THE INVOLVEMENT LOAD HYPOTHESIS: AN INQUIRY INTO VOCABULARY LEARNING

Mayumi Tsubaki

要　旨

2001年に Laufer & Hulstijn が the involvement load hypothesis を提案した。これは，動機付けや認知心理の要素も含んだ語彙習得に関する仮説であり，注目に値する。本論文では，この仮説を心理学の立場や授業中の学習活動において説明及び分析し，それに関する研究について述べた。

Introduction

Vocabulary development is very important, regardless of the purpose of the language learning.　In fact, we need to use our vocabulary knowledge whenever we communicate.　Recent research has revealed that educated native speakers have a vocabulary size of around 20,000 headwords for word families (Goulden, Nation, & Read, 1990).　The researchers counted word families, which include headwords, its inflected forms (e.g., -*ly*, -*ness*, and *un*-) and its derived forms.　Therefore, the actual number of words known seems to be even larger than 20,000.　In fact, native speakers increase their vocabulary an average of 1,000 words a year; they know about 20,000 words by the time they are 20 years old.　Non-native speakers of English do not need to have as large a vocabulary as native speakers do, and it is most important for non-native speakers to learn the most useful words (Nation, 2001).　Still, non-native speakers of English have to learn a great many words, and it is a nearly overwhelming task.　Therefore, language learners and language teaching professionals would like to find ways to increase vocabulary knowledge efficiently.　In 2001, Laufer & Hulstijn (2001) proposed the involvement load hypothesis, and it seems to be beneficial for language learning.　This paper is written to investigate the involvement load hypothesis.　First I will define the hypothesis from the viewpoint of psychology, and then from the viewpoint of Teaching English to Speakers of Other Language (TESOL), especially in Japan.　Afterwards, examples of language learning activities in Japanese university classrooms will be given.　Finally, the research exploring

the hypothesis will be described and criticized.

What is the Involvement Load Hypothesis?
Depth of Processing Model: the Basis of the Involvement Load Hypothesis

The involvement load hypothesis has developed from the depth of processing model. This model was first proposed by Craik & Lockhart in 1972. They suggested that retention in long term memory depends on how deep information is processed during learning. To put it simply, the information that is processed at a deep level stays in memory longer than that which goes through a shallower processing. To take a simple example, a language learner pays attention to whether a target word is written in lower case or capital letters. This visual information is probably processed at such a superficial level that it will not help the word to stay in memory. In order to go beyond that level and make a more robust memory trace, the sound of the word could be added to the processing. Furthermore, when the meaning of words is added, the information is processed at an even deeper level, and, as a result, the information is even more likely to be remembered than in the former cases. One problem with Craik & Lockhart's model is that their perspectives on depth of processing is oversimplified. The memory of words is influenced by other factors in addition to orthographical, acoustic, and semantic information, and the three factors should be classified into more than three levels of memory trace (Baddeley, 1999).

Craik & Tulving (1975) expanded the theory by Craik & Lockhart (1972), asserting that the notion of elaboration should be added. When new information is connected to information that already exists, it is enriched and makes more robust memory traces. To put it another way, the more new information is related to preexisting information, the more likely it will be remembered. The main criticism of Craik & Tulving's proposal is that it has an unclear definition of depth of processing, which can also be called "cognitive effort", "degree of elaboration", and "encoding specificity." It is also hard to operationalize the concept (Zechmeister and Nyberg, 1982, cited in Laufer & Hulstijn, 2001). Nonetheless, cognitive psychologists generally agree that processing activities influence memory performance more than the learners' motivation.

Baddely (1999) argues for "richness" or "breadth" of processing as an alternative concept to that of "depth." The concept of "richness" or "breadth" means that what is coded widely with a variety of information will be better remembered. As I understand it, this concept is similar to elaboration, even though the connection between new information and preexisting information is not mentioned. I believe that the new

information is connected with the information that already existed when it is encoded.   He writes, "In general, information that is encoded in terms of a rich and detailed representation of the world is likely to be more accessible than material that is processed in terms of a simpler or more impoverished scheme" (p. 177).   The enrichment and elaboration of the encoding processes lead to robust memory traces.   Let us think of classroom situations.   I believe a word that language learners encounter for the first time will be retained when they are given or they come up with a variety of properties for the word, such as general meaning, connotation, denotation, association and usages; more than it will be when they only learn its meaning.   This concept of "richness" or "breadth" is also called "elaboration." In this paper, "richness," "breadth" and "elaboration" are used interchangeably.

A Model of Task-Induced Involvement: The Involvement Load Hypothesis

Laufer and Hulstijn (2001) developed the above theoretical notions and constructs, by integrating them and adapting them as needed, and then proposed a construct they call "task-induced involvement." Not only is their proposal grounded on the depth of processing models and elaboration, but it also includes both cognitive and motivational factors.   They refer to their idea as the "involvement load hypothesis," and it posits that the amount of involvement in the task that learners are engaged in will affect the retention of unfamiliar vocabulary.

What is significant about the involvement load hypothesis is that it is comprehensive as it includes motivational factors as well as cognitive factors.   It is evident that motivation plays an important role, and it can be as influential as cognitive factors.   It should also be noted that the hypothesis attempts to draw attention only to vocabulary learning per se in a foreign or second language.   Yet, an additional strength is that this hypothesis could possibly extend to other areas of language learning, such as grammar.

Three Components of Vocabulary Learning

Laufer and Hulstijn (2001) and Hulstijn and Laufer (2001) listed three components of task-induced involvement: need, search, and evaluation.   Need is a motivational construct while search and evaluation come from the cognitive dimension.   Need is the motivation to learn target words.   Search occurs when the learner has to find the meanings of target words or the word form for words indicated by target concepts.   Evaluation involves comparison of a target word with other words.

Not only did the two researchers come up with the three components, but also they

suggested three degrees of value for each component (none, moderate, and strong). Each level is clearly described in Table 1. The depth of processing model stimulated the field of Psychology, but it was vague and needed concrete factors, as supplied by the involvement load hypothesis. Researchers and teachers can employ the three components and factors in their research or teaching situations more easily than they could with the depth of processing model.

Table 1   *The Degrees of the Components in the Involvement Load Hypothesis*

Components	Degrees of the Involvement Load	Explanations
Need	Index 0 (None)	The learner does not feel the need to learn the word.
	Index 1 (Moderate)	The learner is required to learn the word.
	Index 2 (Strong)	The learner decides to learn the word.
Search	Index 0 (None)	They do not need to learn the meanings or forms of the word.
	Index 1 (Moderate)	The meaning of the word is found.
	Index 2 (Strong)	The form of the word is found.
Evaluation	Index 0 (None)	The word is not compared with other words.
	Index 1 (Moderate)	The word is compared with other words in the provided context.
	Index 2 (Strong)	The word is compared with other words in self-provided context.

Let us explore the details of the three degrees, considering classroom situations in Japanese universities. How much the learner feels the need to learn the target word determines which of three indexes of the need exist. When the learners have to learn words in order to pass their English classes or to be engaged in learning tasks in their English class, they are forced to learn words. In such a case, the need index is moderate. If the learner decides on their own to learn a word, the need is strong. In the classroom, it is not very common to have a situation that will lead to a strong need in the English as a Foreign Language (EFL) situation in Japan because English is not used outside of class, and in general learners do not take the initiative to choose which words to learn for communication.

The most common value of the search index in Japan is obviously moderate since they usually, have to find the meanings of words to understand written texts. The learners do not have many opportunities to speak or write English, and they read more often than they listen for vocabulary learning. In fact, they usually have to deal with written text in order to comprehend it. Although it is true that they are exposed to spoken English while listening to English songs and watching movies, but even then they often rely on written input, such as lyrics sheets and subtitles, to catch the orthographical information they might miss when listening. Teaching productive skills is a must to find L2 forms of un familiar words and bring about the strong search index. Luckily, in most EFL environments, the learners have a common first language, and the concept, or a translation, or explanation of the target words can be given easily so that it will be easy for them to write or speak about even complicated topics.

As for the evaluation index, the most moderate index is assumed to be the most common. That is because passages and materials are provided by the teachers in the classroom. Typical examples are "fill in the blanks" and matching, and learners have to choose words from a list that their teachers give them. In order to achieve the strong evaluation index, teachers should have students write essays or make presentations.

It appears to me that to attain a high involvement load that consists of the three components teachers should train the learners to be more independent so that they will choose words that they need to learn. Another thing that teachers can do to have higher involvement load is to have learners use their productive skills. That way, they have to find word forms for the ideas that they need to use and they will have to compare the target words while making sentences using the words. As a result, both the search index and the evaluation index will be strong. The degrees of indexes for the three components of learning activities are presented in Table 2.

Research Related to the Involvement Load Hypothesis

Laufer and Hulstijn (2001) have argued for the involvement load hypothesis by using research results from the investigations of other theories and reinterpreting them. I am certain that a great number of studies support the hypothesis as explained in Laufer & Hulstijn (2001) and Hulstijn and Laufer (2001). However, to my knowledge, only two pieces of research have pursued examination of the hypothesis directly.

The first study that affirmed the importance of the hypothesis was conducted by Hulstijn & Laufer (2001). Participants were advanced English learners enrolled in universities in Israel and the Netherlands and they had each experimental group do the same tasks. The

Table 2 *Task-induced Involvement Load of Common Language Learning Activities in Japanese Universities*

Task	Status of target words	Need Index	Search Index	Evaluation Index
Listening to a song and trying to understand the song.	Filling in the blanks with target words. Looking up the meaning of target words.	1	1	1
Finding words which mean designated ideas by the teacher.	Looking up the word forms of the target words.	1	2	0
Matching target words with the explanation of words. Target words and their example sentences are presented.	Guessing the meanings by reading example sentences.	1	1	1
Retelling a story that they have read before.	Looking up necessary words in a dictionary. Target words are decided by the teacher.	1	1 or 2	2
Talking about movies that they watched before.	Useful words are presented, and they choose words that they need.	1	0	2
Choosing a relatively difficult topic and writing an essay.	Choosing words that they use.	2	2	2

Note: Index 2: Strong   Index 1: Moderate   Index 0: None

participants were assigned to one of the three vocabulary learning conditions to learn 10 target words that were not familiar to them before the experiment.   The first group read a letter to the editor in order to answer comprehension questions.   The target words on the letter were marked in bold and glossed.   Among the three components, need was the only one that existed in the condition since the learners were required to do the task as a class activity, and they had no search or evaluation index condition.   The participants in the first group did not look for the meaning or the form of the target words because the meaning of the target words was glossed.   They did not need to evaluate them against other words.

The second group received the same letter and the same comprehension questions as the

first group.　However, unlike the first group, the 10 target words were eliminated from the reading material and left blank, and translations and explanations of the target words were provided.　The subjects filled in the blanks using the ten target words and five additional words.　Because the learners were required to do the task, the involvement load for the need was moderate (need index 1).　Search did not occur because the students did not have to look for the meanings of the target words.　When they filled in the blanks with 15 choices of words in the context decided by the researchers, they experienced the evaluation index 1.

The third experimental conditions had the highest involvement load of the three.

Table 3　*Task-induced Involvement Load Index of the Three Conditions in the Study*

	Learning activities and time on task	Need	Search	Evaluation	Involvement Index
Task 1	Reading comprehension with marginal glosses Time on task: 40–50 minutes	1 Target words given	0 The meanings of the target words were provided	0 Words not evaluated at all	1
Task 2	Reading comprehension plus "fill in" (fill in 10 gaps with the missing words from the list of 15 words) Time on task: 50–55 minutes	1 Target words given	0 The meanings of the target words were provided	1 Words evaluated in the context that the researcher provided	2
Task 3	Writing a composition using the target words Time on task: 70–80 minutes	1 Target words given	0 The meanings of the target words were provided	2 Words specified to be used in the subjects' original sentences	3

Note: Total Involvement Index was the result of the sums of the three indexes
　　　for each component.

Students wrote compositions with 10 target words in the form of a letter, using the translation and explanation, and using the target words.   The degree of the evaluation index was strong and higher than the other two in that the subjects used the words in their own context.   The three conditions are compared in Table 3.

The third group outperformed the other two.   Based on the result, it is reasonable to assume that this is evidence that can partially support the involvement load hypothesis. Nonetheless, this position seems questionable for several reasons.   For one, the three experimental conditions are different only in evaluation indexes.   In other words, it is safe to say that evaluation influenced the retention of words, but it does not apply to the need nor to the search index.   Related to that, the third condition had the advantage of more information and more time.   The third group students received more information on the target words, and they worked on the task longer than the other two groups.   Naturally, the more information and time the learners have, the more they learn.   The last criticism is that this research assessed vocabulary learning in a translation test, which is not very sensitive to partial learning of vocabulary.   In order to measure partial learning, multiple choice tests or/and recognition tests should have been used (Nagy, Herman, & Anderson, 1985; Waring & Takaki, 2003).

The other study that paid attention to the involvement load hypothesis is a qualitative study by Rott (2005).   Her research is intended as an investigation of how words are processed to strengthen lexical form-meaning connections of four target words during a reading exercise.   Her subjects, native English speakers who learn German as a foreign language, were randomly assigned to the single translation gloss condition or the multiple-choice gloss conditions.   The single-translation gloss condition group saw a reading text glossed with the first language translation when the word was used in the text first.   By contrast, in the multiple-choice gloss condition, they had four choices in the gloss: the correct definitions of the words, two distracters which suited the context, and a "don't know" option.   Here is an example that Rott mentioned.   The target word *kaff* at the first occurrence came up with the four choices: a) shack (distracter), b) valley (distracter), c) village (correct answer), and d) don't know.

The researcher attempted to reveal need, search, and evaluation from think aloud protocols.   She defined the three components for the study as follows.

Need   having a perceived need to find the meaning of the target word
Search   finding meanings of the target words
Evaluation   judging whether the meaning choice of the multiple choice glosses is correct
The subjects reacted to each condition differently.   The multiple-choice gloss facilitated

more need, search, and evaluation than the single translation gloss, and the former also appeared to bring about stronger form-meaning connection and also more robust memory trace than the latter.

The strength of this research explored motivational and cognitive phenomena during the learning activities. When only a quantitative study is conducted, we can only assume what happened in the learners' brain. An obvious shortcoming is that the number of the subjects and target words were small because qualitative research cannot deal with a large number of students. It would be informative if quantitative studies could be done so that the result can be generalized to other groups of language learners.

Conclusion

The involvement load hypothesis should be known to language teaching researchers and professionals far more than it is now, since it covers a variety of factors, and we need to have solid vocabulary learning theories to use in classroom situations. Surprisingly, few studies have so far been attempted to reveal the value of the hypothesis. As a matter of fact, only the two studies described in this paper, which have clear limitations, have contributed to the field as far as I know. Not only that, language teachers must look more carefully into the hypothesis and test it in the classroom. Researchers and teachers might be able to further develop the hypothesis. Indeed, this hypothesis needs to be investigated in studies and classrooms so as to lessen the burden of vocabulary learning.

References

Baddeley, A. D. (1999). *Essentials of human memory*. Hove, UK: Psychology Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Depth of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

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

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11*, 341-363.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*, 539-558.

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

Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

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

Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections

are established and strengthened. *Reading in a Foreign Language, 17*, 95-124.

Waring, R., & Takaki, M.（2003）. At what rate do learners learn and retain new vocabulary? *Reading in a Foreign Language*, *15*, 130-163.