ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 5, No. 9, pp. 1883-1889, September 2015
DOI: http://dx.doi.org/10.17507/tpls.0509.16

Visiting Involvement Load Hypothesis and Vocabulary Acquisition in Similar Task Types

Hassan Soleimani
Department of Applied Linguistics, Payame Noor University, Iran

Mahboubeh Rahmanian
Payame Noor University, Iran

***Abstract*—"Involvement load hypothesis" in second language literature in spite of its inclusiveness has not yet provided a definite answer to vocabulary acquisition. This study investigated *task-induced involvement* suggested first by Laufer and Hulstijn (2001) in similar task types. In order to better examine the involvement hypothesis in incidental learning, applying Nelson Proficiency Test, 70 high proficiency EFL learners from two different English institutes in Isfahan, Iran, were randomly assigned to two experimental groups: gap-filling task completion and reading comprehension task completion. The two input task types had two and one involvement indexes, respectively. The results of repeated measures *ANOVA* considering the performance of the two homogeneous groups in pretest, posttest, and delayed posttest confirmed Laufer and Hulstijn's (2001) involvement hypothesis in similar task types. The study suggests that teachers and language learners can use tasks with higher involvement indexes regardless of their type in order to improve their vocabulary acquisition.**

***Index Terms*—involvement load hypothesis, input task type, output task type, vocabulary acquisition, involvement index**

I. Introduction

There can be found many theories in the literature that strengthened the cognitive approach to learning. *Noticing hypothesis* (Schmidt, 1990, 1994), limited processing ability (VanPatten, 1990), and "pushed output" (Swain, 1985) can be regarded as the most important findings in this field. Among them the theory which induced Laufer and Hulstijn (2001) to develop their innovative and influential *involvement load hypothesis* was Craik and Lockhart's (1972) *depth of processing hypothesis*. Craik and Lockhart (1972) believe that a stimulus can be processed at a deeper level rapidly for a better retention if it is in coordination with the previously-existent information. Laufer and Hulstijn (2001) claim that manipulating tasks in terms of their building blocks of *need*, *search*, and *evaluation* affects the retention of unfamiliar vocabulary items. Although involvement load hypothesis is in line with many previous theories in the literature, it has many ambiguous points and in effect in need of further consideration. Task type can be enumerated as one these challenging points which even though Laufer and Hulstijn (2001) believed in their neutral parts; some researchers (e.g., Hulstijn and Laufer, 2001; Xu, 2009; Walsh, 2009) found the superiority of output over input ones when the two input and output task types were investigated. This study is determined to investigate involvement load hypothesis in tasks with similar types (i.e., input task type) to see if the predictions of Laufer and Hulstijn (2001) can come true.

II. Literature Review

Craik and Lockhart's (1972) depth of processing hypothesis hold memory as a cluster of perceptual processing levels which can be classified into sensory analyses (passing outcomes), pattern recognition, stimulus elaborations (more stable outcomes), and so forth. To put it simply attention, time of processing, and coordination with long-held information can lead to deep processing which , in turn, better retention will be resulted. Laufer and Hulstijn (2001) introduced their innovative construct considering these processing stages that "starting with shallow sensory analysis, and proceeding to deeper, more complex, abstract, semantic analysis" (Solso, 1988, p.133). Their hypothesis holds that we can have better vocabulary retention if more values are devoted to a task in terms of its building blocks. To put it simply need, search, and evaluation components of a task can be of three degrees, that is zero, "moderate", and "strong" which are for their absence, one, and two involvement indexes, respectively. Understandably they state that the more the involvement indexes, the better the retention.

If a task induces learners to put their minds up to select a more appropriate word for a given context moderate evaluation, and if it induces them to compose a sentence strong evaluation occurs. Similarly, the need component can be regarded as moderate (i.e., external driving force) and strong (i.e., internal driving force) in order for a task to be done. Search, on the other hand, deals with finding out different dimensions of a new word consulting a dictionary or teacher.

Although Laufer and Hulstijn's (2001) *task-induced involvement* was proposed to connect the abstract concepts of noticing, elaboration, motivation, or need in tasks, it was the first attempt in this field which also has given existence to

some subsequent researches. Let us point out some empirical evidences which have studied this hypothesis more deeply and can be added to the group of studies that support Laufer and Hulstijn's (2001) theory.

Hulstijn and Laufer's (2001) two parallel experiments were conducted to test their hypothesis. Advanced adult Dutch and Hebrew students did three different kinds of tasks (i.e., reading comprehension with marginal glosses, gap-filling, and composing original sentences with the unfamiliar words). These three tasks have different involvement load indexes since the three components of involvement load have been divided unequally between them. The third task as an output type with an involvement load of three proved to be of higher retention in comparison with the other two input task types with one and two involvement indexes, respectively.

In another paper by Jianping Xu (2009), who worked on 152 freshman university students, three tasks (multiple-choice comprehension questions, blank-filling, and sentence making) were designed. He tried to see his results in the light of time, that is the more the time passage, the less the retention. As a matter of fact, Xu's analyses give us an understanding of Laufer and Hulstijn's theory. Along the same lines, Using 223 high school learners with different levels of second language (L2) vocabulary knowledge, Walsh (2009) compared moderate and strong evaluation. Participants took a delayed test and it was revealed that composing original sentences as an output type with higher involvement index helped in better retention compared to gap-filling as an input one with lower involvement index.

As can be observed, most of the studies which have been done in the area of task-induced involvement deal with different task types with different involvement indexes; however, rare studies have been done with similar task types to see if they can confirm this rooted and comprehensive theory. In the present study we will consider this aspect of Laufer and Hulstijn's hypothesis.

Research Hypotheses

The following null hypotheses were designed in order to unravel this point of the involvement load hypothesis.

Null Hypothesis 1: Using gap-filling task as an input type has no effect on enhancing vocabulary acquisition of Iranian EFL students.

Null Hypothesis 2: Using reading comprehension task as an input type has no effect on enhancing vocabulary acquisition of Iranian EFL students.

III. Method

A. Participants

A hundred and forty male and female Iranian English as foreign language (EFL) students with a mean age of 22 years were selected. Since these intermediate students were from two branches of English language institutes in Isfahan, Iran, they used the same materials. The experiment was done in June-August, 2013 in a six week period during the participants' class time. Nelson Placement Test helped us to have more homogeneous sample groups. In other words, our 140 students were given a placement test in order to let us assign them into two major high and low proficiency groups. Then, randomly, the first 70 students of the high proficiency group were assigned to two different tasks, that is gap-filling and reading comprehension tasks. Therefore, our two experimental groups each containing 35 participants completed a gap-filling task and a reading comprehension task, respectively.

B. Instruments

The materials which were utilized in order to let us conduct our experiment were: Nelson Placement Test and a reading passage with five reading comprehension questions.

1. Nelson Placement Test:

The second version (intermediate) 200 A of Nelson Placement Test consisting of 50 multiple-choice grammar questions was used in order to help us have two homogeneous sample groups from the 140 students of the two English institutes. The K-S normality index in Nelson Placement Test demonstrated the normal distribution of the data (p =.128, $p < .05$). As a result, the parametric independent-sample *t-test* was applied. Levene Test of equality of variances unraveled the homogeneity of the variances between the two groups (p = .538, 2-tailed at $p < .05$). Consequently, no difference was observed between the means of the groups before the treatment.

2. Reading Passage:

Our sample population teachers and some intermediate students of our considered institutes who were not among the participants of our study were given three intermediate reading texts which were of similar readability index, of the participants' general knowledge and vocabulary domain. Furthermore, the reading passages had the single occurrence of target words in order to use Laufer and Hulstijn's (2001) statements for the results explanation and not multiple exposures.

These persons not only selected Walsh (2009) *Child Labor* reading text with 326 words and 7.73 Gunning Fog index readability operating an on-line utility system, but also selected the unknown words of Walsh reading passage for our participants in order to let us administer our pretest. The pretest had 36 vocabulary items and a reliability coefficient of .62 that caused the selection of these ten target words from the passage as two verbs, six nouns, one adjective, and one adverb: *plantation, fair, demonstrations, crops, sweatshop, fiber, partly, blame, march,* and *shrimp.*

C. Data Collection Procedure

Thanks to Nelson Placement Test, we had two homogeneous major groups regarding their level of proficiency. Afterwards, 70 students of the high proficiency group were randomly assigned to two input task types, that is gap-filing task and reading comprehension task. As their proficiency level, age, and used materials were the same, it was assumed that we had a homogeneous sample.

A pretest was administered one week before the treatment (i.e., task completion) in order to be sure that the participants had no prior knowledge concerning the target unknown vocabulary items. Since this experiment was conducted to investigate incidental vocabulary acquisition more profoundly, "learners … typically required to perform a task involving the processing of some information without being told in advance that they will be tested afterwards on the recall of all the words in the list" (Laufer and Hulstijn, 2001, p. 10). Then, the researchers brought the two groups up with two different guidelines in order to help them with completing the tasks. In following the principles of incidental learning the researchers did not provide the participants with any explanation about the meaning of the target words and introduced the treatment or task completion as a reading practice. Afterwards, the first and the second groups were given the gap-filling (Task A) and reading comprehension (Task B) tasks respectively as the treatment in order to make the effect of task on vocabulary acquisition clear.

Walsh reading text with its omitted target lexical items was prepared for the first group in order to do Task A. The participants had to complete the blank spaces using ten L2 glossed words and answer the comprehension questions. The moderate need (extrinsic motivation of filling the gaps of the reading passage), no search (no looking the words up in the dictionary), and moderate evaluation (judging the most appropriate words for the gaps) composed an involvement index of two for this input task type. In accordance with Hulstijn and Laufer's (2001) formulation, we can have (1) need, + (0) search, + (1) evaluation.

The same reading comprehension passage was prepared with the first language (L1) translation of the target vocabulary items. The participants of this task had to answer the multiple-choice questions at the end of the passage. Like the previous task this text contained ten target words in bold print to catch the participants' attention (Schmidt, 1994). The current task was in accordance with involvement load hypothesis, that is its components were manipulated in order to reveal its influence on vocabulary retention. The involvement index for this task was one since it had a moderate need (participants had to know the target words to answer the reading passage comprehension questions), no search, and no evaluation on the part of participants (i.e., (1) need, + (0) search, + (0) evaluation).

The solution which was applied in order to make the values of the need, search, and evaluation components of each task stable was asking the participants not to talk about the target words after the task administration (Walsh, 2009). In addition, their papers were collected after the treatment to prevent them from committing the words to memory or searching the target words after completing the tasks. As time-on-task was not given so much prominence in the present study, a normal time period was devoted to each one, that is fifteen and ten minutes for the first and second tasks. Therefore, the treatment was merely the time when the groups were exposed to the reading passage (i.e. tasks). In other words, they were not provided with the full text of the reading passage at any other time.

Two days and two weeks after the treatment of tasks, immediate and delayed posttests were administered in five minutes. Albeit each of them had the ten target vocabularies, they had the difference of the order of the words. The scoring procedure of zero to one was used for incorrect and correct answers, respectively. In addition, half a score was regarded for semantically close translations of the vocabulary items in order to make a more precise data collection. The experiences of other teachers were also used for approximate answers.

D. Data Analysis

Graphical and statistical tests were applied to reveal if our distribution of data was normal. Upon K-S normality index, the type of statistical tests to be conducted was decided on using their posttest scores. K-S normality index for the first experimental group unraveled that at the $p < .05$ level the assumption of normality was not rejected: $p = .09$ and as a result the parametric repeated measure *ANOVA* was run. Similarly, K-S normality index for the second experimental group showed normal distribution of the data and in effect the parametric repeated measure *ANOVA* was conducted ($p<.05$, $p = .120$).

IV. Results

A. Homogeneity Test Results

The K-S normality index of the students in Nelson Placement Test let us know if our two sample groups were homogeneous. The significant statistic ($p =.128$, $p < .05$) revealed the normal distribution of the data. Hence, the parametric independent-sample *t-test* was utilized. Levene Test of equality of variances demonstrated that the variances of scores across the groups were homogeneous ($p = .538$, 2-tailed at $p < .05$). Therefore, no difference could be observed between the means of the groups before conducting the experiment.

B. Findings of the Study

1. Testing the First Null Hypothesis:

RH0 1: Using gap-filling task as an input type has no effect on enhancing vocabulary acquisition of Iranian EFL students.

The effect of gap-filling task with an involvement load index of two was investigated to see if there was a statistically significant effect on the amount of the acquisition of new vocabulary items. K-S normality index for the first experimental group (i.e. 35 participants of the high proficiency group doing a gap-filing task) revealed that at the $p < .05$ level the assumption of normality was not rejected: $p = .09$. Therefore, the parametric statistic test repeated measure of *ANOVA* was run. Table 1 shows the descriptive statistics for the immediate and delayed posttest scores of the gap-filling group. In general, the descriptive statistics indicated that the gap-filling group obtained the highest mean scores on completing a gap-filling task in the immediate posttest ($M = 7.14$) in comparison with the pretest ($M = 2.22$) and the delayed posttests ($M = 7.05$). In other words, this group containing 35 participants worked better in the immediate posttest than the other two time points of pretest and delayed posttest.

TABLE 1.
DESCRIPTIVE STATISTICS OF THE GAP-FILLING GROUP WITH L2 GLOSS

	M	SD	N
pretest	2.22	1.35	35
immediate posttest	7.14	1.39	35
delayed posttest	7.05	1.60	35

A one-way repeated measures *ANOVA* (Table 2) revealed that there was a significant difference at the $p < .05$ level among the vocabulary mean score of the three tests, Wilks' Lambda = .06, F (2, 33) = 240.45, $p = .000$, with a large effect size ($d = .93$) in accordance with Cohen (1988) criteria.

Therefore, it is implied that the first null hypothesis was rejected. In other words, there was a statistically significant difference in vocabulary mean scores of the gap-filing group with L2 glossing who did three tests at different points of time. This finding could reveal that tasks with an involvement load of two created a difference in the vocabulary acquisition of the participants. In order to pinpoint the location of the differences, the *Scheffé* post hoc test was conducted (Table 3) which demonstrated that each of the differences was significant except the differences between posttest and delayed posttest.

TABLE 2.
REPEATED MEASURES ANOVA OF THE GAP-FILLING GROUP WITH L2 GLOSS

Effect		Value	F	Hypothesis df	Error df	Sig.	Cohen's d
Task B	Pillai's Trace	.93	240.45[a]	2.00	33.00	.000	.93
	Wilks' Lambda	.06	240.45[a]	2.00	33.00	.000	.93
	Hotelling's Trace	14.57	240.45[a]	2.00	33.00	.000	.93
	Roy's Largest Root	14.57	240.45[a]	2.00	33.00	.000	.93

Note. Design: Intercept Within Subjects Design: Task B.
[a] Exact statistic.

TABLE 3.
SCHEFFÉ POST HOC FOR THE GAP-FILLING GROUP WITH L2 GLOSS

(I) Task B		(J) Task B		Mean Difference (I-J)	S E	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
di	1	on2	2	-4.91*	.25	.000	-5.54	-4.28
			3	-4.82*	.27	.000	-5.51	-4.14
	2	D2	1	4.91*	.25	.000	4.28	5.54
			3	.08	.04	.249	-.03	.20
	3	dim	1	4.82*	.27	.000	4.14	5.51
			2	-.08	.04	.249	-.20	.03

Note. Based on estimated marginal means.
[a] Adjustment for multiple comparisons: Bonferroni.
* $p < .05$ level.

2. Testing the Second Null Hypothesis:

RH0 2: Using reading comprehension task as an input type has no effect on enhancing vocabulary acquisition of Iranian EFL students.

In accordance with Laufer and Hulstijn, (2001), Task B had an involvement index of one (moderate need, no search, and no evaluation). K-S normality index for the second experimental group (i.e. 35 participants of the high proficiency group completing a reading comprehension task) revealed that at the $p < .05$ level the assumption of normality was not rejected: $p = .120$. Since the obtained data were normal and other assumptions of parametric techniques were met, the parametric statistic test repeated measure of *ANOVA* was conducted. Table 4 demonstrates the descriptive statistics for the immediate and delayed posttest scores of reading comprehension group. It can be seen that the descriptive statistics for the reading comprehension group had the highest mean score value in the immediate posttest ($M = 3.94$) rather than

the pretest (M = 2.28) and the delayed posttests (M = 2.82). To put it simply, this group containing 35 participants unraveled better results in the immediate posttest than the other two points of time (pretest and delayed posttest).

TABLE 4.
DESCRIPTIVE STATISTICS OF THE READING COMPREHENSION GROUP WITH L1 GLOSS

	M	SD	N
pretest	2.28	1.04	35
immediate posttest	3.94	.80	35
delayed posttest	2.82	1.01	35

The results of one-way repeated measures *ANOVA* (Table 5) revealed a statistically significant difference at the p < .05 level across the three tests, Wilks' Lambda = .12, F (2, 33) = 116.73, p = .000. Moreover, a large effect size (d = .87) was unraveled upon Cohen (1988) criteria.

Consequently, the second null hypothesis was rejected since there was a statistically significant difference in vocabulary means scores of the reading comprehension group with L1 glossing. It can be claimed that tasks with an involvement load of one could bring about a difference in the vocabulary acquisition of the participants. Table 6 which demonstrates the location of the differences applying the *Scheffé* post hoc test indicates that each of the differences was significant except the differences between pretest and delayed posttest.

TABLE 5.
REPEATED MEASURES ANOVA OF THE READING COMPREHENSION GROUP WITH L1 GLOSS

Effect		Value	F	Hypothesis df	Error df	Sig.	Cohen's d
Task B	Pillai's Trace	.87	116.73[a]	2.00	33.00	.000	.87
	Wilks' Lambda	.12	116.73[a]	2.00	33.00	.000	.87
	Hotelling's Trace	7.07	116.73[a]	2.00	33.00	.000	.87
	Roy's Largest Root	7.07	116.73[a]	2.00	33.00	.000	.87

Note. Design: Intercept Within Subjects Design: Task B.
[a] Exact statistic.

TABLE 6.
SCHEFFÉ POST HOC FOR THE READING COMPREHENSION GROUP WITH L1 GLOSS

(I) TaskB		(J) TaskB		Mean Difference (I-J)	S E	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
di	1	on2	2	-1.65*	.14	.000	-2.02	-1.28
			3	-.54	.21	.054	-1.09	.00
	2	Di	1	1.65*	.14	.000	1.28	2.02
			3	1.11*	.13	.000	.77	1.45
	3		1	.54	.21	.054	-.00	1.09
			2	-1.11*	.13	.000	-1.45	-.77

Note. Based on estimated marginal means.
[a] Adjustment for multiple comparisons: Bonferroni.
* p < .05 level.

V. DISCUSSION AND CONCLUSIONS

The research hypotheses examined similar task types to investigate involvement load hypothesis more deeply. Two days and two weeks after the tasks administration, the participants' behaviors were investigated. As can be observed significant differences were found from the treatment of tasks considering the participants' initial learning and long-term retention.

The first research hypothesis that was an input type with an involvement index of two revealed no significant difference between initial and long-term retention (i.e., the two posttests). To put it simply, Task A which induces participants to compare their already known knowledge to the new ones (i.e., moderate evaluation) is superior in initial learning. However, in long-term retention we can see a stable vocabulary acquisition. It seems that the principles of incidental learning did not come true for gap-filling task since there was the persistent effect of Task A with L2 glossing in spite of the interval between the two posttests. In order to explain these contradictory results, Watanabe (1997) mentions repeated exposure which can be in the form of some following confirmatory activities. As these activities can not be seen here, another ambiguous point is added to the field of incidental learning.

On the other hand, the second research hypothesis considered Task B in order to see the differences which it could have on the participants' vocabulary acquisition. Task B which had an involvement index of one induced the participants to read the reading passage and answer its comprehension questions at its end. Initial learning and long-term retention of the target words through the immediate and delayed posttests were investigated. As can be seen a significant decline was resulted between the immediate and delayed posttests in administering Task B. Contrary to the

previous hypothesis, poor vocabulary acquisition was observed which was in line with previous researches in incidental vocabulary acquisition (e.g., Hulstijn and Laufer, 2001; Keating, 2008; Watanabe, 1997, Herman et al., 1987; as cited in Hui-Fang Tu, 2003). The present study like most of the previous researches (e.g., Xu, 2009; Walsh, 2009) confirmed Laufer and Hulstijn's (2001) claim, that is better retention of target words were found for Task A with a higher involvement index.

Even though Laufer and Hulstijn (2001) state that involvement index of the target words is the determinant factor in long-term retention and not its task type, most of the studies, as far as we found in the literature, had worked on input and output type of tasks with different involvement indexes to see the superiority of output types. Task type is a tentative aspect of task-induced involvement that there can not be found many studies in the L2 field about it. Unless we can not bring this hypothesis true for different task types, other explanations (e.g., Swain, 1985) can be resorted to about the superiority of output types. Swain (1995) holds that the output task types can have more cognitive effort on the part of the students and better retention can be resulted. Here, we could state that two tasks with different involvement indexes but of the same type can be in line with the involvement load hypothesis predictions.

Implications and Limitations of the Study

It can be inferred from the current study that regardless of task type, the involvement index is the important factor in determining the better acquisition of the target words. To put it simply, teachers or language learners can use tasks with higher involvement indexes in order to have a more positive impact on the students in terms of vocabulary retention.

This study is exposed to some limitations. First, teachers' attitude, students' psychology (Lee, 2003), and type of teachers' reinforcement (Hulstijn and Laufer, 2001) can be enumerated which we can attribute our results to. Time is another factor which can explain the results. Although Laufer and Hulstijn (2001) know it as an inherent feature of tasks, we did not take it in to account. Some researchers (e.g., Swanborn and de Glopper, 2002) argue that in order for the tasks to be compared, the same time length should be devoted to each. Therefore, more researches need to be done to reveal the impact of task-induced involvement on vocabulary retention especially with regard to the type of tasks.

REFERENCES

[1] Craik. F. I. M. and R. S. Lockhart. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11: 671-684. Retrieved April 11, 2014 from http://dx.doi.org/10.1016/S0022-5371(72)80001-X (accessed 11/4/2013).

[2] Herman. P. A., R. C. Anderson, P. D. Pearson, and W. E. Nagy. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly* 22: 263-284. Retrieved April 11, 2014from http://www.jstor.org/discover/10.2307/747968?uid=2&uid=4&sid=21103781743801(accessed 11/4/2013).

[3] Hu. H.-F. and I. R. Tu. (2003). Effects of task-induced involvement on incidental vocabulary learning in a second language. Unpublished MA thesis. National tsing hua university, Taiwan. Retrieved April 11, 2014 from http://www.hss.nthu.edu.tw/~fl/thesis/tesol/905258.pdf. (accessed 15/4/2013).

[4] Hulstijn. J. H. and B. Laufer. (2001). Some empirical evidence for involvement load hypothesis in vocabulary acquisition. *Language Learning* 51: 539-558. doi: 10.1111/0023-8333.00164(accessed 15/4/2013).

[5] 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. doi: 10.1177/1362168808089922(accessed 11/4/2013).

[6] Laufer. B. and J. Hulstijn. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics* 22 (1): 1-26. doi: 10.1093/applin/22.1.1(accessed 11/4/2013).

[7] Lee, C. (2003). The effect of involvement load on English vocabulary acquisition in an incidental learning situation. Unpublished MA thesis. National tsinghua University, hsinchu, Taiwan.

[8] Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11: 129-158. doi: 10.1093/applin/11.2.129(accessed 13/4/2013).

[9] Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. In J. H. Hulstijn and R. Schmidt (eds) *Consciousness in second language learning.* AILA Review 11: 11-26.

[10] Solso, R. L. (1988). Cognitive psychology (2nd ed.). Boston, Mass.; London: Allyn and bacon.

[11] Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass and C. Madden (eds) *Input in second language acquisition*. (pp. 235-253). Rowley, MA: Newbury house.

[12] Swain, M. (1995). Three functions of output in second language learning. In G. Cook and B. Seidhofer (eds) *Principles and practice and the study of language.* (pp. 125-144). Oxford university press. Retrieved April 16, 2014 from http://www.scribd.com/doc/105840639/Swain-1995-Three-functions-of-output-in-second-language-learning (accessed 8/4/2013).

[13] Swanborn. M. S. L. and K. de Glopper. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning* 52: 95–117. doi: 10.1111/1467-9922.00178 (accessed 12/4/2013).

[14] VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition* 12 (3): 287–301. doi: http://dx.doi.org/10.1017/S0272263100009177(accessed 11/4/2013).

[15] Walsh, M. I. (2009). The involvement load hypothesis applied to high school learners in Japan: Measuring the effects of evaluation. Unpublished MA thesis. Birmingham University, Edgbaston, United Kingdom.

[16] Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition* 19: 287-307. doi: http://dx.doi.org/ (accessed 9/4/2013).

[17] Xu, J. (2009). An experimental study on the effects of different reading tasks on L2 vocabulary acquisition. *English Language Teaching* 2: 69-76. Retrieved April 16, 2014 from http://ccsenet.org/journal/index.php/elt/article/download/3698/3293 (accessed 11/4/2013).

Hassan Soleimani is an assistant professor at the University of Payame Noor, Tehran, Iran, where he teaches computer-assisted language learning and EFL curriculum development for Ph.D. candidates, and research methods and language teaching methodology to graduate and undergraduate students. He has written some books, including An Introduction to Non-parametric Statistics for Applied Linguistics Research (2009), and articles in national and international journals. He also serves as the editorial board member of some journals. His areas of interest include research methodology and statistics, curriculum design, and SLA issues.

Mahboubeh Rahmanian is an instructor at Al-Mahdi University. She teaches pre- and general English and is interested in issues like: applied linguistics, cognitive psychology, and vocabulary learning. She has been teaching English for 9 years at English institutes and has a Master's in Teaching English from Payame-Noor University in Iran. She has been written several articles in applied linguistics field and is the editorial member of an English journal.