This is to certify that the

dissertation entitled

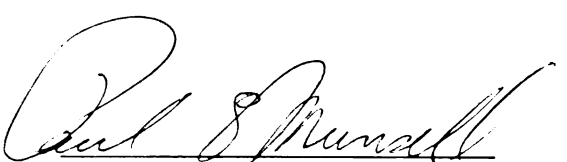
The Learning of English Words by Japanese Seventh-Grade Students by Means of the Keyword Method

presented by

Mamoru Kinjo

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in English



Major professor

Date May 11, 1994

THE LEARNING OF ENGLISH WORDS BY JAPANESE SEVENTH-GRADE STUDENTS BY MEANS OF THE KEYWORD METHOD

By

Mamoru Kinjo

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of English

1994

Dissertation Advisor: Dr. Paul E. Munsell

ABSTRACT

THE LEARNING OF ENGLISH WORDS BY JAPANESE SEVENTH-GRADE STUDENTS BY MEANS OF THE KEYWORD METHOD

By

Mamoru Kinjo

The keyword method is advocated as a way of acquiring vocabulary items of a foreign language (Atkinson, 1975). Research on the effects of the keyword method has shown mixed results. Those in the field of cognitive psychology have overwhelmingly reported success from their laboratories, whereas those in the field of TESOL have mostly reported failure, when the keyword method is used with actual students studying a foreign language as a subject (Fuentes, 1976; Liu, 1978; Willerman & Melvin, 1979). The mixed results seem to want further empirical data. Therefore, the following study was designed to provide such data.

The subjects of the study were 390 seventh-grade Japanese students (12 classes) studying English in Okinawa, Japan. All of the students studied two sets of 15 English words--one set with the keyword method and the other with the rote method. All 390 students took a posttest immediately after their study. Then 187 students (six classes) took the same test after a month, and the remaining 203 students (six classes) took the test after two months.

The students' scores were subjected first to the Wilcoxon matched-pairs signed-ranks test and then to the t-test. The results revealed that at the time of the immediate posttest, the mean score of the words learned by the keyword method was superior to that of the words learned by the traditional rote method.

The follow-up tests showed that all 12 classes had higher means on the words learned with keywords.

Among three groups of students with perfect scores on either set of 15 words on the posttest (100% by rote, the keyword method, and both methods), mean retention was superior for the keyword set on the follow-up tests, without exception in each group.

On five words designated as "hard," the rote method showed a slight advantage at posttest, which disappeared by the first follow-up test. A greater difference was observed at the time of the second follow-up test in favor of the keyword method.

DEDICATION

This dissertation is dedicated to the following people, who have played an important part in preparing me for my studies at Michigan State University. First and foremost my mother,

Sumako Kinjo,

and my two older brothers,

Iwao and Kaoru Kinjo.

And the following educators:

Ritsko Shinzato (at Miyamori Elementary School)

Shunichi Tamamori (at Ishikawa Junior High School)

Dr. Mary E. Green

Dr. Robert E. Shafer

Dr. Daniel Quirk

(all at Arizona State University).

ACKNOWLEDGMENTS

I wish to express my deep appreciation to Dr. Paul Munsell, my major professor, for his guidance, encouragement, and friendship throughout the course of this work spanning almost a full ten years. Without his decade-long, patient assistance, this work could never have been done.

I am equally grateful to my committee members, Dr. Susan Gass, Dr. Teresa Tavormina, and Dr. Marilyn Wilson, for their invaluable contributions. I would also like to thank the outside reader, Dr. John Eulenberg, for his helpful suggestions.

I also would like to express my appreciation to the junior high school teachers and students in Okinawa who participated in the study.

Lorraine Hart, graduate secretary in the English Department, who always helped me in the best possible way.

The staff at the English Language Center, John and all the others, were very kind and helpful whenever I asked for help.

Faouzi Azzouz, graduate student at Michigan State University, offered invaluable time and statistical assistance beyond the call of duty.

Two typists, Cathy Rynbrandt and Susan Cooley Miller, assisted me in ways I cannot recall without gratitude.

Dr. Joseph Cousins at the Office of International Education Exchange, helped me in many ways. His family, Keiko-san, Micah, and Riku, have always given me a warm welcome.

I also thank the Owen Graduate Center and its staff for the kind cooperation and consideration shown to me, which provided me with the best accommodations for the study.

Hiromitsu Ishikawa, who was at the Language Center at Naha, Okinawa, gave me much advice and encouragement in my pursuit of the doctoral degree.

Some international students/friends such as Rajesh Lal and Klengolo Traore, and many American friends, were willing to help me and made my study at Michigan State University pleasant.

I would like to acknowledge with the greatest gratitude the contributions of the three organizations that made my studies in the United States possible.

The United States Civil Administration in the Ryukyus for the four-year full scholarship that sponsored the study at Arizona State University (1968-1972). (The Ryukyus were returned to Japan in 1972.)

The Okinawa Prefectural Government Human Resources Development Foundation for the two-year full scholarship that made possible my graduate study at Michigan State University (1984-1986), and the fellowship given to continue research at MSU (March to June 1991), which led to the acquisition of this degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This chapter states the purpose of this study. It also reviews briefly the vocabulary-acquisition research since the turn of the century to provide a historical context for the present research regarding the keyword method. The chapter also discusses the problems of applying the results of keyword-method experiments to students actually learning a foreign language.

Purpose of the Study

The purpose of this study was to conduct an empirical investigation of certain aspects of the keyword method and, in particular, to see whether the method can enhance the remembering of English vocabulary items. The population used was 390 Japanese seventh graders learning English in Japan. The results were interpreted statistically, and a theoretical account is presented as a possible explanation of the findings.

A Brief History of Vocabulary-Acquisition Research Since the Beginning of the Twentieth Century

One of the historical characteristics of vocabulary acquisition in foreign language teaching during the twentieth century is that it has received widely varying

amounts of attention. In the middle part of the century there was a period of neglect, possibly through the influence of American linguistics represented by as diverse scholars as Fries and Chomsky, whose primary attention centered on syntax.

Reflecting on the neglect of research on vocabulary of the past decades, Carter and McCarthy (1988) observed that "the period 1945-1970 was a limbo for vocabulary as an aspect of language teaching in its own right" (p. 41). They concluded their review of the period by mentioning that "[a]n article such as Higa (1965) is a notable exception . . . in an otherwise unredeemed neglect" (p. 41). Allen (1983) gave one reason for the neglect of vocabulary between 1940 and 1970 as too much emphasis on words in language classrooms in the years before that time. To illustrate these factors, two quotations, one from the 1920s and the other from the 1940s, are provided below.

First, Sapir (1921) warned linguistic students that they "should never make the mistake of identifying a language with its dictionary" (p. 234). Second, Bond, Babcock, Norman, and Andrus (1943) cited 24 categories and 1,145 entries found in Dales's tabulation of studies on vocabulary selection through 1938 as a reason to doubt that in vocabulary research "any stone has been left unturned" (p. 150).

The Vocabulary Control Movement

One significant area of great inquiry during this period of vocabulary neglect is now known as "the vocabulary control movement." Two important concepts have emerged from this movement and deserve a short explanation.

1. A minimum basic vocabulary as suggested in the work Basic English by C. K. Ogden (1930) and I. A. Richards (1943). The underlying principle for the Basic English list was that complex ideas can be expressed by means of simple words. The authors identified 850 words as those elemental, yet sufficient, vocabulary items that can fully meet the basic communication needs of foreign learners of English. Although Basic English is no longer widely used, a concern for the construction of "core or nuclear Englishes for language learning purposes" (Carter & McCarthy, 1988, p. 6) still is an issue of vital importance in the field of ESL or TESOL today.

2. A limited defining vocabulary recommended by Michael West and Harold Palmer. West (1953) published A General Service List (GSL), which contained 2,000 words selected from a corpus of two to five million words. This list was a product of almost three decades of major work on vocabulary selection for teaching purposes. Carter and McCarthy (1988) maintained, with regard to the GSL, that despite some of the criticisms, it "remains one of the most innovative examples of foreign language pedagogy and lexicometric research in this century" (p. 9).

Fries, Lado, and Chomsky

When, during the 1940s, researchers engaged intensely in the vocabulary-control movement and Bond et al. claimed that vocabulary study had hardly left any stone unturned, a new movement was in embryo. Fries (1945) published his Teaching and Learning English as a Foreign Language, in which he stressed that difficulties in learning a foreign language should not be considered to be limited to

the lexicon. He stressed the importance of mastering the sound system and grammatical structure in learning a new language. However, he did not totally neglect the importance of the role of vocabulary in learning a foreign language. This fact is reflected in his student's work; Lado (1955), a decade after Fries's publication, dealt with the difficulty of vocabulary from the standpoint of contrastive analysis of languages.

The field received enormous support for the emphasis on structure with the birth of transformational grammar 12 years later when Chomsky (1957) published his Syntactic Structures. With subsequent developments in the two fields, linguistics and TESOL both came to concur on the primacy of syntax in language study.

The 1970s

It was not until the 1970s that voices against the long neglect of vocabulary research began to be heard. One of these voices was that of Wilkins (1972), who argued, "Without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 111). The massive investigation by Thorndike (1973) using 100,000 students in three different age groups from 15 different countries proved that the knowledge of vocabulary is the surest index to reading comprehension.

It was also in the 1970s that people started "viewing vocabulary learning as a language skill, shifting the responsibility to the learner," rather than uniformly providing a list of words to learn based on frequency (Carter & McCarthy, 1988, p. 42). Even so, Meara (1980), reviewing the literature concerning vocabulary

acquisition, could still title his article "Vocabulary Acquisition: A Neglected Aspect of Language Learning."

The 1980s

In the 1980s, topics in vocabulary acquisition research developed in increasingly divergent directions. Two important publications appeared in 1980, according to Carter and McCarthy (1988). These are Guidelines for Vocabulary Teaching by the Regional English Language Center (RELC) in Singapore and Meara's survey article on vocabulary acquisition, which was already mentioned as part of the literature of the 1970s. As for the RELC publication, a list in one of the articles contains eight "C's" and a "G" of vocabulary teaching. They are quoted here to show the varieties of topics discussed:

> (1) collocation, (2) clines (e.g., scales such as cold/warm/hot expressed diagrammatically), (3) clusters (something akin to the Hallidayan notion of Sets–see Chapter 2), (4) cloze procedures, which reinforce cluster and collocation, (5) context (using features within the text such as definition, word analysis, inference), (6) consultation (using dictionaries and thesauri), (7) cards: students should keep a card index of new vocabulary, (8) creativity: students should be given free rein to describe pictures, etc. and (9) guessing: students should learn this skill.

Two annotated bibliographies were published by Meara. The first volume came out in 1983, covering the years from 1960 to 1980; the second volume was published in 1987, covering the period from 1980 to 1985.

Words were studied increasingly in their collocational possibilities, as seen in the discussion of "prefabricated speech" (Nattinger, 1988), of "composites" (Cowie, 1988), of "institutionalized clauses" (Peters, 1983), and of "sets" (Carter &

McCarthy, 1988). The period illustrated also attempts to integrate important findings from other fields, such as psycholinguistics, as evidenced by the studies of Meara (1984), the Birkbeck Vocabulary Project headed by Meara (1982), Channel (1988), and from the field of linguistics (Carter & McCarthy, 1988).

Renewed interest in traditional word-frequency counts was seen in the 1980s in the compilation of the "University Word List" of about 750 words (Xue & Nation, 1984). Nation (1990) claimed that mastery of the words in their list, together with the 2,000 headwords in the General Service List, should "provide around 95 percent coverage of academic texts" (p. 189).

Global analyses of words in their multiword contexts led some researchers to suggest that a clear distinction between vocabulary and syntax could sometimes collapse (Ard & Gass, 1987; Peters, 1983). Publications by Carter (1987), Carter and McCarthy (1988), and Nation (1990) gave added strength to vocabulary teaching and research in TESOL. By the end of the 1980s, Gass (1989) could say that "it is no longer possible to say that the lexicon is the 'neglected component' of second language acquisition research," but, observing the divergent directions the research had taken, she summarized the state of the art in the 1980s in a single phrase: "lack of focus."

Three Categories Given by Carter and McCarthy (1988)

Although there seems to be a lack of focus in the field of vocabulary acquisition, Carter and McCarthy (1988) in Vocabulary and Language Teaching

mentioned three topics in which lexical-acquisition researchers have taken an active interest. These are (a) memory, (b) word difficulty, and (c) interlanguage. In a passage about word difficulty, Carter and McCarthy mentioned that the difficulty of a word may result from "whether it lends itself to key-word teaching techniques." They further stated:

> One important element in learning new words is the degree of effective *formal* linking learners can establish between a word in the target language and a cognate word in the mother tongue (and because issues of memorization and word difficulty are closely connected). According to research by Craik and Lockhart (1972), oral repetition is not necessarily an effective way of assimilating new words; recalling the *form* of a word is found to be more productive.

The Keyword Method and the Problems of Its Applications to Classroom Settings

The "key-word teaching techniques" that Carter and McCarthy referred to, also called the keyword method, is the subject of this research. The problems of the keyword method are discussed in the following paragraphs.

The keyword method was formally introduced by Atkinson (1975) as a mnemonic technique to enhance the memory of a foreign word. Since the introduction of the keyword method by Atkinson, there has been much research concerning the keyword method among cognitive psychologists within the confines of laboratory settings or scientifically controlled settings. These investigations generally have suggested the effectiveness of the keyword method. In addition to these reports of the experiments, however, there have been some experimental reports with actual students learning a foreign language as subjects. These results

have not turned out to be replications of the success most often reported by keyword advocates. For example, one of the first studies reporting the results of an experiment conducted outside a laboratory setting--that is, in an actual classroom setting with students learning a foreign language--was that of Fuentes (1976). His conclusion was that the keyword method had no effect on increasing the learning of foreign words, as compared with the traditional rote method. Therefore, Fuentes argued, the results of the laboratory experiments do not necessarily transfer directly to classrooms of actual foreign language learners.

Several other experiments have been carried out with actual students of a foreign language as subjects, but the reports have shown mixed success and failure. Consequently, the effects of the keyword method with students learning a foreign language in actual classroom settings have been far from fully explored. Therefore, the present study was designed to provide empirical data about the effect of the keyword method with Japanese seventh-grade students learning English as the study population. Further discussion of the background of the keyword method and specific issues to be studied are presented in the next chapter.

CHAPTER II

REVIEW OF LITERATURE ON THE KEYWORD METHOD

The review of the literature on mnemonics and the keyword method is divided into four sections: the literature before Atkinson's work, Atkinson (1975) and the introduction of the keyword method, the literature of the post-Atkinson period, and that of the applications of mnemonics and the keyword method in TESOL.

Review of the Literature Before Atkinson's Work

Although the present-day keyword method used for acquiring vocabulary in foreign languages was originally promoted by Atkinson (1975), the basic principles of mnemonics and the keyword method date back to the heyday of the ancient Greek orators. In brief, the keyword method uses images as an aid for memory of a vocabulary item; this technique of creating mental images as an aid for memory was widely known to the ancient Greeks. It is said that Simonides of Ceos invented the classic method of Loci, and it was handed down to the Romans. One of the extant discourses on the subject of Loci from the ancient days of Rome is that by Cicero in his *De Oratore*, which dealt with the art of memory as one of the five parts of rhetoric. There have been many treatments of the method by other writers, among which Quintilian's is said to be the clearest, and *Ad Herennium* "the only

complete source, for the classical art of memory both in the Greek and in the Latin world" (Yates, 1966, pp. 1-5).

Furthermore, although the keyword method makes even abstract words in a foreign language concrete and visualizable to assist memory, this distinction between abstract and concrete types associated with Paivio (1969) had been discussed by Cicero, Augustine, and Aquinas. Even the concept of levels of processing (Craik & Lockheart, 1972), which is often given as one explanation of the efficacy of the keyword method, was discussed by Francis Bacon (Gorfein & Hoffman, 1987, p. 53).

The Role of Pictures

On the other hand, the actual use of pictures and images in teaching languages is not as ancient as mnemonics, according to the author of *25 Centuries of Language Teaching*: "Teaching through pictures was done by the ancient Chinese. In the West it is not found consistently before the Renaissance, though there is some evidence of it during the Middle Ages" (Kelly, 1969, p. 15).

The *Orbis sensualium pictus* by Comenius (1654) is said to be "the first fully thought-out scheme of teaching vocabulary with pictures" (Kelly, 1969, p. 17). Comenius recommended five steps in using the book:

1. The pupils were to familiarize themselves with the book itself.

2. They were to make sure they knew the vernacular names of everything depicted in the book.

3. If possible, the teacher was to show them the real thing.

4. The pupils were to copy the illustrations.

5. Finally, they were to color in their own copies and even the etchings in the book.

These steps illustrate Comenius's idea that language learning is an active process employing as many senses as possible. In other words, pictures held the possibility of expanding the context in the course of language learning and teaching. However, "two centuries of almost total neglect of pictures followed Comenius" (Kelly, 1969, p. 18). Then there was a renewed interest in his works at the end of the nineteenth century, which produced a reissue of his writings. However, the effect was negligible, for his work was regarded more as an inspiration for the teacher than as materials for students.

Kelly (1969) also pointed out that, during World War I, the United States Army received many conscripts from minority groups, who did not speak English and were ignorant of many things concerning European countries. Language teachers used photographs cut primarily from the *National Geographic Magazine* to teach them about languages and real things. This was a notable development as the teaching put into practice the basic approach of Comenius to language teaching, taking advantage of pictures as an aid in language teaching.

Recent History

As for the historical origin of the keyword method itself, Desrochers and Begg (1987) placed it, based on their study, "between the epoch of the first celebrated modern mnemonist, Raimond Lulle (1236-1315), and the end of the 19th century"

(p. 56). They mentioned two related figures from the last century, Fenaigle and the Reverend J. H. Bacon. "Fenaigle's (1813) *New Art of Memory*, probably the most complete account of previous mnemonic systems," devoted a whole chapter (Chapter V) to the learning of foreign languages, emphasizing the utility of cognates for memorizing new foreign words. Here, one can see that the value of associating an unfamiliar word form with a familiar one was recognized. As for the Reverend J. H. Bacon, then at the St. Bees College of Cumberland, England, he is known to have advocated what he called the "phonetic link method" for the study of French vocabulary.

Paivio and Desrochers (1981) further mentioned Loisette (1895) and Sayer (1877) as those working along a similar line. They gave an example Sayer used to teach the Latin word *rana* ("frog"), using what he called "the method of assimilation," as seen in the following sentence: "A frog caught in the rain" (p. 780).

Desrochers and Begg (1987) summarized their study of the history of the keyword method up to the end of the nineteenth century by writing that "what remains uncertain is when and by whom the second ingredient of the keyword method, the semantic link between the two concepts, was introduced."

In studies of memory in the early twentieth century, occasional references were found to the utility of the keyword method. Desrochers and Begg (1987) mentioned Estabrooks (1927) as one, and Paivio and Desrochers (1982) referred to Julliot (1919).

In the late 1960s, a series of experiments was conducted to help retarded children. According to Martin (1978), who gave a summary of the results, mnemonic verbal mediators were found to help the children's recall.

Paivio and Desrochers (1981) cited, among the advocates of something similar to the keyword method in the early 1970s, Furst (1972), who labeled his method the "linking word" method, and Lorayne (1974), who named his the "substitute word" method. About the latter author Paivio and Desrocher wrote: "Lorayne proposed specifically that foreign words be learned by associating them with picturable English words that have phonetic resemblance to the foreign word."

It was not until the 1970s, however, that serious empirical investigations on mnemonics were conducted by psychologists. Otto, Butler, Blake, and Ball's (1973) report on the effectiveness of the keyword method was the first of such experimental assessments (Desrochers & Begg, 1987, p. 58).

Atkinson (1975) and the Introduction of the Keyword Method

"The publication of Atkinson's studies," according to Pressley, Levin, and Delaney (1982), "set the stage for all subsequent keyword investigations" (p. 64). Accordingly, one of Atkinson's (1975) articles, "Mnemotechnics in Second Language Learning," is reviewed in detail here.

In the introduction to his article, Atkinson explained that he was involved in pedagogical application of psychological theory to optimize computer-assisted instruction. One of the experimental projects he was engaged in happened to focus

on "the restricted but well-defined problem of optimizing the teaching of foreign language vocabulary" (p. 821). A series of vocabulary learning experiments he conducted with students at Stanford University displayed large between-subject differences, even among this apparently select sample of American university students. Atkinson's search for strategies to help students led to a series of studies on mnemonic techniques. In concluding the introduction, Atkinson mentioned the contribution of his colleague, Michael Raugh, who had been intrigued by the power of mnemonics and convinced him that "this line of research was worth pursuing" (p. 821).

Atkinson then showed how the keyword method works, first explaining the term "keyword" as follows: "By a keyword we mean an English word that sounds like some part of the foreign word." Then two links follow, integrating acoustic and imagery components as quoted here: "The spoken foreign word is linked to the keyword by a similarity in sound (what I call the *acoustic link*) and in turn the keyword is linked to the English translation by a mental image (what I call the *imagery link*)." Atkinson's actual examples of the keyword method then follow. One of them is the following: "The Spanish word for duck is *pato* (pronounced something like "pot-o"). Using the English word *pot* as the keyword, one could image a duck hiding under an overturned flower pot with its webbed feet and tufted tail sticking out below." (See Figure 1.)

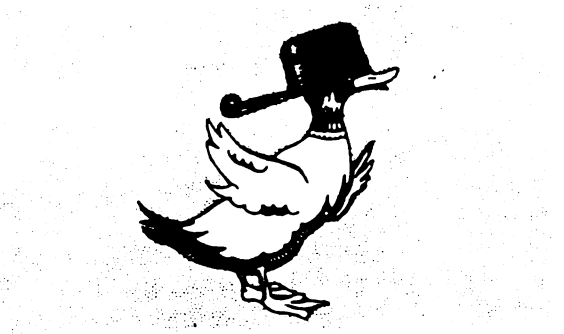


Figure 1: Illustration of how the keyword method works.

Atkinson then reported the results of an experiment teaching 120 Russian words to 52 students at Stanford University. The students learned 40 words a day for three consecutive days, and took a test a day after the last day, and again after six weeks. The results showed that the keyword group recalled 75% of the items, whereas the control group recalled 46% of the items on the first test. On the second test after six weeks, the keyword-method group recalled 43% of the items, and the control group recalled 28% of the items.

Atkinson then discussed the following five questions on the basis of the data available from pilot experiments and other research in progress: (1) Should the experimenter supply the keyword, or can the subject generate his own more effectively? (2) Does supplying the imagery link for the subject facilitate learning? (3) When a foreign word is presented, does the time to retrieve its English translation depend on the method of learning? (4) Are the imagery instructions critical in the keyword method, or can the subject do equally well when told to associate the

keyword and English translation by generating a meaningful sentence connecting the two words? (5) How useful is the keyword method if the subject is asked to retrieve the foreign word when given its English translation?

In the next section, Atkinson deals with the effectiveness of a keyword. Atkinson recommended a committee approach in generation of the keyword, as "[e]xperience indicates that individual experimenters can come up with some pretty bizarre keywords that work for them but for no one else." He then points out that the words students learn through keywords will, when mastered thoroughly, form a direct link between the foreign words and the native translations. Therefore, the keyword method should be as effective as any other ways of learning foreign vocabulary.

In the concluding section, Atkinson stated, "If our instructional applications prove as successful as the experimental work," then the keyword method deserves a role in language-learning curricula.

In the following pages is the review of literature on the keyword method so as to show the general trend of the experiments conducted from 1975 to the present. They are arranged primarily chronologically to highlight the developmental patterns.

Review of the Literature of the Post-Atkinson Period

Since the introduction of the keyword method by Atkinson (1975), there has been active research in this area of teaching, or remembering vocabulary items through the method. Some of the interesting topics of investigation are reviewed in the following paragraphs under these headings:

1. Application of the keyword method to actual classroom situations.
2. Developmental constraints on the usage of the keyword method.
3. The contextual approaches compared with the keyword method.
4. Effects of elaboration on memory.
5. Metamemorial study.
6. Application of the keyword method to teaching content materials in other fields.

These divisions and the order of the topics were created partially on the basis of chronological order of research developments, but studies reviewed in them do not always conform to the chronological restriction because of the topical divisions: Studies done many years apart may appear in a group because of their thematic proximity.

Application of the Keyword Method to Actual Classroom Situations

The experiments by Atkinson (1975) were followed by those of Fuentes (1976); Liu (1978); Levin, Pressley, McCormick, Miller, and Shriberg (1979); Merry (1980); and Pressley, Levin, and Delaney (1982), who investigated the applicability of the keyword method to a real classroom situation with mixed results.

Fuentes (1976) compared the keyword method with a rote-rehearsal group. One hundred thirty-five students taking second-year high school Spanish were asked to learn five new words given ten minutes on the first four days of the school week. He found no difference in performance between the two groups.

Liu (1978) found that the keyword method of teaching vocabulary to first-year high school students in Taiwan proved to be ineffective. However, Pressley, Levin, and Delaney (1982) criticized this report because it provided insufficient details concerning procedures and the results.

Willerman and Melvin (1979) taught 40 French words in a beginner French class. But their results did not indicate the significant differences keyword-method advocates would expect.

Levin et al. (1979) conducted a series of experiments to investigate the classroom application of the keyword method. They confirmed the claim of Fuentes (1976) with high school students, but succeeded in producing striking keyword effects with elementary-school-age children.

Merry (1980) reported five different ways of teaching French words in a real classroom setting. The findings showed that the keyword method showed superiority over all the other methods for both tests, one immediately after the learning session and the other a week later. However, there was no difference between keywords used with pictures and keywords used with verbal images.

Pressley, Levin, and Delaney (1982) provided a comprehensive review of the keyword-method literature, covering more than 200 studies with a view to potential applications of the keyword method to teaching a variety of curricular contents. One of the sections included a review of group-administered keyword studies as follows.

The results of the experiments reviewed showed a mixture of success and failure. Fuentes (1976) and Willerman and Melvin (1979) were criticized; Merry

(1980) was cited as a report in favor of the keyword method. The authors reported "additional failures of the keyword method among high school students according to a self-paced, group-administered format" in their laboratories.

Two general observations were made from the analysis of these mixed results:

1. Based on their findings of the failure with high school students, the authors suggested that less sophisticated learners may benefit more from the keyword method, such as students in elementary schools or junior high schools.

2. The overall results of the application of the keyword method seem to show that laboratory experiments cannot necessarily be directly applied to real classroom settings with actual students as subjects of investigation.

Developmental Constraints on the Usage of the Keyword Method

Various stages in the cognitive development of children interact with the results of the keyword method. Some of the results of studies on this topic are reviewed below. One of the earliest studies of the keyword method with children was conducted by Pressley (1977). His subjects were second- and fifth-grade students. Students in both grades benefited from using the keyword method.

Pressley and Levin (1977) examined the differences between fifth graders and ninth graders on the basis of Rohwer's hypothesis concerning adolescents' change in strategy in learning noun pairs. A majority of the fifth graders reported that they simply rehearsed, whereas few ninth graders relied exclusively on

rehearsal. Pressley and Levin (1978) further reported that their sixth-grade subjects, when provided with the keyword, could generate their own imagery links, but their second-grade subjects could not create images of their own.

A successful way to help second-grade students was reported by Pressley, Levin, and McCormick (1980): they used a sentence variation of the keyword method. In their experiments, second-grade subjects were asked to make up a sentence, integrating in their mind the keyword and a meaning in order to remember the meaning of the Spanish words. When they were asked to remember the sentences constructed, and allowed to utter their created sentences, they showed learning gains of 180%. The authors noted that these second graders benefited from self-generated sentences. In the pilot research the subjects did not benefit from experimenter-provided sentences in the pilot research.

One of the largest learning gains ever reported was a case of facilitation ranging from 200% to 1270% over no-keyword subjects as children tried to acquire foreign verbs (Miller, Levin, & Pressley, 1980).

Spontaneous transfer of the keyword method to other tasks was observed among 18-year-old subjects but not among 12-year-old students (Pressley & Dennis-Rounds, 1980). However, if 12 year olds were told to apply the mnemonic aid they used in a study task to another study task, they could do so.

The youngest subjects studied were children three to six years of age who were required to learn ten new words. The three year olds with the keyword method scored a mean of 5.5, and the control group 0.6. In this experiment, "the lowest

scoring keyword subject outperformed the highest scoring control subject" (Pressley, Samuel, Harshy, Bishop, & Dickenson, 1981).

On the basis of past experiments, Pressley, Levin, and McCormick (1980) suggested three ways that were successful in helping young children learning vocabulary:

1. The experimenter provides interactive pictures.

2. The experimenter provides separate pictures of the keyword and the foreign words and lets the student link (interact) them.

3. The experimenter allows the students to use a sentence variation (student generated) of the keyword.

The Contextual Approaches Compared With the Keyword Method

Many experiments have compared the keyword method and the contextual approaches for their effectiveness as vocabulary-remembering strategies. However, these experimental results have been reported mainly by keyword-method advocates, and the results have tended to report advantages for the keyword method.

Levin et al. (1982) reported 50% more retention of definitions by fourth-grade subjects with the keyword method as compared to subjects with the context approach. In one experiment (Levin et al., 1984), the contextual support for the word "angler" (a person fond of fishing) was the following passage: "The *angler* carried

a lot of things down to the stream. He carried a net and a tackle box, as well as his fishing pole. He hoped to catch a lot of fish that afternoon."

In this instance, the keyword method had an illustration in which two characters had a dialogue relating the keyword to the definition of the vocabulary word. The major figure of the two was an angel who was fishing, in order to provide an acoustic cue for the word "angler."

With college students, tasks of sentence judgment concerning correct usage of the acquired vocabulary items, and sentence creations using them, are tested. The results of both tasks showed the superiority of the keyword method (Pressley, Levin, & Delaney, 1982).

Multiple contextual exposures such as providing synonyms, sentence elaborations, illustrations, and reading of sentences were provided in an experiment to see their effects on vocabulary remembering. This treatment proved inferior to the keyword method: a 62% retention by the keyword subjects and 28% by the context subjects (Pressley, Levin, Kuiper, Bryant, & Michener, 1982).

McDaniel and Pressley (1984) reported that even sentence constructions of both groups of subjects at a later date reflected the initial superiority of the vocabulary acquisition by the keyword method. They reasoned that contexts facilitate meaning discovery, but the discovery of the meaning of a word by itself does not ensure the subsequent retention of the vocabulary item. In other words, inferring the **meaning** of a word is different from remembering the **sound** of a word. The following experiment seems to show some limitations of the keyword method.

McDaniel and Tillman (1985) reported that the keyword subjects' superiority disappeared if students were simply asked to construct synonyms from elaborated definitions. The authors concluded that the keyword method's superiority, relative to the context method, lies in subjects' acquisition of meaning links associating vocabulary items and their definitions.

Reviewing the issues of contextual approaches to vocabulary learning versus the keyword method, therefore, Pressley, Levin, and McDaniel (1987) emphasized the importance of the distinction between getting the meaning from context, on the one hand, and remembering the word form and meaning, on the other. The keyword method enhances the ability to go from meaning to form, or from form to meaning. They thus asserted that "effective vocabulary-inferring processes and effective vocabulary remembering strategies are complementary vocabulary acquisition components, with one's strength being the other's weakness, and vice versa."

Effects of Elaboration on Memory

Some empirical findings have shown that an item of information is stored better in memory as cognitive activities relating to the item increase, and that ambient data encoded at the time of acquisition of an item of information in memory are strong cues for retrieval. The former is known as the levels (or depth) of processing theory (Craik & Lockhart, 1972), and the latter, the encoding specificity principle (Tulving & Thompson, 1973). These findings led researchers to investigate possibilities of introducing some extra materials into learning processes in the form of elaboration for the sake of better acquisition and retrieval of information.

Consequently, some keyword experts have taken an interest in the effect of elaboration. Some of these are reviewed in the following paragraphs.

Shriberg, Levin, McCormick, and Pressley (1982) reported that transformational imagery instruction is much superior to representational imagery instruction. Representational imagery simply represents facts as they are, whereas transformational imagery is so constructed as to enhance the memory of the given facts. For example, the Great Lakes may be represented with the five lakes of their exact shapes with reduction in sizes with their respective names by representational imagery. On the other hand, transformational imagery may have a lake with some villas around it, or on it with the legend: "My homes on the lake." The word "HOMES" functions as the first-letter mnemonics for H--Huron, O--Ontario, M--Michigan, and so on. Or the name of Lincoln's assassin will be remembered by a picture of Lincoln shot by a man standing in a booth, instead of a mere theatrical background as a representational picture might have it.

The success of elaboration techniques in enhancing memory has led some scholars to designate the focus of early research on memory strategies as rehearsal and organization, and the new trend as associative elaborations or elaborative associations (Pressley, Borkowski, & O'Sullivan, 1984).

Facilitation of learning by means of extratextual information is studied in the forms of precise elaboration versus imprecise elaboration, and the elaborative interrogation to acquire confusing facts (Pressley, McDaniel, Turnure, Wood, &

Ahmad, 1987; Pressley, Symons, McDaniel, Snyder, & Turnure, 1988). For example, consider the following three sentences.

The hungry man got into the car.

The strong man helped the woman.

The brave man ran into the house.

These sentences can be confusing as they are, if the task is to remember the relations of the three agents and their actions.

These sentences can be expanded with precise elaboration to make the men and the actions more memorable as follows:

The hungry man got into the car to go to the restaurant.

The strong man helped the woman carry the heavy packages.

The brave man ran into the house to save the baby from the fire.

The relations that emerge from the hungry man going to the restaurant, the strong man carrying heavy packages, and the brave man saving the baby from the fire make it easy to find or remember to whom the car, the woman, and the house belong.

The sentences can be elaborated imprecisely as follows:

The hungry man got into the car to go for a ride.

The strong man helped the woman find the newspaper stand.

The brave man ran into the house to find some ketchup.

Examples of precise elaboration in the first group are more effective for the purpose of remembering these confusing relationships, as they can answer elaborative

interrogation, which can be called a "why" question. The latter group does not facilitate the recall of needed information as well because the sentences are elaborated rather imprecisely.

Metamemorial Study

The natural ability of the learner may often influence the efficiency of study, or study ability in a natural situation. One way to increase the efficiency of study is recognizing the limitations of one's memory and successfully coping with them. This knowledge is an understanding of how one's memory works--metamemory. Some metamemorial studies by keyword-method theorists are reviewed in the following paragraphs.

Pressley, Borkowski, and O'Sullivan (1984) pointed out that "good and poor learners differ in the extent to which they are aware of their own memory limitations." Therefore, they asserted that the keyword method can be a way to make students aware of "metamemory acquisition procedures."

Studies by Pressley, Ross, Levin, and Ghatala (1984) of 10 to 13 year olds investigated the possibility of metamemorial effects on the subjects' cognition of learning strategies. The findings confirmed an increase of the keyword method over the contextual method by means of both (a) giving performance feedback and (b) asking them to think back to decide which strategy made a task easier.

An investigation of memory-strategy monitoring in adults and children by Pressley, Levin, and Ghatala (1984) revealed some differences between the two groups of subjects in their ways of monitoring strategies. A mere description of the

keyword method as well as the repetition method as a possible way of remembering vocabulary did not predispose adults to increase their use of either method, however, when given a description of the keyword method and the repetition method. Children (11 to 13 year olds) tended to try the keyword method. What changed the strategies of adults was (a) an experimenter's recommendation and/or (b) results from practice and test performance. As for children, they accepted inappropriate recommendations even after practice. The authors cited Bjorklund and Zeman (1982) to explain the finding: "Children are often insensitive to their own use of strategies, and the effects of those strategies."

Application of the Keyword Method to Teaching Content Materials in Various Subjects

Shriberg, Levin, McCormick, and Pressley (1982) conducted one of the earliest studies on the application of the keyword method for prose learning. Eighth-grade students were asked to learn the purported accomplishments of some individuals from several short passages. The results showed that gains over the control subjects were 100% by the self-generated keyword-image group and 200% by the imposed-imagery group.

The possibility of teaching content materials other than language was explored by Konopak and Williams (1988). They introduced their own examples of a lesson on minerals for a fifth-grade science class, in which they provided a general guideline to implement the keyword method in the classroom. The authors considered the strength of the keyword method as association between known and

new information. However, at the time of their research, there were no other studies in which the keyword method or mnemonics in general had been used as the basis for teaching a whole content area, such as the history of World War I. Therefore, Scruggs and Mastropieri (1989) developed a 31-page booklet on World War II to teach 30 mildly handicapped adolescents.

In this lesson, for example, to show that the Central Powers were Turkey, Austria, Hungary, and Germany, a picture was shown of people playing "tag" (the first-letter mnemonics for Turkey, Austria-Hungary, and Germany) in Central Park (the keyword for the Central Powers). The result showed that the reconstructive elaboration group outperformed the control students by a factor of nearly 2 to 1 on immediate recall.

Investigations of the possibility of teaching content materials can be seen in many of the following cases: an experiment in which Desrochers, Gelinas, and Wiegand (1989) showed that the grammatical genders of German nouns were taught with the keyword method; teaching complex botany concepts to college students (Rosenheck, Levin, & Levin, 1989); teaching 45 Old English words to college students to compare the keyword method with the contextual approach to vocabulary acquisition (McDaniel & Pressley, 1989); teaching tenth-grade students some biographical information about the United States presidents (Dretzke & Levin, 1990).

Rosenheck et al. (1989) taught botany concepts to college students and concluded from their experiment that "mnemonic superiority found after a two-month

delay verified the staying power of the mnemonic instructions, and mnemonic superiority appears to increase with time."

These experiments seem to show that the keyword method can be applied to many kinds of learning tasks that demand memory as an important part of learning. It can be surmised that the keyword method and many of the variant forms of this method will be explored in teaching many more content materials in the future.

Mnemonics and the Keyword Method in TESOL

Among scholars in the field of second-language teaching, mnemonics is largely a forgotten topic. To illustrate, Blair (1982) could name only Andrew Cohen, Michael Tatton, and Delbert H. Groberg as experimentalists serious about mnemonics. Munsell et al. (1988), reviewing the field as a whole from the perspective of the study of the brain and language learning, pointed out that "the field of language teaching seems relatively silent on the issue of memory, with the exception of Stevick (1976, 1986) and a few others" (p. 270).

Stevick, to whom Munsell et al. referred, was undoubtedly one of the greatest pioneer researchers of memory in the field of second-language teaching, as evidenced by his two publications, *Memory, Meaning and Method* (1976) and *Images and Options in the Language Classroom* (1986). These two books do have much information to inspire those interested in memory in the field of second language. However, even though the basis of the keyword method is using images

as an aid for memory of words, Stevick, in neither of the aforementioned books, discussed the keyword method.

Apparently the first discussion of the keyword method by Stevick appeared in his *Success With Foreign Languages* (1989), as one of his interviewees, Fred, reported that he had finally managed to remember one of the Japanese words that gave him "the hardest time remembering." However, his concluding remarks concerning the keyword method only provided his own very guarded support:

> I do think, though, that it is important to remember that at best, stockpiling through mnemonics builds resources that have a very short shelf-life; as Dexter told us, if the words are not used in context very soon, they will evaporate and the time and energy that went into the mnemonics will have been wasted.

Aside from the study of the keyword method by Fuentes (1976), one of the earliest studies that referred to the keyword method as a vocabulary learning aid and reported the data of empirical research concerning the role of mnemonics for the second-language learner came from Andrew Cohen and Edna Aphek (1980). In their experiment, 26 native-English-speaking students learning Hebrew at a theological seminary in Jerusalem were given a short training in the usage of mnemonic associations and were given a list of associational types they could use to learn 20 new words the subjects selected from given passages. Then they were asked to generate their own associations, writing down each time they used an association for a word across classes over a period of a month.

Cohen and Aphek's report showed five types of associations students generated and included the authors' analyses of those types. Original words

students picked to learn new words and the types of associations used for them were also reported. Another finding revealed that the beginning students used sound as a dominant associational basis, and the advanced students also used structural associations as an aid for memory.

Students reported what kinds of associations they used for each word. There were four groups of recall strategies, according to the kinds of associations they used at the time of the recall, as shown below:

	% of Words	% Correct
1. Used original association	46%	86%
2. Used a new association	13%	78%
3. No longer used an association	31%	61%
4. Used no association at all	10%	72%

Some observations made from the study are as follows:

1. "It might be that almost any learner, no matter what his capabilities, could improve his vocabulary learning through the aid of associations."

2. Generating associations to learn a word "may enhance the learner's attitudes toward the language itself." The processes may stimulate the student's creativity and tap "an innate capacity for enjoying the new language."

3. Some of the subjects expressed a desire to receive more thorough mnemonic instructions.

Blair (1982), in his "Easification," mentioned that "in the field of language teaching surprisingly little experimentation has been done on mnemonic training." He mentioned Michael Tatton and Delbert H. Groberg as the few experimentalists he knew of, and referred to Andrew Cohen as "coiner of the term 'easification' in language learning," and as "the contact person" among those interested in mnemonics in the field. His article explained in detail how textual memorization of Russian expressions could be facilitated through mnemonic aids. At the end of the article, he answered a question about the problem of retrieving words acquired through mnemonics. As the theory answers an important question with regard to the keyword method, the whole passage is quoted below:

> What then of the objection typically raised that association mnemonics force the language learner to take extra, unnatural steps and in the end only hobble his learning? My answer is that none of the learning researchers I know or have read believe this. M. David Merrill, a noted author in the field of instructional science, explains that the brain seems to use what is called the "push-down" principle where once something is learned, regardless of how it was learned, memory access is "pushed down" to the automatic level. Ultimately there is no difference in access time between one thing learned through association mnemonics and a similar thing learned through other strategies.

Cohen (1987) expressed some of the critical issues with regard to the use of mnemonic associations in vocabulary learning, giving particular attention to the keyword method and offering possible solutions. In one section, he observed that mnemonic experiments "have not tended to discuss the nature of associations that the learners actually generate." As one exception, Cohen introduced his new experiment conducted over a 100-day period. The students' reports gave rise to

nine types of associations, and were given in tabular form. These nine associations are more fully developed and explained in Cohen (1990, pp. 26-27).

Cohen referred to differences between laboratory studies and classroom learning by citing three reasons given by Peterson and Swing (1983):

1. Many questions raised in this paper can only be resolved in the natural environment of a classroom.

2. Answers to questions that hold true in the laboratory are not necessarily applicable to the classroom, where there is loss of control of significant variables.

3. Classroom learning is qualitatively different from laboratory learning.

Cohen's conclusion that "there is not as yet enough information about the results of such strategies to recommend that all learners use them" sounds more conservative than Carter's (1987) that "research in this domain points quite conclusively to the value of what has been termed the *key-word technique*" (p. 153). However, Carter did mention that the technique should be considered "as only one among several possible techniques" (p. 154). His view of the keyword method was that an effective usage of context to learn words is not a good possibility among learners in the initial stages of ESL learning, as words may "be assimilated as single (or paired) items." Therefore, the mnemonic aid has some utility. However, the method could be time consuming in cases where finding the right keyword and picture interaction is not easy. There may be some interference with pronunciation

and spelling, but mnemonics, according to Carter, can sometimes "assist with irregular spellings such as 'island = is land' etc."

Nattinger (1988), speaking of memory of words, stated that "form may be more important than meaning in remembering a vocabulary item. . . . We need to teach strategies that take form as the principal path of meaning." As for the keyword method, he believed that "the technique is valuable for students at both advanced and beginning levels of ability."

Nation (1990) thought that "mnemonic techniques make learners independent of the teacher." He also mentioned that one difficulty with the keyword method is that it is often difficult to think of keywords that sound like the foreign word. However, even with this difficulty and time involved, students may benefit by knowing mnemonic strategies, as Cohen (1990) observed:

> In conclusion, then, remember that the mind may close off fruitful options to learning vocabulary based on justifications such as, "I have so many words to learn that I don't have time for gimmicks." The reality is that some of these activities–like learning words through association–can save an enormous amount of time.

Furthermore, Channell (1988) included this suggestion in her concluding section of "Psycholinguistic Considerations in the Study of L2 Vocabulary Acquisition," referring to the works of Cohen and Aphek (1980) and Cohen and Hosenfield (1981): "The prediction that associations which are both semantic *and* phonological in nature are most helpful in aiding recall of vocabulary invites empirical testing." Apparently, the best known and researched approaches used in learning a vocabulary in second language is the keyword method.

Areas Needing Research

Peterson and Swing (cited in Cohen, 1987) pointed out some of the differences that should be noted between laboratory studies and classroom learning. These differences may be illustrated by some of the comments made by the keyword-method researchers in the past. For example, Levin et al. (1982) cited their personal communication with Borton with regard to the encouraging results of a national survey of teacher reactions to keyword-method application to classroom learning:

> The results of a recent national survey of teacher reactions to the present keyword method adaptation were most encouraging (Borton, Note 4). About three-quarters of the upper elementary grade teachers who responded were enthusiastic about the method to the point of requesting an opportunity to explore it further in their classrooms.

Pressley et al. (1990) stated that "no systematic research has been produced on how to prepare teachers to teach strategies." One of the reasons for the above situation was given by Pressley et al. themselves: "Perhaps most regrettably, many experimental psychologists and researchers testing the potency of strategies in laboratory settings have little interest in educational implementation of strategy instruction."

Another marked difference between classroom and laboratory situations is the presumed precision associated with the lab. To illustrate, researchers argue about differences one second can make in an experiment:

> In his reply to Hall and Fuson (1986), Pressley (1987) reported effective use of the keyword method with a presentation rate of 3 s per item, whereas Hall, Owens, and Wilson (1986) reported that their college subjects experienced

> considerable difficulty in the use of the keyword method with a presentation rate of 4 s per item. (Hall & Fuson, 1988, p. 251)

Such scientific pursuit of precision in the laboratory is lopsided from the standpoint of classroom teachers. This point becomes even clearer when one considers a comment by Cohen (1990) on the length of laboratory experiments in the past. He stated that "the longitudinal studies conducted by cognitive psychologists may include a time span of only a week or so" (p. 142). Therefore, he mentioned that "there is room for research conducted for longer periods of time" (p. 142). Therefore, one of the main themes of the experiment was to measure longitudinal effects of the keyword method by testing at intervals of one and two months actual students learning English.

It is also interesting to note with regard to the keyword method here that Raugh and Atkinson (1975) said that "it would be interesting to evaluate the keyword method on a language, such as Russian or Japanese, that has few cognates." They did conduct experiments with a vocabulary in Russian, and the results are available. However, to this researcher's knowledge, no report on an experiment using the keyword method in Japanese or with Japanese junior high school students has yet been conducted. It would be interesting to see whether the positive effects of the keyword method would obtain among Japanese learners of English through an empirical investigation. With the aspects of the keyword method discussed above, then, this study undertook to investigate the following three hypotheses:

Hypotheses

The following hypotheses concerning the effects of the keyword method were tested in this study:

Hypothesis 1: The keyword method used with Japanese students studying English in the seventh grade in Japan will lead to higher scores than the traditional method on the immediate posttest and the follow-up tests designed for this study.

Hypothesis 2: The keyword method will lead to higher scores than the traditional method on the same tests used as a follow-up test one or two months later. The subjects of the study are those students with perfect scores on either half of the immediate posttest.

Hypothesis 3: The keyword method will lead to higher scores than the traditional method on the subset of words on the test designated as difficult words.

Chapter III presents the methods used to investigate these hypotheses.

CHAPTER III

METHODOLOGY

To investigate the effect of the keyword method in vocabulary acquisition, and to test the hypotheses developed for this study and listed at the end of Chapter II, an experiment was designed with the following conditions and procedures.

The Subjects

The subjects were seventh graders in 12 classes, each with about 30 to 40 students, in two kinds of junior high schools in Okinawa Prefecture, Japan: four classes at a national junior high school and eight classes at public schools.

Reason for the Choice of Subjects

The choice of seventh-grade students as subjects was based on an observation made about the effectiveness of the keyword method: "It may well be that supplying the keyword is most helpful to the beginner and becomes less useful as the subject gains familiarity with the language and the method" (Atkinson, 1975, p. 824). Pressley, Levin, and Delaney (1982, p. 79) reiterated the point and further identified such less sophisticated learners as elementary school children and junior high school students.

Therefore, the seventh-grade students fit the description of the ideal population who may benefit most from the keyword method: seventh graders in Japan have just begun formal education in English.

The Classrooms

The experiment was conducted in ordinary Japanese junior high school classrooms in the form of the daily English lesson. A class in a junior high school in Japan usually consists of 30 to 40 students and a teacher. In this case, the teacher who supervised the whole class hour of the experiment was always an English teacher.

The Role of the Teacher in the Experiment

The 12 teachers who participated in the experiment were individually given instructions concerning the experiment; they were expected to take part in only a few activities except for supervision of the class, to distribute the material to the students, and to time the length of each section of the experiment according to the experimental schedule. After the explanation of the keyword method given in the handout discussed below, the teachers took the role more or less of proctors.

The experiment was designed in this way to reduce the teachers' influence on the students' performance. In other words, if a teacher were involved in a vital way in the explanation of the keyword method, the ability of the individual teacher would become an uncontrollable variable. The handout given to each student was

intended as a uniform explanation and guide for the use of the time and materials in the experiment.

The Experiment

The experiment was conducted after the second week in May, the second month of the 1991 academic year in Japan. There were four kinds of tests that provided the data for the study. They were as follows:

1. A pretest of the 432 possible subjects to test the ability to supply a Japanese translation for the 30 words in the study. Of these students, 32 were considered to know too many words and were not included in the tabulations. Ten others were excluded because of irregular participation.

2. A posttest of the 390 students to determine the rate of immediate recall of the 30 words immediately after the learning session. In the posttest and follow-up tests, students selected, rather than supplied a Japanese translation, from a list of 30 Japanese words (translations).

3. A follow-up test 1 of 187 students to determine their retention of the acquired words after one month.

4. A follow-up test 2 of 203 students to determine the retention of the acquired words after two months.

Two kinds of follow-up tests were given. Follow-up test 1 was given after one month, and test 2 was given after two months. One-half (6) of the total (12) classes of students took test 1, and the remainder took test 2. In other words, six of the 12

classes (187 students) took follow-up test 1, and the remaining six classes (203 students) took follow-up test 2.

An Outline of the Experiment

One class hour (50 minutes) was used to introduce 15 English words by the keyword method, as well as 15 words by the rote-memory method, a traditional method in Japanese schools. The pretest began the hour, and the posttest ended the hour. One and two months later, the test was repeated.

Selection of English Words to Be Used in the Experiment

The Ministry of Education in Japan designates 490 English words as essential vocabulary items to be mastered over the course of three years in junior high school. A list of these words can be found in the course guidelines issued by the aforementioned agency in Japan (The Ministry of Education, 1977, pp. 102-107). Most of the words that were used as vocabulary items in the present experiment were derived from this source. However, not all of these 490 English words could be used. Two steps of elimination of a certain group of words preceded actual selection of the words that were used for the experiment. One set to be eliminated consists of English loan words, such as "table" or "radio," which happen to appear on the list of essential English words for junior high school students. Because the students' knowledge of these words might have influenced the experimental data, it was thought best to delete them. The other set of words that was eliminated included such words as articles, pronouns, prepositions, and numbers. As the

experiment was expected to last until almost the end of the first semester, the students might have been exposed to most of these words by the end of the experiment, in addition to the exposure during the experimental treatment.

After these two steps of elimination, the remaining words consisted primarily of nouns, verbs, adjectives, and adverbs. From these words, 30 words that could effectively be matched with a keyword needed to be chosen for the experiment. The number of words to be used in the experiment was decided in consultation with English teachers at junior high schools in Okinawa, as well as with consideration for the number of words reported in previous keyword experiments: 40 words a day for college students (Atkinson, 1975, p. 823) and 18 Spanish words for sixth-grade subjects (Pressley & Levin, 1978, p. 359).

In the actual selection of the 30 words, two considerations were made to prevent a biased collection of words. One criterion was that about half the list should be verbs. The other was that only one word should be chosen for each letter of the alphabet. This manner of selection would produce 26 words, one for each letter. However, this was not done for two reasons. First, there were no words on the list beginning with the letters "X" or "Z." Second, the assistants helping to design keyword prompts could not come up with a good keyword for words beginning with the letter "I." Therefore, only 23 words were selected from the list of 490 words in the guidelines. The remaining seven words came from two sources. One was based on an observation Atkinson (1975) made about the usefulness of the

keyword, and the other was a new list of essential words issued by the Ministry of Education.

The first selection of the remaining words was based on an experiment Atkinson (1975) called a "free-choice procedure" (p. 824). In this study, subjects received a keyword only if they requested one. It was found through the experiment that the likelihood of requesting a keyword would increase if a subject missed the item on the previous trial. This finding seemed to indicate that subjects were likely to use a keyword if they had some difficulty with a vocabulary item. Therefore, to investigate the role of the keyword in helping students with difficult vocabulary items, five closely related and presumed difficult words were added to the list: hen-rooster and bull-cow-ox. These words may be confusing and difficult for Japanese students because the Japanese words that denote these two animals are unmarked with regard to the gender of the animals; thus, hens are called female chickens, and roosters are called male chickens. So it is also with cattle in Japanese.

The remaining two words came from a new list of essential words that the Ministry of Education in Japan designated for the preparation of new textbooks. The new list included 30 new words; 13 words from the original list were deleted from this new list. This mandate took effect in April 1993. However, in the transitional period, words from both lists could be taught to students using some textbooks based on the new guidelines. For this reason, the new list was considered a reasonable source for the two words.

For the word beginning with the letter "I," the new list had only three words: "ideas," "important," and "introduce." The word "idea" is used as a borrowed word in Japan, and the word "introduce" may be well known to students as American English teachers introduce themselves in class at the beginning of the academic year. Therefore, the two aforementioned words were deleted, leaving "important" as the selected word. As for the last word to be selected, a verb was sought because verbs represented less than half the list. Among several verbs found in the new list, *enjoy* seemed to represent interesting cultural and syntactic differences in English and Japanese.

The following are the 30 words chosen for this experiment, in alphabetical order, divided into two groups of 15 words each:

ask, both, cry, daughter, early, enjoy, forget, grow, holiday, important, July, know, lake, hen*, rooster*

****

March, nose, other, put, quickly, ready, send, try, use, visit, want, yard, bull*, cow*, ox*

* = "hard words"; not in alphabetical order.

All 30 words were presented to all of the subjects. However, each subject studied half of the words in the keyword method and the other half in the rote (traditional) method. Therefore, subjects who studied the first 15 words by the keyword method also studied the last 15 items by the traditional method; subjects who studied the last 15 words by the keyword method studied the first 15 words by

the traditional method. This means that, instead of using two groups of subjects, the control and the keyword groups, the 30 vocabulary items could serve as either the control or the keyword words. This division was chosen so as to avoid forcing all of the subjects at one school to memorize 30 words by rote, depriving them of an educational innovation, or a possible benefit from the study. Similarly, the design avoided having all of the subjects at another school use an unfamiliar experimental approach.

Order of Presentation of the Two Groups of Words

In presenting these two groups of 15 words to students, the order was crossed so that if one class in a school was presented the 15 words by the keyword method first, the other class in the school received the same 15 words by the keyword method second. The mixed order of presentation was chosen so as to balance any effects of novelty or fatigue derived from a uniform order of presentation.

Selection of Meaning of English Words

The guidelines issued by the Ministry of Education in Japan provide a list of essential English words for junior high school students. However, this list is accompanied by no information as to which meanings of an English word should be taught to students. Therefore, textbook authors use their discretion in the choice of meanings of English words that are introduced. Consequently, meanings of an English word introduced to students could vary from textbook to textbook. (There

are six different textbooks approved by the Ministry of Education. At the time of this study, three textbooks were being used by different educational districts in Okinawa. Public schools in Japan are not allowed to use any other kinds of textbooks in class except as supplements.)

In choosing the meanings of the 30 words selected for this study, the researcher looked through all three textbooks currently in use in Okinawa and chose the meanings that were most often mentioned when the word first appeared. When all of the textbooks used different Japanese words for an English word (sometimes not necessarily a different meaning), the researcher extended the search of the Japanese word or the meaning to select by looking through the other three textbooks approved by the Ministry of Education. In the few cases when it so happened that equal numbers of the same Japanese words or meanings turned up from the investigation of the six textbooks, the researcher looked for the meaning of the words by referring to dictionaries (about 20) designed for junior high school students. What most of the dictionaries listed as the first meaning (or the Japanese word for it) was selected.

Pronunciation

Junior high school students (seventh graders) in Japan have just begun their formal education in English. Therefore, they are not yet quite sure about pronunciations of new English words. Thus, it was necessary for the 30 selected words to be accompanied by pronunciations written in Japanese so that they could at least have some idea as to how they were pronounced. However, no textbooks

give pronunciations of English words in Japanese. The students are to learn them from the teachers' pronunciations. For this reason, the researcher had to resort to dictionaries designed for junior high school students, for they are the only reliable sources that provide phonetic pronunciations of English words using Japanese writing. However, these dictionaries do not express pronunciations (British and American) or a word identically. In some cases, the word "enjoy" was written in five different ways, all sounding close enough to the English pronunciations. Even dictionaries published by a single agent differ in the Japanese descriptions of pronunciation, probably because of the idiosyncrasies of lexicographers.

Most of the dictionaries use only *katakana* (the most common "alphabetic" Japanese system used to write the pronunciations of loan words) letters or a mixture of *katakana* and *hiragana* letters. The printed letters could be all of equal size or varied, bigger and thick, to show stressed sounds. What is most often seen is *katakana* with the print sizes varied to show the stressed syllable. Therefore, this form of describing pronunciation was adopted for the experiment.

Keyword Production

In compliance with Atkinson's (1975) recommendation that a committee approach should be taken to ensure good keyword production, the researcher formed a three-member committee for the production of good keywords. The committee members consisted of two illustrators of the keywords (described below) and the investigator.

The committee members were first given an introduction to the keyword approach to vocabulary acquisition and were then given the list of 30 English words and asked to produce as many keywords as they could. Then, less than a week later, the committee brought all the keywords they had produced together for discussion. A final list of keywords for the 30 words was selected. The criterion of the choice was what sounded good and was considered memorable.

Occasionally, one or another member came up with a new keyword for a word for which the committee had already finished the illustration. When the three members agreed that the new keyword was an improvement, it was added, deleting the old one from the selections. Several instances of such a refinement process arose until the completion of the printing of the experimental materials.

Illustrators

The two illustrators of the keyword associations on the committee were Yuriko Toyama and Mayumi Machida. They were chosen for their excellent talent and availability. They are students at Okinawa Kokusai University, so they were available for discussions and meetings more often than would have been the case with professional illustrators.

Before the illustrators began their drawings, the group met to discuss what kinds of images or ideas should be included in the pictures. The illustrators then made the illustrations with the results of the discussions in mind. On some occasions, the illustrations were revised when the three committee members agreed

that improvements could be expected from revision. The keywords and illustrations are all provided in Appendix A.

Explanation of the Experiment

The experiment took the form of a 50-minute session that included the following components:

	Time (Minutes)
1. Pretest	5
2. Explanation of the keyword method	5-7
3. Studying 15 words by the keyword method	10
4. Studying 15 words by rote memory	10
5. Posttest	10
Total	40-42

The vocabulary learning materials developed for this one-hour study consisted of six pages of handouts. The content of these handouts is explained in detail below and is provided in Appendix A.

Page one of the handouts. The first page of the handouts contained the following information in simple words that junior high school students could understand:

1. Title: Learning English Words: Using Mnemonics.

2. Explanation of the value of acoustic association in learning an English word, and its integration into a pictorial form to render visual support to vocabulary

memory. Two examples of the keyword technique using specific English words follow, using two English words, "spider" and "hill."

3. Recommendation to students that they will avail themselves of acoustic associations and images to learn English words in the future, and that they are about to learn 15 English words by this method in the class.

4. Explanation of the schedule for the rest of the class hour: ten minutes for learning 15 words by the keyword method, and another ten minutes for learning 15 words by rote. Then, at the end of the class, ten minutes for a test to determine their retention.

5. An example of the matching test format, so as to alert students that mere recognition (and not spelling) will be required.

6. Instructions that students follow the teacher, as he or she tells them what to do.

7. The researcher's name, Mamoru Kinjo, and the name of his university, Okinawa Kokusai University, at the bottom.

Page two. The second page is a pretest, listing all 30 English words without Japanese translations. Students are asked to supply Japanese translations of the English words if they know any of them.

Pages three and four. These two pages contain 15 English words to be learned by the keyword method. The items to be learned by the keyword method are called "Group A." The instructions given in this section concern the amount of time allotted to learn them (ten minutes) and how pronunciation is described.

English pronunciation is given by *katakana.* The thick, slightly larger letters in the description of pronunciation show the stressed syllable in English. English words transcribed in Japanese *katakana* are underlined.

On the third page, there are eight frames (8 cm x 6 cm) (see Appendix A) in which English words are presented by the keyword method; that is, inside this frame is a pictorial representation of an acoustic association. At the bottom of the frame are English words to be learned, *katakana* pronunciations of the words in brackets, and the Japanese equivalents/translations. Under this is another line of a phrase or a sentence in English that integrates the association. These eight frames are numbered top to bottom, one to eight. The next page (the fourth page) follows a similar format except that it contains seven forms of the words to make up the total number of words (15) introduced by the keyword method.

Page five. On the fifth page is a list of 15 English words and their Japanese equivalents to be studied in ten minutes in any way students choose--probably by rote. This group of words is called "Group B." It should be noted, however, that equal numbers of groups used Group B words first with the keyword method and Group A words second with the rote method.

Page six. This last page is a test of all 30 words the students had just studied. The 30 English words appear in alphabetical order (with the exception of "hen" and "rooster" at the end of the first group, and "bull," "cow," and "ox" appearing at the end of the second group). The first 15 words occupy the top half of the page. A column containing the 15 Japanese translations of the first 15 words, in random

order, appears to the right of the English column at the top. The last 15 words are on the bottom half of the page on the left, with Japanese translations in random order in a column to the right. This arrangement was used to save some processing time and effort.

The answer sheets provided the data for the analyses. They were sent to the researcher unscored by the teachers who agreed to conduct the experiment with their classes.

CHAPTER IV

RESULTS

Chapter IV contains the results of the analyses of the data. The first section reintroduces the purpose of the experiment and the three hypotheses under investigation to summarize the purpose and the conditions under which the experiment was conducted. The chapter then presents the data, without statistical analysis, to provide a general outline of the study. Finally, the three hypotheses are examined statistically.

Purpose of the Study

The purpose of this study was to provide an empirical investigation concerning the keyword method in a classroom setting. The subjects were 390 Japanese seventh graders studying English in Japan. The task was to remember the pronunciations and meanings of the words presented in this experiment. To measure important differences between the keyword method and the rote (traditional) method, three hypotheses were investigated.

The Hypotheses

The following were the hypotheses of this study:

Hypothesis 1: The keyword method used with Japanese students studying English in the seventh grade in Japan will lead to higher scores than the traditional method on the immediate posttest and the follow-up tests designed for this study.

Hypothesis 2: The keyword method will lead to higher scores than the traditional method on the same tests among those students with perfect scores on either half of the immediate posttest.

Hypothesis 3: The keyword method will lead to higher scores than the traditional method on the subset of words on the test designated as difficult words.

Description of the Data

The Four Kinds of Data in the Experiment

The data in the experiment were obtained from four tests. They are:

1. The pretest
2. The immediate posttest
3. The one-month follow-up test 1
4. The two-month follow-up test 2

A brief explanation of these four tests is given in the following paragraphs.

The pretest. The pretest (five minutes) was taken by a total of 432 students at the beginning of the experiment. It required students to supply Japanese equivalents for the two sets of 15 English words.

The immediate posttest. The immediate posttest (ten minutes) was taken by all of the subjects in the study at the end of the class hour. It required students to find the correct match for each of the 15 English words from a list of 15 Japanese equivalents for both sets.

3. The one-month follow-up test 1. The one-month follow-up test (15 minutes) was given to the students in 6 of the 12 classes (187 students). It was identical to the posttest except for the time limit.

4. The two-month follow-up test 2. The two-month follow-up test (15 minutes) was given to the students in the remaining six classes (203 students). It was identical to the one-month follow-up test 1.

Explanation of the Pretest

The pretest was administered to ensure that the students had little prior knowledge of the vocabulary items to be learned: 15 items to be learned by the keyword method and another 15 by the traditional (rote) method. The study established a score of 3 as the cutoff point on either of the two groups of 15 words. Thirty-two students exceeded this maximum and were excluded from the data to be reported in this study. (A few had very high scores.) An additional ten students were excluded because of missing data. Among the 390 students in the study were some students who answered one or two words correctly on the pretest.

The following gives an account of these numbers and how they are treated. First of all, the total items that all the students would deal with by means of the keyword method, or the rote method, was 5,850 items (390 students x 15 items). From the group of words (5,850) to be learned by the keyword method, the students already knew a total of 148 items: this is about 2.5% of the total. For the group of words to be learned by the rote method, 161 items of the 5,850 were correctly answered by students, about 2.8% of the total. Both groups, therefore, appeared to

know roughly equal numbers of the English words to start with, and no distinction was made of words as "known" or "unknown."

General Outline of the Experimental Findings

This section provides data of a general nature. The following tables and graphs show general trends of the data with a brief description. The tables contain some abbreviations, which are explained below.

Terms

In the following tables and graphs, the following notation is used for different types of methods and test scores:

K = keyword

R = rote

Kpost (or KP) = the posttest grades of students using the keyword method.

Kfoll (or KF) = the follow-up test (both 1 and 2) grades of students using the keyword method.

Rfoll (or RF) = the posttest grades of students using the rote method.

Rpost (or RP) = the posttest grades of students using the rote method

Kdiff (or KD) = the difference between the posttest and the follow-up test of students using the keyword method.

Rdiff (or RD) = the difference between the posttest and the follow-up test of students using the rote method.

Data From the Entire Population

Table 1 shows the frequency distribution of the students' scores from 0 to 15. The numbers of students are given across the four columns representing four variables: (a) the posttest of students using the keyword method, (b) the posttest of students using the rote method, (c) the follow-up test of students using the keyword method, and (d) the follow-up test of students using the rote method.

The first row of scores shows that, at the time of the immediate posttest, 9 students scored 0 with the keyword method, and 11 students scored 0 with the rote method. The rest of the same row indicates the numbers of those who scored 0 at the time of the follow-up tests: 29 for the keyword method and 33 for the rote method. The next 15 rows provide information about the numbers of students who received the scores indicated.

The bar graphs. The frequency distributions shown in Table 1 are presented in Figures 2 and 3 (bar graphs). Figure 2 represents the frequency distribution of the immediate posttest, and Figure 3 that of the follow-up tests.

In Figure 2, the horizontal line shows the possible scores from 0 to 15, and the vertical line, the frequency counts of each score. The keyword score is shown with the black column (the black column/bar hereafter) and the rote method score, with a white column (the white column/bar hereafter). According to the chart, in the areas showing the scores of 0 to 6, the white columns are taller (more rote students scored in this 0-6-point range). But from 7 to 15 points, the black columns

Table 1: Frequency distributions of students' scores on the posttest and follow-up test.

Score	Posttest				Follow-up Test			
	Number		Percent		Number		Percent	
	KN	RN	K%	R%	KN	RN	K%	R%
0	9	11	2.3	2.8	29	33	7.4	8.5
1	6	7	1.5	1.8	27	37	6.9	9.5
2	9	9	2.3	2.3	39	60	10.0	15.4
3	6	13	1.5	3.3	27	47	6.9	12.1
4	5	3	1.3	0.8	37	41	9.5	10.5
5	12	19	3.1	4.9	32	44	8.2	11.3
6	12	21	3.1	5.4	42	31	10.8	7.9
7	19	13	4.9	3.3	32	29	8.2	7.4
8	21	20	5.4	5.1	29	18	7.4	4.6
9	18	24	4.6	6.2	26	25	6.7	6.4
10	16	28	4.1	7.2	17	8	4.4	2.1
11	30	27	7.7	6.9	22	6	5.6	1.5
12	22	29	5.6	7.4	10	2	2.6	0.5
13	49	34	12.6	8.7	14	5	3.6	1.3
14	55	55	14.1	14.1	3	1	0.8	0.3
15	101	77	25.9	19.7	4	3	1.0	0.8

Note: K = Keyword, R = rote, N = number of students

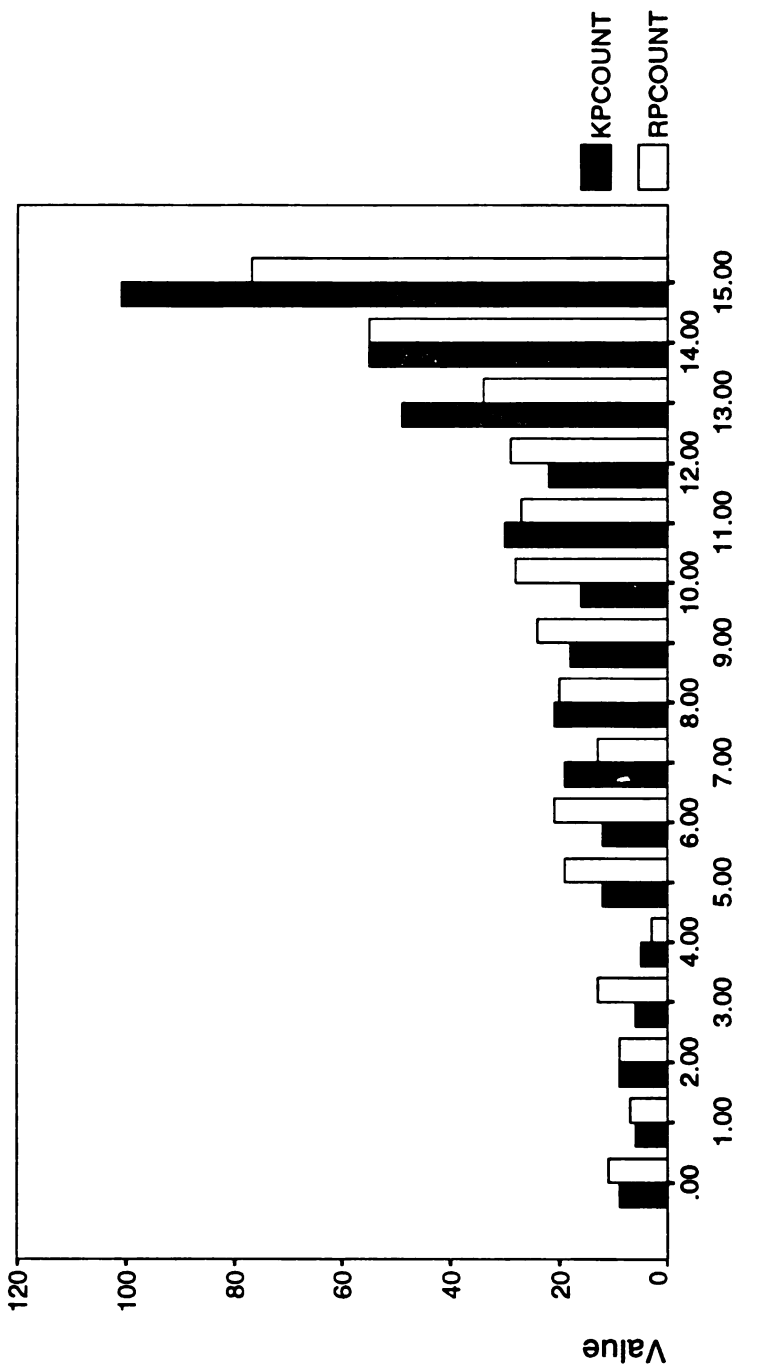


Figure 2: Frequency distribution of the immediate posttest.

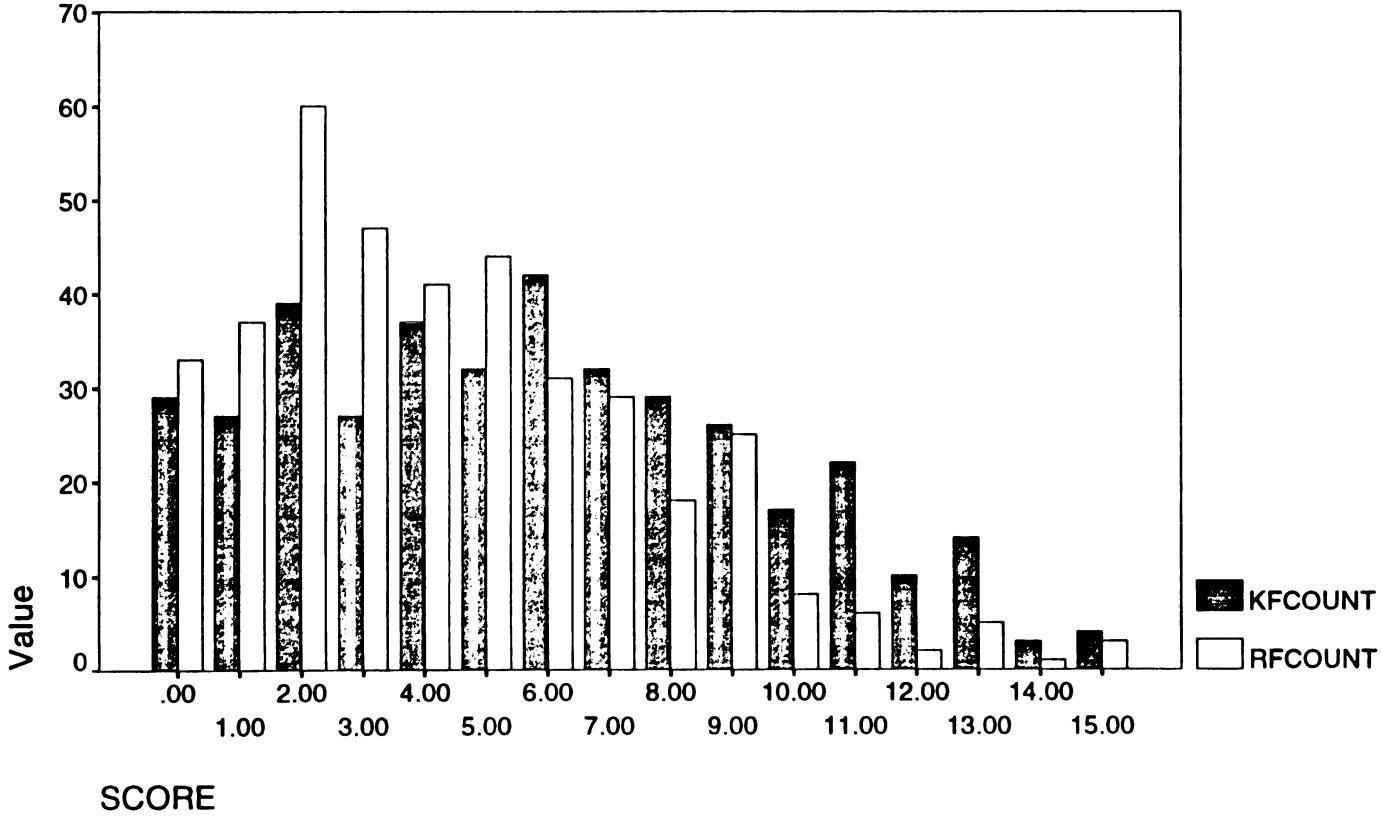


Figure 3: Frequency distribution of the follow-up tests.

are taller except 9, 10, and 12 points, with 14 points being equal (55 and 55). The keyword method seems superior at the upper ranges.

Figure 3 represents the frequency distributions of the follow-up test. The graph shows a different pattern from the results of the immediate posttest. In this graph, the dividing point seems to be around points 5 and 6: in the 0-5-point area, all the white columns are taller; in the 6-15-point area, all the black columns are higher. In other words, the rote method scores tend to be in the lower part and the keyword-method scores in the upper part.

Comparison of the difference between the posttest and the follow-up tests. Tables 2 and 3 and Figure 4 show the differences between the posttest and the follow-up test. In other words, the points shown are the amount of loss that took place between the posttest and follow-up tests at one month or at two months. The chart shows that, generally speaking, the greater loss took place with the rote method and the smaller loss took place with the keyword method. Figure 3 shows that a few students actually had higher scores (1 to 5 points) on the follow-up tests. Three students lost 15 points, meaning they went from 15 on the posttest to 0 on the follow-up test. The tables show that the average loss was about 5 points (5.3 for the keyword method and 5.9 for the rote method).

Comparison of the students whose initial acquisition was equated. According to the data given in Table 4, 101 students had perfect posttest scores with the keyword method, and 77 students with the rote method. A total of 33 students

Table 2: Differences in scores between the posttest and the follow-up test with the keyword method.

Value	Frequency	Percent	Valid Percent	Cum. Percent
-5.00	2	.3	.3	.3
-4.00	1	.3	.3	.5
-3.00	3	.8	.8	1.3
-2.00	4	1.0	1.0	2.3
-1.00	8	2.1	2.1	4.4
.00	25	6.4	6.4	10.8
1.00	22	5.6	5.6	16.4
2.00	32	8.2	8.2	24.6
3.00	30	7.7	7.7	32.3
4.00	40	10.3	10.3	42.6
5.00	27	6.9	6.9	49.5
6.00	44	11.3	11.3	60.8
7.00	41	10.5	10.5	71.3
8.00	29	7.4	7.4	78.7
9.00	34	8.7	8.7	87.4
10.00	21	5.4	5.4	92.8
11.00	9	2.3	2.3	95.1
12.00	7	1.8	1.8	96.9
13.00	8	2.1	2.1	99.0
14.00	2	.5	.5	99.5
15.00	2	.5	.5	100.0
Total	390	100.0	100.0	

Mean = 5.336 Mode = 6.000

Table 3: Differences in scores between the posttest and the follow-up test with the rote method.

Value	Frequency	Percent	Valid Percent	Cum. Percent
-2.00	4	1.0	1.0	1.0
-1.00	7	1.8	1.8	2.8
.00	21	5.4	5.4	8.2
1.00	25	6.4	6.4	14.6
2.00	24	6.2	6.2	20.8
3.00	24	6.2	6.2	26.9
4.00	33	8.5	8.5	35.4
5.00	43	11.0	11.0	46.4
6.00	30	7.7	7.7	54.1
7.00	43	11.0	11.0	65.1
8.00	33	8.5	8.5	73.6
9.00	25	6.4	6.4	80.0
10.00	29	7.4	7.4	87.4
11.00	23	5.9	5.9	93.3
12.000	12	3.1	3.1	96.4
13.00	9	2.3	2.3	98.7
14.00	4	1.0	1.0	99.7
15.00	1	.3	.3	100.0
Total	390	100.0	100.0	

Mean = 5.954 Mode = 5.000

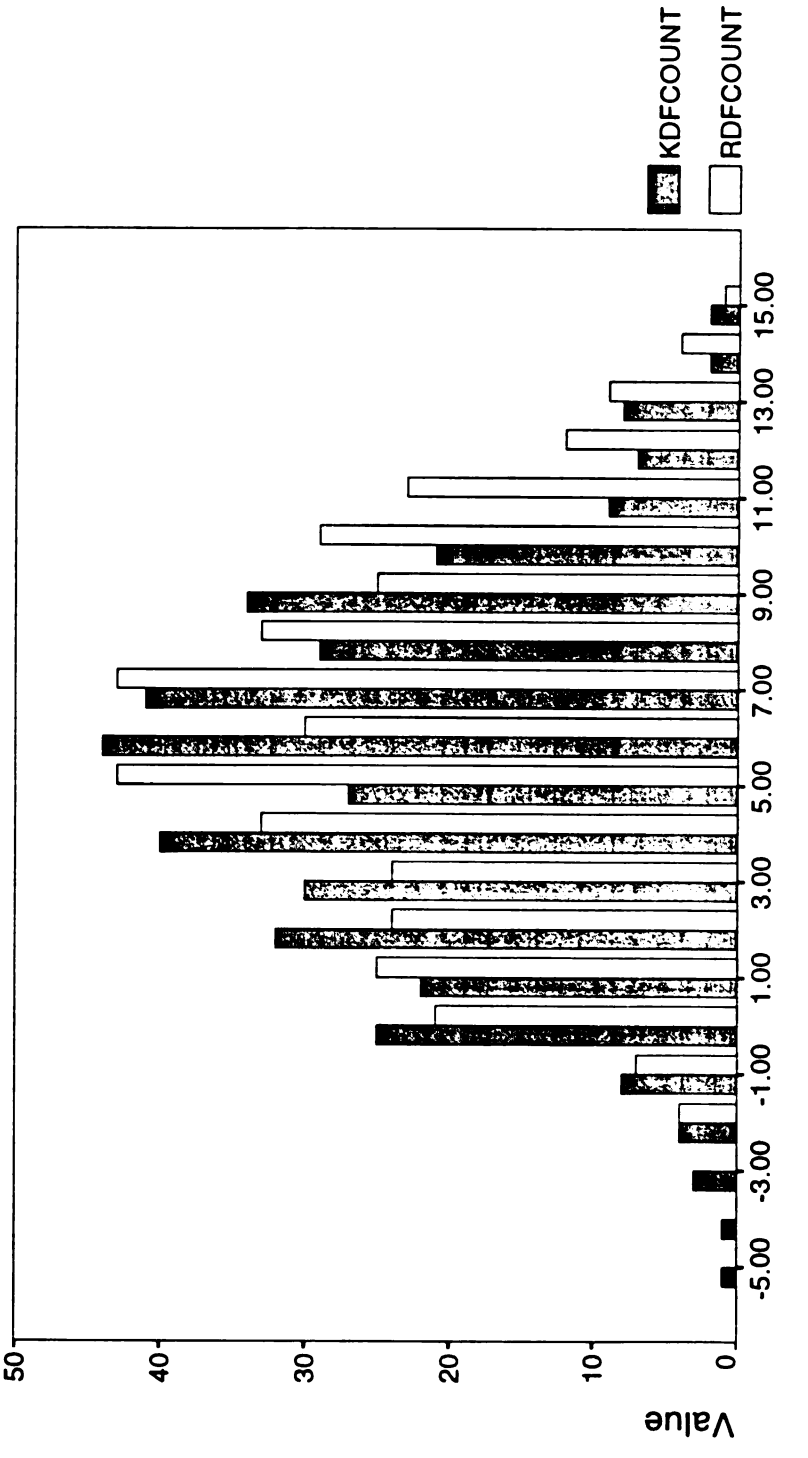


Figure 4: Frequency distribution of difference.

Table 4: The number of students with perfect scores on either half of the immediate posttest or both tests.

Variable	No. of Cases	Mean	Std. Dev.	Std. Error	Difference			t-Value	df	2-tail Prob.
					Mean	Std. Dev.	Std. Error			
Overall--All Correct KPOST										
KPOST	101	15.0000	.000	.000	2.4158	2.725	.271	8.91	100	.000
RPOST		12.5842	2.725	.271						
Overall--All Correct RPOST										
KPOST	77	13.4805	2.075	.236	-1.5195	2.075	.236	-6.43	76	.000
RPOST		15.0000	.000	.000						
Overall--All Correct KPOST and RPOST										
KPOST	33	15.0000	.000	.000	.0000	.000	.000	.00	32	1.000
RPOST		15.0000	.000	.000						

received perfect scores with both methods. Further discussion is given in the analysis section.

The keyword effect on the acquisition of hard words. The data concerning the results of the effects of the hard words are presented in Table 5. Without looking at the detailed parts of the figures, readers can see from the table that, at the time of the posttests, the scores deriving from the rote method are higher. On the other hand, the follow-up test scores show that the scores deriving from the keyword method are all superior. Consequently, there is an apparent reversal in the effects of the two methods.

This concludes a discussion of the general outline of the experimental findings. Further analyses of the data are provided in the following section.

Results of Hypothesis Testing

Hypothesis 1

> The keyword method used with Japanese students studying English in the seventh grade in Japan will lead to higher scores than the traditional method on the immediate posttest and the follow-up tests designed for this study.

To test the first hypothesis of the experiment, the data from the entire population were analyzed by means of the following procedures:

1. The Wilcoxon matched-pairs signed-ranks test.
2. The t-test.
3. Comparison of the difference (lost points) between the posttest and the follow-up test of the keyword method and the rote method.
4. Comparison of 12 classes using the means of each class.

Table 5: Comparison of differences between the posttest and the follow-up tests: hard words.

Variable	No. of Pairs	Corr.	2-tail Sig.	Mean	SD	SE of Mean	Paired Differences			t-value	df	2-tail Sig.
							Mean	SD	SE of Mean			
Overall--Posttest												
HWKPOST	390	.340	.000	.6829	.407	.021	-.0611	.451	.023	-2.67	389	.008
HWRPOST				.7440	.378	.019						
Overall--Follow-up-												
HWKFOLL	390	.302	.000	.3564	.394	.020	.0850	.435	.022	3.86	389	.000
HWRFOLL				.2714	.337	.017						

The following discussion deals with the results of the analyses through these procedures, in the order listed above.

The Wilcoxon Test

The data gathered from the entire study population were first analyzed using the Wilcoxon matched-pairs signed-ranks test (hereafter called the Wilcoxon test) in consultation with advisers at the Statistical Consulting Service at Michigan State University. Although this is not considered a powerful tool, it has the advantage of being useful where the assumption of normal distribution cannot be made with certainty. Because no other studies were available to establish the normal distribution of this population, the Wilcoxon test seemed appropriate as one of the tests.

The Wilcoxon test separates the scores of the students into three groups and gives the numbers of subjects who fall into each category. The three categories are as follows: (a) students' scores with the keyword method that are lower than scores involving the rote method (K LT R); (b) students' scores with the keyword method that are greater than those with the rote method (K GT R); and (c) students' scores with the keyword method that are equal to those with the rote method (K EQ R). The analyses of the posttest scores, follow-up test 1, and follow-up test 2 are provided in the following:

The Results of the Posttest:

%	Cases			
32%	125	-	Ranks	(K LT R)
48%	188	+	Ranks	(K GT R)
20%	77		Ties	(K EQ R)
	390		(2-tailed probability = .0000)	

The analysis by the Wilcoxon test showed that, at the time of the posttest, the keyword method showed a superiority: 48% of the subjects scored higher learning the new vocabulary items with the keyword method, as compared to 32% of the students who did better with the rote method. The net difference between the two treatments was 16%.

Because the 2-tailed probability in the Wilcoxon test was .0000, the observed significance was sufficient to reject the hypothesis of no difference. Therefore, Hypothesis 1, that the keyword method leads to higher scores on the posttest, cannot be rejected in the case of the immediate posttest.

The Results of Follow-Up Test 1:

%	Cases			
28%	51	-	Ranks	(K LT R)
58%	110	+	Ranks	(K GT R)
15%	26		Ties	(K EQ R)
	187		(2-tailed probability = .0000)	

The data from follow-up test 1 showed that, after a month, 28% of the students who took test 1 scored higher with the rote method, and 58% of the students did better with the keyword method. The 2-tailed probability was .0000 ($p < .05$).

The percentage difference between the subjects who did better with the keyword method or the rote method was 30%, in favor of the keyword method. This seems to be an indication that after a month the words learned by the keyword method were retained better.

The data so far have shown that the scores resulting from the keyword method were higher both at the time of the posttest, which tested initial acquisition, and at the time of test 1, which tested the retention of the acquired vocabulary one month afterward, with a 2-tailed probability of .0000 for both times. Furthermore, the percentage difference between the keyword method and the rote method rose from 16% to 30%, in favor of the keyword method, on follow-up test 1. There is yet one more part of the data to consider: follow-up test 2, which would show the results of the scores after two months.

The Results of Follow-Up Test 2:

%	Cases			
25%	51	-	Ranks	(K LT R)
60%	122	+	Ranks	(K GT R)
15%	30		Ties	(K EQ R)
	203		(2-tailed probability = .0000)	

According to the data for follow-up test 2, there was a slight rise in the differences of the percentages between the two populations: It was 30% a month earlier, but rose to 35% a month afterwards. The slight rise was again in favor of the keyword method. Therefore, Hypothesis 1 cannot be rejected in the case of the follow-up tests.

In summary, the posttest suggests that many students did better at the time of the acquisition, as seen in the data: 48% of the students did better with the keyword method, as compared to 32% of the students who did better with the rote method, and the percentage of students who did better with the keyword method kept rising throughout the tests: the posttest and follow-up tests 1 and 2--48% to 58%, and then to 60%. Conversely, there was a declining trend of the percentages of students who scored better with the rote method: 32% to 28%, and then to 25%.

Conclusion. All three test scores of the two populations of the experiment subjected to the Wilcoxon test yielded a 2-tailed probability at the .0000 level for all three tests; the keyword-method groups were superior each time. Therefore, the data suggest that the keyword method is superior as compared with the rote method, and Hypothesis 1 is tentatively accepted.

The t-Test

The next statistical procedure used to test the hypothesis was the t-test. To use a t-test, one needs to assume that the targeted population is normally distributed. This fact is examined through the normal P-P plot of all variables in consideration. Despite the caution expressed above regarding normal distribution in the discussion of the Wilcoxon test, the t-test seems justified with this population, as suggested by Figure 4. The data from the t-test are included in Appendix B.

The keyword posttest average score was 11.1103 and the rote posttest average score was 10.400. The difference between them is .7103. The t-test suggests that there was a significant difference between the keyword post and the

rote post (2-tailed probability = .000). The keyword post was significantly higher than the rote post.

Results of the analyses of follow-up test 1. As for the 187 students with the one–month follow-up test, the t-test indicated that the average of keyword post was different from the average of rote post. The mean of the keyword population was 11.2460, and that of the rote group was 10.1176 (mean difference = 1.128). The students' follow-up test showed that the mean of the keyword group was 5.786, and that of the rote group was 4.524 (mean difference = 1.262). There were differences in both the posttest and follow-up scores (for both, 2-tailed probability = .000), and the keyword-method group showed superior scores each time.

Results of the analyses of follow-up test 2. As for the 203 students with two-month follow-up tests, the t-test suggested that there was a nonsignificant difference in the averages of the posttests for both methods. The mean of the keyword-method group was 10.985, whereas 10.660 was the mean of the rote-method group (2-tailed probability = .170; mean difference = .3251). However, the average of the keyword-method follow-up (5.76) was significantly higher than that of the rote-method follow-up (4.37) (2-tailed probability = .000; mean difference = 1.389).

Comparison of the Difference

Between the posttest and the follow-up tests for all 390 subjects, the keyword group lost, on average, 5.3359 and the rote group 5.9538, for a difference of .618, with a 95% confidence interval between 1.042 and .194. The t-test showed that

there was a significant, if modest, difference between the amount lost by the keyword-method population and the rote-method population (2-tailed significance = .004).

Greater insight was gained into the effect of the keyword method when the overall population was analyzed according to the subpopulations of the one-month follow-up test and the two-month follow-up test.

Results of the analyses of the one-month follow-up test 1. As for the group of students with the one-month follow-up test, the t-test showed that there was no difference in the average number of points lost between the posttest and the one-month follow-up test. The keyword group lost 5.4599 and the rote group lost 5.5936 (difference of 1.064) (a 95% confidence interval between .763 and .495; 2-tailed significance = .675).

Results of the analyses of the two-month follow-up test 2. However, for the population of students with the two-month follow-up test, the t-test indicated that there was a significant difference in the drop between the posttest and follow-up tests. The keyword population lost 5.2217 and the rote population lost 6.2857 (difference of 1.064). The keyword-method population was superior to the rote-method population, with a 95% confidence interval between 1.635 and .493, and 2-tailed significance = .000. This suggests that students with the keyword method were much more likely to score higher when tested after two months than those students with the rote method.

Conclusion. The comparison of posttest means of the total population of 390 students by two different methods, keyword and rote, suggests that the keyword method has an advantage in the initial acquisition of vocabulary items. Furthermore, the analysis of the difference in the amount of loss from the initial acquisition, measured by means of the loss between the posttest and the follow-up tests, suggests that the smaller amount of loss in the means of the follow-up tests of the keyword group is indicative of the superior retention with the keyword method.

Results of the analyses of the comparisons of classes. The whole population was also analyzed as a unit for a comparison of means. All the scores of the total population of 390 students were computed and are shown as 12 units of the data (Appendix C) with their means and standard deviations based on class membership.

The data for the 12 classes are provided in the Appendix. However, what is of interest here is a comparison of the means of the keyword method and the rote method. Therefore, the means of all 12 classes were extracted from the data and listed below for comparison purposes. (See Table 6.)

The first two rows of the table show the scores of the posttest of the keyword method and the rote method. A glance at the scores of the 12 classes shows that out of the 12 classes compared, there are only 3 classes (classes 2, 5, and 8) in which the scores of the rote method on the posttest are superior to those of the keyword method. On the follow-up tests, out of 24 sets of scores, there is no single case in which the rote method showed superiority.

Table 6: Means for the 12 classes on the posttest and the follow-up test for the keyword and rote methods.

Class	1	2	3	4	5	6	7	8	9	10	11	12
KP	11.714	9.800	9.759	11.968	10.895	11.278	10.724	7.179	12.029	11.941	12.297	12.657
RP	9.857	10.400	8.724	8.968	11.658	10.556	10.000	7.857	11.971	10.412	11.459	11.629
KF	4.036	4.433	4.621	5.935	5.000	5.306	6.345	4.036	7.400	7.618	7.618	6.314
RF	2.893	4.367	4.276	4.065	4.079	3.417	4.897	3.286	5.543	5.588	5.588	5.429

In conclusion, using a variety of evaluation procedures, Hypothesis 1 seems to be upheld.

Hypothesis 2

This section provides the analyses of the second hypothesis of the experiment:

> The keyword method will lead to higher scores than the traditional method on the same tests used as a follow-up test among those students with perfect scores on either half of the immediate posttest.

This section of the study compares the scores of the following three groups of subjects to compare the levels of their retention:

1. The 101 students who received perfect scores related to the keyword method.

2. The 77 students who received perfect scores related to the rote method.

3. The 33 students who received perfect scores with both methods.

All 145 of these students received perfect scores through one or both methods. Their initial acquisitions of words are considered to be equated. The data from these students give the following results.

Analyses of the data of the 101 students who received perfect scores related to the keyword method. Out of the 101 students who scored perfectly with the keyword method only, 47 students took follow-up test 1 and 54 students took follow-up test 2. Their data and the results of the t-tests of these 101 students are included in Appendix D.

The means of the three tests show that those of the keyword method are always higher, and the 2-tailed probabilities are always smaller than .05 (.0000, .001, and .0000, respectively). Therefore, the hypothesis of no significant difference is rejected: The keyword method seems to be superior.

When the immediate posttest was given, the advantage of the keyword group averaged 2.7443; after a month, the average advantage was 1.4468, and two months afterwards the average advantage was 2.3519. These figures seem to suggest that the keyword method sustains its advantage over time.

Analyses of the data of the 77 students who scored perfectly with the rote method. Of the 77 students who on the posttest scored 100% on words learned by means of the rote method, 30 students took follow-up test 1 and 47 took follow-up test 2. It is interesting to note that among the 77 students with perfect score on the rote method set at the time of the posttest, the keyword method showed an advantage when the students were tested one month and two months afterwards.

Analyses of the data of the 77 students with the immediate posttest. The 2-tailed probabilities of the analyses show that, at the time of the posttest, the rote method was definitely better (.000), but test 1 showed that the keyword method showed an advantage (.000). Again at the time of test 2, the keyword method was better (.003).

Analyses of the data of 30 students with follow-up test 1. As for the posttest showing the 1.9667 average advantage in favor of the rote method, the

difference shifted to 2.1033 points in favor of the keyword method at the time of test 1 (after one month).

Analyses of the data of 47 students with follow-up test 2. As for the 47 students who took test 2 after two months, they had an average advantage of 1.234 points over the keyword method at the time of the posttest. However, by the time of follow-up test 2, the keyword method had an advantage of 1.5531 points above the rote method.

Analyses of the data of the 33 students who scored perfectly with both methods. Of the 33 students who scored perfectly with both methods, 11 students took test 1, which was the follow-up test after a month, and 22 students took test 2, which was the follow-up test after two months. The results are presented in the following paragraphs. (See also Appendix D.)

Analyses of the data of 11 students with follow-up test 1. These students started with equal scores (15), but after a month, the scores of the 11 students with test 1 averaged 8.7273 for the keyword method and 6.3636 for the rote method. The mean difference was 2.3637, and the 2-tailed probability was .012. As $p < .05$, the hypothesis of no difference was rejected; therefore, the data suggest that the keyword group was superior to the rote group.

Analyses of the data of 22 students with follow-up test 2. For the 22 students who took test 2 after two months, the difference of the means was 2.1819, also in favor of the keyword method, and the 2-tailed probability was .01. The difference was again in favor of the keyword method.

Conclusion. It can be said from the above observations that the data supported the hypothesis, which predicted that even with students whose initial acquisition was very high, the keyword method seems superior for helping in remembering the meaning of English words. Hypothesis 2 thus seems to be upheld.

Hypothesis 3

This section provides the analyses of the data regarding the third hypothesis of the experiment:

> The keyword method will lead to higher scores than the traditional method on the subset of words on the test designated as difficult words.

Raugh and Atkinson (1975) reported an increasing trend for subjects in their experiment to ask for the keywords in learning words that were hard to remember. Therefore, to investigate the keyword effect on hard words in this experiment, five hard words were included in the list. These were two groups of words: (a) hen and rooster and (b) bull, cow, and ox. All 390 students learned one of the groups of words with the keyword and the other with the rote method, or vice versa. The percentages of successful mastery of these words as values and their frequencies were calculated. As there were two groups of words, one of which consisted of two items and the other of three items, the following adjustments in calculations were made in computing the students' scores. All of the students received their percentage of correct answers as their scores. To illustrate, if all the answers were correct (100%), 1 point was given. In the case of one correct out of two, 0.5 point was given as the score. Then, scoring was done as follows: one out of three, 0.33

point; two out of three, 0.67 point; and three out of three, 1 point. (See Table 7; further data on the subpopulation are available in Appendix E.) The data are given below.

Analyses of the posttest score and the follow-up scores on hard words. One of the important findings is that the rote method showed a superiority in acquisition of hard words on the posttest (rote 74% and keyword method 68%). However, the scores of the follow-up tests show that the rote method lost its initial superiority (rote 27% and keyword method 36% correct).

Comparisons of the difference between the two methods. The data show that there is a 15% difference ([74-27] - [68-36] = 15) in the amount of loss between the two methods in favor of the keyword method. In other words, hard words learned by the rote method lost 15% of their advantage over those learned by the keyword method when their retention was measured after a month or two months.

To summarize, the results of the analyses of the data of our samples seem to support the hypothesis that an increased effect of the keyword method is seen in learning words that are hard to remember.

Chapter Conclusions

As for the first hypothesis, the conclusions from the Wilcoxon matched-pairs signed-ranks test, the t-test, and other observations yielded strong support to the hypothesis that the keyword method does lead to greater acquisition and retention of new vocabulary.

Table 7: Results of t-tests for difficult words.

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKPOST	390	.340	.000	.6829	.407	.021
HWRPOST				.7440	.378	.019

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
-.0611	.451	.023	-2.67	389	.008
95% CI (-.106, -.016)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKFOLL	390	.302	.000	.3564	.394	.020
HWRFOLL				.2714	.337	.017

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
.0850	.435	.022	3.86	389	.000
95% CI (.042, .128)					

As for the second hypothesis, the follow-up test scores of the three groups of students with perfect scores on either set of 15 words showed that the keyword method scores were always superior. Even in the case of the 77 students with perfect scores on the rote portion of the posttest, their follow-up scores favored the keyword method. This is a strong indication that the keyword method led to higher retention even among the students who had a perfect score on the rote portion of the posttest.

As for the third hypothesis, the data showed that in the initial acquisition of hard words measured by the immediate posttest, the rote method was superior, but the effects were reversed by the time of the follow-up tests.

CHAPTER V

DISCUSSION

Introduction

This section of the study considers the data in terms of their possible causes and classroom applications. The chapter also considers the limitations inherent in the keyword method itself and limitations concerning the present experiment, and consequent suggestions for future researchers. The first part of the discussion provides possible theoretical causes of the findings by integrating facts and theories found in the experiment, in the review of literature, or in readings not covered in the review of literature.

The first set of interpretations is based on findings from such fields as cognitive psychology, applied linguistics, or TESOL. The specific theories or hypotheses are (a) the levels of processing view (Craik & Lockhart, 1972) and "apperceived input as an activator of selective attention" (Gass, 1988); (b) the tip of the tongue phenomenon (Brown & McNeill, 1966) and the cohort model of word recognition (Meara, 1987); and (c) dual code theory (Paivio & Desrochers, 1981) and the organization-redintegration hypothesis (Desrochers & Begg, 1987).

Possible Causes

Why Does the Keyword Method Facilitate the Immediate Acquisition and Retention of Foreign Words Better Than the Rote Method?

The results of the experiment strongly suggest that the keyword method does help vocabulary acquisition as far as remembering the pronunciation and the meaning of a word is concerned. The keyword method is different from other methods of teaching vocabulary because of two kinds of links, which Atkinson (1975) referred to as the acoustic link and the imagery link. These two devices will be discussed in the following pages in connection with a theory called the levels of processing view, and a notion of apperceived input as an activator of selective attention.

Levels of processing. Craik and Lockhart (1972) proposed a theory now known as the levels of processing theory. According to this view, any incoming stimulus generates some amount of mental processing on the part of the receiver. The mental processing can be on many different levels, from shallow to deep. The level of processing of the target information is hypothesized to influence both the initial acquisition and the strength of retention of the information in memory.

Craik and Lockhart gave as an example of shallow information processing what is known as maintenance rehearsal, and as an example of deep processing, elaborative rehearsal. A familiar form of maintenance rehearsal in the case of vocabulary learning is a simple, mechanical repetition of the words. On the other hand, elaborative rehearsal involves thinking about the word in much more complex

ways during learning. The learners need to use other aspects of their prior knowledge. For example, *train* could be rhymed with *rain*, making a phrase like "the train in the rain." Or a train is classified as a vehicle, and the vehicle is known to be faster than a horse.

If it is true that an item of information is acquired better with the deeper processing of it, we have a possible explanation for the success of the keyword method. In the keyword method, students must process an item of target vocabulary more deeply than in mere repetition, or when merely jotting down the word as many times as necessary until it is acquired. Specifically, the two components of the keyword method, the acoustic link and the imagery link, appear to involve elaborative rehearsal.

Acoustic link. As for the acoustic link of the keyword method, students learning an English word have to find some Japanese words that sound similar. Therefore, it is necessary for them to pay close attention to the individual phonemes and their order in the English word so that they can find Japanese words that properly represent the order of the sounds in the English word. Then, students have to examine which of the Japanese words enable them to replicate the order of the phonemes of the English word most closely, once it is chosen as the keyword. The closer attention the students pay to the word, the more mental investment can be expected. Therefore, these processes of acoustic analysis of the phonemic features of an English word could add to a deeper processing of the word and hence to more effective acquisition and stronger retention.

Furthermore, the final selection of the keyword does not always depend on an exact copy of the phonemes of the English word. In some cases, because of the connotations or individual associations with the word, the second-best word or phrase may be chosen. The choice may be the result of a negative association of the first keyword, or a positive quality of the second best keyword that can override the advantage of the exact ordering of the English phonemes in the Japanese word.

For example, the exact ordering of the phonemes of the English word *invite* can be done with a phrase in Japanese meaning "with a prostitute." If students recite aloud in class the English word followed by its meaning, it would sound in Japanese like "invite (you) with a prostitute." The keyword "with a prostitute" is not, therefore, a good one for teachers to provide in the seventh grade.

Such a process of selection of the keyword entails mental negotiation of the acoustic advantage of one word over the meanings and associations of other keyword candidates. This will again add to the levels of processing of information in the acquisition of vocabulary items.

The selection of the keyword from many acoustically similar Japanese words requires students to draw on their knowledge, and such usage of prior knowledge is considered to be an important process of changing input into intake. Gass (1988) referred to this process as "activators of selective attention," which leads to "apperceived input." Apperception entails application of prior knowledge to the intake of new information.

The importance of the usage of prior knowledge in language learning may be seen in a report of a study by Pressley, Levin, and McCormick (1980). In their experiments using the "sentence variation" technique within the keyword method, their pilot research revealed that experimenter-provided sentences failed to benefit a group of second-grade subjects, whereas self-generated sentences did benefit them. When the students were allowed to utter their constructed sentences aloud, they showed learning gains of more than 180%. The results reported here may illustrate the involvement of different levels of prior knowledge as a cause of success or failure in their learning.

As discussed above, the usage of prior knowledge that is required for input to become apperceived input constitutes one of the crucial factors in successful learning. Thus, it could be said that the keyword method and the process of finding the acoustic link serve as "activators of selective attention" and lead to more apperceived input. This aspect of the usage of prior knowledge with the keyword method should result in some learning gains. What is discussed here about the acoustic link and activation of selective attention applies to an imagery link, as well.

Imagery link. When the Japanese keyword is chosen, the student now has two words to combine in a single image--one coming from the Japanese word and the other from the meaning of the English word. This process of creating an effective image allows for more individual freedom and variation than mere repetition and rote memory may offer. To illustrate, Atkinson (1975) used the English word "pot" as the keyword to memorize the Spanish word meaning "duck." His example

of the keyword image was a duck having an overturned pot on its head, or a duck carrying an overturned pot on its head. However, the image does not have to be that way: it can be of almost any type that can allow for students' creativity in visualizing the combination. Some examples of these possible images could be a pot in which a duck is sitting on its eggs or its ducklings with a legend: a pot-duck party or a pot with a cut-up duck inside: a pot-duck dead. There are as many images as could be created by students; the process of creating various images and selecting one that is the best for each of the students will increase the depth of the processing of the vocabulary item.

Furthermore, the keyword method opens up for students a greater possibility of expressing their feelings while learning vocabulary. For example, a happy student may like to use or create images of an optimistic note, and an angry student may want to vent his or her anger on images of a darker mood, or of something destructive, which may be particularly appealing on certain days, or hours, or because of their personal adversities. Such free accommodations of individual psychological make-up or mood may make it more effective than methods with rather uniform presentations. Students can possibly process deeper what appeals to them than what they are less able to relate to.

As discussed above, the usage of the acoustic link and the imagery link seems to result in cognitive processing that will enhance the acquisition of information. In working on these links, students will also activate their apperception, thus (probably) increasing the efficiency of vocabulary-learning activities. Some of

the positive results of this experiment may reflect the effects of these processes: the levels of processing view and apperception as a motivator.

Why Does the Keyword Method Facilitate the Long-Term Retention of the Acquired Vocabulary Items?

The discussion of this question is divided into two sections. The first section deals with implications stemming from the tip of the tongue phenomenon and the cohort model of word recognition. The second section deals with implications derived from studies of imagery.

The tip of the tongue phenomenon and the cohort model of word recognition. One notable experimental finding is that all the scores of the follow-up tests showed a superiority for the keyword method over the rote method. At the time of the posttests, which measured immediate recall, this was not necessarily so. The rote method produced higher results in some cases (in three of the classes, and with hard words). That is, even when the posttest scores favored the rote method, the superiority had vanished, without exception, in about a month. This result is interpreted to indicate that the keyword method increased long-term retention of the acquired vocabulary items or that words learned by rote were more easily lost.

The following discussion is an attempt to provide some theoretical interpretations. Studies from the tip of the tongue phenomenon and a theory called the cohort model of word recognition (Meara, 1987; see Gass, 1989 for the review of Meara) will provide the basis for the discussion.

Brown and McNeill (1966) studied their subjects when a word was on the tip of their tongues--that is, when they were not quite able to give the word they were sure they knew. These subjects, although unable to provide the word, could supply such information as the first sound, the number of syllables, or the suffix of the word. This means that a word is a composite unit integrating a plethora of information. Horowitz, White, and Atwood (1968) discovered that not all of the components of the word are effective as cues for the retrieval of the word when it is on the tip of the tongue. They identified, among the syllables of the word, the first syllable as the best cue for retrieval of the word, the last as the second, and the middle syllable as the least effective.

If there is some identifiable facilitation in some of the acoustic properties of a word, such as the first syllable of the word, the keyword method that uses the sounds of the foreign word may be taking advantage of this acoustic facilitation of word retrieval. It is most often the case that the keyword copies the first few sounds (phonemes) of a foreign word (in this experiment, all of the keywords [Japanese] included sounds similar to the initial sounds of the English words). The results can possibly therefore be interpreted as this facilitation of acoustic quality at work.

In terms of facilitation of retrieval through the keyword method, a theory called the "cohort model of word recognition" (Meara, 1980; see Gass, 1989, for the review) may provide some insight into a possible advantage of the keyword method over the rote method.

According to the cohort model of word recognition, word recognition proceeds in a linear left-to-right fashion. Identification of the initial phoneme of a word eliminates all the other words that do not begin with that phoneme in the mental lexicon, thus reducing a cohort of possible words to those starting only with that initial sound. Then the identification of the second phoneme will further reduce the cohort to a smaller cohort sharing the two segments in the order given. The process of phoneme identifications, and subsequent cohort reductions, continues until a single candidate emerges.

The effect of cohort reduction by a serial identification of the phonemes is such that, for instance, if the phoneme /ə/ is given as a cue for the name of a country, some readers may already have some countries in mind. With the second phoneme /m/ given, most people can terminate the lexical search for the country and say "America." To give another example, given "fruit" as a cue and /æ/ and /p/ as the first two phonemes, identifying the fruit becomes easy. There are not many fruits that begin with this particular string of phonemes. Either /p/ or /r/ as the next phoneme terminates the lexical search: the word sought is *apple* or *apricot*.

As discussed above, no more than the first few phonemes are necessary to activate enough cohort reductions to zero in on the word(s) sought. Therefore, the following discussion deals with what the cohort model of word recognition might suggest when Japanese students are learning English words through the keyword method.

The keyword method requires a student to use the pronunciation of the English word as a retrieval cue for the meaning. Therefore, the student must first find Japanese words sounding like the target English word. One possible process of the recognition of English words is to process a few initial phonemes of the English word. Using these phonemes to search for the Japanese words, only a few Japanese words that copy the exact serial cluster of phonemes in the word remain in the cohort. With a fourth or fifth phoneme given, fewer Japanese words will remain. Once the Japanese word is chosen and matched to one used to learn the English word, the image combining it with the meaning of the English word should come to mind without much difficulty. The memory of the image can subsequently result in a retrieval of the meaning of the English word.

Even if the meaning of the English word is not retrieved instantly, the student can continue trying to retrieve the meaning of the word. Since the Japanese keyword is available, it can be kept in working memory easily and continue to serve as a cue. With a method such as the rote method, which does not provide learners with heuristic devices, retrieval is often an all-or-nothing process.

Studies of Imagery

One of the important components of the keyword method is an image that integrates the native word and the meaning of a foreign word. If images enhance memory, as seems likely, the keyword method could have some memory-enhancing effect because of its image or picture. The following discussion reviews a few studies concerning the superiority of pictures or images over strictly verbal stimuli.

The discussion relates the dual code theory (Paivio, 1971) and the organization-redintegration hypothesis (Desrochers & Begg, 1987).

In terms of memory, or memorability, pictures and words are known to be very different. To illustrate, in one experiment, subjects inspected 2,560 pictures of complex visual scenes for 10 seconds each and could identify 93% of them in the recognition test. This means that the subjects correctly remembered a total of 2,380 pictures out of 2,560 (Standing et al., 1970). Verbal materials may not show such accurate identification. A comparison of the recognition of pictures and verbal materials is available from a study by Shepard (1967). In this experiment, subjects were shown pictures first one by one; then they were asked to identify the one they had seen from a pair of pictures. Their recognition error rate (misidentifications) was 1.5%. With the verbal materials using sentences to identify in a similar fashion, the recognition error rose to 11.8%.

These experiments seem to show that pictures (images) have different levels of retrievability from verbal materials once they are stored in memory. Interestingly, the different strengths of retrievability from memory between images and verbal materials appear to be reflected in the mental representations of words. To illustrate, Paivio (1971) discovered that concrete words are better remembered than abstract words. Concrete words such as *pen, desk, cat, dog,* and *apple* involve material images of the referents, but abstract words such as *idea, freedom, identification,* and *promise* seem to have few of these concrete, material images.

Paivio (1971) accounted for this difference by postulating the dual code theory. This theory holds that there are two kinds of information coding in memory; one is verbal, and the other is imaginal. Concrete words are stored in these two codes, but abstract words are stored only in a single code, verbal codes. If a word is stored in two codes, the possibility of retrieving the word becomes higher than when it is stored only in a single code. Therefore, concrete words that are stored in two codes are easier to remember and are retrieved with more ease than abstract words.

There is from the field of education some evidence showing that the concrete quality of memorable pictures can be a potent means of helping mentally retarded children develop basic skills of literacy. For example, the letter "f" can be taught using an image of a flower with leaves at the cross bar and a flower at the top (Laubach et al., 1971). Isgar (1975) used a similar technique to teach reading-disabled children and obtained 100% success at letter identification after only 5 to 10 minutes of practice per letter.

The results of these experiments show that pictures and images add something to verbal memory and strengthen retention of certain information. This longer retention becomes even more important when it is considered in connection with Begg's organization-redintegration hypothesis. The hypothesis is discussed below.

According to this hypothesis, whenever an item of information is processed in the brain, it constitutes a unique cognitive and physical event. Each of these

events constitutes a "trace." Traces may be organized into higher-order units such as chunks, schemata, scripts, and so on. The hypothesis further holds that "[b]ecause each trace is assumed to be unitary, access is redintegrative, meaning that *access with any part of a trace makes the entire information content of the trace available for use.*"

The keyword method appears to require students to remember a word in a foreign language using both imagery links and acoustic links, at least in the initial stages of learning the word. If the picture or image part of the keyword method is well integrated into a trace-like unit of the word memory, it may redintegrate the entire information as soon as the keyword image or the picture is retrieved. While it might not be desirable to keep this image cue as a major part of the long-term memory trace, having access to it during the period in which the word is being consolidated in memory clearly seems to provide an advantage in lexical acquisition. With this assumption, it is conceivable to interpret the results of the experiment as reflecting the enhancement of the memory because of images/pictures and redintegrative quality of the memory trace. Rote memory has fewer means of preserving the memory trace.

Summary of the Discussion

The discussion in this section dealt with two questions regarding the keyword method: (a) Why does the keyword method facilitate the immediate acquisition and retention of foreign words better than the rote method? and (b) Why does the keyword method facilitate the long-term retention of the acquired vocabulary items?

The first question was addressed through the frameworks of the levels of processing view and apperceived input as an activator of selective attention. Both views hold that elaboration of information, or relating new information to prior knowledge, generally enhances intake of information. The keyword method normally involves more elaboration than the rote method.

The answer to the second question involved such theories as the tip of the tongue phenomenon and the cohort model of word recognition. Using the initial phonemes of a foreign word as an internal cue for retrieving a Japanese keyword, which then leads to the meaning of an English word, is a possible process. The dual code theory of Paivio (1971) and the organization-redintegration hypothesis (Desrochers & Begg, 1987) both point to the possibility that the retrieval of information could be facilitated either through images, or some tightly chunked information unit such as the unit of information that integrates an image, pronunciation, and the meaning of the target word.

Classroom Applications

In terms of actual classroom applications, the keyword method lends itself to varieties of vocabulary-learning activities, as it allows students to have some choices in finding the keywords and creating some images combining the keyword and the meaning of a target foreign word.

Students may be asked to find good keywords for a foreign word as part of their homework. Then the teacher may use these keywords in some other class period. Or several English words may be given to students, to let the students find

the keywords and images. The answer can be put on uniform-sized sheets of paper like an index card, and placed on the wall of the classroom. Students can look at them during class breaks to study and amuse themselves and learn in the process.

The students may also be instructed to choose the keyword production that strikes them as the best. The creator may even be nominated for some titles like the keyword king or queen of the week, the month, or the year.

Atkinson (1975) reported that an individual can have idiosyncrasy in preference of some peculiar keywords. Therefore, students may be encouraged to ask their family members to provide their favorite keywords for the English words they are studying. They can collect keywords from younger or older brothers and sisters, or from parents. Therefore, the keyword-collection activities and the discussions involved in them will enhance family communication and parents' understanding of what their children are learning in their language classes. This seems to be healthy application of a teaching method called the keyword method.

However, a word of caution may be in order here. Finding the keyword or the combination of the keyword and the foreign word takes some imagination. As a result, it is conceivable that some less imaginative students may never find any of the keywords or combinations, no matter how much time they spend on the task. Therefore, the homework should not be forced, nor should slower students be embarrassed regarding their efforts. Therefore, it may be done as group homework, if group work is desirable.

Another area in which the keyword method seems to be of special help is remembering some difficult words. For example, students learning English may soon encounter groups of words that are confusing because of their homonymy, such as (a) son and sun; (b) vane, vain, and vein; and (c) pair, pear, and pare. Students find learning the correct spellings of these words very difficult. Then trial-and-error processes may make students think that their efforts are wasted and the study of these items very painful; there is no guarantee that they will not be forgotten again. The keyword method can help to overcome or ease the difficulties involved in these groups of words. Some examples this researcher has been using may help make the point easier to understand. The three groups of the homonymous words mentioned above will be used as a sample of the mnemonics:

Son and sun: The question is what goes in between the S and the N?

1. For son--Every father thinks his son is ok (son-ok).
2. For sun--The sun is the life of our universe (the sun-universe).

Vain, vane, vein:

3. For vain--I am vain.
4. For vane--Jane's house has a pretty vane.
5. For vein--I found an octopus with eight veins.

Pair, pear, pare:

6. For pair--God breathed air into His first pair, Adam and Eve.
7. For pear--Bears eat pears.

8. For pare–Paring knives are rare.

As the examples show, the sentence-variation technique within the keyword method has the possibility of making tractable some difficult words. These particular homonyms may seem trivial, and helping students learn some similar pairs may not be our biggest challenge. However, the possibility of helping students feel confident that they will not easily forget the confusing items is not trivial. This variation of the keyword method can give students some confidence that if they try hard enough to find some clues and elaborate their learning methods, there is a way better than mere rote.

In conclusion, the keyword method seems to provide students with a way to retain vocabulary items. This possibility can be explored by students, or teacher, or researchers. However, there are some limitations with regard to the keyword method. They are discussed in the following section.

Limitations of the Keyword Method and the Experiment and Suggestions for Future Researchers

Future researchers may wish to explore theoretically or empirically some areas concerning the validity of the keyword method as a way of teaching foreign language vocabulary. So far, the discussion has dealt with some advantages of the keyword method. In the next section of this paper, therefore, some limitations or disadvantages inherent in the keyword method and in this study will be briefly discussed. These limitations and disadvantages raise empirical or theoretical questions. Among these limitations are the limitations inherent in the keyword

method itself, the language used in this experiment (Japanese), a single exposure, and the keyword effect on hard words.

Limitations Inherent in the Keyword Method

First, the keyword method is a mnemonic device for remembering the meaning of a foreign word. Because of this restricted purpose, the expression "vocabulary acquisition" used by most keyword-method advocates only means the successful remembering of the meaning of the vocabulary item or its sound. In other variations, the keyword method could help students remember some other aspect of information. It is not, however, designed to teach many aspects of words in their natural, valid contexts. When words are learned in context, students can also learn syntax, connotations, collocational possibilities, and many other discourse features that go with the word. However, such a rich possibility is missing from the normal keyword method, clearly one of the great limitations inherent in the method itself. Its great advantage is helping students (perhaps especially beginning students who do not yet have enough knowledge of English to avail themselves of the contexts of the English material) to remember English words, and to remember difficult or confusing words with some assurance they will not forget as easily as they might with the rote method.

A second drawback is that it takes a great deal of time to find a proper keyword in Japanese for some English words. Then there are some English words whose pronunciations almost defy the finding of keywords. This means that there are some words that can be taught with ease with keywords, and some others for

which it is almost impossible to find keywords. Therefore, the advantage of the keyword method cannot extend to every word to be learned.

Language Limitations

There are a few limitations in applying the findings of the present research to classrooms. One obvious limitation stems from the fact that the study was conducted with Japanese seventh graders as subjects. Therefore, the results may not be replicated with other subjects using other languages. This limitation gives rise to the suggestion to conduct further research of a similar kind with other English students of different languages.

Single Exposure

The present study was an experiment with a single exposure of the keyword method. Even though results and the analyses strongly suggested that the keyword method does have a positive effect on students' learning, these results cannot be used to predict the effects of multiple exposures with the keyword method. Therefore, some experiments using multiple exposures of a word using the keyword method, or sustaining the use of the method over a long period of time, can add to the body of knowledge concerning the keyword method.

Length of the Experiment

The present study had two follow-up tests to measure students' retention of the acquired vocabulary. The results showed that the relative advantages of the keyword method over the rote method tended to get greater as time went on, even

though the number of words remembered continued to go down. It would be interesting to conduct experiments to measure effects of the keyword method covering more than three months to see whether there is further widening of the differences. The findings could be of great interest in regard to the claims of the keyword method with regard to retention. Even if the findings are negative, the study could make a contribution by finding the time spans involved in the widening of the differences.

Keyword Effect on Hard Words

The results of the experiment with the hard words in the study showed that, at the time of the first posttest, the scores of the rote method were slightly higher, but the two follow-up tests showed that the keyword method showed an advantage in the retention of hard words. The percentage of difference in scores between the keyword method and the rote method was greater with hard words than with the other words. However, the hard words used for the experiment consisted of only two (hen and rooster) or three (bull, cow, and ox) words. An experiment could be designed with a greater number of hard words, and with the definition of "hard" more carefully specified, to provide better data with which to ascertain the effects of the two methods. A valuable reference on hard words is Higa (1965).

Comparisons of the Verbal Keyword and the Imagery Keyword

Merry (1980) reported that the keyword with verbal images and the keyword with pictures show no significant difference. If mere verbal images are as good as

actual illustrations or pictures, it could save teachers the time and expense of preparing actual productions of the keyword illustrations, which is the hardest part for many teachers in using the keyword method. Consequently, empirical data on this aspect of the keyword method seem desirable.

Comparisons of the Keyword Method With Other Well-Known Methods in TESOL

The present study only compared the keyword method with the rote (traditional) method, but there are many other techniques of teaching foreign languages with which the keyword method could be compared. Some possible candidates for the comparison may be total physical response, language experience, communicative language teaching, or methods involving audio-visual materials. At more advanced stages, where students are able to use the available contexts effectively, the keyword method could be compared with the contextual approaches to teaching vocabulary. Advantages and disadvantages or limitations of the keyword method found in these comparisons would contribute to the body of empirical knowledge available in the field of vocabulary acquisition.

APPENDICES

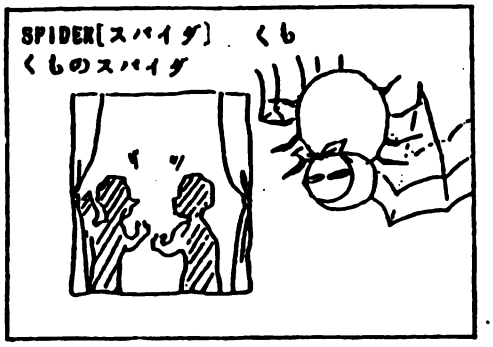
APPENDIX A

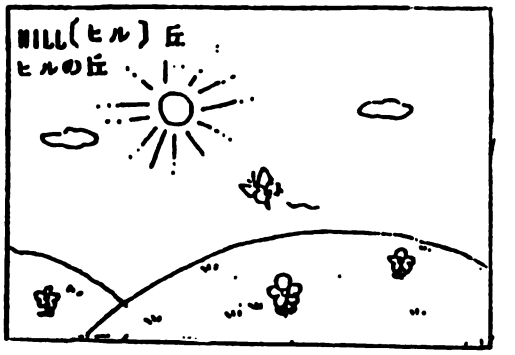
HANDOUTS

1. Explanation of the keyword method and the class activities

英単語の学習・記憶法について

皆さんは英単語を覚えるとき丸暗記していませんか。「ごろあわせ」を利用するとよく覚えられると言われていることを知っていますか。「ごろあわせ」とは3月3日を耳　(3・3ーミミ)の日、8月7日を鼻(8・7ーハナ)の日とするような覚え方です。この「ごろあわせ法」にイラストとして絵を加えると、さらに単語の記憶力がますことが知られています。では、次にこの「ごろあわせ」と絵をつかった英単語の学習と記憶法について、実例をふたつほどご紹介いたします。





上の例にあるように「くも(生物)」は、英語で「SPIDER[スパイダ]」といいます。それで、皆さんは「くも」という英単語を「くものスパイダ」という「ごろあわせ」と絵をみておぼえるようにします。

また「丘」という英語は、絵を見ながら、蝶のとんでいる真昼の丘を想像し「ヒル(HILL:昼)の丘」という「ごろあわせ」で覚えます。

次のページには、15の英単語の「ごろあわせ」と絵を使った覚え方があります。その15の単語を覚えながら、英単語を学ぶときに利用できる「ごろあわせ」の要領(ようりょう)を覚えて下さい。また、そのほかにも15の英単語を丸暗記で覚えてもらいます。最後に、どちらが実際に英単語をよく覚えられるかテストしてみます。

最初の丸暗記で覚える15語をクラスのさいしょの10分、次の「ごろあわせ」でおぼえる15語もまた10分で覚えてもらいます。最後に、今から覚える英単語30について10分間のテストを行います。

テストは下の例のように選択式ですから、特に単語のつづり(スペリング)を覚える必要はありません。

	答え		
例	い	1. BOY [ボイ]	あ、少女
	あ	2. GIRL [ガール]	い、小年

先生が「始め」と合図するまでは次のページを開いてはいけません。

Mamoru　Kinjo
Okinawa Kokusai University

2. Pretest

予備テスト　　中学校　年　組　氏名　　月　日

次のの英語の意味を知っていれば、その意味を表す日本語を答えの欄に書きこんで下さい。

テ　ス　ト　1

答え		
_____	1.	ASK [アスク]
_____	2.	BOTH [ボゥス]
_____	3.	CRY [クライ]
_____	4.	DAUGHTER [ドータ]
_____	5.	EARLY [アーリ]
_____	6.	ENJOY [エンチ"ョイ]
_____	7.	FORGET [ホォゲット]
_____	8.	GROW [グロウ]
_____	9.	HOLIDAY [ホリテ"ィ]
_____	10.	IMPORTANT [インポータント]
_____	11.	JULY [テ"ュライ]
_____	12.	KNOW [ノウ]
_____	13.	LAKE [レイク]
_____	14.	HEN [ヘン]
_____	15.	ROOSTER [ルースタ]

テ　ス　ト　2

答え		
_____	1.	MARCH [マーチ]
_____	2	NOSE [ノウズ]
_____	3.	OTHER [アザ]
_____	4.	PUT [プット]
_____	5.	QUICKLY [クウィックリ]
_____	6.	READY [レテ"イ]
_____	7.	SEND [センド]
_____	8.	TRY [トライ]
_____	9.	USE [ユーズ]
_____	10.	VISIT [ウ"ィス"ィット]
_____	11.	WANT [ワント]
_____	12.	YARD [ヤード]
_____	13.	BULL [ブル]
_____	14.	COW [カウ]
_____	15.	OX [オックス]

3. A group of words to learn by the rote method

グループ　A

次の１５の単語にはごろ合わせや絵はありません。今から１０分間で、いくつ覚えられるかためしてみて下さい。絵とごろ合わせで覚える１５の単語は次のページにあります。しかし、先生の指示があるまでは次のページに移ってはいけません。
カタカナの太い字は、その音が強く発音されることを表しています。

1. ASK［アスク］たずねる
2. BOTH［ボゥス］両方の
3. CRY［クライ］叫ぶ、泣く
4. DAUGHTER［ドータ］娘
5. EARLY［アーリ］早く
6. ENJOY［エンチ゛ョイ］楽しむ
7. FORGET［ホォゲット］忘れる
8. GROW［グロウ］成長する
9. HOLIDAY［ホリテ゛ィ］休日
10. IMPORTANT［インポータント］重要な
11. JULY［チ゛ュライ］7月
12. KNOW［ノウ］知っている
13. LAKE［レイク］湖
14. HEN［ヘン］めんどり
15. ROOSTER［ルレースタ］おんどり

4. A group of words to learn by the keyword method

次の１５の単語を、ごろあわせと絵を利用して今から１０分間で覚えて下さい。（ごろあわせのなかの英語は、「ヒルの丘」のようにカタカナで発音が書かれ、アンダーライン（下の棒線）が引かれています。カタカナの太い字は、その音が強く発音されることを表しています。）

5. A group of words to learn by the keyword method

6. Posttest

中学校　年　組 氏名　　月　日

次の英語と同じ意味になる日本語をさがし、その記号を答えの欄に記入して下さい。

テ　ス　ト　1

答え			
_____	1.	ASK [アスク]	イ．早く
_____	2.	BOTH [ボゥス]	ロ．7月
_____	3.	CRY [クライ]	ハ．忘れる
_____	4.	DAUGHTER [ドータ]	ニ．おんどり
_____	5.	EARLY [アーリ]	ホ．たずねる
_____	6.	ENJOY [エンチ゛ョイ]	ヘ．娘
_____	7.	FORGET [ホォゲット]	ト．知っている
_____	8.	GROW [グロウ]	チ．湖
_____	9.	HOLIDAY [ホリデ゛ィ]	リ．めんどり
_____	10.	IMPORTANT [インポータント]	ヌ．楽しむ
_____	11.	JULY [チ゛ュライ]	ル．両方の
_____	12.	KNOW [ノウ]	ヲ．叫ぶ、泣く
_____	13.	LAKE [レイク]	ワ．成長する
_____	14.	HEN [ヘン]	カ．休日
_____	15.	ROOSTER [ルースタ]	ヨ．重要な

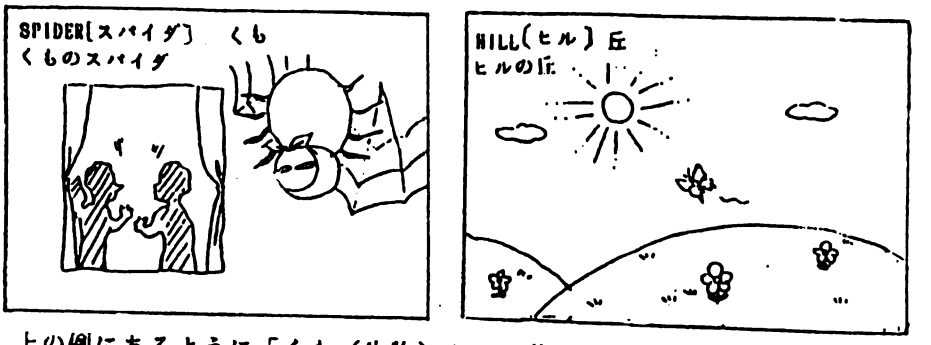
テ　ス　ト　2

答え			
_____	1.	MARCH [マーチ]	ア．庭
_____	2	NOSE [ノウズ]	イ．置く
_____	3.	OTHER [アザ]	ウ．雌牛（めうし）
_____	4.	PUT [プット]	エ．鼻
_____	5.	QUICKLY [クウィックリ]	オ．3月
_____	6.	READY [レデ゛イ]	カ．ためす
_____	7.	SEND [センド]	キ．雄牛（おうし）
_____	8.	TRY [トライ]	ク．他の
_____	9.	USE [ユーズ]	ケ．（食用、荷車用の）雄牛
_____	10.	VISIT [ウ゛ィス゛ィット]	コ．送る
_____	11.	WANT [ワント]	サ．すばやく
_____	12.	YARD [ヤード]	シ．ほしい
_____	13.	BULL [ブル]	ス．使う
_____	14.	COW [カウ]	セ．訪問する
_____	15.	OX [オックス]	ソ．準備ができた

1. Explanation of the keyword method and the class activities

英単語の学習・記憶法について

皆さんは英単語を覚えるとき丸暗記していませんか。「ごろあわせ」を利用するとよく覚えられると言われていることを知っていますか。「ごろあわせ」とは3月3日を耳　（3・3ーミミ）の日、8月7日を鼻（8・7ーハナ）の日とするような覚え方です。この「ごろあわせ法」にイラストとして絵を加えると、さらに単語の記憶力がますことが知られています。では、次にこの「ごろあわせ」と絵をつかった英単語の学習と記憶法について、実例をふたつほどご紹介いたします。



上の例にあるように「くも（生物）」は、英語で「SPIDER[スパイダ]」といいます。それで、皆さんは「くも」という英単語を「くものスパイダ」という「ごろあわせ」と絵をみておぼえるようにします。
また「丘」という英語は、絵を見ながら、蝶のとんでいる真昼の丘を想像し「ヒル（HILL：昼）の丘」という「ごろあわせ」で覚えます。

次のページには、１５の英単語の「ごろあわせ」と絵を使った覚え方 があります。その１５の単語を覚えながら、英単語を学ぶときに利用できる「ごろあわせ」の要領（ようりょう）を覚えて下さい。また、そのほかにも１５の英単語を丸暗記で覚えてもらいます。最後に、どちらが実際に英単語をよく覚えられるかテストしてみます。

最初の「ごろあわせ」で覚える１５語をクラスのさいしょの１０分、次の丸暗記でおぼえる１５語もまた１０分で覚えてもらいます。最後に、今から覚える英単語３０について１０分間のテストを行います。

テストは下の例のように選択式ですから、特に単語のつづり（スペリング）を覚える必要はありません。

例　答え
い　1. BOY［ボイ］　　あ、少女
あ　2. GIRL［ガール］　　い、小年

先生が「始め」と合図するまでは次のページを開いてはいけません。

Mamoru　Kinjo
Okinawa Kokusai University

2. Pretest

予備テスト　　中学校　年　組 氏名　　月　日

次のの英語の意味を知っていれば、その意味を表す日本語を答えの欄に書きこんで下さい。

テ　ス　ト　1

答え

_____1. ASK [アスク]
_____2. BOTH [ボゥス]
_____3. CRY [クライ]
_____4. DAUGHTER [ドータ]
_____5. EARLY [アーリ]
_____6. ENJOY [エンチ゛ョイ]
_____7. FORGET [ホォゲット]
_____8. GROW [グロウ]
_____9. HOLIDAY [ホリテ゛ィ]
_____10. IMPORTANT [インポータント]
_____11. JULY [チ゛ュライ]
_____12. KNOW [ノウ]
_____13. LAKE [レイク]
_____14. HEN [ヘン]
_____15. ROOSTER [ルースタ]

テ　ス　ト　2

答え

_____1. MARCH [マーチ]
_____2 NOSE [ノウズ]
_____3. OTHER [アザ]
_____4. PUT [プット]
_____5. QUICKLY [クウィックリ]
_____6. READY [レテ゛イ]
_____7. SEND [センド]
_____8. TRY [トライ]
_____9. USE [ユーズ]
_____10. VISIT [ウ゛ィス゛ィット]
_____11. WANT [ワント]
_____12. YARD [ヤード]
_____13. BULL [ブル]
_____14. COW [カウ]
_____15. OX [オックス]

3. A group of words to learn by the keyword method

グループ　A

次の１５の単語を、ごろあわせと絵を利用して今から１０分間で覚えて下さい。（ごろあわせのなかの英語は、「ヒルの丘」のようにカタカナで発音が書かれ、アンダーライン（下の棒線）が引かれています。カタカナの太い字は、その音が強く発音されることを表しています。）

① ASK［アスク］たずねる
アスク（明日来）るかたず"ねる

⑤ EARLY［アーリ］早く
アーリ（蟻）　は早く

② BOTH［ボゥス］両方の
両方の　ボス　だ

⑥ ENJOY［エンチ"ョイ］楽しむ
エンジョイ（援助い）いなと楽しむ

③ CRY［クライ］叫ぶ、泣く
クライ（暗い）と叫び、泣く

⑦ FORGET［ホォゲット］忘れる
ボケーット　して忘れる

④ DAUGHTER［ドータ］娘
ドタ、ドタ　はしる娘

⑪ GROW［グロウ］成長する
気グロー（苦労）で成長する

4. A group of words to learn by the keyword method

5. A group of words to learn by the rote method

グループ B

次の１５の単語にはごろ合わせや絵がありません。今から１０分間で、絵やごろ合わせがある時と同じように覚えられるか試してみて下さい。

1、 MARCH [マーチ] ３月

2、 NOSE [ノウズ] 鼻

3、 OTHER [アザ] 他の

4、 PUT [プット] 置く

5、 QUICKLY [クウィックリ] すばやく

6、 READY [レデ"イ] 準備ができた

7、 SEND [センド] 送る

8、 TRY [トライ] ためす

9、 USE [ユーズ] 使う

10、 VISIT [ウ"ィス"ィット] 訪問する

11、 WANT [ワント] ほしい

12、 YARD [ヤード] 庭

13、 BULL [ブル] 雄牛（おうし）

14、 COW [カウ] 雌牛（めうし）

15、 OX [オックス] （食用、荷車用の）雄牛

6. Posttest

中学校　年　組 氏名　　月　日

次の英語と同じ意味になる日本語をさがし、その記号を答えの欄に記入して下さい。

テ　ス　ト　1

答え			
_____	1.	ASK [アスク]	イ．早く
_____	2.	BOTH [ボゥス]	ロ．7月
_____	3.	CRY [クライ]	ハ．忘れる
_____	4.	DAUGHTER [ドータ]	ニ．おんどり
_____	5.	EARLY [アーリ]	ホ．たずねる
_____	6.	ENJOY [エンチ゛ョイ]	ヘ．娘
_____	7.	FORGET [ホォゲット]	ト．知っている
_____	8.	GROW [グロウ]	チ．湖
_____	9.	HOLIDAY [ホリデ゛ィ]	リ．めんどり
_____	10.	IMPORTANT [インポータント]	ヌ．楽しむ
_____	11.	JULY [チ゛ュライ]	ル．両方の
_____	12.	KNOW [ノウ]	ヲ．叫ぶ、泣く
_____	13.	LAKE [レイク]	ワ．成長する
_____	14.	HEN [ヘン]	カ．休日
_____	15.	ROOSTER [ルースタ]	ヨ．重要な

テ　ス　ト　2

答え			
_____	1.	MARCH [マーチ]	ア．庭
_____	2	NOSE [ノウズ]	イ．置く
_____	3.	OTHER [アザ]	ウ．雌牛（めうし）
_____	4.	PUT [プット]	エ．鼻
_____	5.	QUICKLY [クウィックリ]	オ．3月
_____	6.	READY [レデ゛イ]	カ．ためす
_____	7.	SEND [センド]	キ．雄牛（おうし）
_____	8.	TRY [トライ]	ク．他の
_____	9.	USE [ユーズ]	ケ．（食用、荷車用の）雄牛
_____	10.	VISIT [ウ゛ィズ゛ィット]	コ．送る
_____	11.	WANT [ワント]	サ．すばやく
_____	12.	YARD [ヤード]	シ．ほしい
_____	13.	BULL [ブル]	ス．使う
_____	14.	COW [カウ]	セ．訪問する
_____	15.	OX [オックス]	ソ．準備ができた

Explanation of the Keywords

In the following pages is an explanation of the keywords and the phrases/sentences that were used to assist students' remembering of the 30 words tested.

Abbreviations:
K = Keyword
JP = Japanese phrases/sentences
M = Meanings of the phrases/sentences

The following is the format of explanation. First, for each word (1-30), the keyword(s) is (are) given (K: _____).

Second, the Japanese phrases/sentences that incorporate associations of the pronunciations of the English words and the keywords are given in Roman characters, under which English equivalents appear (JP: _____).
English equivalent

Third, the translations or meanings of the phrases/sentences are provided (M: _____).

1. ASK

K: ASU (tomorrow) & KURU (come)

JP:	ASU	KURU	KA	TAZUNERU
	tomorrow	come	if	ASK

M: Ask (if she) comes tomorrow.

2. BOTH

K: BOSU (boss)

JP:	RYOHOU	NO	BOSU	DA
	BOTH	of	boss	are

M: (They) are the bosses of both (groups).

3. CRY

K: KURAI (dark)

JP: KURAI TO NAKU
darkness due to CRY

M: Cry due to dark/ness.

4. DAUGHTER

K: DOTA DOTA (noisily)

JP: DOTA DOTA HASHIRU MUSUME
noisily run DAUGHTER

M: Daughter galloping.

5. EARLY

K: ARI (ant)

JP: ARI WA HAYAKU
ant EARLY

M: Ants (come) early.

6. ENJOY

K: ENJYO (allowance, aid)

JP: ENJYO IINA TO TANOSHIMU
allowance good ENJOY

M: Enjoy good allowance.

7. FORGET

K: BOKETTO (absent-mindedly)

JP: BOKETTO SHITE WASURERU
absent-mindedly FORGET

M: Forget absent-mindedly

8. GROW

K: KIGUROU (anxieties)

JP: KIGUROU DE SODATSU
anxieties by GROW

M: Grow by (overcoming) anxieties.

9. HOLIDAY

K: HORI (dig) and DEII (OK)

JP: KYUJITSU WA (ANA) HORI DEII
HOLIDAY hole dig OK

M: (As for) holiday, digging (a hole) is OK.

10. IMPORTANT

K: IPPON (one) & TANTOU (short sword)

JP: JYUYOUNA IPPON TANTOU
IMPORTANT one short sword

M: One important short sword.

11. JULY

K: JYURAI (up to now)

JP: TANABAT WA JYURAI SHICHI GATSU
the Star Festival has been up to now in JULY

M: The Star Festival (July 7) has been in July up to now.

12. KNOW

K: NOU (the brain)

JP: NOU WA SHITTEIRU
the brains KNOW

M: The brains know.

13. LAKE

K: LEI (spirit, ghost) KURU (come)

JP: LEI KURU IZUMI
the spirit come LAKE

M: The lake that the spirits visit.

14. HEN

K: HEN (strange) & EHEN (hem)

JP: HEN TO NAKU HENNA MENDORI
hem strange HEN

M: A strange hen hemming "Hen!"

15. ROOSTER

K: DOSHITA (what happened to)

JP: ONDORI WA DOSUTA
rooster what happened to

M: What happened to the rooster?

1. MARCH

K: MACHIMASU (wait for)

JP: SAN-GATSU MACHIMASU
MARCH (month) wait for

M: (We'll) wait for March.

2. NOSE

K: NO (of) & ZU (picture)

JP: HANA NO ZU
NOSE of picture

M: Pictures of noses.

3. OTHER

K: AZANA (nickname)

JP: HOKA NO AZANA
OTHER of nickname

M: The other nickname.

4. PUT

K: POTTO (hastily)

JP: POTTO OKU
hastily PUT

M: Put hastily.

5. QUICKLY

K: KUI (eat) & KURI (chestnut)

JP: SUBAYAKU KUI KURI
QUICKLY eat chestnuts

M: Eat chestnuts quickly.

6. READY

K: REDII (a woman)

JP: LEDII WA JYUNBIGA DEKITA
lady ready

M: The lady is ready.

7. SEND

K: SEN (one thousand) & DORU (dollars)

JP: SEN DORU OKURU
one thousand dollars SEND

M: Send one thousand dollars.

8. TRY

K: TORA (tiger) & IRU (is there)

JP: TORA IRU KA TAMESU
the tiger is there if TRY

M: Try if the tiger is there.

9. USE

K: YUZURU (give, share)

JP: YUZUTTE TSUKAU
share USE

M: Share to use.

10. VISIT

K: BIJIN (a beauty)

JP: BIJIN TO HOUMONSURU
beauty with VISIT

M: Visit with a beauty.

11. WANT

K: WAN (bowl) & TO (and)

JP: WAN TO HASHI GA HOSHII
bowl and chopsticks WANT

M: Want a bowl and chopsticks.

12. YARD

K: YADO (hotel)

JP: YADO NO NIWA
hotel of YARD

M: The yard of the hotel.

13. BULL

K: ERABURU (bossy)

JP: ERABURU OUSHI
bossy BULL

M: Bossy bull.

14. COW

K: KAU (buy)

JP: MIRUKU TO MEUSHI O KAU
milk and cow BUY

M: Buy milk and a cow.

15. OX

K: OOKUSU (increase)

JP: (SHOKUYOU, NIGURUMAYOU NO)
(for meat, for transportation) of

OUSHI O OOKUSU
male bovine animal increase

M: Increase male bovine animals for meat and transportation.

APPENDIX B

THE RESULTS OF THE T-TEST

Overall

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KPOST	390	.649	.000	11.1103	4.145	.210
RPOST				10.4000	4.306	.218

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
.7103	3.546	.180	3.96	389	.000
95% CI (.357, 1.063)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KFOLL	390	.621	.000	5.7744	3.764	.191
RFOLL				4.4462	3.203	.162

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
1.3282	3.077	.156	8.53	389	.000
95% CI (1.022, 1.635)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KDIF	390	.342	.000	5.3359	3.697	.187
RDIF				5.9538	3.727	.189

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
-.6179	4.259	.216	-2.87	389	.004
95% CI (-1.042, -.194)					

One-month follow-up

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KPOST	187	.602	.000	11.2460	3.940	.288
RPOST				10.1176	4.314	.315

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
1.1283	3.696	.270	4.17	186	.000
95% CI (.595, 1.662)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KFOLL	187	.680	.000	5.7861	3.827	.280
RFOLL				4.5241	3.391	.248

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
1.2620	2.913	.213	5.92	186	.000
95% CI (.842, 1.682)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KDIF	187	.330	.000	5.4599	3.833	.280
RDIF				5.5936	3.699	.271

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
-.1337	4.360	.319	-.42	186	.675
95% CI (-.763, .495)					

Two-month follow-up

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KPOST	203	.696	.000	10.9852	4.332	.304
RPOST				10.6601	4.292	.301

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
.3251	3.364	.236	1.38	202	.170
95% CI (-.141, .791)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KFOLL	203	.558	.000	5.7635	3.716	.261
RFOLL				4.3744	3.026	.212

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
1.3892	3.226	.226	6.14	202	.000
95% CI (.943, 1.836)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
KDIF	203	.363	.000	5.2217	3.573	.251
RDIF				6.2857	3.731	.262

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
-1.0640	4.124	.289	-3.68	202	.000
95% CI (-1.635, -.493)					

APPENDIX C

THE DATA ON THE 12 CLASSES

The posttest (overall)

- - Description of Subpopulations - -

Summaries of KPOST
By levels of CLASS

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			11.1103	4.1453	390
CLASS	1		11.7143	4.2588	28
CLASS	2		9.8000	4.7879	30
CLASS	3		9.7586	3.5421	29
CLASS	4		11.9677	3.9957	31
CLASS	5		10.8947	4.9472	38
CLASS	6		11.2778	3.9829	36
CLASS	7		10.7241	4.7876	29
CLASS	8		7.1786	4.3890	28
CLASS	9		12.0286	3.4682	35
CLASS	10		11.9412	3.0445	34
CLASS	11		12.2973	3.2047	37
CLASS	12		12.6571	2.5775	35

Total Cases = 390

- - Description of Subpopulations - -

Summaries of RPOST
By levels of CLASS

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			10.4000	4.3057	390
CLASS	1		9.8571	4.0980	28
CLASS	2		10.4000	4.4613	30
CLASS	3		8.7241	5.1610	29
CLASS	4		8.9677	4.2464	31
CLASS	5		11.6579	3.8923	38
CLASS	6		10.5556	4.2456	36
CLASS	7		10.0000	4.5826	29
CLASS	8		7.8571	5.0164	28
CLASS	9		11.9714	2.8126	35
CLASS	10		10.4118	4.4389	34
CLASS	11		11.4595	3.2197	37
CLASS	12		11.6286	4.0663	35

Total Cases = 390

The follow-ups (overall)

- - Description of Subpopulations - -

Summaries of KFOLL
By levels of CLASS

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			5.7744	3.7643	390
CLASS	1		4.0357	3.3387	28
CLASS	2		4.4333	3.5592	30
CLASS	3		4.6207	3.6195	29
CLASS	4		5.9355	4.0409	31
CLASS	5		5.0000	3.6981	38
CLASS	6		5.3056	3.2586	36
CLASS	7		6.3448	3.4772	29
CLASS	8		4.0357	4.1854	28
CLASS	9		7.4000	2.9921	35
CLASS	10		7.6176	3.9236	34
CLASS	11		7.3243	3.5360	37
CLASS	12		6.3143	3.5378	35

Total Cases = 390

- - Description of Subpopulations - -

Summaries of RFOLL
By levels of CLASS

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			4.4462	3.2028	390
CLASS	1		2.8929	2.9732	28
CLASS	2		4.3667	3.9347	30
CLASS	3		4.2759	3.7502	29
CLASS	4		4.0645	3.0543	31
CLASS	5		4.0789	2.9074	38
CLASS	6		3.4167	2.4187	36
CLASS	7		4.8966	3.5591	29
CLASS	8		3.2857	2.8785	28
CLASS	9		5.5429	3.3549	35
CLASS	10		5.5882	2.6756	34
CLASS	11		5.0270	2.9673	37
CLASS	12		5.4286	2.9930	35

Total Cases = 390

APPENDIX D

THE DATA ON STUDENTS WITH PERFECT SCORES ON EITHER HALF OF THE IMMEDIATE POSTTEST OR BOTH TESTS

Data for the 101 students who received perfect scores related to the keyword method.

Variable	No. of Cases	Mean	Std. Dev.	Std. Error	Difference			t-Value	df	2-tail Prob.
					Mean	Std. Dev.	Std. Error			
KPOST	101	15.0000	.000	.000	2.4158	2.725	.271	8.91	100	.000
RPOST		12.5842	2.725	.271						
Follow-up 1 Mo.										
KPOST	47	15.0000	.000	.000	2.7447	2.915	.425	6.45	46	.000
RPOST		12.2553	2.915	.425						
KFOLL[a]	47	6.9574	3.635	.530	1.4468	2.850	.416	3.48	46	.001
RFOLL[a]		5.5106	3.085	.450						
Follow-up 2 Mo.										
KPOST	54	15.0000	.000	.000	2.1296	2.541	.346	6.16	53	.000
RPOST		12.8704	2.541	.346						
KFOLL[b]	54	7.5926	3.189	.434	1.3519	3.546	.483	4.87	53	.000
RFOLL[b]		5.2407	3.126	.425						

[a]Correlation = .651; 2-tail probability = .000.

[b]Correlation = .370; 2-tail probability = .006.

Data for the 77 students who received perfect scores related to the rote method.

Variable	No. of Cases	Mean	Std. Dev.	Std. Error	Difference			t-Value	df	2-tail Prob.
					Mean	Std. Dev.	Std. Error			
KPOST	77	13.4805	2.075	.236	-1.5195	2.075	.236	-6.43	76	.000
RPOST		15.0000	.000	.000						
Follow-up 1 Mo.										
KPOST	30	13.0333	2.539	.464	-1.9667	2.539	.464	-4.24	29	000
RPOST		15.0000	.000	.000						
KFOLL[a]	30	8.1333	3.224	.589	1.8333	2.520	.460	3.98	29	000
RFOLL[a]		6.3000	2.744	.501						
Follow-up 2 Mo.										
KPOST	47	13.7660	1.684	.246	-1.2340	1.684	.246	-5.02	46	000
RPOST		15.0000	.000	.000						
KFOLL[b]	47	6.9574	3.310	.483	1.5532	3.374	.492	3.16	46	.003
RFOLL[b]		5.4043	2.939	.429						

[a]Correlation = .654; 2-tail probability = .000.

[b]Correlation = .422; 2-tail probability = .003.

Data for the 33 students who received perfect scores related to the both methods

Variable	No. of Cases	Mean	Std. Dev.	Std. Error	Difference			t-Value	df	2-tail Prob
					Mean	Std. Dev.	Std. Error			
KPOST	33	15.0000	.000	.000	.0000	.000	.000	.00	32	1.000
RPOST		15.0000	.000	.000						
Follow-up 1 Mo.										
KPOST	11	15.0000	.000	.000	.0000	.000	.000	.00	10	1.000
RPOST		15.0000	.000	.000						
KFOLL[a]	11	8.7273	3.771	1.137	2.3636	2.580	.778	3.04	10	.012
RFOLL[a]		6.3636	2.976	.897						
Follow-up 2 Mo.										
KPOST	22	15.0000	.000	.000	.0000	.000	.000	.00	21	1.000
RPOST		15.0000	.000	.000						
KFOLL[b]	22	7.1364	2.949	.629	2.1818	2.702	.576	3.79	21	.001
RFOLL[b]		4.9545	2.497	.532						

[a]Correlation = .732; 2-tail probability = .010.

[b]Correlation = .518; 2-tail probability = .013.

APPENDIX E

THE DATA ON HARD WORDS

Overall

- - - t-tests for paired samples - - -

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKPOST	390	.340	.000	.6829	.407	.021
HWRPOST				.7440	.378	.019

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
-.0611	.451	.023	-2.67	389	.008
95% CI (-.106, -.016)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKFOLL	390	.302	.000	.3564	.394	.020
HWRFOLL				.2714	.337	.017

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
.0850	.435	.022	3.86	389	.000
95% CI (.042, .128)					

One-month follow-up

- - - t-tests for paired samples - - -

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKPOST	187	.331	.000	.7130	.379	.028
HWRPOST				.7264	.400	.029

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
-.0134	.451	.033	-.41	186	.685
95% CI (-.078, .052)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKFOLL	187	.304	.000	.3690	.398	.029
HWRFOLL				.2968	.371	.027

Paired Differences Mean	SD	SE of Mean	t-value	df	2-tail Sig
.0722	.454	.033	2.17	186	.031
95% CI (.007, .138)					

Two-month follow-up

- - - t-tests for paired samples - - -

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKPOST	203	.362	.000	.6552	.430	.030
HWRPOST				.7603	.356	.025

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
-.1051	.448	.031	-3.34	202	.001
95% CI (-.167, -.043)					

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
HWKFOLL	203	.299	.000	.3448	.391	.027
HWRFOLL				.2479	.303	.021

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
.0969	.417	.029	3.31	202	.001
95% CI (.039, .155)					

APPENDIX F

LETTER OF APPROVAL FROM UCRIHS AND

PRINCIPAL CONSENT FORM

MICHIGAN STATE UNIVERSITY

OFFICE OF VICE PRESIDENT FOR RESEARCH
AND DEAN OF THE GRADUATE SCHOOL

EAST LANSING • MICHIGAN • 48824-1046

April 17, 1991

Mamoru Kinjo
E-714 Owen Hall

RE: AN INVESTIGATION INTO THE KEYWORD EFFCTIVENESS AMONG JAPANESE SEVENTH GRADERS IN LEARNING A FOREIGN LANGUAGE VOCABULARY, IRB#91-176

Dear Mr. Kinjo:

The above project is exempt from full UCRIHS review. The proposed research protocol has been reviewed by another committee member. The rights and welfare of human subjects appear to be protected and you have approval to conduct the research.

You are reminded that UCRIHS approval is valid for one calendar year. If you plan to continue this project beyond one year, please make provisions for obtaining appropriate UCRIHS approval one month prior to April 15, 1992.

Any changes in procedures involving human subjects must be reviewed by UCRIHS prior to initiation of the change. UCRIHS must also be notifed promptly of any problems (unexpected side effects, complaints, etc.) involving human subjects during the course of the work.

Thank you for bringing this project to my attention. If I can be of any future help, please do not hesitate to let me know.

Sincerely,

David E. Wright, Ph.D., Chair
University Committee on Research Involving
Human Subjects (UCRIHS)

DEW/deo

cc: Dr. Paul Munsell

Safety of students involved in the study

The techniques and purpose of this statistical study on vocabulary learning and retention/acquisition by Professor Mamoru Kinjo at Okinawa Kokusai University have been explained to me. I recognise that the activities involved in this study are safe and are not incompatible with normal classroom procedures.

研究教材クラス導入の児童生徒への安全性について

沖縄国際大学の金城守講師による「英単語の学習・記憶法について」の研究教材のクラス導入について、氏よりその目的、方法、手順等の説明がありました。 それらは通常の教育、指導活動の学習過程の範囲、又はその延長線上にあるものと見なされ、児童生徒の心身に危害を加えるものではないことを承認いたします。

氏名 _____ 中学校 職名 _____ 月 日

NAME SCHOOL TITLE Principal DATE

Mamoru Kinjo

金城 守

沖縄国際大学

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen, V. F. (1983). *Techniques in teaching vocabulary.* Oxford: Oxford University Press.

Ard, J., & Gass, S. M. (1987). Lexical constraints on syntactic acquisition. *Studies in Second Language Acquisition, 9,* 233-252.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist, 30,* 821-828.

Blair, R. W. (1982). Easification. In R. W. Blair (Ed.), *Innovative approaches to language teaching* (pp. 219-225). Rowley, MA: Newbury House Publishers.

Bond, O. F., Babcock, J., Norman, H., & Andrus, L. (1943). Teaching the romance languages: French, Spanish, Italian. *Review of Educational Research, 13,* 142-161.

Brown, R., & McNeill, D. (1966). The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behaviour, 5,* 325-377.

Carter, R. A. (1987). *Vocabulary.* London: Allen & Unwin.

Carter, R., & McCarthy, M. J. (Eds.). (1988). *Vocabulary and language teaching.* London: Longman.

Channell, J. (1988). Psycholinguistic considerations in the study of L2 vocabulary acquisition. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 83-96). London: Longman.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Cohen, A. D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition, 9,* 43-62.

Cohen, A. D. (1990. *Language learning: Insights for learners, teachers, and researchers*. New York: Newbury House.

Cohen, A., & Aphek, E. (1980). Retention of second language vocabulary over time: Investigating the role of mnemonic associations. *System, 8,* 221-235.

Cohen, A., & Aphek, E. (1981). Easifying second language learning. *Studies in Second Language Acquisition, 3,* 221-236.

Cohen, A., & Aphek, E. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition, 9.1*, Cambridge University Press.

Cohen, A. D., & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. *Language Learning, 31,* 285-313.

Cowie, A. P. (1988). Stable and creative aspects of vocabulary use. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 126-139). London: Longman.

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

Curtis, M. E. (1987). Vocabulary testing and vocabulary instruction. In M. C. McKeown & M. E. Curtis (Eds.), *Nature of vocabulary acquisition* (pp. 37-51). Hillsdale, NJ: Erlbaum.

Desrochers, A., & Begg, I. (1987). A theoretical account of encoding and retrieval processes in the use of imagery-based mnemonic techniques: The special case of the keyword method. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 56-77). New York: Springer-Verlag.

Desrochers, A., Gelinas, C., & Wieland, L. D. (1989). An application of the mnemonic keyword method to the acquisition of German nouns and the grammatical gender. *Journal of Educational Psychology, 81,* 25-33.

Dretzke, B. J., & Levin, J. R. (1990). Building factual knowledge about the U.S. presidents via pictorial mnemonic strategies. *Contemporary Educational Psychology, 15,* 152-169.

Estabrooks, G. H. (1927). A handy memory trick. *Journal of Genetic Psychology, 34,* 615-619.

Fenaigle, G. von. (1813). *The new art of memory.* London: Sherwood, Neely, & Jones.

Fries, C. C. (1945). *Teaching and learning English as a foreign language.* Ann Arbor: University of Michigan Press.

Fuentes, E. J. (1976). An investigation into the use of imagery and generativity in learning a foreign language vocabulary (Doctoral dissertation, Stanford University). *Dissertation Abstracts International, 37,* 2694A. (University Microfilms No. 76-25, 997)

Furst, B. (1972). *Stop forgetting.* New York: Doubleday.

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9,* 198-217.

Gass, S. M. (1989). Vocabulary in a second langauge. *Studies in Second Language Acquisition, 11,* 121-122.

Gorfein, D. S., & Hoffman, R. R. (1987). *Memory and learning: The Ebbinghaus Centennial Conference.* Hillsdale, NJ: Erlbaum.

Hall, J. W., & Fuson, K. C. (1986). Presentation rates in experiments on mnemonics: A methodological note. *Journal of Educational Psychology, 78,* 233-234.

Hall, J. W., & Fuson, K. C. (1988). The keyword method and presentation rates: Reply to Pressley (1987). *Journal of Educational Psychology, 80,* 251-252.

Higa, M. (1965). The psycholinguistic concept of "difficulty" and the teaching of foreign language vocabulary. *Language Learning, 15,* 167-179.

Horowitz, L. M., White, M. A., & Atwood, D. W. (1968). Word fragments as aids to recall: The organization of a word. *Journal of Experimental Psychology, 76,* 219-226.

Isgur, J. (1975. Establishing letter-sound associations by an object-imaging-projection method. *Journal of Learning Disabilities, 8,* 16-20.

Julliot, Ch.-L. (1919). *Lëducation de la mémoire.* Paris: Flammarion.

Kelly, L. G. (1969). *Twenty-five centuries of language teaching.* Rowley, MA: Newbury House.

Konopak, B., & Williams, N. L. (1988). Using the keyword method to help young readers learn content material. *Reading Teacher, 41,* 682-686.

Lado, R. (1955). Patterns of difficulty in vocabulary. *Language Learning, 6,* 23-41.

Laubach, F. C., Kirk, E. M., & Laubach, R. S. (1971). *The new streamlined English series: Teacher's manual.* Syracuse, NY: New Readers Press.

Levin, J. R., Johnson, D. D., Pittelman, S. D., Hayes, B. L., Levin, K. M., Shriberg, L. K., & Toms-Bronowski, S. (1984). A comparison of semantic- and mnemonic-based vocabulary-learning strategies. *Reading Psychology, 5,* 1-15.

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.

Levin, J. R., Pressley, M., McCormick, C. B., Miller, G. E., & Shriberg, L. K. (1979). Assessing the classroom potential of the keyword method. *Journal of Educational Psychology, 71,* 583-594.

Liu, I.-M. (1978). Appraisal of the keyword method in second-language learning. *Proceedings of the National Science Council (Taiwan), 2,* 437-442.

Loisette, A. (1895). *Memory: A scientific practical method of cultivating the faculties of attention, recollection and retention.* New York: Funk & Wagnalls.

Lorayne, H. (1974). *How to develop a super-power memory.* New York: Signet.

Martin, C. J. (1978). Mediatational processes in the retarded: Implications for teaching reading. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 9, pp. 61-84). New York: Academic Press.

McDaniel, M. A., & Pressley, M. (1984). Putting the keyword method in context. *Journal of Educational Psychology, 76,* 598-609.

McDaniel, M. A., & Tillman, V. P. (1985). *The keyword method versus learning from context: Where's the link?* Manuscript submitted for publication, Department of Psychology, University of Notre Dame, South Bend, IN.

McDaniel, M., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. *Journal of Educational Psychology, 81,* 196-207.

McKeown, M. G., & Curtis, M. E. (Eds.). (1987). *The nature of vocabulary acquisition.* Hillsdale, NJ: Erlbaum.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts, 13,* 221-246.

Meara, P. (1982). Word associations in a foreign language: A report on the Birkbeck Vocabulary Project. *Nottingham Linguistic Circular, 11*(2), 29-38.

Meara, P. M. (Ed.). (1983). *Vocabulary in a second language* (Vol. 1). London: Center for Information on Language Teaching and Research.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, A. Howart, & C. Criper (Eds.), *Interlanguage* (pp. 225-235). Edinburgh: Edinburgh University Press.

Meara, P. (Ed.). (1987). *Vocabulary in a second language* (Vol. 2). London: Center for Information on Language Teaching and Research.

Merry, R. (1980). Keyword method and children's vocabulary learning in the classroom. *British Journal of Educational Psychology, 50,* 123-136.

Miller, G. E., Levin, J. R., & Pressley, M. (1980). An adaptation of the keyword method to children's learning of foreign verbs. *Journal of Mental Imagery, 4,* 57-61.

Ministry of Education, Japan. (1977). *Chyugakkou gakushyu shidou youriyou [Guidelines for junior high schools],* 102-107.

Munsell, P. E., Rauen, M. G., & Kinjo, M. (1988). Review essay: Comprehensive survey of recent conclusions. *Language Learning, 38*(2), 261-278.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York: Newbury House.

Nattinger, J. (1988). Some current trends in vocabulary teaching. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 62-82). London: Longman.

Ogden, C. K. (1930). *Basic English: A general introduction.* London: Kegan Paul, Trench & Trubner.

Ott, C. E., Butler, D. C., Blake, R. S., & Ball, J. P. (1973). The effect of interactive-image elaboration on the acquisition of foreign language vocabulary. *Language Learning, 2,* 197-206.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76,* 241-263.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, & Winston.

Paivio, A., & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology, 73,* 780-795.

Peters, A. (1983). *The units of language acquisition.* Cambridge, NY: Cambridge University Press.

Peterson, P. L., & Swing, S. R. (1983). Problems in classroom implementation of cognitive strategy instruction. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Educational applications* (pp. 267-287). New York: Springer.

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. (1987). Are keyword method effects limited to slow presentation rates? An empirically based reply to Hall and Fuson (1986). *Journal of Educational Psychology, 79,* 333-335.

Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1984). Memory strategy instruction is made of this: Metamemory and durable strategy use. *Educational Psychologist, 19,* 94-107.

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. (1977). Developmental differences in subjects' associative learning strategies and performance: Assessing a hypothesis. *Journal of Experimental Child Psychology, 24,* 431-439.

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., & Chatala, E. S. (1984). Memory strategy monitoring in adults and children. *Journal of Verbal learning and Verbal Behavior, 23,* 270-288.

Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research, 52,* 61-92.

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., & McCormick, C. B. (1980). Young children's learning of foreign language vocabulary: A sentence variation of the keyword method. *Contemporary Educational Psychology, 5,* 22-29.

Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. C. McKeown & M. E. Curtis (Eds.), *Nature of vocabulary acquisition* (pp. 107-127). Hillsdale, NJ: Erlbaum.

Pressley, M., McDaniel, M. A., Turnure, J. E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effects on intentional and incidental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 291-300.

Pressley, M., Ross, K. A., Levin, J. R., & Ghatala, E. S. (1984). The role of strategy utility knowledge in children's strategy decision making. *Journal of Experimental Child Psychology, 38,* 491-504.

Pressley, M., Samuel, J., Hershey, M. M., Bishop, S. L., & Dickinson, D. (1981). Use of a mnemonic technique to teach young children foreign language vocabulary. *Contemporary Educational Psychology, 6,* 110-116.

Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L., & Turnure, J. E. (1988). Elaborative interrogation facilitates acquisition of confusing facts. *Journal of Educational Psychology, 80,* 268-278.

Pressley, M., Woloshyn, V., Lysynchuk, L. M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review, 2,* 1-58.

Raugh, M. R., & Atkinson, R. C. (1975). A mnemonic method for learning a second language vocabulary. *Journal of Educational Psychology, 67,* 1-16.

Richards, I. A. (1943). *Basic English and its use.* London: Kegan Paul.

Rohwer, W. D., Jr. (1973). Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 8). New York: Academic Press.

Rosenheck, M. B., Levin, M. E., & Levin, J. R. (1989). Learning botany concepts mnemonically: Seeing the forest and the trees. *Journal of Educational Psychology, 81,* 196-203.

Sapir, E. (1921). *Language.* New York: Harcourt, Brace, & Co.

Sayer, T. A. (1877). *Aids to memory.* London: Daldy, Isbister & Co.

Scruggs, T. E., & Mastropieri, M. A. (1989). Reconstructive elaborations: A model for content area learning. *American Educational Research Journal, 26,* 311-327.

Shepard, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156-163.

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*(2), 238-247.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single trial learning of 2560 visual stimuli. *Psychonomic Science, 19,* 73-74.

Stevick, E. W. (1976). *Memory, meaning and method.* Rowley, MA: Newbury House.

Stevick, E. W. (1986). *Images and options in the language classroom.* Cambridge, UK: Cambridge University Press.

Stevick, E. W. (1989). *Success with foreign languages: Seven who achieved it and what worked for them.* Englewood Cliffs, NJ: Prentice-Hall.

Thorndike, R. L. (1973). *Reading comprehension education in fifteen countries.* New York: Wiley.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Journal of Experimental Psychology, 80,* 352-373.

West, M. P. (1953). *A general service list of English words.* London: Longman.

Wilkins, D. A. (1972). *Linguistics and language teaching.* London: Edward Arnold.

Willerman, B., & Melvin, B. (1979). Reservations about keyword mnemonic. *Canadian Modern Language Review, 35,* 443-453.

Xue-G-Y., & Nation, P. (1984). A university word list. *Language Learning and Communication, 3*(2), 215-229.

Yates, F. A. (1966). *The art of memory.* Chicago: University of Chicago Press.