

Connections to existing knowledge: The effectiveness of methods of vocabulary acquisition

Brendan Rodda

本論文は日本の大学の一年生の中での単語習得の方法の比較である。主に、コンテキストと言う意味内容的な方法とキーワードと言う記憶術的な方法の間の比較である。二つの方法とも以前の研究の反復練習より、いい結果であったことが示されている。更に、本論文はその二つの方法の組み合わせ、及び対象となる単語の日本語訳の変わりに簡単な英語の説明を使用したコンテキスト方法を調査する。この実験は、以前の方法を比べた実験の不均衡を正そうとする。結果は、キーワードを使用したグループがコンテキスト方法を使用したグループより有意差があつて単語を習得した。

This study compared the effectiveness of methods of vocabulary acquisition among first-year students at a Japanese university. The main comparison was between a semantic-based method – called the context method – and a mnemonic one – called the keyword method. Both of them have achieved significantly better results than rote learning in earlier studies. This study also tested a combination of the two methods as well as a version of the context method using simple English explanations of target words rather than Japanese translations of the words. This study aimed to redress some of the irregularities of earlier comparisons of the methods. It found that students using the keyword method were able to acquire and retain significantly more words than students using the context method.

Many researchers have investigated methods that are designed to facilitate vocabulary acquisition. Unfortunately, research findings have been inconsistent and have sometimes lacked validity in typical language learning situations. Perhaps those points explain the neglect of vocabulary acquisition methods among teachers and learners – faced with the inconsistent results and arcane conditions of research studies, teachers and learners tend to ignore the field. This is a pity because some of the methods offer significant advantages over the usual practice of reading the new word together with its translation in order to memorize it or, more simply, acquiring the word through exposure in communication. Despite the confusing results of the research, it is clear that by using the methods, learners can save considerable time and effort.

The aim of this study is to compare two methods of vocabulary acquisition – the context method and the keyword method – in ways that avoid the imbalances and irregularities of some earlier studies. The context method – also known as semantic-context or semantic processing – involves exposure to example sentences using the target word. It is a straightforward method that somewhat mimics the way children learn their first language. However, it differs from natural acquisition in that it concentrates in time a number of exposures to the new word and encourages learners to notice the word and think about it. The keyword method is more complicated. It is a mnemonic technique that involves matching a target word in the L2 with a similar sounding word (the keyword) from the L1. The meanings of the two words are then combined in one mental image (the keyword image). For example, an English speaker trying to learn the Japanese word *mado* – meaning *window* – could use *mud* as the keyword and then imagine a window covered with

mud. The method aims to assist acquisition of vocabulary through creating associations for the form and meaning of the target word with existing semantic and phonological knowledge.

Although early findings tended to support the keyword method, various studies in the 1990s cast some doubt on its effectiveness. Ellis and Beaton (1993) found that the much simpler rote rehearsal method – vocalized or subvocalized repetition of the target word – produced significantly better results for production of target words in a foreign language, although the keyword method was more effective for receptive learning. However, rote rehearsal subjects in the study produced the target words during the learning phase while the keyword method subjects did not and this might have influenced results for productive learning.

Wang and Thomas (1995) found that retention of vocabulary learnt by the keyword method is not durable in comparison to the context method. They attributed their findings to the lack of practice in their study. They set up the study so that subjects in the delayed recall groups did not have any practice sessions or do other tests between the initial learning session and the delayed post-test. Furthermore, they used incidental instruction techniques so as not to encourage subjects to practice recall of their own accord. They claimed that these procedures provide a purer test of the learning methods, unconfounded by other factors such as practice and over-learning. Although the keyword method resulted in greater learning when practice was allowed, in two experiments without practice the context method group had greater retention of target words two days after learning. Wang and Thomas concluded that retention of words learnt by the keyword method is less durable and practice is necessary for effective retention using the method. Gruneberg (1998) criticized Wang and Thomas's study for its lack of relevance to actual situations of vocabulary acquisition. He claimed that in most learning situations, learners review recently learned words, so practice almost always occurs anyway, making Wang and Thomas's findings against the keyword method trivial. Another point that compromised the external validity of their findings is that normally language learners study words with the intention of remembering them, whereas Wang and Thomas's study used incidental learning. Retention might differ considerably under such varied conditions.

Brown and Perry (1991) achieved more external validity in their comparison of the context and keyword methods by using genuine learners as subjects and by providing review of the target words in the form of tests. Furthermore, they gave thorough instruction and practice in the use of the methods to the subjects in their study. The study compared three groups – one using the context method, one the keyword method and one a combination of both methods. They found that the combined method group performed significantly better than the keyword group. The context method group was between the other two, with no significant difference to either group. Brown and Perry referred to Craik and Tulving's expanded depths-of-processing theory (1975, cited in Brown and Perry, 1991) to explain the results. In this theory, acquisition is improved when cognitive processing is deeper and is best when there is elaboration – that is, further processing at another level. Brown and Perry ascribed the keyword method to the sensory level because of its reliance on imagery. They claimed that the depths-of-processing theory explains the greater effectiveness of the context method because this method operates at the semantic level, which is a deeper level of processing than the sensory level, according to the theory. The combined method is most effective because it adds elaboration to the deeper processing.

Despite these negative results for the keyword method, Gruneberg (1998) claimed that about 50 studies have found that it facilitates foreign language vocabulary acquisition, with only about five studies failing to find an advantage. He cites various field studies that have shown that learners have learnt up to 200 words a day, with very high retention after a delay. As these are field studies, there are problems with external validity and internal validity is sometimes compromised. Nonetheless, the findings suggest that the keyword method can be powerful in practice. Experimental studies, too, have found advantages for the keyword method. McDaniel and Pressley (1989) found it to be more effective than the context method. Recently, Sagarra and Alba (2006) compared it to semantic mapping, which involves connecting the L2 target word to L1 words with similar meanings. They found the keyword method to be much more effective, although they only used nouns with a high degree of concreteness and imageability as their target words. Interestingly, although they found the method to be superior to a semantic-based method, they followed Brown and Perry (1991) in using depths-of-processing theory to explain the results. Sagarra and Alba claimed that the keyword method works on both the semantic and the sensory levels and therefore provides deeper cognitive processing than semantic mapping.

As mentioned in the above summaries, researchers tend to attribute the effectiveness of the keyword method to the power of imagery. This view was supported by Shapiro and Waters (2005) who found that the method was much more effective with words that had high imageability.

With results from Brown and Perry (1991) and Ellis and Beaton (1993) favoring combined approaches, it might well be that the most effective method of foreign language vocabulary learning is a combination of context method, keyword method and rote rehearsal. Wang and Thomas have also stated that "the most optimal training program may include a variety of strategies." (1999, p.285)

Purpose

This study compared the effectiveness of the context method and the keyword method for long-term vocabulary retention in ways similar to the Brown and Perry study (1991). However, there were two main differences. Brown and Perry provided subjects in the keyword method groups with keywords but had them think of their own keyword images. This places a greater cognitive load on these subjects compared to those in the context method group. To avoid this disparity, subjects in the keyword method groups in this study were provided with keywords and keyword images for each target word. Secondly, Brown and Perry tested long-term retention of the words at nine days after the final learning session. Here subjects were tested 13 days after the final learning session. This period gives a slightly better indication of the effectiveness of the methods over the long term. Like Brown and Perry but unlike Sagarra and Alba (2006), the methods were compared using concrete and abstract nouns and verbs in order to test the methods in the normal circumstances of language learning, giving the results greater generalizability.

Research questions

1. Do the keyword method, the context method and a combination of the two methods differ in effectiveness for long-term retention of foreign language vocabulary?

The combined method was expected to be superior, in keeping with the results of

other studies. The version of the keyword method used here was expected to make it more effective than the context method. Advantages for the context method that appeared in earlier studies might have been the result of using more demanding versions of the keyword method. When used well, the keyword method connects strongly the form and meaning of the target word to existing semantic and phonological knowledge. The context method, on the other hand, requires learners to establish new knowledge largely independent of existing knowledge, which seems to be a more difficult feat.

2. In regard to the context method, is the use of L1 translations of L2 target words more effective than the use of simple explanations in the L2 of the target words?

This is a relevant point for classes that follow an English-only approach. It was expected that the context method would be more effective when L1 translations are used because, again, there is a stronger connection with existing knowledge.

Another question arose in response to Sagarra and Alba's study (2006), after the experiment for this study was completed. As such, it was not rigorously investigated but some provisional results are indicated.

3. Is the keyword method less effective with abstract words in comparison with the context method?

Method

Design

A between-subjects design was used with four treatment conditions: combined keyword method/context method, keyword method, context method with Japanese translation and context method with English explanation. The two dependant variables were an immediate post-test and a delayed post-test. Subjects were not assigned to groups randomly; for practical reasons, class groups at a Japanese university were used for the treatment groups.

Participants

Four first-year classes in the English department at a Japanese university participated in the study. Students in the classes were at a low-intermediate to mid-intermediate level. Only results from students who received all instruction and testing were included in the study. Three students who had prior knowledge of the words on the final list were excluded. The final numbers of students were 21 in the combined keyword/context group, 19 in the keyword group, 24 in the group using the context method with Japanese translations and 24 in the group using the context method with English explanations.

Materials

Subjects were presented with 40 words – 20 nouns and 20 verbs, including 8 phrasal verbs. These words were selected from among the words that first-year students at the university in the previous year had indicated they did not know. However, in a questionnaire administered after the final learning session, several subjects indicated on a 5-point Likert-type scale that they had at least a fair knowledge of one or more of the 40 words before the experiment. Three of the subjects indicated knowledge of many of the words so these subject's results were

excluded from the final results of the study. Words that received from other subjects a score of 3 or more on the 5-point scale were eliminated from the final list of words considered for the results of the study. This final list consisted of 24 words – 13 nouns and 11 verbs, including 2 phrasal verbs.

Subjects in the keyword and combined methods groups received keywords and keyword images for each word. Each keyword was chosen for its phonological similarity to the target word. Imageability of keywords was a secondary factor in the selection. Most of the keywords were Japanese but a few were English words or names well-known to the subjects (e.g. *photo* for *fort*). Some of the target words had two keywords (e.g. *egg* and *soup* for *excerpt*). Most of the keywords were nouns but some verbs (e.g. *suwaru*, meaning *sit*) and adjectives were also used.

Subjects in the two context method groups received three short and relatively easy example sentences for each word. The first two were declarative sentences and the third was a question that was designed to prompt subjects to think about the meaning of the word in relation to their own experience or ideas.

Subjects in the keyword, the combined methods and the context (Japanese) groups received Japanese translations for each of the words. Subjects in the context (English) group did not receive a translation of the words, instead receiving a simple English explanation of each word. Subjects in all groups heard each word pronounced four times.

Procedure

Before the first learning session, the keyword, context (Japanese) and context (English) groups received 30 minutes of instruction and practice in the method that they were to use. The combined methods group received 40 minutes of instruction and practice in the keyword and context methods.

The target words were presented to the treatment groups in two learning sessions on consecutive days. In the first session, the 20 nouns were presented and in the second session the 20 verbs. Subjects listened to each word being pronounced four times and were told to repeat after each pronunciation. This was done in order to control for the effect of rote rehearsal – that is, vocalized or subvocalized repetition of the target word. Some subjects might otherwise have used rote rehearsal by themselves, possibly interfering with the results. The pronunciation and repetition of each word took 10 seconds and subjects had a further 30 seconds in which to use their method(s) to learn each word.

After the second session, subjects answered a questionnaire and then took the first post-test (Test 1). This test was a measure of productive knowledge of the target words. Subjects were given the Japanese translation for each target word – in a different sequence from the learning sessions – and asked to write the corresponding English word that they had studied.

The delayed post-test was administered to the groups 13 days after the final learning session. In order to discourage subjects from reviewing the words, they were not told in advance of the delayed post-test. The first part of the test (Test 2A) was the same as the immediate post-test. Subjects' answer sheets were collected before beginning the second part of the test. The second part (Test 2B) was a test of receptive knowledge. All of the target words were given – again, in a different sequence – and subjects were asked to write a Japanese translation. In the case of the context (English) group, subjects were allowed to answer with a Japanese translation or an English explanation.

Each answer on Tests 1 and 2A was given one point if it was correct or if there was only a minor spelling mistake (e.g. *salk* instead of *sulk*). An answer received half

a point if it had a more serious spelling error but was close to the correct word (e.g. *thron* instead of *thorn*). Synonyms were not accepted. Test 2B was graded by a native-speaker of Japanese. Each answer received one point if it was the translation or explanation that was given in the experiment materials or if it was a synonym. Answers that were fairly similar in meaning to the English words were given half a point.

Results

To assess the effects of the four treatments, a one-way analysis of variance (ANOVA) was conducted on results from all tests. The effects of the vocabulary learning methods differed significantly in Test 1 ($F(3,83) = 23.77, p < .001$), Test 2A ($F(3,83) = 10.95, p < .001$) and Test 2B ($F(3,83) = 14.57, p < .001$). Mean scores and standard deviations for all treatment groups on all tests are presented in Table 1.

Table 1: *Descriptive Statistics for Tests*

<i>Treatment</i>	<i>N</i>	<i>Test 1</i>		<i>Test 2A</i>		<i>Test 2B</i>	
<i>Group</i>		M	SD	M	SD	M	SD
Keyword	18	7.31	3.99	3.44	2.73	6.81	4.23
Combined Methods	21	5.98	3.8	3.21	2.91	6.9	4.41
Context (Japanese)	24	2.56	1.86	1.33	1.22	3.77	2.58
Context (English)	24	0.85	0.92	0.48	0.65	1.35	1.36

Post hoc paired ANOVA tests showed no significant difference between the combined methods group and the keyword group in any test. However, in every test, both the combined methods and keyword groups had significantly higher acquisition than both context groups ($p < .01$). The context (Japanese) group had significantly higher acquisition ($p < .01$) than the context (English) group in every test.

To determine the effect of treatment condition on the learning of abstract words, the six most abstract words on the noun list were selected and scores for these six words were calculated for the keyword group and the context (Japanese) group on Tests 2A and 2B. ANOVA tests showed that on both tests the keyword group had significantly higher acquisition of the abstract words than the context (Japanese) group ($p < .005$).

Discussion

The main research question investigated here was whether the keyword, the context and combined methods differed in effectiveness. The results do not support the initial expectation that the combined method would be superior. In each test, there was no significant difference between the keyword group and the combined methods group. It might be that the participants in the combined methods group – with only 30 seconds to attend to the keyword, keyword image and three example sentences for each target word – did not have enough time to benefit fully from both

methods. It could also be that the use of the context method did not provide any benefit beyond that provided by the keyword method.

The finding that the keyword method led to greater acquisition than the context method is consistent with the initial expectation. The marked difference between the methods was apparent from the immediate post-test. Interestingly, the rate of forgetting for the keyword group was slightly higher than the rate for the two context groups but the initial benefits of the keyword method were such that the method remained significantly more effective in the long term.

The second research question involved a comparison of the use of Japanese translations and simple English definitions. The findings clearly support the expectation that an L1 translation is more effective than a definition in the L2. This is in keeping with the findings of several researchers in psycholinguistics (e.g. Fox, 1996) that suggest that in the brains of language learners, lexical representations of L2 words tend not to connect directly with semantic representations and instead connect with them via the existing connections from words of the L1. In effect, the context (Japanese) group was encouraged to follow this path from L2 words to semantic representations via L1 words by using an L1 translation, while the context (English) group was encouraged to make a new and direct connection from the L2 to semantic representations independent of the L1.

Mediation of the acquisition of L2 vocabulary through the L1 goes a step further with the keyword method. In this case, not just the meaning but also the sound of the L2 word is connected to L1 words. Furthermore, through the keyword image, those two connections are themselves connected to each other. The keyword method takes what is most natural and apparently most effective in language learning on the semantic level and extends the process by mimicking it on the phonological level also and then binds the two processes together. Researchers tend to attribute the effectiveness of the method to its use of imagery, in the belief that visual stimuli in general have a greater impact on memory than other stimuli. However, the finding that the keyword group had significantly greater acquisition of abstract nouns than the context group suggests that imagery might not be such a central part of the keyword method.

The overall findings of this study are better explained by the amount of connection to existing knowledge. The keyword method creates connections to existing knowledge of sounds and meanings and emerges as the most effective. The context (Japanese) method has a connection to existing knowledge in the L1 only on the semantic level and it has less effectiveness than the keyword method but more effectiveness than the context (English) method, in which the connection to the established semantic representations of the L1 is diminished. Given that rates of forgetting were comparable for all methods, the keyword method's advantage lies not in the durability of retention but in the power to prompt initial acquisition. The findings suggest the possibility that initial acquisition is more likely to succeed when there is a greater amount of connection to existing knowledge.

Conclusion

This study has provided evidence of the effectiveness of the keyword method for vocabulary acquisition. There is now a sizeable body of research that supports the method's effectiveness in comparison to semantic-based approaches and rote memorization. Although the method remains largely unknown or unused by language teachers and learners, this body of research provides strong grounds for its introduction into language learning programs. The results of this study counter the claims that the method is not suitable for learning abstract words and that

intermediate learners do not benefit from it as much as they do semantic-based methods (Brown and Perry, 1991) or rote rehearsal (Van Hell and Candia Mahn, 1997).

The use of classes as treatment groups is one limitation of this study. Slight differences between the language levels of classes might have had some influence on results. Replication of the findings with random assignment of subjects or a within-subjects design would provide greater validity. Future research might also compare more rigorously the effectiveness of the keyword method for acquisition of concrete and abstract words. Only a small number of items were used in the comparison undertaken here so no strong conclusions can be drawn.

Although the context method was relatively ineffective here, its good results in other studies suggest that it has a part to play in facilitating vocabulary acquisition. It might be that it is better suited to later stages of a word's acquisition. Used then, it can add durability to the acquisition and provide deeper understanding of meaning and usage. At the initial stage, though, the keyword method is a more powerful means of facilitating acquisition.

References

- Brown, T., & Perry, F. (1991). A comparison of three learning strategies for ESL Vocabulary Acquisition. *TESOL Quarterly*, 25 (4), 655-670.
- Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 46A (3), 533-558.
- Gruneberg, M., (1998). A commentary on criticism of the keyword method of learning languages. *Applied Cognitive Psychology*, 12, 529-532.
- McDaniel, M.A., & Pressley, M. (1989) Keyword and context instruction of new vocabulary meanings: effects on text comprehension and memory. *Journal of Educational Psychology*, 81 (2), 204-213.
- Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *Modern Language Journal*, 90 (2), 228-243.
- Shapiro, A.M., and Waters, D.L. (2005). An investigation of the cognitive processes underlying the keyword method of foreign vocabulary learning. *Language Teaching Research*, 9 (2), 129-146.
- Van Hell, J., & Candia Mahn, A. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, 47 (3), 507-546.
- Wang, A.Y., & Thomas, M.H. (1995). Effect of keywords on long-term retention: Help or hindrance? *Journal of Educational Psychology*, 87, 468-475.
- Wang, A.Y., & Thomas, M.H. (1999). In defense of keyword experiments: A reply to Gruneberg's commentary. *Applied Cognitive Psychology*, 13, 283-285.