

A Pilot Study Comparing Vocabulary Retention in Bilingual and Monolingual Tasks

Andrew OBERMEIER

2 言語使用と 1 言語使用のタスクによる語彙習得の比較

オーバマイヤー・アンドリュー

Accepted July 2, 2008

Abstract : This pilot study compared the retention effects of two vocabulary learning tasks in an English as a Foreign Language (EFL) classroom: a monolingual English task and a bilingual task translating English to Japanese. According to this partial trial of the Involvement Hypothesis (Laufer & Hulstijn, 2001), the English task would result in better retention because deriving the meaning from an English definition involves a greater level of *evaluation*. In both of the treatment tasks of a pretest-posttest design, the participants matched English words to their definitions. The only difference was the language of the definition in the matching task, Japanese or English. The results were contrary to the hypothesis and the *t*-test was non-significant: $t(16) = 0.48, p = .6$. A Rasch Analysis of the pretest and posttest items was informative, however, showing that some of the words were more memorable merely due to distinctive features of the words themselves. Thus, future studies on the issue of monolingual and bilingual task involvement were meaningfully informed by this pilot study.

Key Words : EFL Vocabulary Teaching, Involvement Hypothesis, Rasch Analysis

抄録 : 本試験的研究は、英語教育における 2 言語使用と 1 言語使用のタスクの英語学習の語彙習得に対する効果を比較した。Laufer & Hulstijn (2001) は単語の意味や使い方に付いて判断することが習得につながるもの Involvement Hypothesis を示した。そこで、本試験的研究では、外国語である英語で行うタスクが母語の日本語で行うタスクより語彙習得に効果的であると仮説を設定した。対象者は、日本語と英語で難易度の高い英語の単語を単純の英語の意味説明と対応させる作業を行った。結果は、仮説に反して、両群におけるポストテスト間に有意差は見られなかった ($t(16) = 0.48, p = .6$)。これは、作業で取り上げた単語の特徴による印象が影響していると推測される。本研究による結果から、作業において取り上げる単語に配慮する必要性が示唆され、この点を考慮した研究が必要であると考えられた。

索引語 : 英語教育の語彙指導 ; 関わり仮説 ; Rasch 分析

Introduction

The belief that foreign language learners can effectively develop their vocabulary knowledge through abundant exposure, and without word focussed tasks, has had a strong influence on the foreign language teaching profession recently. Input-centered methods like extensive reading are backed by the faith that exposing learners to many words many times will result in better vocabulary learning. The rationale behind this seems to draw evidence from first language acquisition: children do not intently set out to increase their vocabulary, but they simply absorb new words by constantly encountering them in their daily lives. However, the differences between first and foreign language acquisition are abundant. For children learning a first language, word learning is highly connected to their cognitive development and survival. The first language is acquired much more rapidly than a foreign language, and with a snowballing effect: as it develops into a highly sophisticated network of meanings and understandings, learned words can quickly associate with abundant, relevant, first language knowledge. When a person learns a foreign language, the learner's intention plays a central role. As they gain competence and build their vocabulary, frequency becomes a less reliable teacher, and language learning must be intentional and systematic.

Word learning is incremental through repeated exposures, and in foreign language learning these increments are particularly slow: it has been estimated that it takes 8 to 10 exposures to learn a new word. In an analysis of a corpus from a selection of extensive readers, Nation & Wang (1999) showed that after the most frequent words have been learned, subsequent progress becomes much more difficult. Since new words appeared more sporadically, only 108 new words were repeated the necessary 8 to 10 times throughout a span of 200,000 words (9 books). Since so many exposures to new words are required, waiting for them to appear naturally in foreign language learners' reading is unrealistic. For learners seeking to develop their vocabulary beyond the level of the most frequent words, learning from natural contexts becomes impossible. Cobb's (2007) corpus analysis of extensive readers confirmed this. The contents of extensive reading materials do not have sufficient lexical coverage to take learners beyond the bounds of the most frequent 2000 words. The inefficiency of a "natural" input-based approach to learning points to a need for word-focussed activities that can support vocabulary learning in a systematic way.

This study sought to explore the effectiveness of word focussed tasks in foreign language vocabulary teaching. Specifically, it focussed on the difference between tasks that used the first language and those that used the target language.

Literature Review

Laufer & Hadar (1997) examined the differences of monolingual, bilingual, and "bilingualised" dictionary use in the comprehension and production of new words by EFL learners. The study tested participants on the comprehension of 15 low-frequency words, 5 of which were given with definitions from a monolingual learner's dictionary, 5 with translations from a bilingual dictionary, and 5 with hybrid "bilingualised" definitions that included both the bilingual translations and the monolingual definitions. The participants were tested on their

comprehension of target words and on their ability to use these words in their own sentences. Participants were asked to use the definitions they had from the three different dictionary types to help them to do two tasks: a comprehension task, and a task using the target words to create an original sentence. The results of an ANOVA showed significant differences in word learning across all three of the dictionary types. This study also found that dictionary using skill varied widely, regardless of the participants' level of proficiency. The most popular method of dictionary use among learners is to simply get the translation and move on. "Bilingualised" dictionaries help the reader to fuller understandings of words by also offering monolingual definitions and example sentences.

Since combining target and native languages in definitions proved to be significantly more helpful than the translations or monolingual definitions alone, it appears that the interaction of native and target languages might be a powerful factor to help learning. Perhaps the focus of this present study, combining languages within tasks, is an effective strategy for teaching. Next I will outline a framework for understanding vocabulary learning tasks.

Building on an extensive review of research on foreign language grammar acquisition, Laufer & Hulstijn (2001) proposed a construct of involvement for analyzing vocabulary learning tasks in motivational and cognitive terms. They developed the construct as an attempt to stimulate researchers to operationalize traditional labels such as noticing, attention, elaboration, and motivation into components that can be specified by tasks. The framework assesses three aspects: 1) how much a task instills a *need* to learn a word, 2) the *search* the task requires in order for learners to find a word's meaning, and 3) the *evaluation* the task requires when learners use them to make sense of the messages they speak, write, read, or hear.

According to these criteria, a reading comprehension task in which words are glossed in the margin will produce a lower involvement load than one which requires learners to look words up in a dictionary because the first task does not require that learners search for unknown words. Similarly, a cloze exercise in which the words to be filled into the blanks are provided in an accompanying list requires a moderate level of evaluation, but a similar task without the words provided produced a higher level of evaluation.

Hulstijn & Laufer (2001) tested their proposed framework with a study in which they compared the learning effects of three tasks with different involvement loads: 1) reading comprehension with marginal glosses, 2) reading comprehension plus fill-in, and 3) writing a composition incorporating the target words. They predicted that Task 1 would have the lowest retention scores because it induced a moderate need but no search or evaluation, and that Task 2 would be more effective since the need remained the same but the fill-in element induced moderate levels of evaluation. Task 3 was predicted to be the most effective because it induced a strong evaluation aspect as learners had to evaluate ways to use the target words in creating an original text.

The articles reviewed above confirm that the more involved a learner becomes with the words they learn, the better the chances for retention. Increasing cognitive and motivational involvement benefits learning. Cognitive and motivational elements of tasks can be analyzed according to the need, search, and evaluation elements that are required in carrying them out. In using dictionaries, involving both the first language translation and target language definitions and example sentences enhances learning.

Integrating the native and target languages together into a task might be similarly effective. Contrasting monolingual and bilingual tasks in regards to their different levels of need, search, and evaluation might yield interesting results. The research question for this study is as follows: how do bilingual and monolingual tasks affect the retention of new words? The hypothesis is that since monolingual tasks require more evaluation, they will lead to better retention.

Procedure

Fifty-five students at a small Japanese university were the participants in this study. Most of these students were candidates for English teaching licenses, so their motivation was moderately high. Although there is not a uniform proficiency test given at this institution, the students that do take the TOEFL average scores of around 460 on the paper version, and 48 on the iBT. Most students were second and third year English majors taking 9 to 12 classes per week in English skills, education, literature, and linguistics.

The two tasks examined in this present study had similar involvement loads. In both the monolingual and bilingual matching tasks, participants had to evaluate whether a word matched a definition. The only difference was that in one task, the definition was a Japanese translation, and in the other, the definition was a simplified explanation in understandable English. There was a slight need as well, as the participants needed to match the words to the definitions to complete the exercise (see appendix). There was no search in either of the tasks. Although the tasks seem quite similar according to their levels of evaluation, need, and search, the great difference between them is that in one case, the evaluation is carried out in English and in the other case, in Japanese. Since more extensive evaluation must be carried out to complete the task using English, I predicted that the English task would lead to better learning.

A week before the experiment, the 55 participants took a pretest to see if they could write the translations for 20 target words: laconic, harbinger, pachyderm, kibitzer, idiosyncrasy, immanent, nadir, macaque, quixotic, jalopy, ebullient, fastidious, abominable, aberrant, serendipity, cacophony, bachic, aardvark, gamut, obituary. The pretest showed that the participants knew none of the words. The next week, each participant received a two-sided sheet of paper, on the front side of which there was a list of the target words alongside their translations in Japanese. On the back side was a "vocabulary learning exercise" in which the 20 target words were used in two contrasting matching exercises: the monolingual task matched the target words to English definitions, and the translation task matched them to Japanese translations (see Appendix 1).

Results

A quantitative analysis of the data revealed the Item Difficulties for each of the items, showing the probability that a learner will get an item correct. The mean of the English Treatment Item Difficulties ($m = 50.8$) was slightly higher than that of the Japanese Treatment ($m = 48.2$), meaning that the English Treatment words were more difficult to remember. The independent-samples t test was non-significant, $t(16) = 0.48$, $p = .6$. There appears to be a slight difference in favor of the learning achieved when learners have a definition in their native language. This finding is contrary to my hypothesis, and contrary to the finding of Hulstijn & Laufer (2001) in favor of the use of bilingual information when learning words.

Item Difficulty Measures

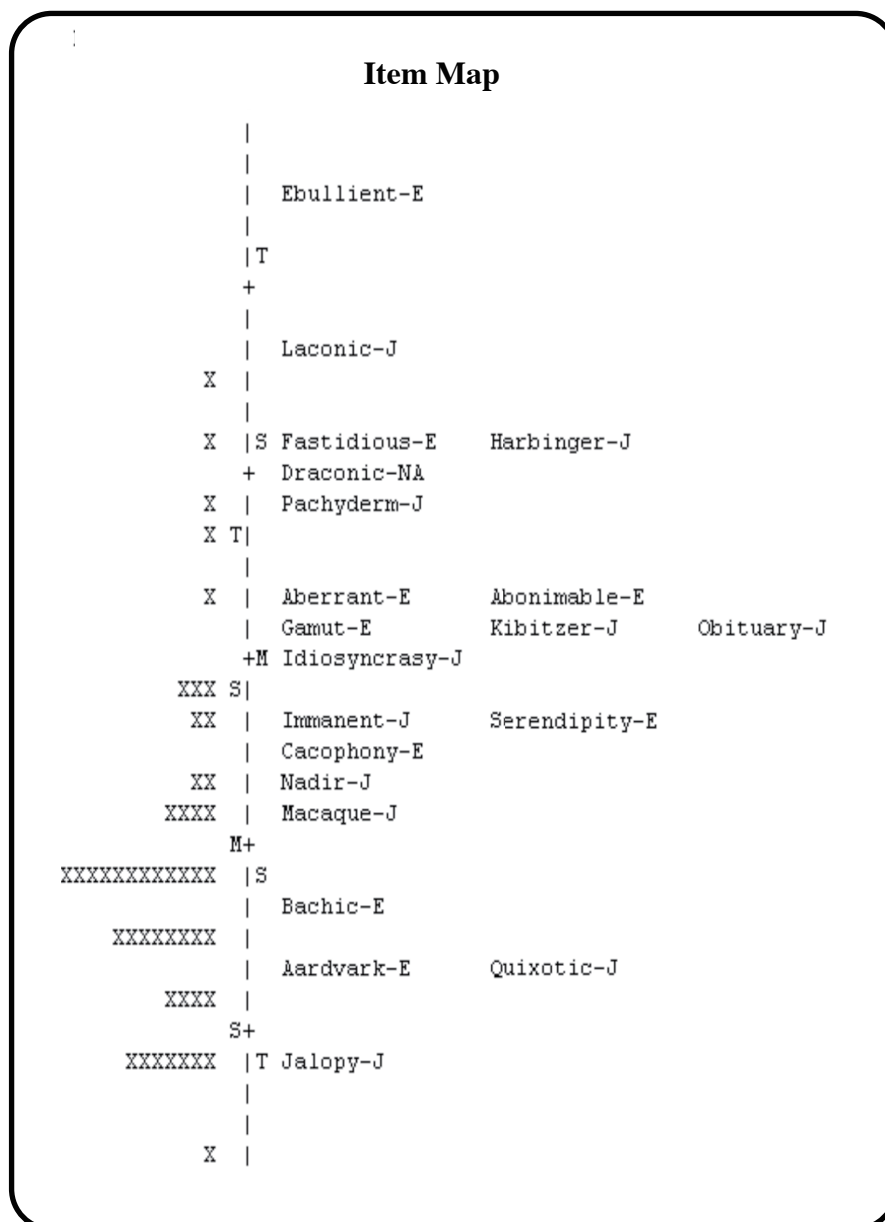
English Treatment		Japanese Treatment	
Ebullient	74.7	Laconic	66.7
Fastidious	61.6	Harbinger	61.6
Abonimable	53.1	Pachyderm	57.7
Aberrant	53.1	Kibitzer	51.7
Serendipity	46.9	Idiosyncrasy	49.2
Cacophony	44.8	Immanent	46.9
Bachic	36.9	Nadir	43.8
Aardvark	34	Macaque	41.7
Gamut	51.7	Quixotic	34.1
Obituary	51.7	Jalopy	28.7
Mean: 50.8		Mean: 48.2	
SD: 11.7		SD: 11.9	

Rasch software (Winsteps) was used to determine the Item Measures for each of the words, and the Person Abilities for each of the participants. A visual representation of the item difficulties and participant abilities are arranged in the Item Map below. The "X's" on the left represent participants, and the words on the right represent the items from the posttest. The easiest items are on the bottom of the map, and they correspond to the least able participants, who are represented by the "X's" adjacent to them. The hardest items are on the top of the map, and correspond to the people who performed best on the test. After each word in the chart there is a "-E" or a "-J", which symbolizes whether the treatment was the English or Japanese task. Therefore, one can derive from the map that "Ebullient" was the hardest word. In fact, since there are no "X's" at this level, it can be seen that nobody was able to remember the word "Ebullient" on the posttest. Since "-E" is attached to "Aardvark", we know that the treatment for it was the English task. Likewise, since "-J" is attached to "Laconic", we know that it is associated with the Japanese task. Since my hypothesis was that words which were learned with the English task would be remembered more readily than those learned with the Japanese task, I expected that there would be mostly "E's" toward the bottom of the map and mostly "J's" near the top. This would mean that the words learned with the English task were easier for participants to remember than those learned with Japanese. As can be seen from the Item Map, the words did not fall into the arrangement I expected. The items associated with the English and Japanese treatments seem to alternate with each other at random. Due to the failure of the pilot experiment to produce a meaningful result, the pretest/posttest was examined for potential flaws.

Pretest/Posttest Instrument Analysis

- (1) The bulk of the items fall in the difficulty range between "Macaque-J" and "Ebullient-E," but there are relatively few participants corresponding to these words. This means that only a few people were able to get

these items correct on the posttest. In other words, it did not matter whether the words were associated with the English task or the Japanese task, but the items were all simply too difficult. This implies that perhaps an easier form of item such as multiple choice would have enabled participants to demonstrate partial retention on posttest performance. In the next version of this study, it may help to have varying degrees of difficulty in item formats.



- (2) The Item Map also reveals that the items "Bachic-E", "Aardvark-E", "Quixotic-J" and "Jalopy-J" correspond to many people. This means that most of the participants were able to remember these words and answer the items correctly on the posttest. These are very vivid words for many different reasons. "Aardvark" is one of the uncommon English words that begin with "double-a." In fact, some of the participants asked me while reading through the definitions whether I had misspelled it or not. The Japanese translation for "Jalopy" was

also very distinct, being written in Japanese slang; perhaps this is why it was more memorable than the other words.

- (3) The strong images of an unusually spelled rare animal and a beat-up car were the best remembered items in this study; these words were probably memorable due to word features, and not the experimental treatment. "Bachic," describing drunken behavior, was also quite memorable, as was "Quixotic" but I cannot devise a conclusion for why these words were easy to remember. The three most difficult words, "Ebullient," "Laconic," and "Fastidious" are all adjectives, and all rather abstract.

Conclusion

Analyzing tasks in regard to cognitive and motivational dimensions can help teachers to understand the amount and nature of learners' involvement with the words they learn in performing them. The higher the involvement, and the more it integrates these factors, the higher the quality of word learning, and therefore the better the retention. This framework can also help in designing tasks by enhancing understanding of the different factors that affect word learning.

This study was conducted to investigate whether performing a monolingual matching element in the target language would enhance word retention. It sought to prove the hypothesis that adding monolingual definitions to a matching task would raise the level of involvement, and thus enhance retention. There was not a significant finding, but perspectives can be gained which will assist in a further revision of this research. First, a more thorough range of tasks should be implemented by participants in order to ensure better retention overall; the posttest difficulty was so overwhelming that the influence of the experimental variables was undetectable. Second, variance for word difficulty needs to be eliminated. This can be done by changing the study to a counterbalanced design so that all the words get both treatments. In sum, this research was an illuminating first step to understanding how to design and test tasks of vocabulary retention.

References

- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11(3), 38-63.
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539-558.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary learning in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Laufer, B., & L. Hadar. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualized" dictionaries in the comprehension and production of new words. *Modern Language Journal*, 81, 189-196.
- Nation, P., & Wang Ming-Tzu, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 355-379.

Appendix 1: Vocabulary Learning Research Exercise (Experimental Treatment)

In the column labeled "answer", write the letter of the definition that matches the keyword.

English Matching

Answer	Keywords	Definitions
	Gamut	A. markedly different from an accepted norm
	Aardvark	B. riotously drunken merrymaking, drunk and wild
	Fastidious	C. loud, confusing disagreeable sounds
	Cacophony	D. giving careful attention to detail; hard to please; excessively concerned with cleanliness
	Bacchic	E. An animal of the grasslands of Africa that feeds on termites.
	Abonimable	F. Good luck in making unexpected and fortunate discoveries.
	Ebullient	G. The entire range of a group
	Aberrant	H. exuberant, high-spirited (joyously unrestrained)
	Serendipity	I. atrocious, abominable, awful, dreadful, painful, terrible, unspeakable

Japanese Matching

	Pachyderm	A. 騎士気取りの; 空想的な, 非実際的な.
	Immanent	B. (略歴つき) 死亡記事; 死の; 死を記録した.
	Idiosyncrasy	C. アジア, アフリカ産のサル的一种
	Kibitzer	D. 天底 ((zenith (天頂) と正反対の点)); どん底
	Jalopy	E. 先ぶれ (する); 先駆者 (forerunner); 前兆 ((of)).
	Laconic	F. 厚皮動物 ((ゾウ・カバなど)); 鈍感な人
	Macaque	G. [話] ぼろ自動車
	Nadir	H. 特質, 特異性; (変った) 性癖, 奇行;
	Obituary	I. [米話] おせっかい屋.
	Harbinger	J. 簡潔な, 口数の少ない.
	Quixotic	K. 内在する; 【神学】(神が宇宙に) 遍在する.

Appendix 2: Vocabulary Learning posttest

Write the Japanese definitions of the words below.

- | | | |
|-----|--------------|-------|
| 1. | Serendipity | _____ |
| 2. | Aardvark | _____ |
| 3. | Abonimable | _____ |
| 4. | Aberrant | _____ |
| 5. | Bacchic | _____ |
| 6. | Cacophony | _____ |
| 7. | Draconic | _____ |
| 8. | Ebullient | _____ |
| 9. | Fastidious | _____ |
| 10. | Gamut | _____ |
| 11. | Harbinger | _____ |
| 12. | Immanent | _____ |
| 13. | Jalopy | _____ |
| 14. | Kibitzer | _____ |
| 15. | Idiosyncrasy | _____ |
| 16. | Laconic | _____ |
| 17. | Macaque | _____ |
| 18. | Nadir | _____ |
| 19. | Obituary | _____ |
| 20. | Pachyderm | _____ |
| 21. | Quixotic | _____ |