

## Teaching English Vocabulary to Cantonese-speaking Students with the Keyword Method\*

See Shing YEUNG and Rex M. HEYWORTH  
The Chinese University of Hong Kong

Effectiveness of four instructional conditions for teaching 18 English target words in a 35-min session to Cantonese-speaking Chinese students in Hong Kong was examined in two experiments. The four conditions were: (a) The context method which placed the target word in a linguistic context exemplifying its meaning, (b) the keyword method which provided a Cantonese sound similar to the target English word and a picture showing the referent of the target word and that of the Cantonese keyword interacting with each other, (c) the context + keyword method which applied both strategies a and b, and (d) rote-memory control. In Experiment 1 with 101 low-ability subjects in Form 4, the keyword method and the context + keyword method were found to be effective in enhancement of immediate recall and delayed recall two weeks and ten weeks after initial acquisition. Retention rates in conditions b and c were better than those in conditions a and d. In Experiment 2, a comprehension subtest of a standardized English examination was used to categorize 240 Form 3 subjects from 8 classes of 2 schools into high- and low-ability groups. Analysis with a 4 (conditions) X 2 (abilities) X 2 (immediate test and 2 weeks' delayed test) ANOVA with repeated measures indicated condition by ability interaction. High-ability subjects did not differ significantly in either of the tests; but in the delayed test, low-ability context + keyword subjects outperformed the context subjects and the control subjects, and keyword subjects outperformed the context subjects. Retention rates of low-ability subjects in conditions b and c were better than those in conditions a and d at both levels of ability. The keyword method proved to be a potential effective mnemonic aid in a classroom situation to supplement the context method in vocabulary instruction to low-ability Chinese learners of English as a second language, and it has the potential to minimize the difference between high- and low-ability learners.

In teaching vocabulary of a second language, Stahl and Fairbanks (1986) have categorized the various methods as (1) methods which provide definitional and contextual information, (2) methods which increase depth of processing, (3) methods which increase exposure, (4) group discussion

methods, and (5) the mnemonic keyword methods. Their meta-analysis of these methods indicates that the context method and the keyword methods are the most effective methods in vocabulary instruction. The present study compares the effectiveness of (1) the context method, (2) the

\* This article is an abstract of the MA (ED.) thesis of the first author under the supervision of the second author.

keyword method, (3) the context+keyword methods (i.e., 1 and 2 combined), and (4) control, in teaching English vocabulary to Cantonese-speaking students of Hong Kong.

### The Context Method

Teaching lexical items through linguistic context has been a well established method which is in current curriculum usage (Pressley, Levin, & Miller, 1982). This method has been commonly used in Hong Kong where English is taught as a second language.

The context method of vocabulary teaching is one in which the lexical item to be learned is placed in a linguistic context such that the learner relates the new word with his pre-existing knowledge (Stahl & Fairbanks, 1986). For example, Gipe and Arnold (1979) put the target word in a three-sentence passage, one of which was the definition of the target word. The other two sentences were written in simple structure with words which the subjects were familiar with, and they helped to exemplify the target word.

Beheydt (1987) proposes that the context of a target word should consist of known words to the learner and the target word should be embedded in the context such that the concepts associated with the target word may be evoked. In the present study, when the word "mallet" was taught with the context method, the three sentences were:

*mallet*      A mallet is a wooden hammer.  
                 Tom used a mallet to break nuts.  
                 He didn't use an iron hammer.

Gipe and Arnold (1979) found that the context method was more effective than three other commonly used vocabulary teaching methods under investigation. Other researchers have also found facilitation of vocabulary learning and retention by means of the context method (Crist & Petrone, 1977; Eubanks & Ferguson, 1982; Gipe, 1980; Wittrock, Mark, & Doctorow, 1975).

However, the usefulness of the context method relies on the learner's previous possession of a vocabulary sufficient for the cognitive processing and for the establishment of a linguistic context, and it also relies on their competence in syntactic parsing. For learners who are weak in either, or in both, of these aspects, its effect may be doubtful.

### The Keyword Method

Retrieval seems to occur by means of a search-like process (Ackerman, 1986). The more search is constrained and directed by associative structure, the more successful the retrieval attempt (Ackerman, 1986). Atkinson (1975) claims that the keyword method provides a definite route for retrieval of the response, and the probability of successful retrieval would be high. In second-language instructions, the keyword method, as coined by Atkinson (1975), uses both a verbal acoustic link and an imagery link between a keyword in the learner's mother tongue and the lexical item in the second language.

In this study, one of the examples is using the keyword method to learn the English target word "mallet". The keyword is the Chinese sound 錘/*mèi* which is similar to the sound of the first syllable of the target word, and this is the acoustic link. The imagery link is formed by creating an interacting image showing a person 負/*mèi*, which means carrying on the back, a mallet. The picture of a mallet for the context and the control conditions is shown in Figure 1. A corresponding picture showing the interacting image for the keyword and the context+keyword conditions is shown in Figure 2.

Interacting images were used because it was found that interacting images are more effective than two separate images (Bower, 1970; Wollen & Lowry, 1971).

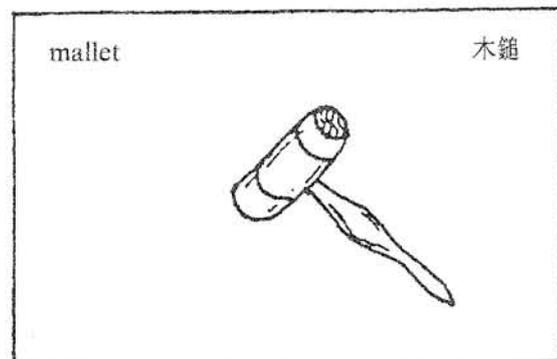


Figure 1. Picture of example "mallet" in Context condition and Control

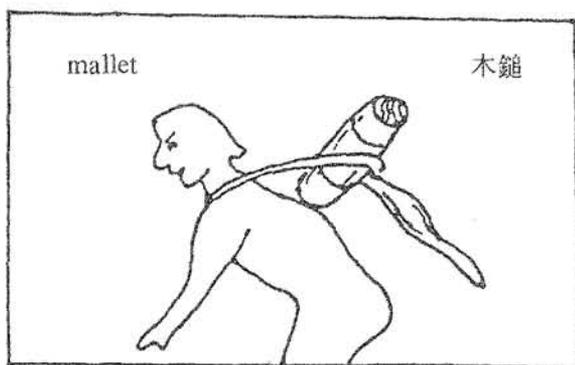


Figure 2. Picture of example "mallet" in Keyword and Context + keyword conditions

The effectiveness of the keyword method has been tested with subjects of different ages and levels; for example, at Grade 2 and Grade 6 (West, Stanovich, Feeman, & Cunningham, 1983), Grade 5 (McGivern & Levin, 1983), Grade 8 (Jones & Hall, 1982; Shriberg, Levin, McCormick, & Pressley, 1982), Age 12 and Age 18 (Pressley & Dennis-Rounds, 1980), and with undergraduates (Pressley, Levin, Nakamura, Hope, Bispo, and Toye, 1980; Raugh & Atkinson, 1975).

Although the keyword method has been generally proved to be effective in the laboratory, it will be necessary to test its effectiveness in the classroom in order to confirm its application (Merry, 1980; Pressley & Levin, 1978). It will also be worth assessing the effectiveness of the keyword method with subjects whose mother tongue is a logographic/ideographic language, like Chinese, learning a morpho-phonemic language, like English.

General research findings show that subjects instructed with the keyword method outperformed control subjects and subjects instructed with the context method (Levin, McCormick, Miller, Berry, & Pressley, 1982; McDaniel & Tillman, 1987; Pressley, Levin, Kuiper, Bryant & Michener, 1982; Pressley, Levin & Miller, 1982). Merry's (1980) pilot experiment also showed that the keyword method with interacting pictures provided was particularly

effective when used in a group setting with low reading age subjects.

Referring to the results of their experiments, Johnson, Adams, and Bruning (1985) argued that the keyword method could only be useful for learning concrete words; but Pressley, Levin, and Miller (1981) found remarkable keyword effects for abstract words as well as for concrete ones. Johnson, Adams, and Bruning (1985) also concluded that the keyword method could only be effective for retention of vocabulary for a short time. This study is an attempt to test the short-term and long-term effects of the keyword method with a wide range of randomly selected words.

The keyword method may be employed in several variations, but it was generally found that imposed imagery strategies, in which the learner was provided with the image, were more effective for younger learners (Levin, 1976). It was also found that the use of a well-structured imposed imagery strategy could decrease the difference between high- and low-ability learners in vocabulary learning (McGivern & Levin, 1983). In this study, imposed imagery is used.

### The Context + keyword Method

Memory schemata and mnemonic devices operate rather differently (Battig & Bellezza, 1979). In second-language vocabulary learning, during the storage phase, the context method operates by activating the learner's schema but the keyword method operates by transforming the target word into some other form so as to aid memory (Bellezza, 1981).

For example, a learner may be taught the word *brawl* using the context:

*brawl*      To brawl means to quarrel noisily.  
                  Tom and Mary both wanted the  
                  ticket for Alan's concert. They  
                  brawled over it and it was torn.

The procedure the learner follows to retrieve the meaning of the word "brawl" may be hypothetically summarized as in Figure 3. Figure 4 illustrates the hypothetical procedure of retrieval in the keyword condition.

|                      |                               |                                    |                                   |                                  |
|----------------------|-------------------------------|------------------------------------|-----------------------------------|----------------------------------|
| STEP 1               | STEP 2                        | STEP 3                             | STEP 4                            | STEP 5                           |
| Learner sees "brawl" | Recognizes it has been learnt | Searches for information via links | Recalls information about "brawl" | Reproduces definition of "brawl" |

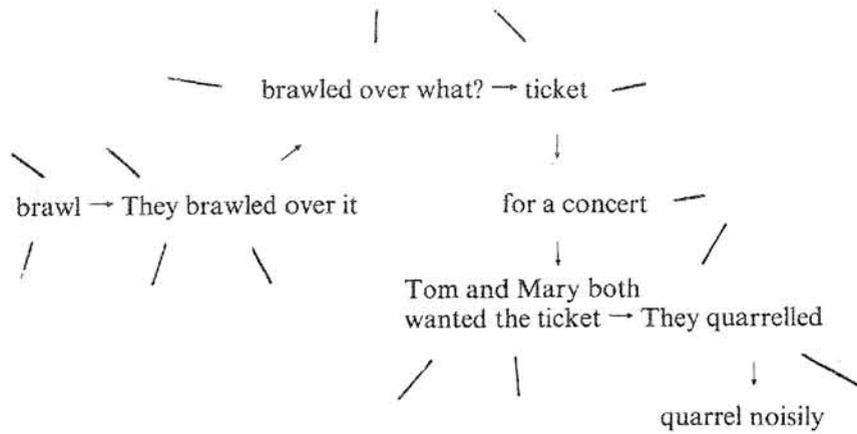


Figure 3. Procedure of retrieval in Context condition

|                      |                   |                           |                 |                                      |
|----------------------|-------------------|---------------------------|-----------------|--------------------------------------|
| STEP 1               | STEP 2            | STEP 3                    | STEP 4          | STEP 5                               |
| Learner sees "brawl" | Vocalizes "brawl" | Transforms into Cantonese | Recalls keyword | Forms image representing the keyword |

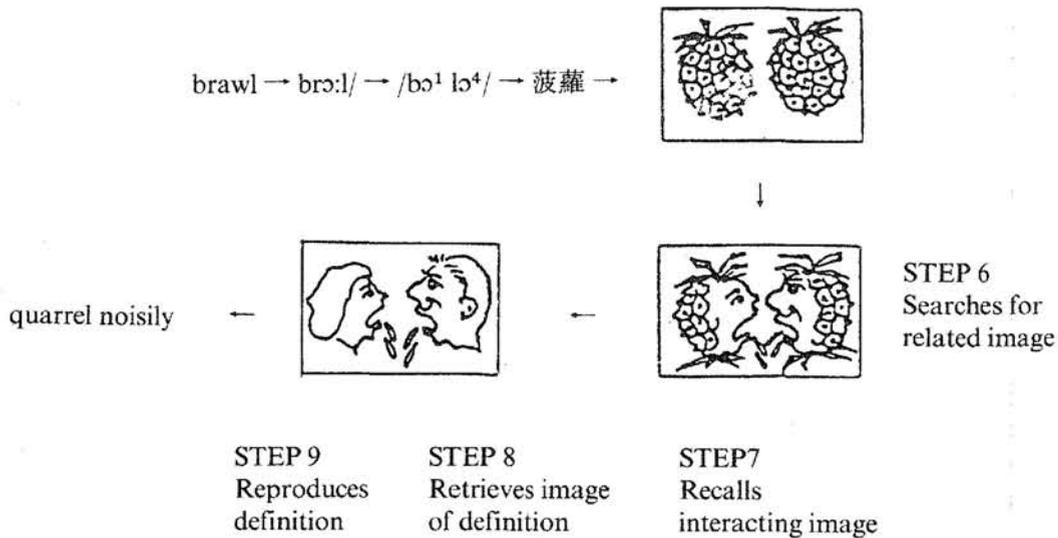


Figure 4. Procedure of retrieval in keyword condition

In the context condition, although the learner may start successfully in Step 1 and Step 2, there is no guarantee for the successful search in Step 3, which subsequently leads to Steps 4 and 5. On the other hand, because the keyword method provides a direct link between the stimulus and the response during the storage phase, the route from the stimulus to the response during retrieval is definite.

Accordingly, in the condition of using the context + keyword method, i.e., the combined use of both the context method and the keyword method, the learner may choose between either learning by schema activation or by the mnemonic keyword method.

Restle (1964) suggests that subjects might learn items by actively sampling from a pool of memorization strategies. Some researchers suggested that mnemonic strategies could be used in conjunction with other methods of learning (Gruneberg & Sykes, 1978). Levin (1986) has proposed an integration of two or more strategies with different theoretical strengths such that the strengths of both kinds of strategies may be fully exploited. A combined use of the context and keyword methods may yield better results than either of them when used alone.

### **Individual Differences**

In two experiments, McDaniel and Pressley (1984) taught university students 61 obscure English words in four conditions, namely, (a) the keyword method with imposed keyword but no picture, (b) the context method with linguistic context but no definition, (c) the context + keyword method with no picture and no definition, and (d) control. The results indicated superiority of the keyword method with low-ability subjects but not with high-ability subjects. Low-ability keyword subjects recalled significantly more word meanings than low-ability context subjects, but respective high-ability subjects did not differ statistically. Also, the difference between high- and low-ability keyword subjects was not statistically significant. It was also interesting to find that low-ability context+keyword subjects recalled significantly more than low-ability context subjects.

High-ability learners may have "a bag of tricks" for learning vocabulary (Atkinson, 1975) and thus instructions of memory strategies may benefit low-ability learners more. High-ability learners may apply their own strategies unless they believe that

the newly introduced strategy is more effective than their own.

## **Experiment 1: Teaching Low-ability Students Design and Procedure**

### **Subjects**

The subjects were Cantonese-speaking Chinese students studying in Secondary 4 in a vocational training school of Hong Kong. All were boys with ages ranging from 15 to 17. Four classes of students, i.e., 101 subjects took part. Because of absentees, the scores of 98 subjects were analysed in a comprehension pretest, vocabulary pretest, immediate vocabulary posttest, and delayed vocabulary posttest two weeks after acquisition, and the scores of 95 subjects were analysed in the delayed vocabulary posttest ten weeks after acquisition.

The objective of this study is to assess the effectiveness of each method in authentic classroom vocabulary learning. Treatment was thus given in intact classes.

The subjects came from 38 secondary schools and all of them had finished Secondary 3. All subjects had failed in English in the JSEA which was a public examination so as to qualify for Secondary 4 places.

### **Research Design**

A repeated measures analysis of variance (ANOVA) design was used with treatment (four conditions of instruction) as the between-subjects factor, and the series of vocabulary tests as the within-subject factors.

### **The Material**

#### *The Target Lexical Items*

Three hundred and seventy-nine English words were taken from 912 words which have appeared five times in a million in print (Carroll, Davies & Richman, 1971; Thorndike & Lorge, 1944). The criteria for selecting these lexical items were: That they were not formed by combining two or more other words; that they were not proper nouns, or abbreviations, or vulgar words, or words confined to colloquial use; that they were not transformations from more frequent words; that the meaning could

not be detected with knowledge about prefix/suffix indications; that they were unfamiliar words, i.e., they did not appear in the lexicon recommended for the First Certificate in English; that they were not discipline-biased or culture-biased.

#### *Length and Parts of Speech*

From the collection of 379 words, 18 of them (six nouns, six verbs, and six adjectives/adverbs) were randomly selected to constitute the target lexical items. In each of these parts of speech, two items were one-syllable words, two were two-syllable, and the other two were of three or more syllables. A list length of 18 target words was used to avoid both ceiling and floor effect (Pressley, 1977). In addition, 27 words were randomly drawn to form nine distractors in each of the three posttests.

#### *The Keywords*

A list of the target words was given to 30 Secondary 5 students who were not involved in the experiment. They noted down the sound of each word in any way they thought was best. The transcriptions were then chosen arbitrarily in terms of their imageability to form the keywords. Sixteen of the keywords were Cantonese and two were English. Two "blind" raters indicated what the image could be. All the 18 keywords were different so as to avoid interference, as warned by Erkstrand, Wallace, and Underwood (1966).

#### *The Picture Stimulus*

Each picture was 42cm × 30cm in size, illustrating the meaning of the lexical item (see Figure 1 and 2). The target word was printed in lower-case letters on the top left corner. Each letter was 3cm to 6cm tall. The lines were 3mm thick. On the top right corner of the picture, the Chinese translation was printed 3cm × 3cm with lines 3mm thick. Two sets of pictures were used. The first set of pictures each showing the referent of the target word was used in the pronunciation practice and in the context and control conditions. The second set of pictures each showing an interacting picture of the referent of the target word and that of the keyword was used in the keyword and the context + keyword conditions.

#### **Instruments**

The instruments were: (1) Reading comprehension subtest: A previous JSEA English reading comprehension passage (322 words) with 10 multiple-choice questions, (2) vocabulary pretest: A randomized list of 18 target words, and (3) three vocabulary posttests, namely an immediate vocabulary posttest, two weeks' delayed vocabulary posttest, and ten weeks' delayed vocabulary posttest, each including a list of the same 18 target words and nine different distractors, in a different randomized order.

#### **Procedure**

The reading comprehension subtest took about 15 min. The vocabulary pretest, instructions, and immediate vocabulary posttest took about one 35-min session. The procedure was timed by the use of an audio cassette tape-recording. The delayed vocabulary posttests took about 15 min each.

*The reading comprehension subtest.* All subjects were tested one week before instruction.

*The vocabulary pretest.* The subjects wrote down the meanings of all of the words they knew within 4 min. Any other way to show understanding would be accepted.

*Pronunciation practice.* Each group listened to a one-minute introduction, then, to the same presentation of 18 lexical items one by one in the same random order at the rate of 8s per item during which the subjects repeated each item orally twice for educational reasons. At the same time, the experimenter raised the appropriate picture in front of the class.

*Instructions with the four methods.* Handouts for each method were distributed. Instructions in English were given in two repeated practices. During Practice 1, a three-minute instruction with an example appropriate for each particular group was given. Then the tape presented the lexical items one by one at the rate of 15 s per item. Subjects were told that they would be tested after learning them. In accordance with the voice, the experimenter raised the relevant picture for that method.

In the context method group, each subject was given the list of 18 target words each with the definition of its meaning in one sentence. Similar to that of Gipe and Arnold (1979), two more sentences exemplified its meaning. These three sentences together formed a short meaningful passage. The

sentences were of simple sentence structure as stipulated in the English syllabus for primary schools of Hong Kong and contained words which appear in the Cambridge English lexicon. Two English language teachers had checked, amended, and confirmed that average Secondary 3 learners would have no difficulty in understanding the three-sentence context and that the context could show the meanings of the target words. The pictures were the same as those in the pronunciation practice.

The keyword method group followed a similar procedure except that the handout contained only a list of target words with a one-sentence definition. The pictures for this group showed the image of the target word and a line-drawing of the keyword image interacting with each other. The subjects were told that they would be taught a new method to remember the English word meanings, and that they would be tested to see how well they could use this method.

In the context+keyword group, the same procedure and the same pictures as in the keyword group were used. In the rote-memory group, the same procedure as in the context group was followed. The subjects were told that they would be tested on how well they could remember English word meanings, and they were instructed to study the definitions carefully.

During Practice 2, the procedure as in Practice 1 was repeated, with another example, in the context group, the keyword group, and the control group. For the context+keyword group, however, Practice 2 included the use of the context method with appropriate handouts.

*Immediate vocabulary posttest.* Immediately after the presentation, all the handouts were collected. Each subject was given a test paper on which 27 English words were printed. The subjects were told to write down the definition or translation of each word next to it. They were told to mark a cross against each item that had not been presented. They were told to complete the test within 5 min and they could work on the items in any order they

liked. After each ninth item, the subjects were prompted with a printed instruction to try to use the method they had just learned. At the end of 5 min, the tape told the subjects that they could ask for two more minutes if they had not finished.

*Two weeks' delayed vocabulary posttest.* The same procedure as for the immediate vocabulary posttest was followed. The subjects were told to complete the test within 10 min. After the ten-minute period, The experimenter allowed an extra of 2 min if they had not finished. The subjects were not told about another test eight weeks later.

*Ten weeks' delayed vocabulary posttest.* The same procedure as for the two weeks' delayed vocabulary posttest was followed.

## Results

### *The Vocabulary Pretest*

As expected, all subjects scored zero, i.e., all 18 target words were new to all subjects before instruction.

### *The Comprehension Subtest*

One-way analysis of variance (ANOVA) was conducted and no significant difference was found among the mean scores of the subjects in the four conditions at .05 level. It was thus assumed that the subjects had similar abilities in comprehension and vocabulary competence. Scores in all the other tests were therefore compared directly among the conditions.

### *The Vocabulary Posttest*

ANOVA with repeated measures was conducted with method (four conditions) as the between-subjects measure and the scores of the immediate vocabulary posttest, two weeks' delayed vocabulary posttest, and ten weeks' delayed vocabulary posttest as within-subject measures. The results are listed in Table 1.

**Table 1 Results of ANOVA with Repeated Measures in Experiment 1  
(a) Between-subjects effects**

| SOURCE       | SUM OF SQUARES | df | MEAN SQUARES | F      | P   |
|--------------|----------------|----|--------------|--------|-----|
| Within cells | 2593.66        | 91 | 28.50        |        |     |
| Constant     | 24287.64       | 1  | 24287.64     | 852.15 | .00 |
| Method       | 2194.67        | 3  | 731.56       | 25.67  | .00 |

**(b) Within-subject effects**

| SOURCE          | SUM OF SQUARES | df  | MEAN SQUARES | F      | P   |
|-----------------|----------------|-----|--------------|--------|-----|
| Within cells    | 548.47         | 182 | 3.01         |        |     |
| Retest          | 1740.98        | 2   | 870.49       | 288.86 | .00 |
| Method X retest | 138.95         | 6   | 23.16        | 7.68   | .00 |

For the between-subjects effects, the main effect of method was found to be significant at .05 level. This indicated that different methods of instruction yielded different facilitative effects on vocabulary meaning recall.

For the within-subject effects, the results indicated significant main effect of the repeated posttests, and also significant effect of interaction between method and the repeated posttests at .05 level.

When repeated measures ANOVAs were conducted with the immediate vocabulary posttest and the two weeks' delayed vocabulary posttest, or with the two weeks' and ten weeks' delayed vocabulary posttests as the within-subject factors separately, similar results were obtained at .05 level.

These results indicated significant difference between the mean scores in the immediate

vocabulary posttest, two weeks' delayed vocabulary posttest, and ten weeks' delayed vocabulary posttest; and also significant difference was found in the effect of instructions on long-term retention of word meanings two weeks and ten weeks after instruction with respect to initial acquisition of word meanings.

One-way ANOVA was conducted to determine the effects of the four conditions on immediate vocabulary-meaning recall, and recall of vocabulary meanings two weeks and ten weeks after initial acquisition respectively. The means and standard deviations as functions of the four conditions are given in Table 2. The percentage of lexical items retained after a period of 14 days, and 70 days, is given respectively in Table 2. Figure 5 illustrates the respective retention rates.

Table 2 Mean Vocabulary-meaning Recall Scores as a Function of Condition in the Vocabulary Posttests in Experiment 1

|                             | CONTEXT<br>METHOD | KEYWORD<br>METHOD | CONTEXT+KEYWORD<br>METHOD | ROTE-MEMORY<br>CONTROL | <i>F</i> | <i>P</i> < |
|-----------------------------|-------------------|-------------------|---------------------------|------------------------|----------|------------|
| IMMEDIATE POSTTEST          |                   |                   |                           |                        |          |            |
| <i>n</i>                    | 27                | 21                | 23                        | 27                     |          |            |
| <i>M</i>                    | 10.30             | 13.52             | 15.35                     | 11.78                  | 10.95    | .0001      |
| <i>SD</i>                   | 3.51              | 3.09              | 3.05                      | 3.37                   |          |            |
| TWO WEEKS' DELAYED POSTTEST |                   |                   |                           |                        |          |            |
| <i>n</i>                    | 27                | 21                | 23                        | 27                     |          |            |
| <i>M</i>                    | 4.07              | 8.90              | 12.61                     | 5.89                   | 30.78    | .0001      |
| <i>SD</i>                   | 2.54              | 3.83              | 3.77                      | 3.23                   |          |            |
| RETAINED                    | 39.51%            | 65.83%            | 82.15%                    | 50.00%                 |          |            |
| TWO WEEKS                   |                   |                   |                           |                        |          |            |
| 10 WEEKS' DELAYED POSTTEST  |                   |                   |                           |                        |          |            |
| <i>n</i>                    | 26                | 21                | 21                        | 27                     |          |            |
| <i>M</i>                    | 3.19              | 9.19              | 11.52                     | 4.81                   | 27.15    | .0001      |
| <i>SD</i>                   | 2.42              | 4.45              | 4.50                      | 2.83                   |          |            |
| RETAINED                    | 30.97%            | 67.97%            | 75.05%                    | 40.83%                 |          |            |
| TEN WEEKS                   |                   |                   |                           |                        |          |            |

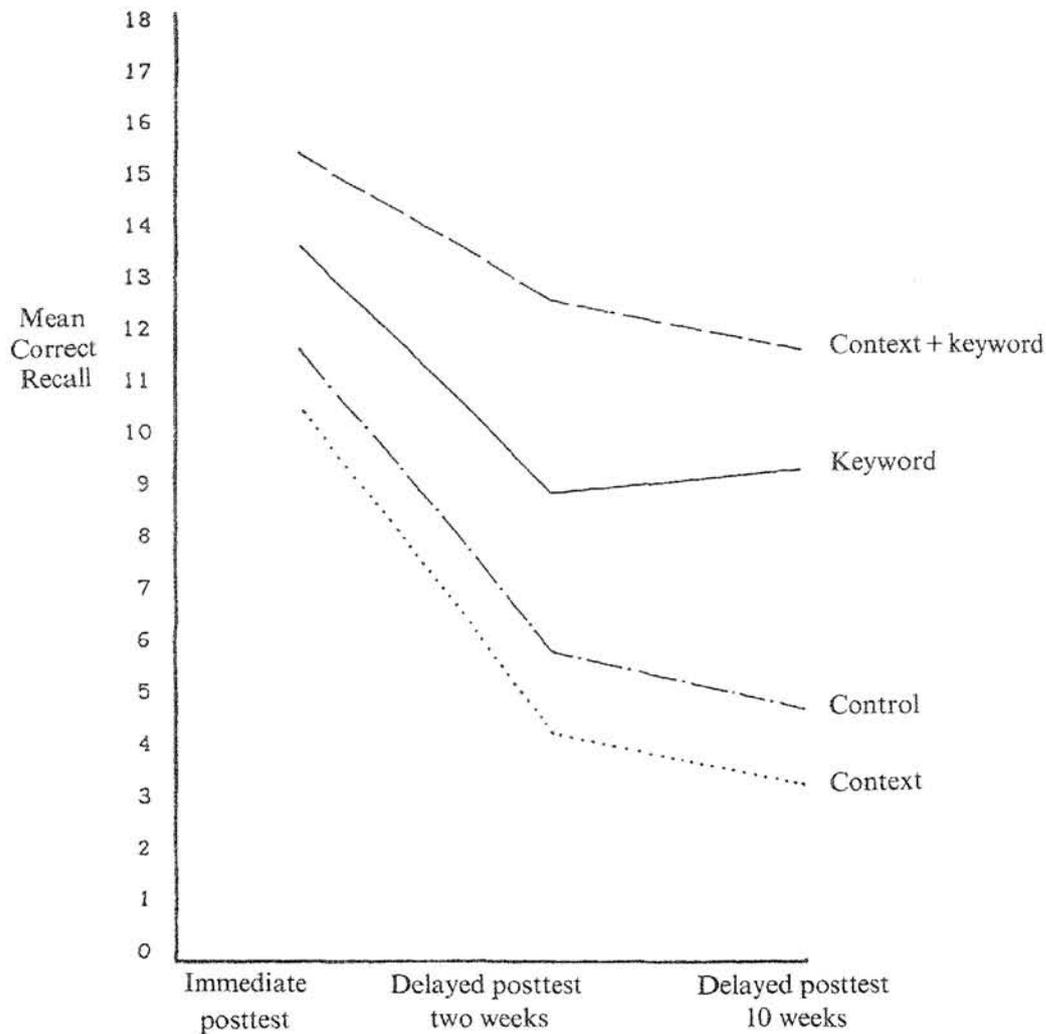


Figure 5. Mean correct recall of vocabulary meaning of four conditions at three vocabulary posttests in Experiment 1

### Discussion

In Experiment 1 where 18 new words were taught, each subject in the context method had to manipulate 54 sentences. Although these sentences were written in simple structure, the weak learners might need a lot of effort to manipulate each sentence. On the contrary, the control subjects had the advantage of sparing more time for repetition of the target words and their meanings.

#### *The Context Method vs the Control*

The results showed that subjects in the context method group did not perform better than those in the control group in immediate recall. Furthermore, the subjects' scores in the context group and the control group were not significantly different in the delayed vocabulary posttest either 14 days or 70 days after instructions, either.

### *Facilitative Effects of the Keyword*

Difference in the rate of decline in retention may indicate the effectiveness of particular methods in facilitating retention. The context+keyword method was found to be superior to the context method and control both in immediate recall and long-term retention of word meanings. The retention rates were much higher than those in the context method and the control groups.

There was no significant difference between the keyword method and the context+keyword method in immediate recall but difference between the two methods in long-term retention for two weeks was significant. However, this difference disappeared ten weeks after initial acquisition. But methods which used the keyword were found to be more effective than the context method and control in facilitating long-term retention.

Other research findings have also indicated superiority of the keyword method to the context method and to rote-memory control (e.g. Levin, McCormick, Miller, Berry & Pressley, 1982; McDaniel & Tillman, 1987; Pressley, Levin, Kuiper, Bryant & Michener, 1982; Pressley, Levin & Miller, 1982). However, the difference of effects on high- and low-ability second-language learners could be further investigated (Delaney, 1979). Experiment 2 was conducted to compare the effectiveness of the methods with different abilities.

### *A Combined Use of the Context and Keyword Methods*

With the aid of the keyword method, subjects in the context+keyword method group could select from the 54 sentences a certain number of sentences they thought could be helpful to their memorizing the meanings. Subjects in the context+keyword group could have benefited from their own selection of exemplifications in the context with a deeper processing of the word meanings. At the same time, they benefited from the memory enhancing keyword method.

Effectiveness of the keyword method in facilitating long-term memory of word meanings could be explained with Atkinson's (1975) model of the use of both an acoustic link and an imagery link so as to enhance retrieval from memory. Although the context method might facilitate understanding, and a firmer storage of information (Gipe & Arnold,

1979), it might not enhance retrieval when the subject encountered the lexical item later again. The problem of memory may lie in both storage and retrieval of information. The keyword method could provide some readers with an easier access to the word meanings stored in long-term memory.

The context method may be better in enhancing storage of word meanings by processing them at the semantic level, whereas the keyword method may be better in enhancing retrieval of word meanings through the acoustic and imagery links between the keywords and the target lexical items. Subjects in the context+keyword group might have benefited from the advantages of both methods.

## **Experiment 2: Teaching Students of Varying Abilities**

### **Design and Procedure**

#### **Subjects**

The subjects were Cantonese-speaking Chinese students studying in Secondary 3 in average-standard secondary schools in Hong Kong, ages ranging from 13 to 16. Eight classes of students from two schools, i.e., a total of 240 took part, with two classes assigned to each of the four conditions.

The subjects with scores in the comprehension subtest at or above the mean were classified as high-ability and the others as low-ability.

#### **Research Design**

A  $4 \times 2 \times 2$  factorial design was used with treatments (4 conditions) and abilities in English comprehension (high vs low) as between-subjects factors, and the vocabulary posttests (immediate and delayed) as the within-subject factor.

The vocabulary pretest, immediate vocabulary posttest, and delayed vocabulary posttest scores were the dependent measures.

#### **The Material**

The same 18 target words as in Experiment 1 were used. The same keywords and pictures as in Experiment 1 were used except for two of them which had English keywords in Experiment 1, and which yielded poor results. In Experiment 2, the keywords for these two target words were

Cantonese, and the pictures were changed accordingly. The same procedures were followed as in Experiment 1.

#### Instruments and Procedure

The instruments and procedures were similar to those in Experiment 1 except that the ten weeks' delayed vocabulary posttest was omitted.

#### Results

Only the subjects who scored zero in the vocabulary pretest were included in the following statistical analyses.

#### *The Comprehension Subtest*

The overall mean score of the comprehension subtest was 4.46. The subjects who scored 4.46 or above were classified as high-ability subjects and the others were classified as low-ability subjects.

#### *The Vocabulary Posttests*

Reliability of the immediate vocabulary posttest,  $\alpha = .86$ , and reliability of the delayed vocabulary posttest,  $\alpha = .78$ .

ANOVA was conducted involving a repeated measures design. The results are listed in Table 3.

**Table 3 Results of ANOVA with Repeated Measures in Experiment 2**  
(a) Between-subjects effects

| SOURCE           | SUM OF SQUARES | df  | MEAN SQUARES | F       | P   |
|------------------|----------------|-----|--------------|---------|-----|
| Within cells     | 5101.23        | 232 | 21.99        |         |     |
| Constant         | 42109.00       | 1   | 42109.00     | 1915.08 | .00 |
| Method           | 124.85         | 3   | 41.62        | 1.89    | .13 |
| Ability          | 313.15         | 1   | 313.15       | 14.24   | .00 |
| Method X ability | 225.00         | 3   | 75.00        | 3.41    | .02 |

#### (b) Within-subject effects

| SOURCE                    | SUM OF SQUARES | df  | MEAN SQUARES | F      | P   |
|---------------------------|----------------|-----|--------------|--------|-----|
| Within cells              | 1717.39        | 232 | 7.40         |        |     |
| Retest                    | 4984.79        | 1   | 4984.79      | 673.39 | .00 |
| Method X retest           | 324.35         | 3   | 108.12       | 14.61  | .00 |
| Ability X retest          | 10.03          | 1   | 10.03        | 1.35   | .25 |
| Method X ability X retest | 20.16          | 3   | 6.72         | .91    | .44 |

For the between-subjects effects, the main effect of method was found to be nonsignificant at .05 level; but both the main effect of ability and the effect of interaction between method and ability were found to be significant at .05 level. This indicated that different methods of instruction did not differ in overall facilitative effects on long-term

retention of vocabulary meaning when the subjects of both ability levels were pooled, but subjects of different abilities scored differently in vocabulary recall tests, and different methods differed in strengths of facilitation at different levels of ability.

For the within-subject effects, results of the repeated measures ANOVA yielded significant main

effect of the repeated posttests; and significant effect of interaction between method and the repeated posttests was also found at .05 level. These results indicated significant difference between the mean scores in the immediate vocabulary posttest and the delayed vocabulary posttest and also significant difference in the effect of instructions on long-term retention of target words with respect to initial acquisition of their meanings.

Neither the two-way interaction effect between

ability and repeated testing nor the three-way interaction effect among method, ability, and repeated testing was found statistically significant at .05 level.

The means and standard deviations at two levels of ability as functions of the four conditions in the immediate vocabulary posttest and the delayed vocabulary posttest are given respectively in Table 4. Figure 6 illustrates the retention rate of each of the four conditions at two levels of ability.

**Table 4 Mean Vocabulary-meaning Recall Scores at Two Levels of Ability as a Function of Condition in Two Vocabulary Posttests in Experiment 2**

| Ability                            |           | (a)<br>CONTEXT<br>METHOD | (b)<br>KEYWORD<br>METHOD | (c)<br>CONTEXT +<br>KEYWORD<br>METHOD | (d)<br>ROTE-<br>MEMORY<br>CONTROL | <i>F</i> | <i>Most<br/>effective<br/>(By Tukey)</i> |
|------------------------------------|-----------|--------------------------|--------------------------|---------------------------------------|-----------------------------------|----------|--|
| <b>IMMEDIATE POSTTEST</b>          |           |                          |                          |                                       |                                   |          |  |
| High                               | <i>n</i>  | 20                       | 23                       | 31                                    | 35                                | 3.68     | n.s.                                     |
|                                    | <i>M</i>  | 15.10                    | 12.48                    | 12.61                                 | 15.23                             |          |  |
|                                    | <i>SD</i> | 4.77                     | 4.39                     | 4.42                                  | 3.21                              |          |  |
| Low                                | <i>n</i>  | 37                       | 38                       | 27                                    | 29                                | 1.79     | n.s.                                     |
|                                    | <i>M</i>  | 12.08                    | 10.87                    | 13.30                                 | 11.38                             |          |  |
|                                    | <i>SD</i> | 4.37                     | 4.65                     | 3.95                                  | 4.22                              |          |  |
| ALL Ss                             | <i>n</i>  | 57                       | 61                       | 58                                    | 64                                | 2.47     | n.s.                                     |
|                                    | <i>M</i>  | 13.14                    | 11.48                    | 12.93                                 | 13.48                             |          |  |
|                                    | <i>SD</i> | 4.70                     | 4.58                     | 4.19                                  | 4.15                              |          |  |
| <b>TWO WEEKS' DELAYED POSTTEST</b> |           |                          |                          |                                       |                                   |          |  |
| High                               | <i>n</i>  | 20                       | 23                       | 31                                    | 35                                | 2.42     | n.s.                                     |
|                                    | <i>M</i>  | 5.40                     | 7.35                     | 7.97                                  | 7.14                              |          |  |
|                                    | <i>SD</i> | 2.54                     | 3.93                     | 3.05                                  | 3.65                              |          |  |
| Low                                | <i>n</i>  | 37                       | 38                       | 27                                    | 29                                | 9.18     | c > a,d<br>b > a                         |
|                                    | <i>M</i>  | 4.11                     | 6.18                     | 8.00                                  | 4.14                              |          |  |
|                                    | <i>SD</i> | 2.62                     | 3.95                     | 4.07                                  | 2.56                              |          |  |
| ALL Ss                             | <i>n</i>  | 57                       | 61                       | 58                                    | 64                                | 10.03    | c > a,d<br>b > a                         |
|                                    | <i>M</i>  | 4.56                     | 6.62                     | 7.98                                  | 5.78                              |          |  |
|                                    | <i>SD</i> | 2.65                     | 3.95                     | 3.53                                  | 5.78                              |          |  |
| <b>RETAINED TWO WEEKS</b>          |           |                          |                          |                                       |                                   |          |  |
| High                               |           | 35.76%                   | 58.89%                   | 63.20%                                | 46.88%                            |          |  |
| Low                                |           | 34.02%                   | 56.85%                   | 60.15%                                | 36.88%                            |          |  |
| ALL Ss                             |           | 34.70%                   | 57.67%                   | 61.72%                                | 42.88%                            |          |  |

> Significantly superior at .05 level

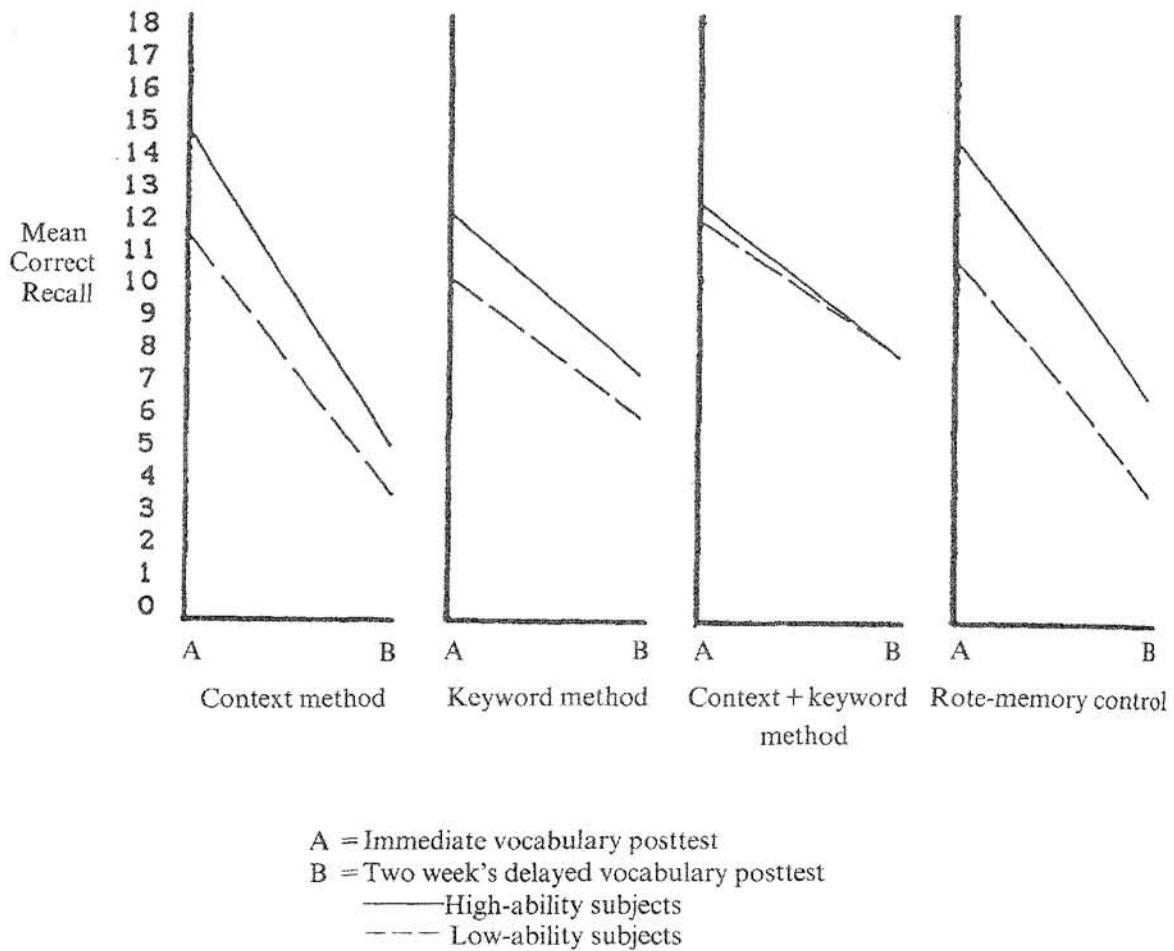


Figure 6. High- and low-ability subjects' mean correct vocabulary-meaning recall scores of four conditions at two vocabulary posttests in Experiment 2

*Errors of Recall*

Willerman and Melvin (1979) speculated that subjects might remember the keywords instead of the vocabulary meanings. Here, the number of errors made (i.e., number of attempts — number of correct responses) were compared among the four conditions at each ability level. One-way ANOVAs revealed that for both high- and low-ability subjects, no statistical difference was found in the number of errors made among the four conditions.

**Discussion**

The main effect of treatment was found only in the low-ability group, not in the high-ability group. Interaction between method of instruction and ability was found statistically significant at .05 level.

*The Ability Factor*

The results indicated no superiority of method in immediate recall of vocabulary meaning for

subjects of either high- or low-ability. These results were beyond expectation.

The results in the delayed vocabulary-meaning recall posttest seemed to support Atkinson's (1975) suggestion that high-ability learners are able to use effective strategies on their own in order to aid memory of vocabulary meanings. High-ability subjects did not differ significantly in long-term vocabulary-meaning recall scores whether they be taught with any memory strategy, or no strategy at all.

Perhaps high-ability context subjects had been able to make use of the linguistic context to aid long-term memory. On the other hand, high-ability subjects in the control group could have made use of self-devised mnemonic strategies (Rohwer, Raines, Eoff & Wagner, 1977). Several other studies have also indicated that some control subjects may use self-devised elaborative strategies (Kemler & Jusezyk, 1975; Martin, Boersma & Bulgarella, 1968). Low-ability control subjects might not have been able to use their own mnemonic strategies as effectively as did high-ability subjects.

#### *The Context Method vs the Control*

Neither of the keyword conditions yielded superior scores in either the immediate or delayed vocabulary posttest in high-ability subjects, and the context condition did not yield superior scores as compared to the control condition, either. This result, therefore, failed to find superior effect of the context method to the control, which Gipe and Arnold (1979) found in the high-ability subjects in their experiment. Perhaps the information contained in the sentences had not been particularly helpful for retention by high-ability subjects.

Consistent with the findings in Experiment 1, the context subjects did not perform better than the control subject in the delayed vocabulary posttest. Low-ability subjects of the context condition might not have been able to use the linguistic context to aid memory.

Superiority of the keyword group to the context group and superiority of the context+keyword group to both the context group and the control found in the delayed vocabulary posttest were consistent with the findings in Experiment 1. Low-ability subjects may find memorizing the linguistic context a burden rather than help.

#### *Retention Rate*

The retention rates for high-ability and low-ability context+keyword subjects were much higher than those of respective subjects in the context group and the control group. The retention rates of the keyword subjects were also comparatively high at both ability levels, though lower than those of respective context+keyword subjects.

The results of both experiments did not support McDaniel, Pressley, and Dunay's (1987) findings that the effect of the keyword method may be found in the short term only. In Experiment 2, superiority in recall was found in the two weeks' delayed vocabulary posttest rather than in the immediate posttest.

#### *The Context+keyword Method*

It seems quite reasonable to take Atkinson's (1975) explanation that the keyword method facilitates retrieval from memory through a more direct route from the stimulus of the given word to the desired response of the meaning. Although the context method may facilitate storage of vocabulary meaning through deeper semantic processing (Gipe & Arnold, 1979), after some time, when the learner tries to retrieve it, the linguistic context itself does not provide direct access from the stimulus to the desired response.

However, the keyword method by itself did not seem to be more useful than rote-memory since the former did not yield significantly higher recall scores than the latter either immediately or two weeks after initial acquisition. The combined use of the context and keyword methods seemed to be more useful. It may be argued that subjects taught with the keyword method only were totally unfamiliar with the strategy and the effect of the keyword method might not have reached its optimal level. Turner (1983) found that when subjects were taught vocabulary-learning strategies, they seemed to be so preoccupied with executing the strategies that they failed to learn the actual lexical items. In this study where the subjects were given only two examples before they executed the keyword method, a problem similar to Turner's (1983) might have occurred.

In contrast, the context+keyword subjects had the choice between the use of the familiar context method and the newly introduced keyword method.

It may be assumed that low-ability context+ keyword subjects have benefited from both the advantage of storage enhancement facilitated by the context method and the advantage of retrieval enhancement facilitated by the keyword method.

## Conclusion

General findings of Experiment 2 suggested that the currently used context method might not benefit low-ability learners in vocabulary meaning recall. In contrast, the context+ keyword method was found to be more effective in facilitating long-term retention of English vocabulary meanings for low-ability subjects. However, the keyword method alone did not yield as promising results as those found in previous research literature (e.g. Atkinson & Raugh, 1975; McDaniel & Tillman, 1987; Pressley, Levin & Miller, 1982).

The results of both Experiments 1 and 2 indicated that a combined use of the context method and the keyword method might be useful in helping low-ability Cantonese-speaking learners to remember English vocabulary meanings. There are reasons, therefore, to recommend an incorporation of the keyword method into the classroom as a mnemonic strategy supplementing the context method currently used as a predominant vocabulary-learning method. Furthermore, the combined use of the context method and the keyword method may be expected to be more effective when implemented with low-ability learners. It may also be expected that through the incorporation of the keyword method in vocabulary learning, the difference between high- and low-ability learners may be decreased (McGivern & Levin, 1983).

## References

- Ackerman, B. P. (1986). Developmental differences in the use of conceptual features in retrieving episodic information from memory. *Child Development*, 57, 1109-1122.
- Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, 30, 821-828.
- Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 126-133.
- Battig, W. F., & Bellezza, F.S. (1979). Organization and levels of processing. In C. R. Puff (Ed.), *Memory organization and structure*. New York: Academic Press.
- Beheydt, L. (1987). The semantization of vocabulary in foreign language learning. *System*, 15, 55-67.
- Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research*, 51, 247-275.
- Bower, G. H. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, 9, 529-533.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *Word frequency book*. New York: American Heritage.
- Crist, R. L. & Petrone, J. M. (1977). Learning concepts from contexts and definitions. *Journal of Reading Behavior*, 9, 301-303.
- Delaney, M. D. (1979). *Individual differences and materials differences which may affect keyword method utility*. Paper presented at the Annual Meeting of the American Educational Research Association. CA: San Francisco. (ERIC Document Reproduction Service No. ED 186 441)
- Erkstrand, B. R., Wallace, W. P., & Underwood, B. J. (1966). A frequency theory of verbal-discrimination learning. *Psychological Review*, 73, 566-578.
- Eubanks, A. C. & Ferguson, W. F. (1982). *Learning of nouns under three treatments with students of English as a second language*. Paper presented at the Mid-South Educational Research Association. LA: New Orleans. (ERIC Document Reproduction Service No. ED 236 914).
- Gipe, J. P. (1980). Use of a relevant context helps kids learn new word meanings. *Reading Teacher*, 33, 398-402. (From ERIC Document Reproduction Service Abstract No. EJ 225 171).
- Gipe, J. P., & Arnold, R. D. (1979). Teaching vocabulary through familiar associations and contexts. *Journal of Reading Behavior*, 11, 281-285.
- Gruneberg, M. M., & Sykes, R. N. (1978). Knowledge and memory: The feeling of knowing and reminiscence. In M. M. Gruneberg, & P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 189-196). London: Academic Press.
- Johnson, C. W., Adams, M., & Bruning, R. (1985). Keywords and vocabulary acquisition; Some words of caution about words of assistance. *Educational Communication and Technology Journal*, 33(2), 125-138.
- Jones, B. F., & Hall, J. W. (1982). School applications of the mnemonic keyword method as a study strategy by eighth graders. *Journal of Educational Psychology*, 74, 230-237.
- Kemler, D. G., & Jusezyk, P. W. (1975). A developmental study of facilitation by mnemonic instruction. *Journal of Experimental Child Psychology*, 20, 400-410.

## Teaching English Vocabulary with Keyword Method

- Levin, J. R. (1976). What have we learned about maximizing what children learn? In J. R. Levin & V. L. Allen (Eds.), *Cognitive learning in children* (pp. 105-134). New York: Academic Press.
- Levin, J. R. (1986). Four cognitive principles of learning-strategy instruction. *Educational Psychologist*, 21, 3-17.
- Levin, J. R., McCormick, C.B., Miller, G. E., Berry, J. K., & Pressley, M. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies for children. *American Educational Research Journal*, 19, 121-136.
- Martin, C. J., Boersma, F. J., & Bulgarella, R. (1968). Verbalization of associative strategies in normal and retarded children. *Journal of General Psychology*, 78, 209-218.
- McDaniel, M.A., & Pressley, M. (1984). Putting the keyword method in context. *Journal of Educational Psychology*, 76, 598-609.
- McDaniel, M. A., Pressley, M., & Dunay, P. K. (1987). Long-term retention of vocabulary after keyword and context learning. *Journal of Educational Psychology*, 79, 87-89.
- McDaniel, M. A., & Tillman, V.P. (1987). Discovering a meaning versus applying the keyword method: Effects on recall. *Contemporary Educational Psychology*, 12, 156-175.
- McGivern, J. E., & Levin, J. R. (1983). The keyword method and children's vocabulary learning: An interaction with vocabulary knowledge. *Contemporary Educational Psychology*, 8, 46-54.
- Merry, R. (1980). The keyword method and children's vocabulary learning in the classroom. *British Journal of Educational Psychology*, 50, 123-136.
- Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition*, 7, 273-282.
- Pressley, M. (1977). Children's use of the keyword method to learn simple Spanish vocabulary words. *Journal of Educational Psychology*, 69, 465-472.
- Pressley, M., & Dennis-Rounds, J. (1980). Transfer of a mnemonic keyword strategy at two age levels. *Journal of Educational Psychology*, 72, 575-582.
- Pressley, M., & Levin, J. R. (1978). Developmental constraints associated with children's use of the keyword method of foreign language vocabulary learning. *Journal of Experimental Child Psychology*, 26, 359-372.
- Pressley, M., Levin, J. R., Kuiper, N. A., Bryant, S. L., & Michener, S. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies: Additional comparisons. *Journal of Educational Psychology*, 74, 693-707.
- Pressley, M., Levin, J.R., & Miller, G.E. (1981). How does the keyword method affect vocabulary comprehension and usage? *Reading Research Quarterly*, 16, 213-226.
- Pressley, M., Levin, J. R., & Miller, G. E. (1982). The keyword method compared to alternative vocabulary-learning strategies. *Contemporary Educational Psychology*, 7, 50-60.
- Pressley, M., Levin, J. R., Nakamura, G. V., Hope, D. J., Bispo, J. G., & Toye, A. R. (1980). The keyword method of foreign vocabulary learning: An investigation of its generalizability. *Journal of Applied Psychology*, 65, 635-642.
- Raugh, M. R., & Atkinson, R. C. (1975). A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, 67, 1-16.
- Restle, F. (1964). Sources of difficulty in learning paired-associates. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford, CA: Stanford University Press.
- Rohwer, W. D. Jr., Raines, J. M., Eoff, J., & Wagner, M. (1977). The development of elaborative propensity in adolescence. *Journal of Experimental Child Psychology*, 23, 472-492.
- Shriberg, L. K., Levin, J. R., McCormick, C. B., & Pressley, M. (1982). Learning about "famous" people via the keyword method. *Journal of Educational Psychology*, 74, 238-247.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56, 72-110.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University.
- Turner, G. (1983). Teaching French vocabulary: A training study. *Educational Review*, 35, 81-88.
- West, R. F., Stanovich, K. E., Feeman, D. J., & Cunningham, A. E. (1983). The effect of sentence context on word recognition in second- and sixth-grade children. *Reading Research Quarterly*, 19(1), 6-15.
- Willerman, B., & Melvin, B. (1979). Reservations about the keyword mnemonic. *Canadian Modern Language Review*, 35, 443-453.
- Wittrock, M. C., Mark, C., & Doctorow, M. (1975). Reading as a generative process. *Journal of Educational Psychology*, 67, 484-489.
- Wollen, K. A., & Lowry, D. H. (1971). Effects of imagery on paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, 10, 276-284.

---

Mr. See Shing YEUNG was a Master of Arts in Education graduate of the Faculty of Education, The Chinese University of Hong Kong.

Dr. Rex M. HEYWORTH is Lecturer of the Faculty of Education, The Chinese University of Hong Kong.