to tap partial vocabulary knowledge gains. Results indicated that the process of reading and retelling a text promotes incidental vocabulary learning and that generative processing enhances vocabulary learning with greater levels of generative processing leading to greater vocabulary gains for unknown words.

Rott (2004) investigated whether L2 readers' sensitivity towards a new lexical form is heightened if they are repeatedly pushed to produce output and are immediately provided with relevant input in input-output cycles. Fourth-semester learners read three texts, with four target words each, under the following conditions: (a) cued-output task, (b) self-selected output task, and (c) un-enhanced (control) reading. Results showed that four input-output cycles did not contribute to retaining more robust form-meaning connections (FMCs) than the normal reading condition. In all three conditions, FMCs varied in strength and completeness, requiring different cues for retrieval.

Newton (2001) argued that rather than removing difficult words, teachers should consider a number of cooperative options for exposing learners to new words during task-based interaction. He examined data from a number of classroom tasks where learners had to deal with new words

during task performance without access to a dictionary or teacher's intervention. The results suggested that not only rich language use results from negotiating new words, but that the meanings of many of these words are retained in the days after the task performance.

Kim's (2011) study consisted of two experiments investigating the involvement load hypothesis in vocabulary learning. Experiment 1 compared the performance of 64 adult English as a second language (ESL) learners from a range of countries at two different proficiency levels (i.e., matriculated undergraduate students vs. students in an Intensive English Program) to ascertain the effectiveness of three vocabulary tasks with different levels of task-induced involvement. Experiment 2 investigated whether two tasks hypothesized to represent the same level of task-induced involvement would result in equivalent initial learning and retention of target words by 20 adult ESL learners at two different levels of proficiency. The results of Experiment 1 showed that a higher level of learner involvement during the task promoted more effective initial vocabulary learning and better retention of the new words. The findings of Experiment 2 indicated that when different tasks had the same involvement load, they resulted in similar amounts of initial vocabulary learning and retention of new words.

These studies show that researchers seem to have a consensus as to the role of task based instruction in vocabulary learning. Despite this consensus on the general role of task-based instruction in vocabulary learning, there seems to be a paucity of research on how difficult task types can affect L2 vocabulary, particularly in an EFL context. In response to this paucity, the present study addresses the effects of Input-oriented and Output-oriented tasks on Iranian EFL learners' vocabulary comprehension and production.

RESEARCH QUESTIONS

More specifically the present study aims to answer the following questions:

Q1. Which task type (Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task) is more effective on L2 vocabulary comprehension?

Q2. Which task type (Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task) is more effective on L2 vocabulary production?

METHODOLOGY

Participants

A sample of 113 Iranian male and female EFL learners participated in the study. They were studying English at Sokhane Ashna and Shokohe Andisheh institutes in Tehran. The subjects were selected from the intermediate level of proficiency, and their age ranged from 18 to 30. After the administration of the Michigan test of English language proficiency, 33 participants were excluded from the study because of either a different language proficiency test score, or not writing their names on the papers. There remained 80 learners to take part in the study.

Instruments and materials

To conduct the present study, the following instruments were employed: a Michigan general language proficiency test was administered to homogenize the participants and to validate the posttests. The sample of Michigan test used in this study contained 25 multiple-choice items. It was an already established and verified in terms of reliability.

The target words for this study were selected from reading passages taken from Cambridge ESOL Examinations. To make sure that the participants had no prior knowledge of the target words, and to minimize the effect of their prior vocabulary knowledge, a vocabulary pretest containing 100 items was given to all the participants. These lexical items were contextualized in 100 English sentences. Each item contained one of the target words and required students to supply the Persian equivalent of the italicized words in the sentence. Those words the meaning of which the participants did not know were selected for inclusion in the posttest, and the familiar words were excluded from the posttests.

The instructional materials were ten reading passages taken from Cambridge FCE (First Certificate in English) examination. The passages were in an appropriate difficulty level to roughly match the learners' ability**.** Each passage contained 10 target words. Each of the four groups read the reading passages under one of the four conditions of Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task.

At the end of the experimental period, a vocabulary comprehension posttest was administered to compare the effects of Multiple-choice cloze, Gap-filling, Word formation, and Code-mixing tasks on L2 vocabulary comprehension. It was a 30-item vocabulary test in the multiple-choice format. A 30-item fill-in-the-blanks vocabulary test was also used to measure the participants' productive knowledge of vocabulary. In the fill-in-the-blank test, the first letter of each word was given with its translation in Persian. This was done to ensure that the learners could produce the target words and to prevent the possibility of learners providing either partial synonyms or other words that fitted the context without necessarily being the intended words.

Procedures

Initially, a total number of 113 participants were selected. To homogenize the participants, a Michigan test of English language proficiency was administered. It was an already established and verified test in terms of reliability. As a result, 33 participants, who had scored more than one standard deviation away from (above or below) the mean, were excluded from subsequent statistical analyses, and there remained 80 approximately homogenous participants to take part in this study. The participants were divided into four groups. Each group was randomly assigned to one of the treatment conditions as follows:

Group A: Multiple-choice cloze task
Group B: Gap-filling task
Group C: Word formation task
Group D: Code-mixing task

To make sure that the participants had no prior knowledge of the target words and to minimize the effect of prior knowledge, a vocabulary pretest containing 100 items was administered prior to the treatment. The participants were required to supply the Persian equivalent of the italicized English words in 100 sentences. Each sentence contained one of the target words which had been extracted from the reading passages the learners were supposed to receive as treatment. The words which turned out to be familiar to more than three participants were excluded from the subsequent vocabulary comprehension and production post-tests.

In the next phase, the treatment began. Each group of participants received their treatment under one of the following conditions. The four task types consisted of the Multiple-choice cloze task (Group A); Gap-filling task (Group B); Word formation task (Group C), and Code-mixing task (Group D).

The participants of Group A, the Multiple-choice cloze group, received ten cloze texts. Each text contained ten gaps and was followed by ten four-option multiple choice questions. The Participants of this group were required to fill the gaps in each text by choosing the right words from a list of ten four-option multiple choice items.

The Gap-fill group (Group B) was provided with ten cloze texts. In each text ten target words were deleted, leaving ten gaps. The target words and their English explanations were provided in random order as a list on a separate page along with five distractors in each text. The task for this group was to read the text and complete the ten gaps with the most appropriate words from a list of 15 words.

Group C was treated with Word formation task. They received ten texts. Each text contained ten gaps, each gap corresponding to a word. The 'stems' of the missing words were given along with the text and had to be transformed to the target and missing words. The types of word formation involved not only the addition of affixes (e.g. 'honest' to 'dishonesty'), but also internal changes ('strong' to 'strength') and compounding (e.g. 'rain' to 'raindrop'). The misuse of capital letters and other punctuation marks was ignored, but correct spelling was important.

Ten cloze texts were given to the participants of Group D, the Code-mixing task group. In each text, ten words were deleted, leaving ten blanks and missing words. The Persian equivalent of each missing word was provided in parentheses after each blank. These Persian equivalents were the clear hint to the missing words. The task for the learners was to fill each of the blanks with the English equivalent of the Persian words given in parentheses.

The experimental period lasted for 8 weeks, of which 5 weeks were allocated to the treatment, two weeks to the Michigan test and the pretest, and one week to the posttests. It needs to be noted, however, that not all the class time was used for the treatment each session. Since the learners were taking their general English course, only a third of each class time every week (about 45 minutes) was allocated to the experiment.

At the end of the experimental period, two post-tests were administered. A 30-item vocabulary test in multiple-choice format was used to measure the participants' vocabulary comprehension knowledge. Another 30-item vocabulary test in fill-in-the-blanks format was utilized to measure the participants' vocabulary production. To validate the post-tests, and to avoid creating learner sensitivity toward the target words, the Michigan proficiency test was given to a group of 30 students with characteristics similar to the target groups (EFL learners studying at the intermediate level) concurrently with the post-tests. To check the validity of the posttests, the correlation between these tests and a Michigan test was checked. The validity index of the comprehension and production posttests turned out to be .83 and .82, respectively. The reliability of the post-tests was also estimated through the KR-21 formula. The reliability index of the vocabulary comprehension and production post-tests turned out to be .42 and .66, respectively.

The validated post-tests were then administered to the 80 approximately homogeneous participants. The obtained data were then summarized and submitted to statistical analyses. Two separate one-way ANOVA procedures were run on the immediate posttest of vocabulary comprehension and production.

RESULTS AND DISCUSSIONS

Investigation of the first Research Question

The first research question sought to investigate the effect of various task types including Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task on L2 vocabulary comprehension. To do so, a one-way ANOVA was used. Descriptive statistics are summarized in *Table 1*.

Table 1: Descriptive Statistics for the ANOVA on Vocabulary Comprehension

	N	Mean	Std. Deviation	Std. Error
Code-mixing	20	22.4500	2.64525	.59150
Word formation	20	24.7500	2.51050	.56137
Multiple-choice	20	21.5000	2.72416	.60914
Gap-filling	20	20.2000	2.83957	.63495
Total	80	22.2250	3.11783	.34858

Table 1 indicates that the highest mean on the vocabulary comprehension test belongs to the Word formation group ($\bar{x}$ = 24.75), followed by the Code-mixing group ($\bar{x}$ =22.45). The third highest mean belongs to the Multiple-choice cloze group ($\bar{x}$ =21.50), and the Gap-filling group ($\bar{x}$ = 20.20) has the lowest mean.

In order to see whether the observed mean differences among the selected groups are statistically significant, the one-way ANOVA procedure was used. The results of the ANOVA procedure are summarized in *Table 2*.

Table 2: the ANOVA Procedure on Vocabulary Comprehension

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	221.050	3	73.683	10.239	.000
Within Groups	546.900	76	7.196		
Total	767.950	79			ω^2=.25

In *Table 2*, the observed F-value and the significance level ($F_{(3.76)} = 10.23$, $p < .05$) indicate that there are statistically significant differences among the four groups. Therefore, it can be safely claimed that there are significant differences among the effect of these four task types on vocabulary comprehension.

At the same time, the index of the strength of association ($\omega^2 = 0.25$) indicates that 25 percent of the total variance in the dependent variable (vocabulary comprehension) is accounted for by the independent variable (task types). This means that the remaining 75 percent of the variance is left unaccounted for.

In order to locate the differences among the study groups, a post hoc Scheffe test was utilized. The results are summarized in *Table 3*. As it can be observed in *Table 3*, the mean score of the Word formation group is significantly better than the mean scores of the Multiple-choice and the Gap-filling groups, suggesting that the participants of the word formation group have outperformed their counterparts in the two groups, but the mean scores of the latter two groups do not differ significantly from each other.

Table 3: Multiple comparisons for the ANOVA on vocabulary comprehension

					95% Confidence Interval	
(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Code –mixing	Word formation	-2.30000	.84830	.070	-4.7254	.1254
	Multiple-choice	.95000	.84830	.741	-1.4754	3.3754
	Gap-filling	2.25000	.84830	.080	-.1754	4.6754
Word formation	Multiple-choice	3.25000*	.84830	.004	.8246	5.6754
	Gap-filling	4.55000*	.84830	.000	2.1246	6.9754
Multiple-choice	Gap-filling	1.30000	.84830	.507	-1.1254	3.7254

*. The mean difference is significant at the 0.05 level.

The result further indicates that although there is a difference between the means of the Code-mixing group and the Word formation group, the difference is statistically insignificant. Moreover, the result shows that there are no statistically significant differences between the Code-mixing group and the Multiple-choice group, as well as between the Code-mixing group

and the Gap-filling group. The following graphic representation of the results *(Figure.1*) shows the differences among these groups more conspicuously

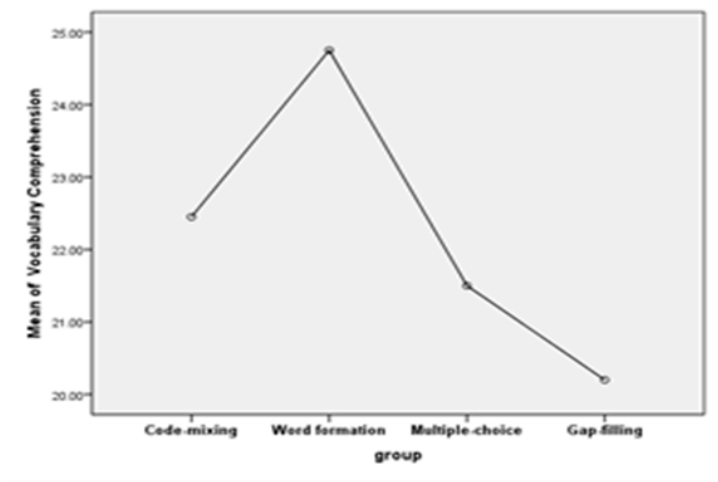

Figure1: Means Plot on the Vocabulary Comprehension

Investigation of the second Research Question

The aim of the second research question was to examine the effect of various task types including Multiple-choice cloze task, Gap-filling task, Word formation task, and Code-mixing task on L2 vocabulary production. To this end, another one-way ANOVA procedure was run. The descriptive statistics are presented in *Table 4.*

Table 4: Descriptive Statistics for the ANOVA on Vocabulary Production

	N	Mean	Std. Deviation	Std. Error
Code-mixing	20	22.5500	4.09717	.91616
Word formation	20	22.9000	3.89196	.87027
Multiple-choice	20	19.3500	4.01674	.89817
Gap-filling	20	18.2500	3.29074	.73583
Total	80	20.7625	4.27013	.47742

Table 4 indicates that the highest mean ($\bar{x}$ =22.90) belongs to the Word formation group, followed closely by the mean of Code-mixing group ($\bar{x}$ =22.55). The Multiple-choice group has the third position ($\bar{x}$ =19.38). The participants of the gap-filling group have the lowest mean ($\bar{x}$ =18.46), which is noticeably lower than the other groups. To see whether or not the observed differences among the groups are statistically significant, the one-way ANOVA procedure was used. The obtained results are presented in *Table 5.*

Table 5: The results of the ANOVA on vocabulary production

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	321.437	3	107.146	7.277	.000
Within Groups	1119.050	76	14.724		
Total	1440.487	79			ω^2=.19

As it can be seen in *Table5,* the F-value and the significance level ($F_{(3.76)}$ =7.27, $p < 0.05$) are indicative of significant differences among the means. Moreover, The index of the strength of association ($\omega^2 = 0.19$) shows that 19 percent of the total variance in the dependent variable (vocabulary production) is accounted for by the independent variable (task types), and that the remaining 81percent is left unaccounted for.

Another post hoc Scheffe test was utilized to locate the differences among the groups. The results of the multiple comparisons are summarized in *Table 6.*

Table 6: Multiple Comparisons of Means for the ANOVA on Vocabulary Production

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Code -mixing	Word formation	-.35000	1.21344	.994	-3.8194	3.1194
	Multiple-choice	3.20000	1.21344	.082	-.2694	6.6694
	Gap -filling	4.30000*	1.21344	.008	.8306	7.7694
Word formation	Multiple-choice	3.55000*	1.21344	.043	.0806	7.0194
	Gap -filling	4.65000*	1.21344	.004	1.1806	8.1194
Multiple-choice	Gap -filling	1.10000	1.21344	.844	-2.3694	4.5694

*. The mean difference is significant at the 0.05 level.

Based on *Table 6,* the difference between the Code-mixing group and the Gap-filling group is statistically significant, with the Code-mixing group outperforming the Gap-filling group. It is worth noting that there are no statistically significant differences between the Code-mixing group and the Word formation group, as well as between the Code-mixing group and the Multiple-choice group.

Moreover, the Scheffe test indicates that the mean score of the Word formation group is significantly better than the mean scores of the Multiple-choice and the Gap-filling groups, but the mean scores of the latter two groups do not differ significantly from each other. In other words, the participants of the Word formation group have outperformed those of the Multiple-choice and the Gap-filling groups, suggesting that Word formation task is the most effective task on vocabulary production.

The following graphic representation *(Figure 2)* shows the differences among the groups more conspicuously.

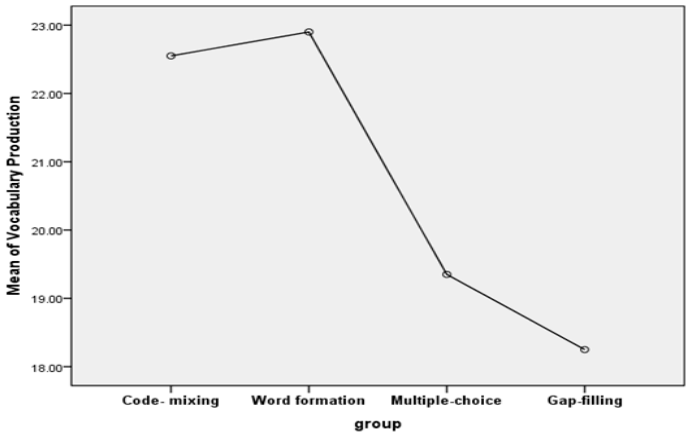

Figure 2: Means Plot on the Vocabulary Production Test

DISCUSSION

Considering the findings of the study, we can conclude that there is a large gap between receptive knowledge from input-oriented tasks and output-oriented tasks in a way that productive tasks are more effective than receptive tasks on both vocabulary comprehension and production. This result is in contradiction with Krashen's input hypothesis, avowing that input alone is sufficient for the development of both receptive and productive second language knowledge. In addition, based on Krashen's natural order hypothesis, speech emerges. In other words, even productive knowledge develops naturally from receptive knowledge. This implies that input-oriented activities alone are not only capable of creating language knowledge, but also even more effective than output-oriented activities on the development of both receptive and productive language knowledge. Drawing on the distinction between receptive and productive knowledge of vocabulary, the finding of this study seems to support the role of output.

At the same time, there are studies that are in line with the findings of this study. In an experimental study, Ellis and He (1999) found that the interaction between output and dialogic interaction could be a beneficial factor for learners to acquire productive as well as receptive vocabulary knowledge. Similarly, de la Fuente (2002) found that only negotiated interaction that incorporated output promoted both receptive and productive learning of words, and increased productive word retention.

CONCLUSIONS AND IMPLICATIONS

Based on the findings of this study, it can be concluded that in spite of the teacher-centered tradition of instruction in Iranian context, applying output-oriented tasks gives sufficient room to students to show themselves. In other words, if treated with output-oriented tasks, students become more engaged in the learning of the target language elements, and they take more active roles in the learning process. It can be concluded that employing output-oriented tasks can keep students motivated, and facilitate students' learning processes.

In the Iranian context, language instruction is mainly based on the input. Target language elements are mostly learnt through the receptive skills. There is no sufficient room for students to show their abilities through productive skills. On the other hand, employing output-oriented tasks necessitates learners to produce the intended language items. The production of the target language items through output-oriented tasks may have a greater level of appeal to learners due to their novelty. This sense of motivation and enthusiasm may partially account for the dominance/superiority of the output-oriented tasks over the input-oriented tasks.

Based on the obtained results, we can claim that this study lends support to the output hypothesis proposed by Swain (1985). The output hypothesis suggests that successful second language acquisition requires not only comprehensible input, but also comprehensible output, language produced by the learner that can be understood by other speakers of the language. It has been argued that when learners have to make efforts to ensure that their messages are communicated (pushed output), this puts them in a better position to notice the gap between their productions and those of proficient speakers, fostering acquisition (Richards & Schmidt, 2002).

The findings of the present study can have implications not only for teachers and learners, but also for materials developers. The knowledge of the effects of diverse task types on vocabulary learning can help teachers make more informed decisions as to which task types to choose to engage their learners in. Moreover, the results may be helpful in improving learners' autonomy. Given the superiority of output-oriented tasks over input-oriented tasks, learners can be allowed to experience greater levels of autonomy by being actively engaged in productive activities and assume more responsibility for their own learning rather than being passive bench-bound recipients of information.

Furthermore, the results of the current study might provide useful insights for the developers of instructional materials and syllabus designers in their selection of effective word-focused tasks in EFL General English materials. It can hardly be denied that adequate and sufficient vocabulary knowledge leads to good comprehension. Thus, given the importance of vocabulary in EFL General English classes, any word-focused task that helps learners to develop their vocabulary knowledge would certainly be welcomed.

And finally, this study could possibly lay the groundwork for a great deal of research to touch on the effect of different word-focused tasks on various aspects of vocabulary knowledge.

All in all, despite the apparent areas of discrepancy as to which task type is more beneficial than which other types, there seems to be almost a consensus that, overall, tasks are effective and conducive to language learning in general and vocabulary learning in particular. Yet, the discrepancies among the findings of various studies as to the effectiveness of different task types on vocabulary learning, coupled with areas of gap between the findings of this study and those of other similar studies warrant more studies in the future.

LIMITATIONS AND DELIMITATIONS OF THE STUDY

In the present study, the following limitations and delimitations should be taken into account:

1. There are various types of tasks considered as pedagogical tools in task-based language teaching approach. It was impossible to compare all task types in one study. Therefore, this study was confined to a limited number of selected tasks. This implies that the findings of this study may not be generalizable to other task types.
2. Due to time and administration limitations, only 80 language learners participated in the present study. Therefore, care must be exercised in generalizing the findings.
3. The students who participated in this study were male and female. This means that the participants' gender was not a variable.
4. The proficiency level of the participants was constrained to intermediate level. Thus, care should be taken in generalizing the result to learners of other proficiency levels.
5. The present study investigated the effect of different types of task on students' vocabulary comprehension and production; this means that the other language skills (e.g. reading comprehension, listening, speaking, etc.) were not of concern here.
6. While participants were carrying out the selected tasks, their performance could be affected by many different factors including their personality type, level of anxiety or motivation, and their learning styles. Additionally, it was impossible to impose complete control on the students' probable practice outside the classroom.
7. The vocabulary pretest (one hundred-item check-list) used in this study was not verified in terms of reliability. Therefore, more trustable findings will be reached using vocabulary pretest verified in terms of reliability.

REFERENCES

Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly*, *25(4),* 671 704.

de la Fuente, M .J. (2002). Negotiation and oral acquisition of L2 vocabulary. The roles of input and output in receptive and productive acquisition of words. *Studies in Second Language Acquisition, 24,* 81-112.

Ellis, R. (2000). Task-based Research and Language Pedagogy. *Language Teaching Research, 4 (3),* 193-220.

Ellis, R. & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. Studies in Second Language Acquisition, 21, 285–301.doi:10.1017/S0272263199002077

Joe, A. (1998). What effect do text-based tasks promoting generation have on incidental vocabulary acquisition? Applied Linguistics, 19, 357–377. doi:10.1093/applin/19.3.357

Kim, Y, J. (2011). The Role of Task-Induced Involvement and Learner Proficiency in L2 Vocabulary Acquisition. *Language Learning, 61(1),* 100-140.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications* (Vol. 1, p. 985). London: Longman.

Laufer, B., & Hulstijn, J. (2001). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics, 22 (1),* 1-26.

Maftoon, P., & Haratmeh, M. S. (2012). Effects of Input and Output-oriented Tasks with Different Involve-ment Loads on the Receptive Vocabulary Knowledge of Iranian EFL Learners. *Iranian Journal of Research in English Language Taeching*, *1 (1),* 65-88.

Nation, I. S. P. (1990).*Teaching and Learning Vocabulary.* New York: Newbury House.

Newton, J. (1995). Task-based Interaction and Incidental Vocabulary Learning: A case study. *Second Language Research, 11(2),* 159-176.

Newton, J. (2001). Options for vocabulary learning through communication tasks. *ELT journal, 55(1),* 30-37.

Nunan, D. (1989). *Designing Tasks for the Communicative Classroom*. Cambridge, UK: Cambridge University Press.

Rashtchi, M., & Keyvanfar, K. (2007). *ELT Quick'n'easy* (3rd ed.). Rahnama Publication: Tehran.

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

Richards, J.C., & Schmidt, R. (2002). *Dictionary of language teaching & applied linguistics* (3rd ed.). Essex: Longman.

Rott, S. (2004). A comparison of output interventions and un-enhanced reading conditions on vocabulary acquisition and text comprehension. *Canadian Modern Language Review, 61(2),* 169-202.

Song, W. (2011). *Learning vocabulary without tears: a comparative study of the jigsaw and information gap tasks in vocabulary acquisition at school* (Unpublished master thesis). Kristianstad University Sweden, School of Education and Environment. Retrieved February 7, 2013, from: http://hkr.diva-portal.org/smash/record.jsf?pid=diva2:439234.

Swain, M. (1985).Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and Second Language Acquisition* (pp. 235-253). New York: Newbury House.

Swain, M. (1996). Three Functions of Output in Second Language Learning. In G. Cook & B. Seidlhover (Eds.), *principles and practice in the Study of Language. Studies in Honor of H.G.Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics*, 16(3), 371-391.

Widdowson, H. G. (1987). Aspects of Syllabus Design. In M. Tickoo (ed.): *Syllabus Design: The State of the Art.* Singapore: Regional English Language Centre.

Willis, J. (1996). *A Framework for Task-based Learning*. Harlow, UK: Pearson Education.