International Journal of English Linguistics; Vol. 8, No. 5; 2018
ISSN 1923-869X E-ISSN 1923-8703
Published by Canadian Center of Science and Education

A Verification of Involvement Load Hypothesis on Chinese Adult English Learners

Shuyun Huang[1]

[1] School of Foreign Languages, Huaiyin Normal University, Jiangsu, China

Correspondence: Shuyun Huang, School of Foreign Languages, Huai'an, Jiangsu, 223300, China. E-mail: huangsmaggie@hytc.edu.cn

Received: April 21, 2018 Accepted: May 14, 2018 Online Published: May 23, 2018

doi:10.5539/ijel.v8n5p125 URL: https://doi.org/10.5539/ijel.v8n5p125

Abstract

The present research designed six tasks with various distributions of involvement components: need, search and evaluation to verify the predictability of Involvement Load Hypothesis on Chinese adult English learners. The results showed that the vocabulary exercises did facilitate the incidental vocabulary acquisition, but the exercise with higher involvement load did not necessarily benefit the students more than the exercise with lower involvement. Three components of involvement did not reveal the same effect on incidental vocabulary acquisition. And the superiority of exercise with higher involvement load existing in the immediate vocabulary test did not survive in the delayed vocabulary test. In the delayed vocabulary test there were not any statistically significant differences among six groups. The further analysis reported besides the cognitive processing aroused by the tasks, other critical factors also worked on the incidental vocabulary acquisition: inference skill and repetition of occurrence.

Keywords: Involvement Load Hypothesis, incidental vocabulary acquisition, task-induced involvement, processing capacity

1. Introduction

There has been a long-running debate about which method is more effective for the learners to acquire vocabulary: direct instruction or incidental learning (Chen & Truscott, 2010; Hulstijn & Laufer, 2001; Laufer & Roitblat, 2011). Incidental vocabulary acquisition offers the language learners a rich context than the traditional meaning-form match. "A large amount of vocabulary is learned without an overt intention, in other words, incidentally" (Ender, 2016). It is the autonomous learning adjusting to the various needs and language proficiency of the learners. At the same time it is pedagogically efficient to practice two activities: vocabulary acquisition and reading, writing or listening. Many studies have been conducted to investigate different aspects of incidental vocabulary acquisition, including the types of multimodal input; the frequency of the target words; the work of marginal glosses and dictionary; the modification of input like typographical enhancement (Chen & Truscott, 2010; Ellis &Le, 2016; Folse, 2006).

One criticism of incidental vocabulary acquisition for foreign language learner is that large exposures perceived by the native language learners do not accessible for L2 learners, especially for intermediate and even low proficiency learners. Some aspects of vocabulary learning are more amenable to conscious learning like the properties of a word's semantic and conceptual features.

What needs clarification is that incidental vocabulary acquisition actually can be both implicit and explicit cognitive processes. Due to the limited implicit processes, "it seems clear that explicit processing methods such as consulting a dictionary, inferring the meaning based on different associations or a combination of these two approaches are more promising for learning" (Ender, p. 558). Therefore task-induced reading and listening direct the learner's attention on the language forms without the destruction of top-down understanding.

Wesche and Paribakht (2000, p. 200) grouped vocabulary output tasks. They discerned that whether a given task succeeds in eliciting the anticipated mental response largely depends upon a learner's perception of the required outcome, relevant prior knowledge (including English proficiency), and his or her motivation and perseverance.

Table 1. Typology of text-based vocabulary exercises

typology	purpose	example
Attention	draw learners' attention to the target item	ask the learners to underline or circle the target words every time they appear
Recognition	recognize the target items and their meanings	match the word with its definition or synonym; recognize the meaning from multiple-choice
Manipulation	require structural analysis of target words to rearrange/organize given elements	give derivations of words; use stems and affixes to construct words
Interpretation	involve semantic and syntactic analysis	show words' collocation (syntagmatic or paradigmatic)
Production	require retrieval and production of target word in appropriate original contexts	construct a sentence or composition with target items; answer questions with them

2. Literature Review

2.1 Incidental-Intentional and Implicit-Explicit

The underpinning of incidental vocabulary acquisition is that the learners can acquire the new words without the intention to do so, and it is the by-product of the main cognitive activity. The feature of unintentional retention is always associated with implicit learning, which is totally unconscious and unintentional learning of language items in a meaningful context. The counterpart of the implicit learning is explicit learning. The dichotomy intentional-incidental concentrates on the learner's aim of accomplishing language activity. The later pair explicit-implicit focuses on the process involved in the activity. "Thus, the dichotomy incidental-intentional lies beside but also somehow across the pair implicit-explicit" (Ender, p. 538). When he is processing the reading for the aim to understand the general idea and finish the comprehension questions, the learner manages to discover the form-meaning correlations of the new words with conscious attention. That explicit learning can happen without the intention to retain the process in the long term memory. In conclusion, intentional learning is only triggered by explicit processing of input; while incidental learning embodies implicit process, explicit process or the combination of both.

Gass (1999) proposed three conditions for the incidental learning: "(a) there is recognized cognation in the native language, (b) the learner has heard or read the word inform numerous times, and (c) the learner knows the word information (L2 word association)" (Gass, p. 323). In other words if a word satisfies all these three criteria, there is high likelihood for incidental vocabulary acquisition. Otherwise the absence of these positions requires the deliberate intentional learning.

2.2 Processing Effect

The Levels of Processing Effect identified by Craik and Lockhart (1975) described memory recall of stimuli as a function of the depth of mental processing. Information is processed through an ordered hierarchy of levels beginning with physical features to phonemic to semantic. They proposed that a shallow encoding dealing with superficial characteristics such as physical analysis leads to poor recall in incidental learning. Deeper encoding makes the recall easily accessible in incidental learning. It is the method and depth of processing that affect how an experience is stored in memory. That is in what way the input is elaborated depends on the perceptual analyses carried out. For example if the stimuli are analyzed in a shallow sensory level, the input gives rise to very transient memory traces. On the other hand, stimuli in higher levels processing yield a deeper recoding and a longer trace.

Lockhart and Craik (1990) further expanded the processing effect by highlighting the two stages in the language learning. In the first stage, the orthographic and phonological features of the input are analyzed. In the second stage semantic and conceptual features are retrieved with deeper analysis. "In this model, it is critical for acquisition not only initial attention, noticing and processing of words, but also their subsequent retrieval and consolidation of the semantic encoding of the word features in memory (Hu & Nassaji, p. 29).

2.3 The Involvement Load Hypothesis

Laufer and Hulstijn (2001) developed three dimensional involvement load hypothesis and proposed a motivational-cognitive construction to operationalize depth of processing and degree of elaboration. "Retention of words when processed incidentally is conditional upon the following factors in a task: need, search, and evaluation" (Laufer & Hulstijn, p. 14).

The need component is the motivational, noncognitive dimension of involvement. Two degrees of prominence are suggested for need: moderate and strong. Need is moderate when it is imposed by an external agent like the

learners. Need is strong when it is intrinsically motivated, that is self-imposed by the learners themselves. For instance, learners feel the desire to look up a word in a dictionary when they read an instruction of a washing machine.

Search is a cognitive dimension of involvement, which is an attempt to find the meaning of an unknown L2 word or the attempt to find the L2 word form expressing a concept (e.g., trying to find the L2 translation of an L1 word) by consulting a dictionary or another authority.

Evaluation is also a cognitive dimension of involvement with a distinction of moderate and strong. The learners evaluate the syntagmatic and paradigmatic features, genre and register preference in a new context. The moderate evaluation is required in the exercises like cloze. The strong evaluation is the output tasks like writing, translating or oral presentation.

Each of the above three factors can be absent or present when the learners are processing a word in a natural or artificial designed task. The combination of three factors with the degree of prominence constitutes involvement load. The underlying idea of Involvement Load Hypothesis is that the retention of news words is conditional upon the degree of involvement when processing the words. "In other words, it is conditional upon who sets the task, whether the new word has to be searched, and whether it has to be compared, or combined with other words. The greater the Involvement load, the better the retention" (Hulstijn and Laufer, p. 545).

A group of researchers verified the efficiency of involvement hypothesis (Hu & Nassaji, 2016; Kim, 2008; Zou, 2017). Liu (2013) proved the more effectiveness of Chinese to English translation task than English to Chinese translation task in both immediate and delayed vocabulary test. The students in the English to Chinese group "directly resorted to their mother tongue Chinese to find equivalents, which was conducted in a lexical level" (333) with glossed target words. Comparatively in the process of Chinese to English, the learners first consulted their Chinese mental lexicon for the target words, then again experienced searching English mental lexicon for the image. More searching component is involved in the Chinese to English group.

Lee and Pulido (2017) evaluated the interactions between the factors of topic interest, L2 proficiency, and gender in L2 vocabulary learning by reading. The findings support the proposition topic interest significantly assists the vocabulary acquisition. The intrinsic motivation is more strongly induce the learners' involvement in the tasks. It was also accepted language proficiency works as an indicator of vocabulary learning. While the postulation of gender difference on L2 learning.

Wang (2015) designed four tasks with different load index: reading only (0+0+0), vocabulary note (1+0+0), gap filing (1+0+1), and productive task (1+0+1). The results were in accordance with the premise that the word gains increase as more tasks are performed. And more importantly "a single exercise after reading an article is likely to help learners retain very few words" (162). It is suggested that reading activity with at least three other tasks (like vocabulary note, gap filling) would guarantee incidental vocabulary acquisition.

Zou (2017) inspected the component of evaluation in the Involvement Load Hypothesis, examining how three approaches (cloze, sentence writing and composition writing) promote word acquisition from the perspectives of information organization in cognitive science: chunking and hierarchical organization. The failure of moderate evaluation group is attributed to the lack of chunking. The composition writing group outperformed the sentence writing group because of the hierarchical organization and pre-task planning.

The present research intends to verify the predictability of Involvement Load Hypothesis on the Chinese adult English learners and to closely observe the efficiency of vocabulary acquisition by reading. Three research questions are listed:

1). Do all six vocabulary exercises facilitate the incidental vocabulary acquisition and retention?

2). Do the vocabulary exercises with the different amount of involvement load have the different effects on vocabulary acquisition and retention?

3). Do all three components: need, search and evaluation have same effect on vocabulary acquisition and retention?

3. Research Design

Six groups of sophomores from a college in Jiangsu Province attended the test lasting for three weeks. A reading material about the various friendship between women was selected with its clear text structure and plain language. The Flesch-Kincadi measure of grade is 7.2 and the ratio of unfamiliar words is controlled under 2%. Being different form the previous researches where the targets words were presented in original form, ten target words of the present study were replaced by the disguised forms. The participants of all six groups were offered

the reading material passage and eight comprehension questions to help the students understand the passage.

In the present research the incidental vocabulary acquisition is driven by the vocabulary tasks, not intrinsically desired by the learners themselves. Therefore the need index of all six group was 1. The search was set into two levels. The absence of search means the news words had been glossed or explained by the teachers or other authorities. In the present research group 1, 3 and 5 (all with search index 0) were provided with glosses, including the part of speech, the English paraphrase, the Chinese translation and one illustrative phrase or sentence which was all taken from Oxford Advanced Learner's English-Chinese Dictionary (Four Edition). The presence of search requires the learners to find the appropriate meaning for the L2 word. In this study group2, 4 and 6 (all with the search index 1) consulted ten polysemous target words in the dictionary. Due to the employment of nonsense words, the researcher copied the words explanations from the dictionary with the replacement of each entry into its disguised form. The evaluation is the retrieve and appreciation of the target words in the original context. The tasks in Group 1 and 2 did not require the learners to evaluate the context. Group 3 and 4 both embodied moderate evaluation by processing the higher semantic knowledge in the given context. The last two exercises demanded the learners to produce original sentences largely dependent on the intricate processing and elaboration with strong evaluation.

Six types of vocabulary exercises were different in amount of involvement load and the distribution of the three components (need, search and evaluation).

Table 2. A summary of the six tasks and involvement load index

	Components of involvement load			Amount
Tasks	need	search	evaluation	index
1. answer true or false questions	1	0	0	1
2. read with dictionary	1	1	0	2
3. complete sentences	1	0	1	2
4. read with dictionary & conduct cloze	1	1	1	3
5. make original sentences	1	0	2	3
6.read with dictionary & make sentences	1	1	2	4

In the immediate vocabulary test, the participants in six groups were all unexpectedly given a list of ten target words and other ten distracters. They were asked to write down the Chinese lexicalization for each word. In the two-week delayed vocabulary test, same words were presented in the order different from the immediate vocabulary test. The participants were also required to write down the Chinese meaning for each word.

4. Results

The results in the immediate vocabulary test partially proved the underlying assumption that all six tasks facilitated the incidental vocabulary acquisition and retention. The mean of immediate vocabulary test followed such ascending order from group 1 to group 6: group 1(Mean=5.05) < group 2 (Mean=5.90) < group 3 (Mean=6.38) < group 4 (Mean=6.75) < group 5 (Mean=8.28) < group 6 (Mean=9.03). However the Post Hoc Tests revealed not all mean differences between any two groups reached to statistically significant difference. And the difference was not found in the following pairs: task 2 vs. task 3 and task 4 and vs. 5. The results proved that two tasks with the same accumulating involvement load but diversified distribution of three components exhibited the same contribution to the immediate vocabulary acquisition. For example Task 2 and Task 3 shared the same involvement load index 2 with different distribution of search and evaluation. It is the same case in another pair Task 4 vs. Task 5.

The statistics also suggested that the superiority of vocabulary exercise with higher involvement load did not definitely come into better incidental vocabulary retention. In the immediate vocabulary test there was no statistical difference in such pairs: Task 1 vs. Task 2 (sig=.590 >.05), Task 3 vs. Task 4 (sig=.981>.05), and Task 5 vs. Task 6 (sig=.741>.05). In terms of the components distributed in each pair, need and evaluation were at the same level and search presented differently. In the first pair task 1 and task 2 were in the same levels of need (＋) and no evaluation (－) with the only difference of component search. In the second pair task 3 and task 4 shared the same levels of need (＋) and moderate evaluation (＋) with the only difference of component search. In the third pair task 5 and task 6 were in the same levels of need (＋) and strong evaluation (＋＋) with the only difference of component search. Therefore the failure of search was to explain the absence of difference between two tasks in three pairs. The analysis drew conviction that the component of search did not work on the vocabulary acquisition no matter in what combination with other two components.

Table 3. Post Hoc Tests of immediate vocabulary test

		Mean Difference	Sig
Task 6 (1+1+2)	Task 1 (1+0+0)	3.97368	**.000**
	Task 2 (1+1+0)	3.12354	**.000**
	Task 3 (1+0+1)	2.65132	**.000**
	Task 4 (1+1+1)	2.27632	**.000**
	Task 5 (1+0+2)	.74854	**.714***
Task 5 (1+0+2)	Task 1 (1+0+0)	3.22515	**.000**
	Task 2 (1+1+0)	2.37500	**.000**
	Task 3 (1+0+1)	1.90278	**.006**
	Task 4 (1+1+1)	1.52778	**.052***
Task 4 (1+1+1)	Task 1 (1+0+0)	1.69737	**.018***
	Task 2 (1+1+0)	.84722	**.608***
	Task 3 (1+0+1)	.37500	**.981***
Task 3 (1+0+1)	Task 1 (1+0+0)	1.32237	**.126***
	Task 2 (1+1+0)	.47222	**.950***
Task 2 (1+1+0)	**Task 1 (1+0+0)**	**.85015**	**.590***

* The significant level is <0.05.

The statistical differences existed in the following four pairs: Task 1 vs. Task 5 (sig=.000), Task 3 vs. Task 5 (sig=.006), Task 2 vs. Task 6 (sig=.000), and Task 4 vs. Task 6 (sig=.000). The superiorities of task 5 (1+0+2) to task 1(1+0+0), task 5 (1+0+2) to task 3(1+0+1), task 6 (1+1+2) to task 2 (1+1+0), and task 6 (1+1+2) to task 4 (1+1+1) would be contributed to the work of strong evaluation. In other words evaluation has positive influence on incidental vocabulary retention when it present in the strong level like making original sentences or writing a composition with target words, etc.. Another two pairs did not show the difference Task 1 (1+0+0) vs. Task 3 (1+0+1) (sig= .126>.05) and Task 2 (1+1+0) vs. Task 4 (1+1+1) (sig=.608>.05), which concluded that the moderate evaluation failed to facilitate the acquisition.

The data of two-week delayed vocabulary test reported that the vocabulary retention underwent statistically decrease in every group during the interval of two weeks between immediate vocabulary test and delayed one. Group 6 received the largest fall (Mean Difference=5.6) between the scores in the immediate test and delayed one. The decrease was smallest in group 1 (Mean Difference=2.3). ANOVA analysis declared that there was no statistically significant difference in six groups. After two weeks, the superiority of the tasks in the immediate vocabulary test did not survive in the delayed vocabulary test. And the Post Hoc Tests also verified there were not any differences between any two groups. Even group 6 with the largest involvement index 4 did not perform better in the incidental vocabulary retention than group 1 with lowest involvement (Sig=.507>.005).

Table 4. One-Way ANOVA of delayed vocabulary test

	Sum of Squares	Mean Square	F	Sig
Between Groups	19.719	3.944	1.451	.207*
Within Groups	581.808	2.719		

* The significant level is <0.05.

5. Discussion

5.1 Word Frequency

The data in the immediate vocabulary test indicates that those vocabulary exercises do push the learners to locate more consciousness on the target words and facilitate the vocabulary acquisition. Two weeks later, the participants achieved much lower score in the delayed test. A more important conclusion is that no matter how effectively they facilitate the vocabulary acquisition in the immediate test, six tasks fail to present difference on incidental vocabulary retention in the delayed vocabulary test. The task 6 with highest involvement index 4 does not bring out more vocabulary retention than task 1 in the delayed vocabulary test. The phenomenon seems to suggest that there could be some hidden influences on the vocabulary retention except the involvement load.

One reason for the low acquisition rate in the delayed vocabulary test is that all ten target words appeared only once. And the employment of pseudo-word ensures there is no chance for the learners to meet the words during

two weeks' interval. Therefore no repetition is found in the present research. Large numbers of studies have indicated repeated encounters are necessary to incidental vocabulary acquisition because itA is a gradual process in which gains are made in small increments with repeated encounters. The learning items would leave strong durability and be easy to accessible in the future only by conscious repetition or more elaborate cognitive processing.

Web (2007) declared that the knowledge of word meaning developed more slowly than other properties like syntax and orthography. A valuable finding is suggested that seven is the threshold for the productive knowledge. The results of Chen & Truscott (2010) not only support Web's perception but further develop it into a more comprehensive one. Productive knowledge was found to be more beneficial from repetition than receptive knowledge. But this relation was reversed by the time of the posttest of two weeks later due to the greater "vulnerability of productive skills" (710). A contrast was found in the pattern of development for different vocabulary knowledge. Orthographic feature was sensible to three exposures but not at all from four. Syntax displayed a more steady increase with the repetition. As to semantic property the gains developed during three and seven exposures.

Zeeland & Schmitt (2013) tested the increase of word knowledge by listening. They announced both immediate and delayed vocabulary test justified the assistance of higher occurrences toward the acquisition of form and grammar knowledge. "It appears that immediate, short-term knowledge of form and grammar starts developing with relatively few exposures, yet it needs to be hear considerably more than fifteen times for this knowledge to fully develop and be retained" (621). And even more than fifteen occurrences guarantee the knowledge of meaning.

5.2 Inference of Word Meaning

It is detected that there was no statistically significant difference in such pairs Task 1 and Task 2, Task 3 and Task 4, and Task 5 and Task 6. It revealed that the component search did not work on incidental vocabulary acquisition. In the present research search was operationalized as the attempt to find an appropriate meaning of an unknown L2 word by dictionary work. Search is expected to attract the learner's attention to the target words and enhance the memory trace in the brain. However the non-correlations between the inference and acquisition of word meaning are detected in other studies (Hulstijn et al., 1996; Ellis & Le, 2016). Hulstijn et al. (1996) refused the claim that "meaning inferred" yields higher retention than "meaning given" and explained the inconsistency that "measures were taken to guarantee that the Meaning Inferred conditions did indeed attempt to infer meanings of the unfamiliar words" (335). Ellis and Le (2016) reported that the "contextual clues that were embedded in the texts were designed to help the learners work out the meanings of the target words. However these clues were implicit in nature" (148). Their study denied the benefit of meaning-inference training. The second language learners feel difficult to infer the meanings of new words. They prefer to lexicalize the foreign language words in their first language or native language. "Even the increasing number of exposures up to seven makes little contribution to the acquisition of meaning for NL words" (Chen & Truscott, p. 711).

In the present research, the non-search groups (group 1, 3 and 5) were completing the tasks with glosses of new words, including the part of speech, English paraphrase, Chinese meaning, and one illustrative sentence or phrase. While the search groups (group 2, 4 and 6) were provided with complete dictionary works for the polysemy words. It is justified that effect of direct and definite explanation outperformed the multiple-choice explanation on vocabulary learning.

The interview after the test affords the explanation. The students in search group did not infer the meaning intentionally during reading process. Language learners did not feel an urge to interrupt the flow of reading by investing considerable time and mental effort to infer the meaning of unknown words. The quality analysis indicates that even if some student does devote time to analyze the context and infer the word meaning, the big possibility of wrong inference does exist. If the inference is wrong during the reading, it is no way to expect the learners to come out with right response towards the target words in the immediate and delayed vocabulary test. The researcher cannot easily analyze the reasons of the student's wrong choice in the test as the false reasoning or as the random and careless selection due to the decay of memory trace or ignorance. The first explanation actually proves the search works on the incidental vocabulary acquisition and retention, the cognitive processing leaving deep traces in the brain and the traces surviving during the two weeks' interval.

A close case study of word *bacilent* explains this wrong inference. It indicated that nearly 47% participants failed to give the right explanation in the posttest. This word occurred in the context: "Meeting different needs, friends range from those as soul sisters mentioned above to that of the most *bacilent* and ordinary playmates". This sentence expresses the theme of the passage that friends are various ranging from the intimate one to those who

we just say hello to those we even do not know her name. The learners in the group 2, 4 and 6 (all search group) were provided with such explanations:

bacilent (adj.)

cold, not hot 凉的，不热的

eg. The coffee is not bacilent enough to drink.

Calm, unexcited 冷静的，镇定的

eg. She always remains bacilent and calm in a crisis.

Not showing friendliness or interest 冷淡的，冷漠的

eg. She was bacilent about the proposal.

Without fully understanding the theme of the passage, the learners failed to deeply appreciate the difference of these three explanations, and circled the wrong explanation like calm, unexcited (冷静的，镇定的) or even cold, not hot (凉的，不热的) rather than the right explanation (not showing friendliness or interest 冷漠，冷淡). Then in the vocabulary test even if they would recall the meaning of *bacilent*, the respondents would not score. They responded with the wrong Chinese equivalent they inferred.

5.3 Limited Processing Capability

The present study suggested that evaluation did not work in moderate level and it influenced incidental vocabulary acquisition only in strong level. According to the Involvement Load Hypothesis, evaluation means a comparison of given word with other words, or a contrast of a specific meaning of a word with the other meanings, or a combination of the word with other words in order to assess whether a word does or does not fit its context. If the task demands the learners to complete the sentences with the target words like task 3 and task 4, it induces moderate evaluation. If the task requires retrieval and production of target words in an original context like task 5 and task 6 it motivates strong evaluation. The failure of evaluation at moderate level could be attributed to the cognitive processing capability. Evaluation works on incidental vocabulary acquisition with the condition that the tasks actually invoke the learners' language knowledge and meta-cognition strategy. Task 3 provided the learners with incomplete sentences and required the learners to supply the missing word with a mixture of target words and interferers. The learners in the interview after the treatment recalled that some sentences were too difficult that they were unable to infer what the semantic loss was, like the sentence "The kids were always ____ awkward questions" or "Even after the divorce they shared the ____ of the children". And the newly acquired target words aggravated learning burden. Although the meaning of each word was printed on the sheet of explanation, those words were still strange to them. The word knowledge stopped in the phase of form recognition, a superficial receptive knowledge. The syntax, semantics and free automaticity of receptive knowledge to productive knowledge have not possessed by the learners.

It was the same case with task 4, a cloze exercise, which was even more demanding than task 3. It is a common but challenging exercise to the students. Cloze calls for the overall understanding of the passage even with the loss of information. At the same time cloze is based on the detailed appreciation of the specific context. Only two processing methods bottom-up and top-down work together effectively can the learners pick out an appropriate word for each blank. The absence of any processing would lead to the wrong choice. And the distractors in task 4 also enhance the obstacle to the evaluation. The students in task 3 and 4 complained the tedious exercise tasks and they felt deep depression and anxiety to complete the all the tasks in the given time.

In contrast group 5 and 6 with strong evaluation made original sentences with the given words. The requirement directly leads the learners' attentions to the ten target words. Although the task intended to push the leaners into deepest processing, the quality analysis demonstrated that most learners mechanically duplicated the sentences of the reading material or samples in the glosses. They just made subtle change for part of the sentence. Little processing burden eases the pressure brought about by the test. And during the copy activity, the learners review the word knowledge like grammatical functions, collocations and spelling. The deeper elaboration leaves stronger traces in the mind for the initial learning of the word knowledge.

5.4 Vocabulary Knowledge Test

The present research employs the productive knowledge of meaning as the evidence proving the acquisition of vocabulary acquisition. The learners were demanded to write down the Chinese equivalence for a list of twenty words, a mixture of ten target words and other ten interferers. There is no hint about the interrelation of the vocabulary test and the reading material or vocabulary tasks the learners have completed.

Such vocabulary tests call for the productive word knowledge, which is high level in the framework of word. Word knowledge is a multi-dimensional notion like what Nation proposed in 2001: form (spoken form, written form, and word parts); meaning (concepts, reference, and associations); use (grammatical functions, collocations, and constraints on use, such as register and frequency). And in each domination there is the further classification as receptive knowledge and productive knowledge. In an integrated framework advanced by Qian (2002), besides the depth and breadth of vocabulary knowledge, an important aspect is the automaticity of receptive-productive knowledge, which refers to all fundamental processes through which access to word knowledge is achieved for both receptive and productive purposes. The automaticity of lexical knowledge would include the following processes: encoding and decoding phonological and orthographic features, accessing to structural and semantic features, integrating lexical-semantic representation. The simplified vocabulary test in the present research is involved with recognition of form and recall of meaning, which is not sufficiently sensitive to reveal the incremental process of word knowledge. Zeeland and Schmitt (2013) found that 'learner start developing knowledge of a word (form and grammar recognition) long before they master the form-meaning link' (609). The recognitions of form and grammar knowledge are relatively easily than the construction of the form-meaning link. The participants demonstrate receptive knowledge of form in 45.8%, of grammar in 33.7%, while only of 8.5% in meaning.

The retrospect of vocabulary posttest manifests that the learners felt extremely distress to recall the meaning without any context. The target words were examined out of context in the immediate test and the delayed test. Rott (2004) criticized the common approach of vocabulary test---present the target words in L2 and require the learner to choose or write the correct word meaning in L1 as being not sensitive to record partial word gain and provide little insight into word storage process. Rott proposed three levels test with different retrieval cues to explore clearly which stage the word knowledge stays. The scholars are suggested to devise more broad and diversified vocabulary tests to discover the multi dimensions of the word knowledge and mobile increase of the receptive/productive knowledge (Ellis& Le, 2016; Ender, 2014; Zeeland & Schmitt, 2013).

6. Conclusion

The present research aims to corroborate the Involvement Load Hypothesis on Chinese adult English learners and report the relationship between incidental vocabulary acquisition and task induced reading. The results prove that Involvement Load Hypothesis does detect the active relationship between reading and incidental vocabulary acquisition. However it is also concluded that the Involvement Load Hypothesis is designed to be an idealized conceptualization of cognitive processing on vocabulary learning. The immediate vocabulary test indicates that the learners acquire seventy percent of target words, but in the delayed test it is found that the learners forget most of newly acquired words after two weeks' interval. If no repetition is made during time interval, the memory trace of those words would decay. No matter how high the involvement load is induced by a task previously, the incidental vocabulary couldn't be retrieved without the reoccurrence. "Isolated text and even a book or two do not reflect the value of reading for vocabulary growth. Only a flood of reading will ensure repeated exposures to words, reveal the linguistics and pragmatic properties of the new words, and reinforce the learner's memory of already familiar words" (Laufer & Roitblat, 2011).

The failure of search in the present study suggests that the readers seldom stop reading for the new words and most learners choose to ignore them. The learners especially with intermediate or low language proficiency could not succeed in inferring meaning accurately and exactly. The ambiguous or even wrong inference impedes the overall understanding of the reading material and hinders the direct and explicit connection of form and meaning either in L1 or L2. The clear and direct glosses are more advisable for foreign language learners to achieve the aim of incidentally acquiring vocabulary by reading activity.

Another noticeable point is that the abilities for the learners to process the information broadly and deeply is not increasing unlimitedly. The more demanding vocabulary task like cloze with distractors has surpassed the range of the learners' cognitive ability. They would suffer from the psychological anxiety and motivational deficiency due to the difficulty of tasks. The supposed benefits from the task turn into the hurdles for the learners from focusing on the reading. In practice the teachers are expected to conduct tasks with the balance of the processing ability and the task complexity in order to maximize the effect of reading on vocabulary acquisition. An additional amendment is advocated. To have a close look at dynamic process and other dimensions of the word knowledge, the researchers are suggested to design more vocabulary tests rather besides form-meaning association in the future researches.

References

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, *31*(5), 693-713.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology*, *104*(3), 268-294.

Ellis, R., & Le, C. (2016). The effects of inference-training and text repetition on Chinese learners' incidental vocabulary acquisition while listening. *Chinese Journal of Applied Linguistics, 39*(2), 137-152. https://doi.org/10.1515/cjal-2016-0009

Ender, A. (2014). Implicit and explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics*, *37*(4), 536-560. https://doi.org/10.1093/applin/amu051

Folse, S. K. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, *40*(2), 273-293. https://doi.org/10.2307/40264523

Gass, S. (1999). Incidental Vocabulary Learning. *Studies in Second Language Acquisition*, *21*, 319-333. https://doi.org/10.1093/applin/amq031

Hu, H. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement Load Hypothesis versus technique feature analysis. *System*, *56*, 28-39. https://doi.org/10.1016/j.system.2015.11.001

Hulstijin, J., Hollander, M., & Geridanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, *80*(3), 327-339. https://doi.org/10.1111/j.1540-4781.1996.tb01614.x

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

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, *58*(2), 285-325. https://doi.org/10.1111/j.1467-9922.2011.00644.x

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1-26. https://doi.org/10.1093/applin/22.1.1

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research*, *15*(4), 391-411. https://doi.org/10.1177/1362168811412019

Lee, S., & Pulido, D. (2017). The impact of topic interest, L2 proficiency, and gender on EFL incidental vocabulary acquisition through reading. *Language Teaching Research, 21*(1), 118-135. https://doi.org/10.1177/1362168816637381

Liu, C. (2013). The effect of different translation tasks on incidental vocabulary acquisition. *Applied Linguistics*, *36*(3), 326-337. https://doi.org/10.1515/cjal-2013-0022

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology*, *44*(1), 87-112. https://doi.org/10.1037/h0084237

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*(3), 513-536. https://doi.org/10.1111/1467-9922.00193

Rott, S. (2004). A comparison of output interventions and un-enhanced reading conditions on vocabulary acquisition and text comprehension. *The Canadian Modern Language Review*, *61*(2), 169-202. https://doi.org/10.3138/cmlr.61.2.169

Wang, J.-H. T. (2015). Task-induced involvement load: the accumulation effects on vocabulary acquisition. *Chinese Journal of Applied Linguistics, 38*(2), 150-165. https://doi.org/10.1515/cjal-2015-0009

Webb, S. (2007). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language Teaching Research*, *11*(1), 63-81. https://doi.org/10.1177/1362168806072463

Wesche, M. B., & Paribakht, T. S. (2000). Reading-based exercises in second language vocabulary learning: an introspective study. *The Modern Language Journal*, *84*(2), 196-213. https://doi.org/10.1111/0026-7902.00062.

Zeeland, V., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 Listening: a dimensions approach. *System*, *41*(3), 609-624. https://doi.org/10.1016/j.system.2013.07.012

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