The Role of Mediation in Knowledge Acquisition and Retention: Learning Foreign Vocabulary Using the Keyword Method

by

Robert J. Crutcher

Institute of Cognitive Science

University of Colorado
Campus Box 345
Boulder, CO 80309-0345

ICS Technical Report 90-10

THE ROLE OF MEDIATION IN KNOWLEDGE ACQUISITION AND RETENTION: LEARNING FOREIGN VOCABULARY USING THE KEYWORD METHOD

by

Robert J. Crutcher

Institute of Cognitive Science
University of Colorado, Boulder
1990

Abstract

The role of mediation in acquiring and retrieving foreign vocabulary definitions was examined in two experiments. Subjects learned Spanish vocabulary using the keyword method (Atkinson & Raugh, 1975), a mnemonic technique for learning foreign-English pairs by relating the foreign word to a keyword, a concrete English word similiar in sound to the foreign word, and then relating the keyword to the English equivalent by creating an interactive image between the keyword and English equivalent.

Previous theoretical accounts concerning the role of mediation in memory suggest two general retrieval models to describe how the English equivalents might be retrieved from the Spanish words. According to a mediation model, mediators used to learn new material (e.g. the keywords) play a central role in the retrieval process: to retrieve new information from memory, the mediatiors used to learn the information must be accessed first and then used to retrieve the new material. According to a direct model, mediators are not thought to play an important role in the retrieval process. Thus, the English equivalents are accessed via a direct link or simple association between the Spanish word and its English equivalent.

Experiment 1 showed that performance decrements in retrieving the vocabulary definitions related to failure to retrieve the keyword from the Spanish word and/or failure to retrieve the English equivalent from the keyword, supporting the mediation model. Experiment 2 strengthened the case for the mediation model by showing that practice retrieving the vocabulary definitions improved not only retrieval of the definitions but retrieval of the mediatiors as well. In both experiments, additional evidence for retrieval of keywords in recalling vocabulary definitions was provided by retrospective verbal reports. Despite strong support for the mediation model in general, some results in the second experiment suggested that a simple mediation model (i.e. keyword and interactive image always retrieved into working memory) may be insufficient to completely account for continued retrieval of the English equivalents. The implications of these results for research on the keyword method and mediation are discussed.

ACKNOWLEDGEMENTS

I would like to thank the many individuals who have supported and helped me in completing this project. First, tremendous thanks to my advisors, Anders Ericsson and Gary Bradshaw, for their insight and support throughout all phases of the project. They managed to skillfully balance criticism with encouragement to elicit my best effort. Many thanks.

I also would like to thank Walter Kintsch and Lewis Harvey, for their suggestions and contributions. My thanks to Alice Healy and Lyle Bourne and to all the individuals associated with our research group on knowledge and skill retention. Many thanks to Jon Roberts for his programming help and friendship. Thanks to Lori Meiskey for her friendship, her kindness, and our many interesting discussions. Thanks, also, to the staff members in the psychology department who have helped in so many ways. Thanks to Will Morse for help running experiments and encoding protocols. My special thanks to Susan Davies for her friendship and support when I needed them most.

Finally, I would like to thank my parents, Signa and Jim, without whose love, help, and encouragement, this project would never have been inititated. I thank them for their patience and their understanding throughout my wanderings and their support when I have taken the less-worn path, for it has indeed "made all the difference."

This research was supported in part by the United States Army Research Institute Contract Numbers MDA903-86-K-0155 and MDA903-90-K-0066.

CONTENTS

CHAPTER

I. INTRODUCTION....1
Background....1
The Keyword Method....2
Retrieval Models....6
General Methodology....11
Overview of the Experiments....14
II. EXPERIMENT 1....17
Method....17
Subjects....17
Apparatus....17
Materials and Design....17
Procedure....18
Results....22
Proportion of Words Recalled....23
Recall Analyses for Individual Items....26
Retrieval Times....28
Facilitation Effects....34
Verbal Reports....38
Discussion....44
III. EXPERIMENT 2....45
Method....48
Subjects....48
Materials and Apparatus....48
Procedure....49
Results....50
Proportion of Words Recalled....50
Retrieval Times....54
Verbal Reports....54
Discussion....58
IV. GENERAL DISCUSSION....62
BIBLIOGRAPHY....66
APPENDIX
A. STIMULI FOR EXPERIMENT 1....70
B. STIMULI FOR EXPERIMENT 2....71
C. ENCODING INSTRUCTIONS FOR VERBAL PROTOCOLS....72

TABLES

Table

1. Summary of Effects in Analysis of Variance for Delayed Session Recall Results (Experiment 1).....25

2. Summary of Effects in Analysis of Variance for Immediate Session Latency Results (Experiment 1).....30

3. Summary of Effects in Analysis of Variance for Delayed Session Latency Results (Experiment 1).....33

FIGURES

Figure

1. Mediation hypothesis versus direct hypothesis.....3

2. Example vocabulary item for keyword method.....5

3. Retrieval models for the keyword method.....7

4. The three retrieval tasks.....13

5. Computer display for acquisition phase.....20

6. Computer display for dropout phase.....21

7. Proportion recall as a function of retrieval task and delay interval for Experiment 1.....24

8. Proportion of vocabulary items for which both subtasks were successfully performed (Experiment 1).....27

9. Mean retrieval time as a function of retrieval task, test block, and verbal condition (immediate session only) for Experiment 1.....29

10. Mean retrieval time as a function of retrieval task and delay interval for Experiment 1.....32

11. Task arrangements to test facilitation effects.....35

12. Mean retrieval time for different retrieval tasks as a function of the facilitating task (block 3 only).....37

13. Encoding categories for the verbal reports.....41

14. Proportion of vocabulary items for different categories of mediation (Experiment 1).....42

15. Mean retrieval time as a function of type of mediation reported (Experiment 1).....43

16. Computer displays for Experiments 1 and 2.....51

17. Mean proportion of items recalled for the vocabulary task as a function of practice type (Experiment 2).....52

18. Mean proportion of items recalled for the subtasks as a function of practice type (Experiment 2).....53

19. Mean retrieval time for the vocabulary task as a function of practice type (Experiment 2).....55

20. Mean retrieval time for the subtasks as a function of practice type (Experiment 2).....56

21. Proportion of vocabulary items for different categories of mediation as a function of practice type (Experiment 2).....57

22. Mean retrieval time as a function of type of reported mediation (Experiment 2).....59

CHAPTER I

INTRODUCTION

Background. An important part of memory research is the study of mediation. Klatzky (1975) defines mediation as the use of long-term memory "to relate previous information to information currently being processed." For example, to memorize the German word "Fleisch," which in English means "meat," instead of simply rehearsing the words together, a natural-language mediatior may be used to relate the two words by noticing that Fleisch sounds similar to the English word "flesh," which is related in meaning to meat. People frequently report using mediators to memorize material. Furthermore, researchers attribute the effectiveness of mnemonic techniques to the mediators experimenters provide or ask subjects to generate in using the mnemonic (Solso, 1988; Bellezza, 1981).

There is general agreement that in learning unfamiliar material, the use of mediators is related to ease of learning and improved retention (Adams & McIntyre, 1967; Bellezza & Poplawsky, 1974; Bellezza, Poplawsky, & Aronovsky, 1977; Montague, 1972; Montague, Adams, & Kiess, 1966; Prytulak, 1971;). An important issue, however, is what role mediators play in the encoding and retrieval of material from memory.

According to one view, which has been called the "mediation hypothesis" (Adams & McIntyre, 1967; Bellezza, 1986; Bellezza & Poplawsky, 1974; and Bellezza, Poplawsky, & Aronovsky, 1977), mediators play a functional role in the formation of the memory representation and in the later use of that representation. For example, consider acquisition of an A-B pair (e.g. Fleisch and meat), using a mediator M (flesh) to learn them. According to the mediation hypothesis, in studying A and B, the mediator M becomes part of the actual memory representation relating A and B (see Figure 1). Therefore, to retrieve B from A, it is necessary to first access M. It follows that when M cannot be accessed, B cannot be accessed either. A number of studies show that unless the mediatior is recalled, the target material (i.e. B) is unlikey to be recalled (e.g. Bellezza & Poplawsky, 1974; Montague, Adams, & Kiess 1966; and Montague, 1972).

However, when the mediator is recalled the probability of recalling the target material is quite high. This relation between the retention of mediators and target material has been interpreted as evidence that the mediators play an important role in the encoding and retrieval process.

According to an alternative view, which has been called the "association hypothesis," (e.g. see Adams & McIntyre, 1967; Underwood, 1972), or what we will call the "direct hypothesis," the mediators that subjects generate or are provided with play no functional role in the formation of the memory representation. They are simply epiphenomena that accompany learning but play no role in the formation of the memory trace or the retrieval process. What is actually learned is a simple association between A and B (see Figure 1). Subjects may also learn an association between A and M or between M and B but these associations do not facilitate A-B recall. By this account, recall of the mediator M is related to the recall of B because both are associated to A. The ease with which the mediator M or the target B can be related to the stimulus A is simply an index of the "meaningfulness" of the stimulus A (Underwood, 1972).

If the mediators subjects use to encode new material are simply epiphenomena and not part of the memory representation, then they are hardly worth bothering about. On the other hand, if they play a functional role in learning and retrieval, then theories of memory need to account for them. In this research, we focused on trying to understand the role of mediation in a mnemonic technique known as the keyword method. Specifically, we examined whether the mediators used in the keyword mnemonic play a functional role in the retrieval process.

The keyword method. The keyword method (Atkinson & Raugh, 1975; Raugh & Atkinson, 1975) is a two-step mnemonic technique for learning foreign vocabulary items and their English equivalents. The first step requires the learner to relate the foreign word to a keyword, a concrete English word that sounds and/or looks like the foreign word or to a salient part of the foreign word. The second step requires relating the keyword and the English equivalent by forming an interactive image between the referents of the two words. For example, to learn the Spanish word doronico, meaning leopard, the English word door might serve as a keyword.

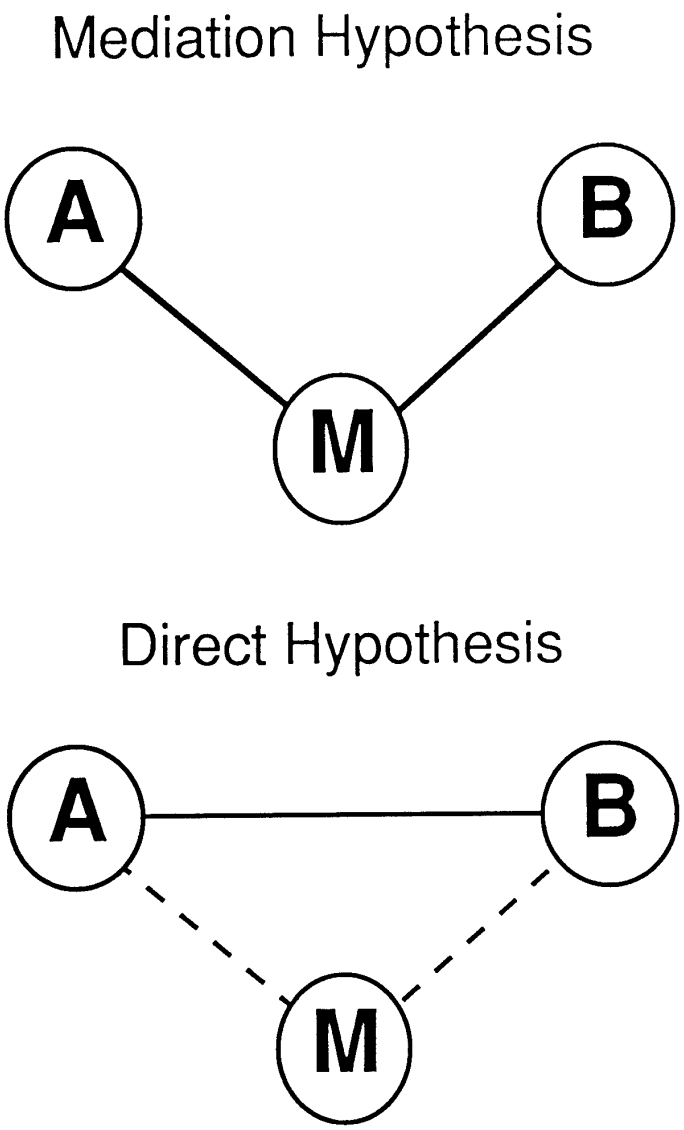


Figure 1. Mediation hypothesis versus direct hypothesis.

The first step is to notice that door sounds similar to the first part of the Spanish word doronico. The second step is to form an interactive image between door and leopard: an image of a leopard going through a door for example (see Figure 2).

The keyword method is ideal for investigating the role of mediation in memory. First, as the keyword method has proven itself an effective technique for learning and remembering new vocabulary, it is worth understanding the mechanisms that underlie its effectiveness. Second, most explanations of the keyword method's effectiveness assume that the keyword and the interactive image are important mediating structures in the retrieval process, but little direct and detailed evidence supports the use of these structures in the retrieval process. Third, the keyword and interactive image are examples of two common types of mediatiors: natural language mediatiors and imagery. Understanding how these mediators are used in the keyword method should improve our understanding of their use in other domains as well. Finally, there is evidence from our own studies, as well as those of other researchers (Hall, Wilson, & Patterson, 1981; Pressley, Levin, Digdon, Bryant, & Ray, 1983; Pressley, Levin, Kuiper, Bryant, & Michener, 1982), that people spontaneously employ vocabulary-learning strategies quite similar to the keyword method.

There is considerable research on the keyword method (see Pressley, Levin, & Delaney, 1982; Levin & Pressley, 1985, for reviews), much of which has demonstrated the effectiveness of the keyword strategy relative to other vocabulary learning strategies (e.g. Atkinson & Raugh, 1975; McDaniel & Pressley, 1984; Pressley, Levin, Kuiper, Bryant, & Michener, 1982; Pressley, Levin, Hall, Miller, & Berry, 1980; Pressley, Levin, & Miller, 1982; and Raugh & Atkinson, 1975). In addition, a number of studies have assessed the adaptability of the method to different settings (e.g. Jones & Hall, 1982; Levin, 1981; Levin, Pressley, McCormick, Miller, & Shriberg, 1979; Raugh, Schupback, & Atikinson, 1977; Turnure & Lane, 1987), and to different types of materials: for example, English vocabulary (Levin, 1981; Levin, McCormick, Miller, Berry, & Pressley, 1982); technical terms (Jones & Hall, 1982; and Levin, 1981); names, facts, events, and accomplishments (Jones & Hall, 1982); and English states and capitols (Levin, Shriberg, Miller, McCormick, & Levin, 1980). However, less attention has been

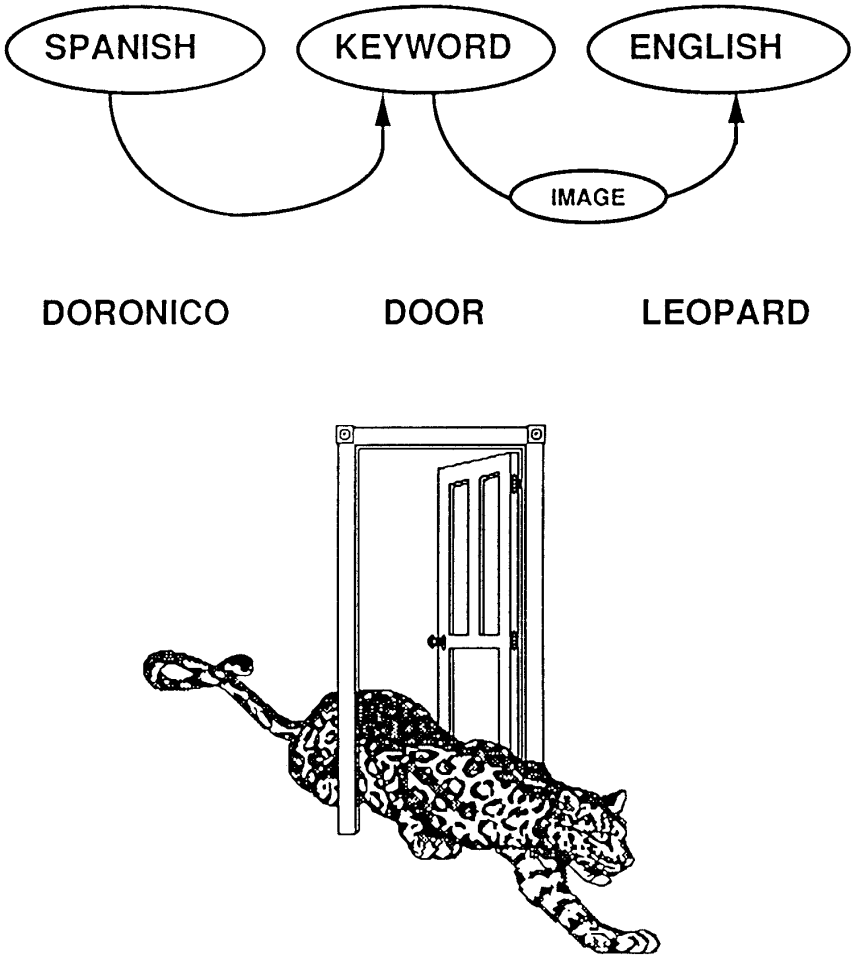


Figure 2. Example vocabulary item for keyword method.

devoted to developing and testing detailed theoretical models of the encoding and retrieval processes involved in the keyword method. Such models would be important not only to explaining how the keyword mnemonic works but to understanding the role of mediation in memory.

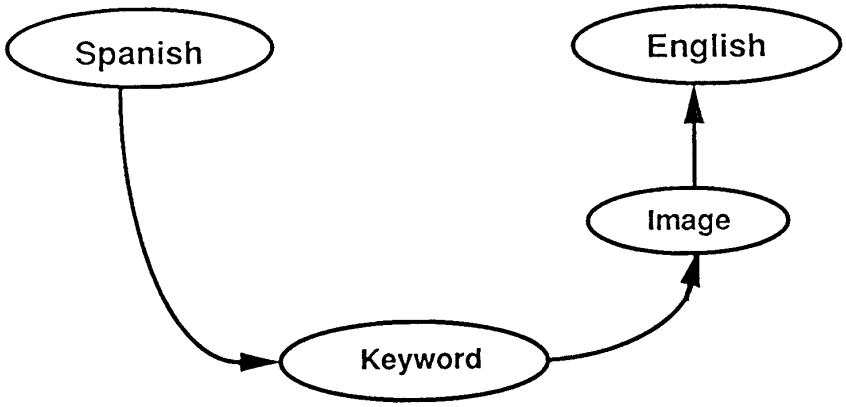
Retrieval Models. Given our discussion of the mediation issue there appear to be two general classes of models to explain retrieval of the English equivalent of a foreign word learned using the keyword method: "direct models" and "mediation models".

A "direct model" assumes that the keyword and interactive image play no functional role in the retrieval of the English words from the foreign words (see Figure 3). Rather the English word is retrieved directly from the foreign word: for example, via a simple association. Subjects may learn or associate the mediators with the foreign words as a result of instructions to learn the foreign-English pairs using the mediators, but these associations are separate from the memory representation relating the foreign word and its English equivalent and are not used to retrieve the English equivalents. Therefore, any relationship between recall of the mediators and the English equivalents is purely correlational or epiphenomenal. Whether the keyword is recalled or not does not directly influence the recall of the English equivalent.

In contrast to the "direct model," a "mediation model" assumes that the keyword and interactive image mediators play an important role in retrieving the English equivalents from the foreign words (Figure 3). According to this model, retrieval of the English equivalents is only possible via the mediating keyword and interactive image. No direct link connects the foreign word and English word. Furthermore, retrieval of the English equivalent occurs in a specific sequence. This sequence begins when the foreign word is used to retrieve the keyword. Next the keyword cues the retrieval of the interactive image. Finally, the English word is recovered from the interactive image.

We have not specified the nature of the retrieval processes that recover each cognitive component (e.g. the retrieval of the keyword from the Spanish word). These processes could be implemented in a variety of ways: for example, via simple associations or relational links. Some of these have been suggested in previous accounts of the keyword method, which

Mediation Model



Direct Model

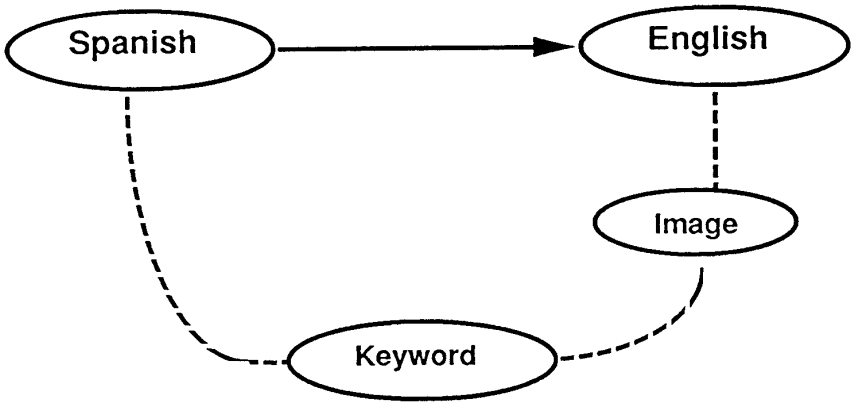


Figure 3. Retrieval models for the keyword method.

we will discuss in a moment. However, for our current purposes, it seems premature to specify the details of retrieving the keyword and interactive image before establishing that these structures are indeed part of the retrieval process.

Current accounts of the retrieval process in the keyword method are more or less detailed versions of the basic mediation model we have just described (Atkinson and Raugh, 1975; Clark and Pavio, 1987; Desroachers & Begg, 1987; Pavio & Desroachers, 1981; Turnure & Lane, 1987). Atkinson and Raugh, for example, have proposed a retrieval account based on different types of associative links between the foreign word and its English equivalent. In this model, the Spanish word is connected to the keyword by an acoustic link and the keyword is connected to the English word by an imagery link. Little detail is provided concerning the nature of these links. However, Atkinson and Raugh assume they are easily learned and provide a simple means of retrieval while the subject learns the direct association between the foreign word and its English equivalent. Because the acoustic and imagery links are stronger than the newly-formed direct link, retrieval is at first governed by the mediated pathway instead of by the direct link.

A retrieval account based on Dual Coding Theory (e.g. Clark and Pavio, 1987; Pavio & Desroachers, 1981) provides more detail concerning the nature of the linkage between the foreign word, keyword, and English word. Recovery of the keyword depends on perceptually-driven logogens: wordlike entities composed of visual and phonemic features. When one of these features is present in a stimulus it tends to activate any logogens containing these features. Unfamiliar verbal material (e.g. a nonsense syllable or a foreign word) shares certain features with logogens in memory and will activate logogens having features perceptually similiar to the presented stimulus. For example, the syllable "nos" might activate the logogen "nose" or the Spanish word "doronico" would tend to activate the logogen "door." In short, a word that shares visual or phonemic features with a foreign word is more likely to be retrieved than a word sharing few of these features. Hence, when the Spanish word is presented as a cue, the keyword is more likely to be retrieved than the English word. As for the second link, according to the Dual Coding Theory, the keyword and English word are associated by two different systems, verbal and imaginal, which

connect the keyword and English word logogens as well as their corresponding imagistic representations. This redundancy increases the likelihood of retrieving the English word from the keyword.

Desroachers and Begg (1987) have suggested an account based on an organization-redintegration hypothesis. The central concept of the organizatgion-redintegration hypothesis is the memory trace, a record of how one or more items were interpreted at encoding. If items are studied in meaningful interaction, they are stored in a unitary trace rather than separate traces. Perception of the physical similarity between the foreign word and keyword increases the probability that these two items will form a unitary memory trace, just as the construction of a composite mental image between keyword and English equivalent increases the probability that these will form a unitary memory trace. Unitary traces store relational information such that given a part of the information stored in the trace, the rest is recovered with some probability. Information stored in a unitary trace is more likely to be recovered than information stored in separate traces.

However, redintegration is not guaranteed. Depending on the type of discriminations originally made at study, the probability of redintegration given a specific cue varies. A discrimination that simply related the first letter of the foreign word to the keyword might not be specific enough to ensure recovery of the correct keyword given the foreign word. On the other hand, if the foreign word corresponded to all but the last letter of the keyword, the probability of redintegration would be quite high. Therefore, the keyword is redintegrated with some probability and decoded from the trace content. This keyword then serves in discriminating a particular interpretation of its meaning, which is followed by the redintegration of the appropriate composite image. It is then necessary to discriminate the target item from the composite structure. Discrimination is determined by the specificity of the initial interpretation at study.

From another framework, Turnure and Lane (1987) have proposed a mediated retrieval account similar to the one we have suggested. Rather than focusing on the details of how each component is retrieved, they suggest a set of information processing steps to retrieve the English definition of the Spanish word. This sequence begins by first

discriminating a part of the foreign word, which is then used to access the keyword. The keyword is then used to retrieve the interactive image, and the image is decomposed into a "Word-image" (corresponding to the keyword) and a "Meaning-image" (corresponding to the English word). The English word is then retrieved and used to respond.

Though these models differ from one another in specific details, all are consistent with the basic mediation model in that they implicate the keyword and interactive image in the retrieval process. The critical difference between the direct and mediation accounts of the retrieval process in the keyword method is whether or not retrieving the English equivalents depends on first retrieving the keyword and interactive image. The mediation model claims that accessing the English equivalent of the foreign word depends on first accessing the mediating keyword and interactive image. The direct model makes no such claim. Though previous keyword experiments have provided some evidence for the mediation model, this evidence is not definitive. We now discuss this evidence and suggest what type of evidence would permit a more definitive choice between the mediation and direct models.

One type of evidence offered to support the mediation model comes from a study by Sweeney and Bellezza (1982), in which they performed an analysis of the conditional probability of recalling the English definition given that the keyword was or was not recalled. Their results indicated that recall of the English definitions was correlated with recall of the keywords. Though consistent with the mediation model, this result can also be explained by the direct model. One simply assumes that the correlation between recall of the keywords and English definitions is the result of both being associated to the foreign word.

More direct support for mediation comes from verbal report procedures, such as post-session questionnaires, in which subjects report the use of keywords or images (e.g. Delaney, 1978; Pressley, Levin, Digdon, Bryant, & Ray, 1983; Pressley et al, 1982; Pressley, Levin, Nakamura, Hope, Bispo, & Toye, 1980; and Raugh & Atkinson, 1975). However, these procedures usually ask subjects to report what strategies they used to learn the vocabulary items (e.g. whether they used keywords) rather than what they were aware of in retrieving the items. Thus, the evidence for the role of mediators in retrieval is indirect, based on whether or not

subjects used the mediators at encoding. These reports would be strengthened by obtaining evidence for the role of mediators at retrieval as well as at encoding. However, there are additional problems with the verbal reports used in previous studies of the keyword method. One problem is that these reports are gathered after multiple trials or delays of more than a few seconds, jeopardizing the reliability and validity of the reports (see Ericsson & Simon, 1980; 1984 for a discussion of these issues). In addition, most of these reports are not validated by any convergent measure, such as response latencies.

In short, though some support for the mediation model is provided by previous studies, stronger evidence is necessary to rule out the direct retrieval model as a plausible account. Specifically, evidence for the intermediate cognitive structures and processes that mediate retrieval of the English equivalents is needed. For example, retrospective verbal reports collected after each retrieval would provide a more direct assessment of what subjects processed in retrieving a given vocabulary item than global post-questionnaire procedures. In addition, convergent evidence for intermediate structures and processes (e.g. response latencies) would increase our confidence in the mediation account. The latter goal is particularly important because an experimental result using a single dependent measure (e.g. recall or verbal report) is difficult to interpret unambiguously.

In the research described here, we adopted just such an approach: in addition to using performance measures, such as cued recall, we also employed more direct process measures, such as verbal reports and response latencies, to examine the intermediate cognitive structures and processes that underlie the retrieval process for vocabulary learned using the keyword method. We now describe our general methodological approach and two experiments in which we used this methodology to test the mediation model.

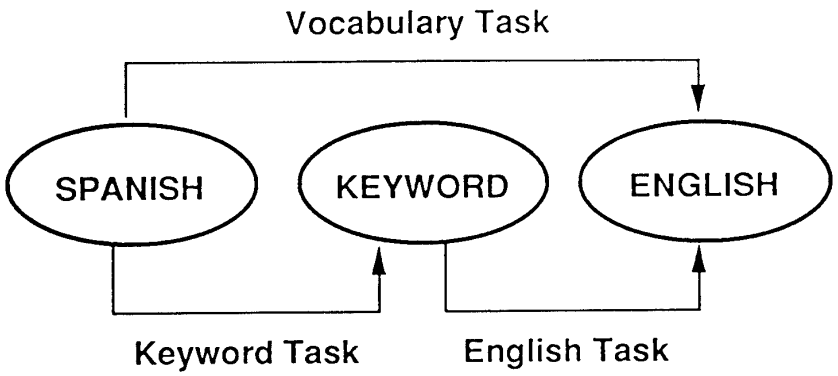
General Methodology. In all of the experiments reported here, subjects learned Spanish-English vocabulary pairs using the keyword method. Keywords were supplied and subjects generated their own interactive images. Subsequent to learning the vocabulary items, subjects were tested on the items using a cued-recall procedure: that is subjects would see the Spanish

word and respond with the English equivalent. This procedure is called the vocabulary retrieval task (or vocabulary task) and corresponds to the general testing procedure used in keyword studies. However, in addition to testing subjects on the vocabulary retrieval task, we also tested them on two additional retrieval tasks. The first we call the keyword retrieval task (or keyword subtask), in which subjects see the Spanish word and respond with the appropriate keyword. The second we call the English retrieval task (or English subtask), in which subjects see the keyword and respond with the corresponding English word. The three retrieval tasks are displayed in Figure 4. If, as the mediation model assumes, retrieving the English equivalent from the Spanish word requires first retrieving the keyword and then using the keyword to retrieve the English equivalent via the interactive image, then the keyword and English subtasks together should involve the same cognitive structures and processing steps accessed by the vocabulary retrieval task, and performance on these subtasks should relate to performance on the vocabulary task.

This task decomposition has a clear advantage over previous recall procedures used in keyword studies. Previous studies have simply compared retention performance for keywords and English equivalents, without describing the processing steps in which the keyword is involved. Consequently, correlations between recall of keywords and English definitions do not necessarily imply that the keyword was actually used to retrieve the English definition. The current approach provides a way to test not only whether recall of the keyword relates to recall of the Spanish word, but whether recall of the English definition using the keyword relates to recall of the English definition using the Spanish word.

In addition to the recall performance, response latencies for the three retrieval tasks provided an additional means to test the mediation model. If each of the subtasks accesses a subset of the cognitive structures and processes accessed by the vocabulary retrieval task, then: (1) it should take longer to perform the vocabulary retrieval task than either of the subtasks; and (2) whatever influences the latencies for the vocabulary retrieval task should influence the latencies for the subtasks and vice versa.

A third and important type of data collected in these experiments was verbal reports. These reports were of two types: think-aloud reports



Retrieval Task	Display		
Vocabulary	doronico	__	?
Keyword	doronico	?	__
English	__	door	?

Figure 4. The three retrieval tasks.

and retrospective reports (see Ericsson & Simon, 1980; 1984 for a complete discussion of verbal reports). The think-aloud reports were used during the acquisition phase: subjects thought-aloud as they studied the vocabulary items. These reports are not discussed in the studies reported here as our focus is on the retrieval processes in the keyword method. The other type of verbal report collected was retrospective verbal reports. Immediately after retrieving an item, subjects would provide a retrospective report of whatever they could remember thinking during the trial. These reports provided an independent and converging source of evidence for testing the mediation account. Whereas with the recall and latency data, the evidence for mediation is based on performance comparisons of the subtasks and the vocabulary retrieval task, the verbal reports provide a direct source of evidence for the cognitive structures accessed in retrieving the English defintions of the Spanish words. If subjects are indeed accessing the keyword and interactive image as part of this process, then these reports should provide evidence for their retrieval. Also, as we shall discuss in detail later, response latencies provided a means to validate the content of these reports.

Overview of the Experiments. In the first study, subjects learned a set of Spanish-English vocabulary items using the keyword method and were tested immediately and after a 1-week or 1-month delay. A number of predictions concerning the the recall, latency, and verbal report results followed from the mediation model. The first set of predictions were for the recall results. According to the mediation model, performing the vocabulary task depends on performing both subtasks correctly. Therefore, recall performance on the vocabulary task was expected to be lower than recall performance on the two subtasks. Comparing the vocabulary task to each of the subtasks individually, performance on the vocabulary task was predicted to be lower than one or both of the subtasks. On the other hand, if performance on the vocabulary retrieval task exceeded performance on either of the subtasks, the direct model would be supported.

A second prediction made by the mediation model was that whenever the vocabulary retrieval task was successful for an individual vocabulary item, both subtasks for that item would be successful; whereas, when the vocabulary retrieval task for an item was unsuccessful, the likelihood of

successfully performing both subtasks for that item would be quite small. According to the direct model, performing both subtasks correctly or incorrectly would not be expected to predict performance on the vocabulary retrieval task.

In addition to the predictions concerning the recall results, a number of predictions concerning retrieval times were made. As each of the subtasks taps only a subset of the cognitive structures and processes that are part of retrieving the vocabulary definitions, the mediation model predicted that retrieval times for the vocabulary retrieval task would be longer than those for the two subtasks. The direct model made no such differential prediction concerning the retrieval times, as each of the three retrieval tasks relies on a separate association between a stimlus/response pair. In addition, the mediation model predicted that the response latency for a given retrieval task (e.g. English subtask) would be facilitated when preceded by a retrieval task that shared specific intermediate processing steps with the first task. Thus performing the vocabulary retrieval task prior to the English subtask was expected to facilitate performance on the English subtask, as both tasks require using the keyword to retrieve the English definition. The direct model, on the otherhand, would predict no such facilitation effects based on intermediate processing steps, though it might predict facilitation effects based on stimulus similarity (e.g. between two tasks using the same stimulus, such as the vocabulary and the keyword tasks).

Finally, according to the mediation model, the retrospective verbal reports for the vocabulary retrieval task were expected to provide evidence for accessing the keyword and interactive image to retrieve the English definition. According to the direct model, subjects were expected only to report retrieving the Spanish word and the corresponding English definition.

The purpose of the second study was to investigate whether, once subjects began to practice retrieving the vocabulary items without benefit of subtask practice, the mediation model would remain an adequate account of the retrieval process for the vocabulary items. According to the direct model, as soon as subjects begin practicing the vocabulary task alone, only the direct link between the Spanish word and English word should be accessed. However, according to the mediation model, even though subjects

perform only the vocabulary retrieval task, in order to retrieve the English equivalent, it is still necessary to access the mediating cognitive structures (i.e. retrieve the keyword and then use it to retrieve the English equivalent via the interactive image). Therefore, retrieval should continue to be mediated.

CHAPTER II

EXPERIMENT 1

Method

Subjects. A total of 24 subjects at the University of Colorado received course credit for participating in this experiment. The experiment was restricted to subjects who had no previous training in Spanish.

Apparatus. The study and test materials were presented to each subject individually via a computer display controlled by an IBM PC. Latencies for the study trials were recorded via a button box connected to the computer. For the cued recall trials, subjects answered orally and response latencies were registered via a microphone and a voice actuated relay connected to the computer. In addition, a cassette tape-recorder was used to record subjects' responses and verbal reports.

Materials and design. The stimuli were 42 Spanish-English pairs and their associated keywords (e.g. doronico-door-leopard). Twenty-nine of the items were taken from two previous keyword studies (Hall, Wilson, & Patterson, 1981; and Pressley, 1977). An additional 11 items were selected using criteria similar to those employed by Hall, et al (1981) and Pressley (1977). The stimuli were selected so that both the keyword and English word were concrete, imageable nouns. In most cases, the keyword corresponded to the first syllable of the Spanish word. A complete list of the stimuli appears in Appendix A.

Three lists were constructed, using the same 42 three-word stimuli for each list. The first 6 items on each list were "controls," which were always presented before the other 36 items. The purpose of these controls was twofold: (1) To reduce facilitation effects that might occur for an item that appeared at the end of one test and at the beginning of a subsequent test; and (2) To provide additional warm-up trials between different phases of the experiment. Piloting of the materials showed that during transitions from one part of the experiment to

another, some subjects had difficulty performing the new task or tasks on the first few trials. The 6 controls were the same for each list and were not analyzed.

Subjects performed three different retrieval tasks for each item within a block of testing. These were the vocabulary, keyword, and English retrieval tasks, in which subjects saw the Spanish word and retrieved the English word, saw the Spanish word and retrieved the keyword, or saw the keyword and retrieved the English word (see Figure 4). Task order was counterbalanced for the 36 target items and the 6 controls on each list. As there were 6 possible task orders, each task order was used for 6 target items and 1 control. Task order assignment was random, except that, across the 3 lists, no item was assigned to the same task order twice. Half the items on each list were verbal (subjects concurrently or retrospectively reported their thoughts); the other half were silent. Assignment of items to the verbal and silent category was appropriately counterbalanced across subjects. Finally, two different randomized presentation orders were used.

Procedure. The experiment was conducted in two sessions. The first session was 3 hours long; the second session was an hour long and was held after either a 1-week or 1-month delay. Upon arrival at the lab, subjects were randomly assigned to either the 1-week or 1-month delay group and to one of the list conditions. The first session consisted of four phases: an introductory phase, an acquisition phase, a dropout phase, and a test phase.

In the introductory phase, the experimenter gathered background information concerning each subject's previous foreign language experience. The subjects then received instructions and practice on giving think-aloud and retrospective verbal reports (see Ericsson and Simon, 1980, and 1984, for details). For the think-aloud reports, subjects were asked to report what they were thinking as they performed a task. For the retrospective reports, subjects were asked to report what they remembered thinking from the time they initiated a task until they reported a solution or stopped working on the task. Subjects were asked to report only what they could actually remember thinking and not to explain or add anything to their reports that they did not actually remember thinking. To ensure that subjects clearly understood the verbal report instructions, they were asked

to perform mental addition and other warm-up tasks, providing concurrent as well as retrospective reports for each task.

Following the verbal report instructions and practice, the experimenter explained the details of the keyword method and subjects practiced learning and retrieving sample items. The practice tasks were the same as those used for the acquisition and dropout phases described below.

In the acquisition phase, each vocabulary item was presented individually. A display consisting of the Spanish word on the left, the keyword in the middle, and the English word on the right appeared on the computer screen (see Figure 5). Subjects had been instructed to study the items as follows. First, look at the Spanish word and pronounce it as best you can. Then look at the keyword and try to relate it to the Spanish word by noticing the simlarity in sound or appearance. Then form an interactive image between the keyword and English word. For half the items, a cue instructing subjects to think aloud while studying the items preceded the onset of the three words. Subjects had 20 seconds to study each display, at which time, the program went on to the next trial.

Following the acquistion phase was a dropout phase, in which only the Spanish words were displayed and subjects were asked to respond with the correct English equivalents. Tests were conducted using a voice key. For each trial, a Spanish word appeared on the left of the screen, in the same location as it had during acquisition, and a question mark appeared on the far right of the screen, where the English word had previously appeared (see Figure 6). Subjects responded by saying aloud the correct English word as soon as they thought of it. The experimenter, sitting behind and to the right of the subject, entered whether the response was correct or incorrect via the computer keyboard. For incorrect responses or responses not made within a 15 second time limit, feedback was provided, consisting of the original three words, displayed as before. After 5 seconds, the feedback display disappeared, and the next item was presented. After all 42 items had been tested, correctly-recalled items were eliminated and missed items were retested in a second cycle of testing. Following this second test, correctly-recalled items were eliminated from further testing and incorrectly-recalled items were tagged for retesting. This procedure

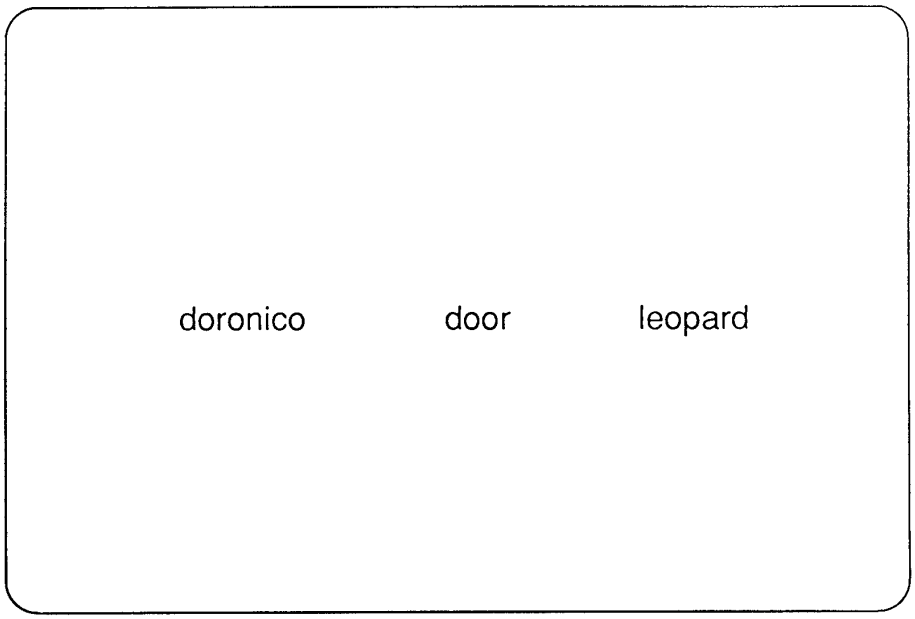


Figure 5. Computer display for acquisition phase.

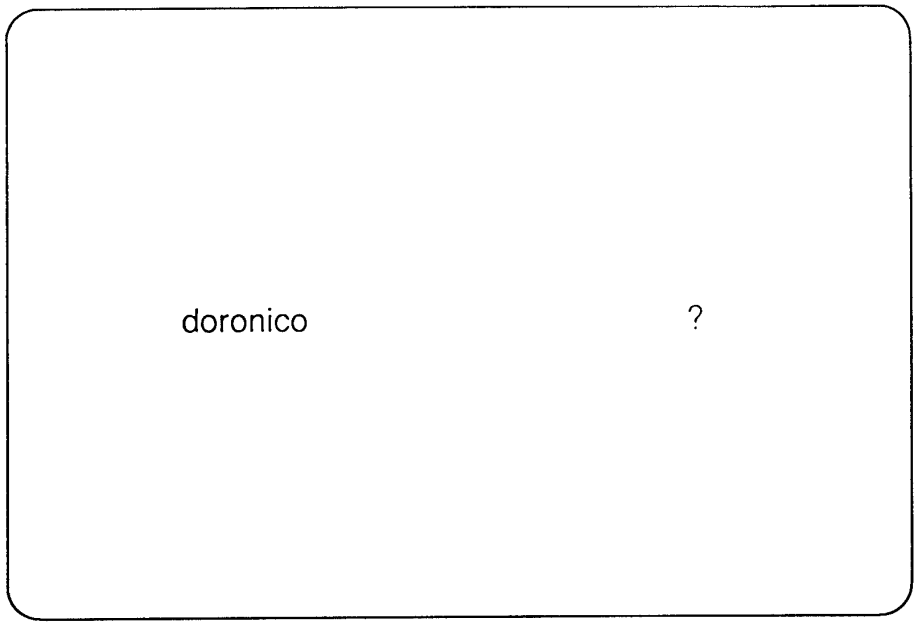


Figure 6. Computer display for dropout phase.

was repeated until each item had been correctly recalled once. No verbal reports were taken during the dropout phase.

After completion of the acquisition and dropout phases, which required about 2 hours, the test phase began. As retrospective reports were used during this phase, the experimenter first reviewed the instructions for these reports. After instructions on the new tasks, subjects practiced
performing the new tasks and giving retrospective reports. Throughout the test phase, subjects performed three different retrieval tasks on each item: (1) The vocabulary retrieval task (given the Spanish word, say the English word); (2) The keyword retrieval task or keyword subtask (given the Spanish word, say the keyword); and (3) The English retrieval task or English subtask (given the keyword, say the English word). For each task, the appropriate cue (Spanish word or keyword) was displayed where it had originally appeared during study and a question mark appeared in the location where the target word (keyword or English word) had originally appeared. Figure 4 summarizes the three tasks. After the subject's response or 15 seconds, the program went on to the next trial. No feedback was provided concerning accuracy. For the verbal items (half the trials), a visual cue followed the offset of the stimulus display, asking subjects to give a retrospective report.

The test phase for session one consisted of 2 blocks of 3 tests each. On each test, all 42 vocabulary items were tested, but for a given item only one of the 3 possible retrieval tasks was performed. After three tests (one test block), all retrieval tasks had been performed on all items. Half the subjects returned after 1-week and half after 1-month to complete 2 additional blocks of testing.

Results

Voice key errors occurred on 2.4% of the trials and were excluded from the analyses. The 6 control items on each list were not analyzed. The ANOVAs for the recall and latency analyses were performed on subject means computed by averaging across the 6 task orders. For the recall analyses, a data point was a proportion correct score based on 18

observations; for the latency analyses, a data point was a mean RT computed using the correct retrieval trials from 18 observations. The immediate (blocks 1 and 2) and delayed (blocks 3 and 4) results were analyzed separately for the ANOVAs. In addition, as two subjects were unable to complete the fourth block of testing, the ANOVAS for the delayed results were restricted to block three.

Proportion of words recalled. For the immediate recall results mean proportion correct for the vocabulary, keyword, and English retrieval tasks, respectively, was .94, .94. and .97 items. Given the obvious ceiling effect, no further analyses of these results were performed.

For the delayed recall results, mean proportion of items correctly recalled as a function of retrieval task and delay condition (1-week or 1-month) is displayed for the silent and verbal items in Figure 7. An analysis of variance was performed with delay as a between-subjects factor, and task and verbal condition as within-subjects factors. A summary of the analysis appears in Table 1. An a priori comparison confirmed that performance on the vocabulary task was less than the average of the two subtasks (.63 versus .78 of the items correct), $F(1,22) = 67.28$, $p < .0001$, $MSe = .01$, as predicted by the mediation model. Post hoc comparisons of the vocabulary task to each of the subtasks showed that recall was lower for the vocabulary task than for the keyword task (.63 versus .91), $F(1,22) = 77.61$, $p < .0001$, $MSe = .02497$, but that the difference between the vocabulary and English retrieval tasks (.63 versus .64) was not significant, $F(1,22) < 1$. Taken together, these results were consistent with the mediation model's prediction that recall for the vocabulary task would be less than that for the subtasks.

In addition to the main effect of task, there was an interaction of task with verbal condition, $F(2,44) = 17.52$, $p < .0001$, $MSe = .001$, and delay, $F(2,44) = 38.39$, $p < .0001$, $MSe = .02$. Both interactions were due to a ceiling effect for the keyword retrieval task. Additional post hoc comparisons using a Bonferroni procedure, as recommended by Meyers (1979) and Maxwell (1980), did not alter any of the basic results for the task comparisons. Recall for the vocabulary task was still lower than recall for the keyword retrieval task at 1-week (.86 versus .95, $F(1,11) = 9.27$, $p < .01$, $MSe = .00864$) as well as at 1-month (.40 versus .88), $F(1,11) =$

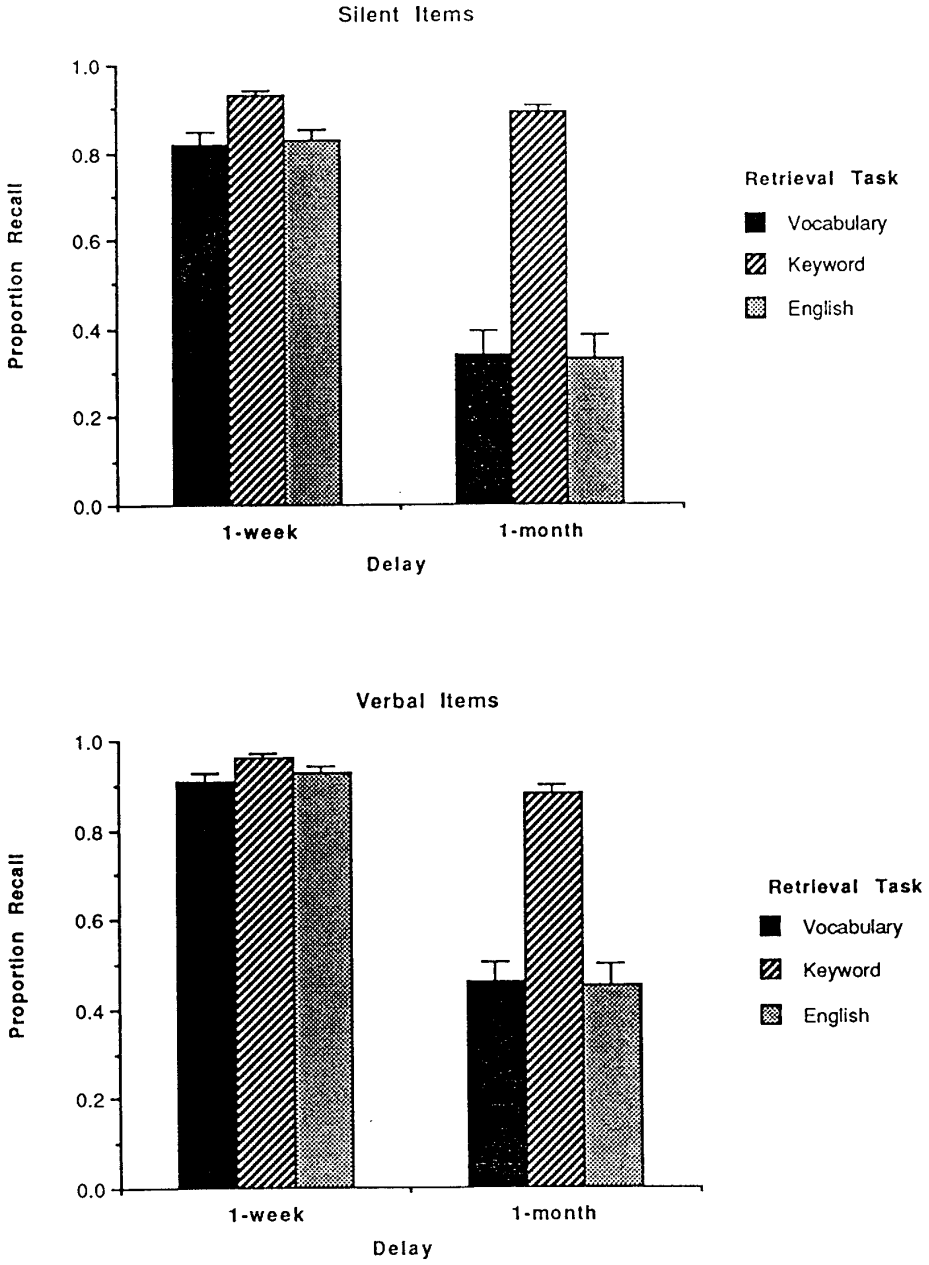


Figure 7. Proportion recall as a function of retrieval task and delay interval for Experiment 1.

Table 1

Summary of Effects in Analysis of Variance for Delayed Session Recall Results

Source	df	F	p <	MSe
Delay	1,22	37.25	.0001	.11
Task	2,44	69.37	.0001	.02
Verbal Condition	1,22	27.79	.0001	.01
Delay x Task	2,44	38.39	.0001	.02
Delay x Verbal	1,22	.11	.74	.01
Task x Verbal	2,44	17.52	.0001	.001
Task x Delay x Verbal	2,44	1.78	.18	.001

68.81, $p < .0001$, MSe = .04130), whereas the difference between the vocabulary and English retrieval tasks was still not reliable at 1-week (.86 versus .88, $F(1,11) < 1$) or at 1-month (.40 versus .39, $F(1,11) < 1$). Recall performance on the vocabulary task was below that for the keyword task for silent items (.58 versus .91, $F(1,11) = 32.74$, $p < .0001$, MSe = .04032) as well as for verbal items (.68 versus .92), $F(1,11) = 22.58$, $p < .0001$, MSe = .03561), whereas recall performance on the vocabulary task was still not reliably less than that for the English retrieval task for the silent items (.58 versus .58, $F(1,11) < 1$) or the verbal items (.68 versus .69, $F(1,11) < 1$).

Recall analyses for individual items. The task comparison results from the ANOVA were consistent with the mediation model, as performance on the vocabulary retrieval task was less than that for the subtasks, and performance on the vocabulary task never exceeded performance on either of the subtasks, which would have supported the direct model. However, the mediation model made a stronger prediction. If retrieval of the English equivalent from the Spanish word depends on successfully retrieving the keyword and then using the keyword to retrieve the English definition, it follows that when the vocabulary retrieval task is successful, both of the subtasks should be successful; whereas, when the vocabulary retrieval task is unsuccessful, the likelihood of success on both subtasks should be quite small. The direct model makes no such differential prediction, as successfully retrieving the English word from the Spanish word does not require retrieving the keyword and using it to retrieve the English word.

To test this prediction we assessed performance on the 3 retrieval tasks for each vocabulary item. For each vocabulary item in a test block for a given subject, we scored whether both, neither, or one of the subtasks was correctly performed. For both the correctly and incorrectly retrieved vocabulary items, we calculated the proportion of items for which both subtasks were correctly performed. Only the delayed results were analyzed, as performance was at ceiling for the immediate test session. Furthermore, recall was near ceiling for the 1-week subjects even at delay, so their data data should be interpreted cautiously.

Figure 8 displays the results. As predicted, when the vocabulary retrieval task was successful, the proportion of items for which both

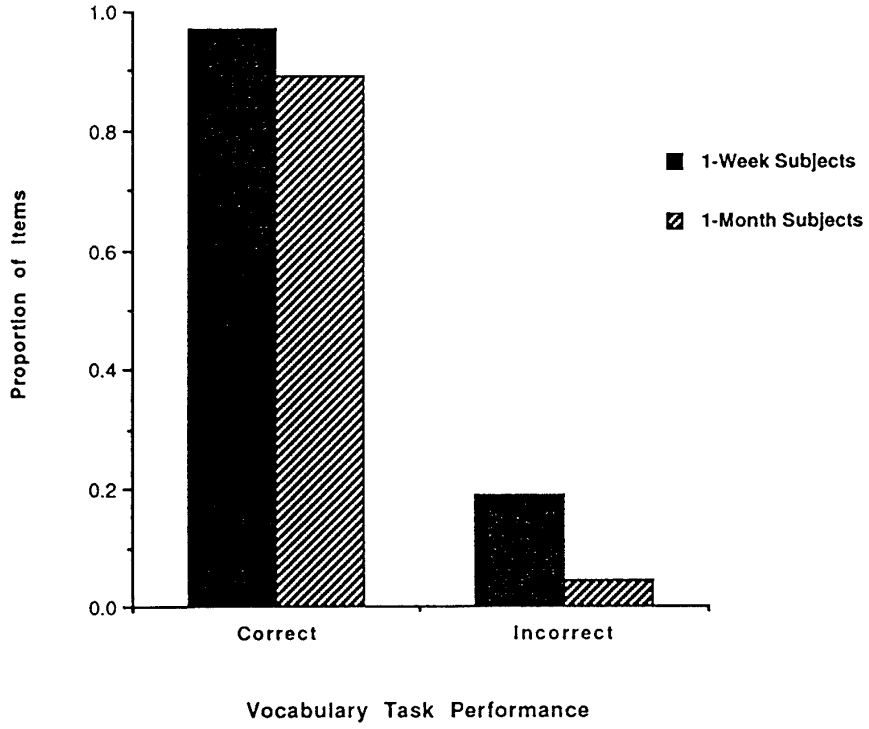


Figure 8. Proportion of vocabulary items for which both subtasks were successfully performed (Experiment 1).

subtasks were successful was quite high: .97 of the items for the 1-week subjects and .89 of the items for the 1-month subjects. On the other hand, when the vocabulary retrieval task was unsuccessful, the proportion of items for which both subtasks were successful was quite low: .19 of the items for the 1-week subjects and less than .05 of the items for the 1-month subjects.

Retrieval Times. Both the overall analyses and the item analyses for the recall results results demonstrated that performance on the vocabulary retrieval task was related to subtask performance, supporting the mediation model. However, the case for the mediation model would be strengthened by time course evidence relating the vocabulary retrieval task and the two subtasks. The mediation model predicted that the time necessary to perform the vocabulary retrieval task would be longer than that required to perform either of the subtasks.

In general, response latencies are distributed log-normally when subjects are permitted to respond beyond several seconds (Ratcliff, 1978), which was the case here. Therefore, the latency analyses were performed on log-transformed times, computed by first subtracting a constant (610 msecs) and then taking the log (base 10) of the result. However, for ease of interpretation, the means were reconverted to raw form (by taking the antilog of the mean and adding back 610 msecs) for the tables and graphs. As with the recall data, separate analyses were performed for the immediate and delayed results. For the delayed results, a few data points were missing, as none of the 18 observations used to compute the RTs were correct. These missing values constituted less than 2% of the data and were replaced with subject means.

Mean retreival time as a function of retrieval task and block for the immediate test session is displayed separately for the silent and verbal items in Figure 9. The statistical results are summarized in Table 2. An a priori comparison confirmed that latencies for the vocabulary retrieval task were longer than those for the subtasks (2041 versus 1642 msecs, $F(1,23) = 173.8$, $p < .0001$, $MSe = .00739$), as predicted by the mediation model. Post hoc comparisons of the vocabulary task to each of the subtasks individually confirmed that the vocabulary retrieval task was

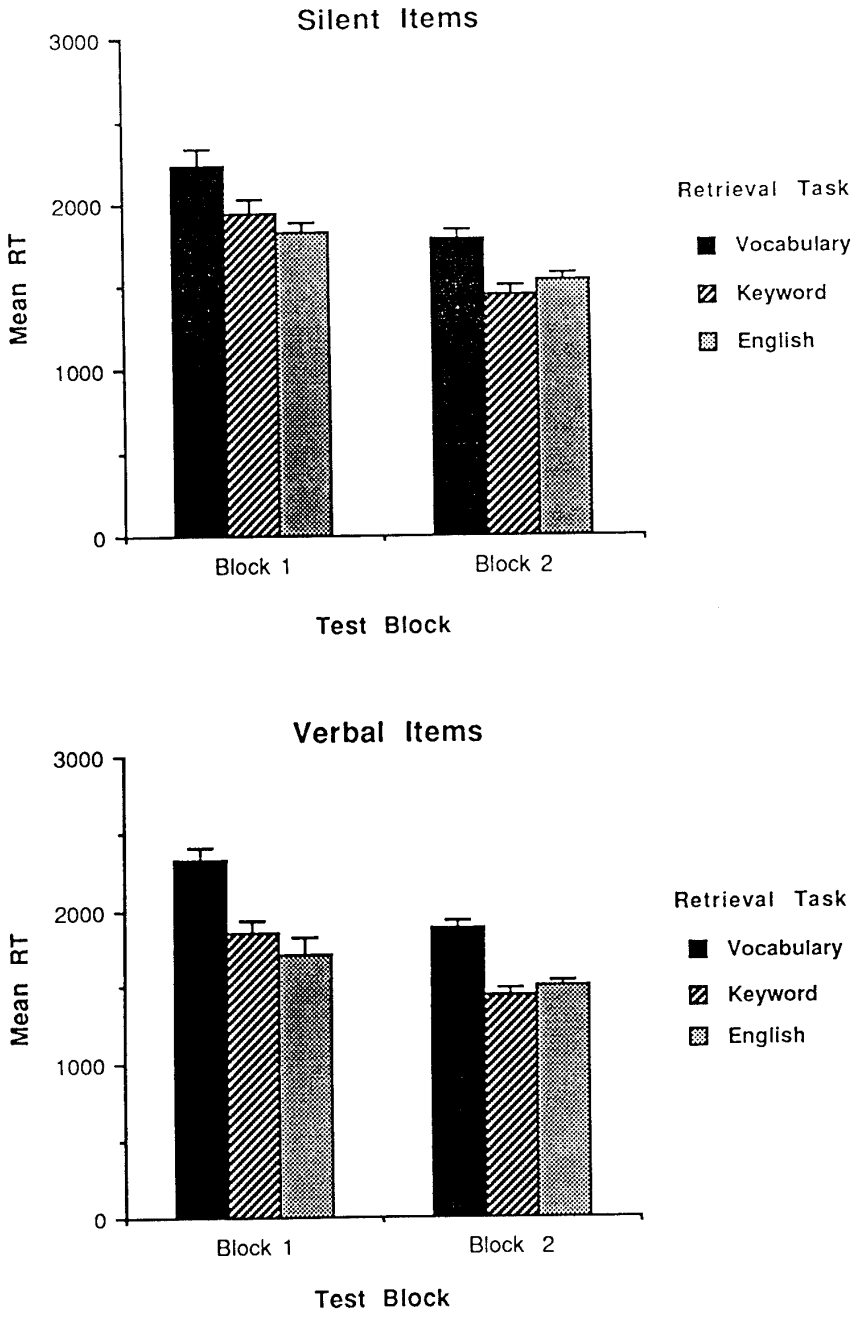


Figure 9. Mean retrieval time as a function of retrieval task, test block, and verbal condition (immediate session only) for Experiment 1.

Table 2

Summary of Effects in Analysis of Variance for Immediate Session Latency Results

Source	df	F	p <	MSe
Block	1,23	91.62	.0001	.02
Task	2,46	49.02	.0001	.01
Verbal Condition	1,23	.20	.66	.01
Block x Task	2,46	9.55	.0001	.001
Block x Verbal	1,23	1.25	.27	.01
Task x Verbal	2,46	4.09	.023	.01
Task x Block x Verbal	2,46	1.78	.18	.001

slower than the keyword subtask (2041 versus 1653 msecs), $F(1,23) = 66.34$, $p < .0001$, MSe = .0137) and slower than the English subtask (2041 versus 1633 msecs), $F(1,23) = 149.3$, $p < .0001$, MSe = .00683.), as predicted by the mediation model.

Given the interaction of task with block, $F(2,46) = 9.55$, $p < .0001$, MSe = .001, and verbal condition, $F(2,46) = 4.09$, $p < .02$, MSe = .01, further post hoc analyses of the task comparisons were performed. Latencies for the vocabulary task were longer than those for the keyword subtask for both levels of block: 2282 versus 1900 msecs, $F(1,23) = 40.46$, $p < .0001$, MSe = .00749, for the first block; and 1835 versus 1452 msecs, $F(1,23) = 68.51$, $p < .0001$, MSe = .00929, for the second block. Latencies were also longer for the vocabulary task than for the English subtask for both levels of block: 2282 versus 1767 msecs, $F(1,23) = 86.08$ $p < .0001$, MSe = .00712, for the first block; and 1835 versus 1514 msecs, $F(1,23) = 119.15$, $p < .0001$, MSe = .0035, for the second block. The task comparisons for the task by verbal interaction revealed a similar pattern of results. Latencies for the vocabulary retrieval task were always longer than those for either of the subtasks at both levels of verbal condition: for the silent items: 1995 versus 1671 msecs, $F(1,23) = 38.29$ $p < .0001$, MSe = .00835, for the vocabulary vs keyword tasks comparison; and 1995 versus 1665 msecs, $F(1,23) = 79.3$ $p < .0001$, MSe = .00422, for the vocabulary vs English tasks comparison; for the verbal items: 2089 versus 1634 msecs, $F(1,23) = 81.36$, $p < .0001$, MSe = .00753, for the vocabulary vs keyword tasks comparison, and 2089 versus 1602 msecs, $F(1,23) = 81.94$, $p < .0001$, MSe = .00882, for the vocabulary vs English tasks compairson.

For the delayed results (block 3 only), mean retreival time as a function of retrieval task and delay condition (1-week or 1-month) is displayed in Figure 10. The statistical results are summarized in Table 3. Again, retrieval times for the vocabulary retrieval task were longer than those for the subtasks (2882 versus 2170 msecs), $F(1,22)=72.35$, $p < .0001$, MSe = .01180. Post hoc comparisons of the vocabulary retrieval to each of the subtasks individually confirmed that the vocabulary task took longer than either of the subtasks: for the vocabulary versus keyword tasks comparison, 2882 versus 1941 msecs, $F(1,22) = 74.74$, $p < .0001$, MSe = .01733; and for the vocabulary vs English tasks comparison, 2882 versus 2439 msecs, $F(1,22) = 2.79$, $p < .0001$, MSe = .00981.

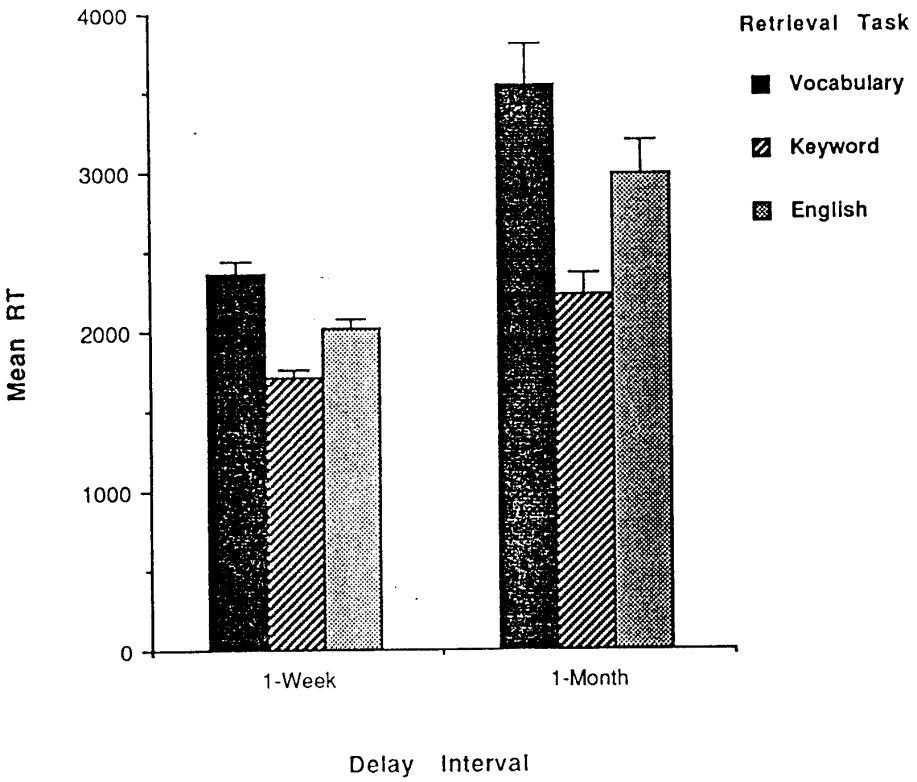


Figure 10. Mean retrieval time as a function of retrieval task and delay interval for Experiment 1.

Table 3

Summary of Effects in Analysis of Variance for Delayed Session Latency Results

Source	df	F	p <	MSe
Delay	1,22	31.80	.0001	.05
Task	2,44	42.71	.0001	.02
Verbal Condition	1,22	2.02	.169	.02
Delay x Task	2,44	.91	.411	.02
Delay x Verbal	1,22	0.00	.988	.02
Task x Verbal	2,44	1.69	.197	.01
Task x Delay x Verbal	2,44	0.42	.659	.01

Facilitation effects. The pattern of latencies for the different retrieval tasks was consistent with the mediation model: the time necessary to carry out the vocabulary task was longer than that required to perform either of the subtasks. However, the latency analyses were performed on task latencies that were computed by averaging across the 6 different task orders to minimize possible order effects. We therefore performed a more focused analysis on the specific task orders themselves. According to the mediation model, we might expect systematic facilitation for a specific retrieval task depending on the retrieval task preceding it.

For example, if practice on the vocabulary retrieval task facilitated performance on the subtasks or practice on the subtasks facilitated performance on the vocabulary retrieval task, this would constitute strong evidence for the mediation model. A direct model would have great difficulty explaining such results because it assumes the three retrieval tasks are unrelated to one another. Actually, a proponent of the direct model might argue that the vocabulary task is related to the keyword task in that both use the Spanish word as a cue. However, this account would not be able to explain how the vocabulary retrieval task and the English retrieval task might facilitate one another.

In looking for specific facilitation effects, there were three comparisons we examined. In each case, we compared the retrieval times for a specific retrieval task when it followed either of the other two retrieval tasks. First, we compared retrieval times for the vocabulary task when it followed the keyword subtask to retrieval times for the vocabulary task when it followed the English subtask. Second, we compared the retrieval latencies for the keyword subtask when it followed the vocabulary task to when it followed the English retrieval task. Third, we compared the latencies for the English subtask when it followed the vocabulary retrieval task to when it followed the keyword subtask. Figure 11 illustrates the 3 comparisons.

The first comparison was of little interest, as both subtasks should facilitate the vocabulary retrieval task, according to the mediation model, though the amount of facilitation might differ. The second comparison was also of little interest for two reasons: First, both of the facilitating tasks involve accessing the keyword. The vocabulary retrieval

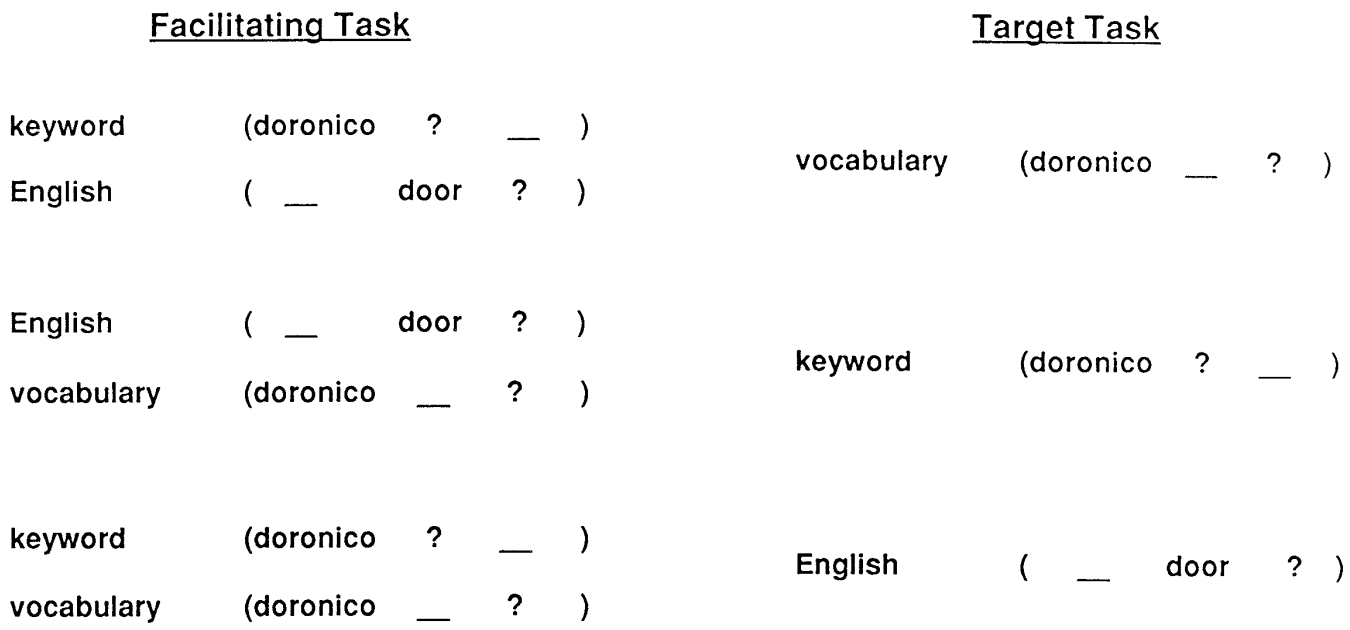


Figure 11. Task arrangements to test facilitation effects.

task requires accessing the keyword as part of accessing the English definition; the English subtask requires processing the keyword cue to retrieve the English equivalent. In both cases, the keyword is accessed as part of the first task (the vocabulary or English task) and generated as a response in the second task (keyword task). Therefore, either the vocabulary retrieval task or the English subtask might facilitate the keyword subtask.

The third comparison, though, provided a clear-cut test between the mediation and direct models. According to the mediation model, the English retrieval task should be faster after practicing the vocabulary retrieval task than after practicing the keyword retrieval task because the vocabulary and English retrieval tasks share intermediate processing steps, whereas the keyword and English retrieval tasks do not. According to the direct model no such facilitation effect should be expected because this model posits no intermediate processing steps, and the possibility of stimulus effects is controlled for because both the keyword and vocabulary retrieval tasks employ the Spanish word as a retrieval cue.

To test these possible facilitation effects, planned comparisons were performed on the retrieval times for a specific task preceded by each of the other two tasks. As these analyses were performed on retrieval times for correct items, only the 1-week subjects' data were analyzed. In addition, the analyses were restricted to the second test of block three (the first delay block) for two reasons: (1) Test two was the only point at which a vocabulary item had been tested on one but not the other of the two retrieval tasks; and (2) Block three was the only block immune to cross-block facilitation effects (e.g. on block one, subjects had just finished practicing the vocabulary retrieval task for all vocabulary items as part of the dropout phase). All analyses were performed on subject means computed for the different tasks and task orders (collapsed across verbal report, as a preliminary analysis showed no effect due to verbal report).

Mean retrieval times for each of the three retrieval tasks as a function of the preceding retrieval task are presented in Figure 12. As predicted by the mediation model, the English subtask was faster after practice on the vocabulary retrieval task than after practice on the keyword retrieval task. The difference between the facilitated and

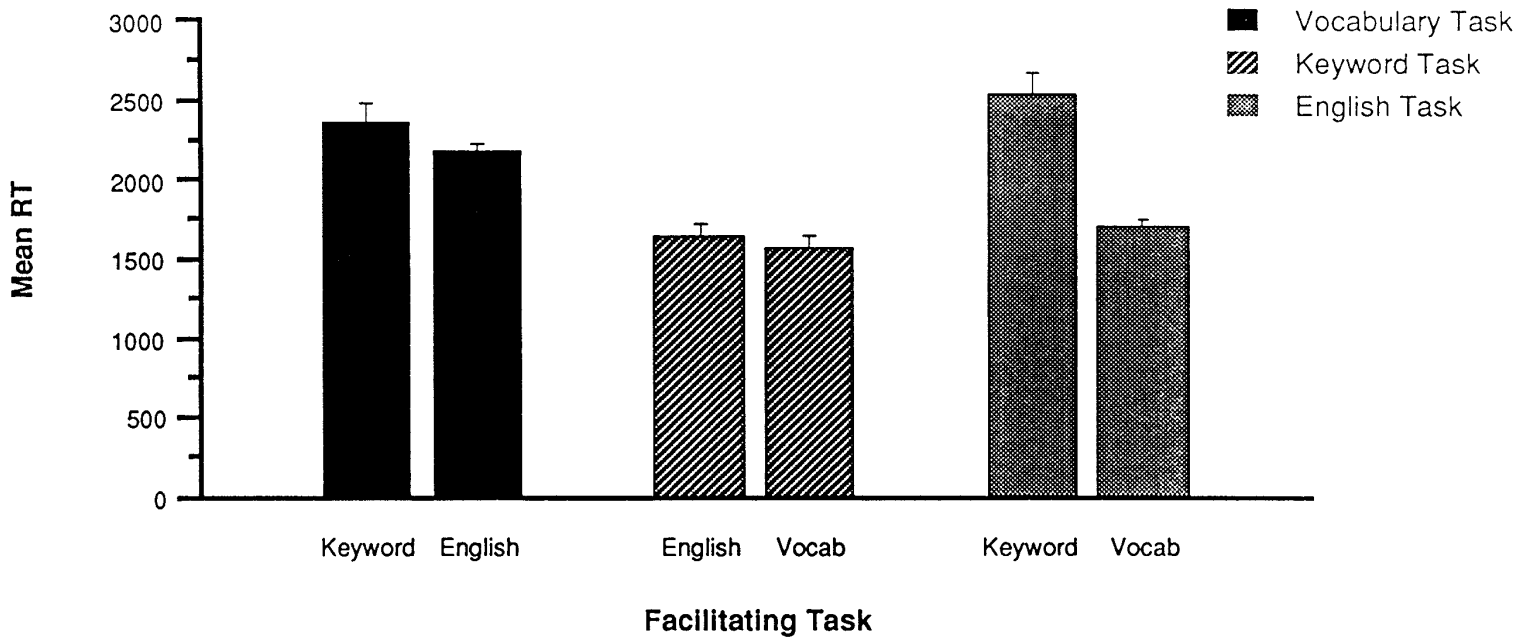


Figure 12. Mean retrieval time for different retrieval tasks as a function of the facilitating retrieval task (block 3 only).

unfacilitated English subtask was greater than 800 milliseconds and was statistically reliable, $F(1,11)=36.2$, $p < .0001$, MSe=.010. Whether the keyword subtask was preceded by the English subtask or by the vocabulary retrieval task did not reliably affect the latency for the keyword retrieval task, $F < 1$. The vocabulary task was faster after practice on the English subtask than after practice on the keyword subtask (2180 versus 2367 msecs); however, the difference was not statistically significant, $F(1,11) = 2.04$, $p < .18$, MSe = .007.

Verbal reports. All of the results reported so far provide evidence for mediated retrieval in the keyword method by comparing performance on the subtasks and the vocabulary task. The retrospective reports that we collected provide further evidence for the retrieval of the keyword and interactive image as part of the vocabulary retrieval task itself.

According to the mediation model, the keyword and interactive image are activated in working memory in retrieving the English definition of the Spanish word. Therefore, subjects should report the keyword and/or interactive image as part of retrieving the English equivalents of the Spanish words. On the other hand, according to the direct model, we would expect the majority of the reports to mention only the Spanish word and the English word.

The encoding scheme for the verbal reports was developed prior to collecting the data. The encoding categories were developed by the experimenter, who did not participate in the actual encodings of the protocols. Two encoders were employed to test the reliability of the encoding scheme. Both encoders used the set of rules in Appendix C. Agreement was 92 per cent or higher for each of the five variables encoded. After a general discussion of the encoding rules with the experimenter, with no reference to specific protocols, the encoders recoded the protocols on which they had disagreed. Agreement was 99% this time.

We will briefly describe the encoding scheme used for the verbal reports; a complete description appears in Appendix C. For each retrospective report, five variables were encoded, corresponding to the following 5 components: the Spanish word, the Spanish-keyword link, the keyword, the keyword-English word link, and the English word. For each

variable a number coded whether the corresponding component or link was present ('1'), absent ('0'), or insufficient information to decide ('9'). The encodings for the Spanish word, the keyword, and the English word were straight-forward: if the word was mentioned in the retrospective report or had appeared as part of the computer display (e.g. the Spanish word appeared as a cue in the vocabulary retrieval task), it was encoded as present. Pilot studies suggested that the link between the Spanish word and keyword was difficult to encode using retrospective reports. Therefore, for the current experiments, this link was simply encoded as a '9.'

The rules for encoding the keyword-English word link were more involved than those for the other components and are explained in detail in Appendix C. However, we will briefly review these rules as well.

The category of 'absent' applied when both the keyword and English word were mentioned but there was no mention of any relationship or link between the two words (e.g. 'I thought of 'door' then 'leopard'). If even a vague relationship was mentioned (e.g. "I simply correlated 'door' with 'leopard' " or "I got 'leopard' from 'door'"), the 'absent' category was not used to encode the keyword-English word link.

The category of 'present' for the keyword-English word link was reserved for reports in which the correct keyword and English word were mentioned and a semantic or imagistic link relating the keyword and English word was mentioned. It was not sufficient to specify, for example, "I just remembered my image of the chicken and the pole." The subject had to specify a relationship between the chicken and the pole (e.g. "I remembered my image of the chicken on the pole."). Though this link was coded as a '1' (for a relation specified between the keyword and English word) or as a '3' (if reference to an image was made--e.g "I remembered my picture of the chicken on top of the pole"), for the current analyses these two categories were combined.

The category of 'ambiguous' was reserved for those reports that could not be encoded using the above 3 categories: for example, when subjects used non-specific relations such as 'I correlated chicken with pole.' About 9% of the reports fell into this category. In addition, for about 0.4% of the items subject reported the wrong keyword. Neither of these categories was analyzed.

Based on the above encoding scheme, reported mediation was categorized as follows: (1) keyword+link: keyword and link between keyword and English word reported; (2) keyword: keyword but no link between keyword and English word reported; and (3) unmediated: No mention of the keyword and/or any link to the English word. Categories one and two are mediated retrieval and category three is unmediated retrieval. Examples of each of the encoding categories are displayed in Figure 13.

Proportion of items (for all 4 blocks) falling into each of the above categories of reported mediation are displayed in Figure 14. Subjects invariably reported some type of mediation in retrieving the vocabulary items: the keyword and interactive image were reported .69 of the time, and the keyword alone was reported .23 of the time. For a small proportion of retrievals subjects reported no mediation at all (about .08). These results are strikingly consistent with the results for the recall analyses. In both cases, the vast majority of correct vocabulary retrievals involved mediation.

Given the converging support provided by the recall and latency analyses, our confidence in the verbal report analysis was quite strong. However, we had yet another means to validate the verbal reports. For each retrieval, we collected a response latency in addition to a retrospective verbal report. If, these reports accurately reflect the processing steps that subjects carried out to retrieve the English definitions, then we should expect a systematic relationship between the content of the reports and the response latencies. Specifically, the latencies should be increasingly longer as a function of the number of medating steps necessary to retrieve the English definition from the Spanish word.

Mean retrieval times for the different types of mediation reported are displayed in Figure 15. Analyses were performed on subject means for each mediation type. For 7 of the subjects, there were no observations for one or more of the mediation categories. Therefore, the analyses were performed on data from the 17 subjects who had observations for each of the 3 categories. Two questions were asked: (1) Were retrieval times faster when there was no reported mediation versus when there was mediation reported (either the keyword+link or simply the keyword)? and (2) Were the retrieval times faster when just the keyword was reported compared to when the keyword and a specific mediating link between the keyword and English

Encoding Category	Example Protocol
Keyword + Link	Saw the Spanish word 'doronico'/ Remembered 'door'/ And then the 'leopard' leaping through the 'door'
Keyword	Saw the Spanish word/ Thought of the keyword 'door'/ And then 'leopard'
Unmediated	Saw 'doronico'/ And thought 'door'

Figure 13. Encoding Categories for the verbal reports.

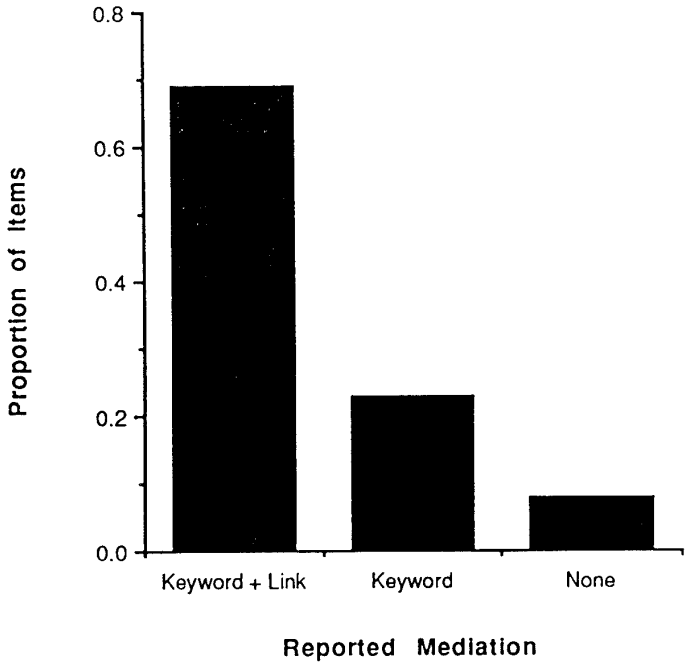


Figure 14. Proportion of vocabulary items for different categories of mediation (Experiment 1).

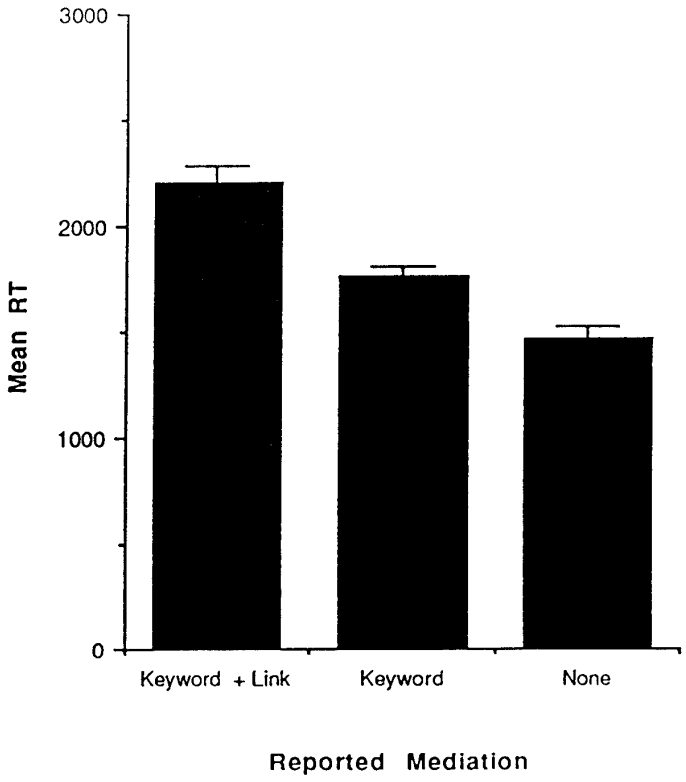


Figure 15. Mean retrieval time as a function of type of mediation reported (Experiment 1).

word were mentioned? To answer these questions, planned comparisons were performed on the mean retrieval times for the different categories of mediation. Retrieval times were reliably faster when no mediation at all was reported (1463 msecs) than when some type of mediation was reported (1965 msecs), $F(1,16) = 37.64$, $p < .0001$, $MSe = .012$. Furthermore, retrieval times were reliably faster when only the keyword was reported (1758 msecs) than when the keyword and a link between the keyword and English word was reported (2210 msecs), $F(1,16) = 54.74$, $p < .0001$, $MSe = .00326$. Finally, a post hoc comparison of the times for the keyword mediated (1758 msecs) and unmediated retrievals (1463 msecs) was also significant, $F(1,16) = 14.27$, $p < .002$, $MSe = .00980$.

Discussion

All of the results in the first experiment strongly supported the mediation model, and none of the results supported the direct model. As predicted by the mediation model, overall recall for the vocabulary items was less than recall performance on the subtasks. More importantly, for individual vocabulary items, if subjects were able to recall the English definition given the Spanish word, they were nearly always able to perform both subtasks, whereas if subjects were unable to recall the definition, they were almost never able to perform both subtasks. The latency results revealed that subjects were slower retrieving the English definitions of the Spanish words than they were performing either of the subtasks. The latency analyses also demonstrated that performance on the English subtask could be facilitated if the vocabulary retrieval task occurred on the previous subtest. The verbal reports for the vocabulary task revealed that subjects almost always reported retrieving the keyword (and usually the interactive image as well) as part of the process of retrieving the vocabulary definition. Finally, response latencies increased as the number of reported mediating steps increased.

CHAPTER III

EXPERIMENT 2

The first experiment provided considerable evidence to support a mediated retrieval model for vocabulary items learned using the keyword method. However, in the first experiment, subjects not only practiced retrieving the vocabulary definitions for each vocabulary item; they also practiced the two subtasks. One limitation of this procedure might be that by having subjects continue to perform the subtasks along with the vocabulary task, we encouraged subjects to retrieve the vocabulary items by simply carrying out the two subtasks. A closer analog to the usual vocabulary-learning situation would be for subjects to learn a vocabulary item and then simply practice the vocabulary task without practicing the subtasks. In fact, an important issue is whether, once subjects have practiced retrieving the vocabulary items in the absence of the subtasks, the mediated model is still an adequate account of the retrieval process. Alternatively, it might be that as soon as subjects begin to practice the vocabulary task itself, the direct model becomes a better account of the retrieval process.

According to the direct model, the Spanish word and its English definition are connected by a direct link between the two words. Therefore, as soon as subjects begin retrieving the vocabulary definitions, the direct link connecting the Spanish and English words should be strengthened, whereas the other cognitive structures (e.g. the keyword and the interactive image), which were simply activated by requiring subjects to perform the subtasks, would no longer be accessed and should no longer be strengthened. In fact, if we assume a decay process for unaccessed cognitive structures, these structures should actually weaken with time. If we retested subjects on the two subtasks as well as the vocabulary task after a period of practice on only the vocabulary task, we should expect performance on the vocabulary task to have improved but performance on the subtasks to have declined (e.g. longer latencies or perhaps even retrieval failure after a sufficient delay).

On the other hand, according to the mediation model, the only way to retrieve the English definition from the Spanish word is via the

mediating links between the Spanish and English words. There is no direct link between the Spanish word and English word. Consequently, retrieving the English definition should strengthen the mediating links between the Spanish word and the English word. Thus, if we retested subjects on the two subtasks and the vocabulary task after practice on the vocabulary task, we would expect performance on both the vocabulary task and the subtasks to have improved.

A third possibility, not directly examined in the current experiment, has been suggested by previous researchers. According to this account, retrieval is at first mediated, but is later modified as a function of practice. For example, in addition to the mediated retrieval path, a direct link is established between the Spanish word and English word. According to this account, the Spanish word and English word are at first connected only by the mediating keyword and interactive image links. However, as the vocabulary definition is repeatedly retrieved, a direct link between the Spanish word and English word is added so that in addition to the mediated retrieval route between the Spanish and English words, a second direct retrieval route is established between the Spanish word and English word. At first this link would be too weak to support retrieval of the English definition, so the mediated pathway continues to support retrieval. However, with practice, the direct link would become strengthened and able to support retrieval. Some keyword researchers have suggested exactly this sort of possibility (e.g. see Atkinson and Raugh, 1975).

Until now, no experiments have directly investigated the effect of practice retrieving the English definitions of foreign words learned by the keyword method. In this second experiment, we investigated whether either the direct or mediation model provides an adequate account of the retrieval process after a limited period of practice retrieving the vocabulary items.

To investigate this issue, we again had subjects learn a list of Spanish-English vocabulary items using the keyword method and we then tested all vocabulary items on our 3 three retrieval tasks (the vocabulary, keyword, and English retrieval tasks). However, a critical change in the second study was the introduction of a practice phase, in which different items received different types of practice, prior to the final test phase in which all vocabulary items were retested on the three retrieval tasks.

During the practice phase, some items received practice only on the vocabulary retrieval task (the vocabulary-practice condition). For other items, only the two subtasks were practiced (subtask-practice condition). Finally, some items were not practiced at all (no-practice condition). The purpose of the vocabulary-practice condition was to examine the effect of continued practice retrieving the English definitions of the Spanish word. Depending on whether the mediation model or direct model correctly characterizes retrieval, the effect of such practice on the final testing of the three retrieval tasks should be quite different. The subtask- and no-practice conditions served as controls for the vocabulary-practice condition.

The specific predictions made by the two models were as follows. Both the direct and mediation models predicted improved recall on the vocabulary task following vocabulary practice compared to no-practice. However, the direct model predicted that recall on the vocabulary task would be better after vocabulary practice than after subtask practice because subtask practice does not access the direct link between the Spanish and English words, whereas vocabulary practice does. On the other hand, according to the mediation model, either subtask practice or vocabulary practice requires retrieving similar mediating structures. Therefore, final performance on the vocabulary retrieval task would not be expected to differ after subtask versus vocabulary practice.

The second prediction for the recall results focused on final recall for the subtasks. According to the direct model, retrieval of the vocabulary definitions does not involve accessing the mediating links (e.g. the keyword) involved in the subtasks. Therefore, recall on the subtasks at final test was expected to be no better after vocabulary practice than after no practice. In either case, subjects are not accessing the memory structures necessary to performing the two subtasks. However, according to the mediation model, vocabulary practice requires accessing the intermediate links between the Spanish word and English word. Therefore, recall performance on the subtasks should be better following vocabulary-practice relative to no practice at all.

A similiar set of predictions was made for the latency data. Focusing first on the final test for the vocabulary task, the direct model predicted that latencies would be faster following practice on the

vocabulary task than after practice on the subtasks, whereas the mediation model predicted no such difference. For final performance on the subtasks, the direct model predicted latencies for the subtasks would be no faster following vocabulary practice than following no practice. The mediation model, on the other hand, predicted that the subtask latencies would be faster following vocabulary practice than following no practice.

In addition to the effect of vocabulary retrieval practice on the recall and latency data for the various retrieval tasks at final test, we considered more direct evidence for mediation by analyzing the reported mediation for the vocabulary retrieval task, as we did in the first experiment. According to the direct model, the retrospective verbal reports for the vocabulary task were predicted to show a high incidence of unmediated retrievals, with few or no reports of mediating keywords or images. However, according to the the mediation model, reported mediation was expected to occur on almost every trial, as found in the first experiment.

Method

Subjects. The subjects were 18 undergraduate students at the University of Colorado, recruited by an ad placed in the campus newspaper. Subjects were paid for their participation in the experiment. The experiment was restricted to individuals who had no previous training in Spanish.

Materials and apparatus. The stimuli were 45 Spanish-English pairs along with their keywords. Forty-one of the Spanish-English pairs and keywords from the first experiment, along with four new Spanish-English pairs and keywords, were used to create the set of 45 stimuli used in the second experiment (see Appendix B for the complete list of items). The four new items were selected using the same criteria described in experiment 1. There were no "control" items in this experiment. Time constraints forced the elimination of the controls. Moreover, further analyses of the data for experiment 1 revealed no differences between the control items and the other items.

As there were three practice conditions (vocabulary-practice, subtask-practice, and no-practice), three lists were created, so that each item appeared in each of the practice conditions across the three lists. Three of the six possible task orders were used in this experiment and these were counterbalanced across the items in each list so that every item in the experiment appeared in each practice condition and each task order twice, yielding two replications of the design.

Procedure. The procedure was similar to that in Experiment 1, with a few minor changes. As in the first experiment, there were three phases: acquisition, dropout, and test. However, in the current experiment, a practice phase was interposed between the dropout and test phases. A subset of the 45 items were practiced during this phase as described below. The final test phase consisted of a single test block of three subtests, so that each item was tested on all three retrieval tasks. Subjects first went through the acquisition, dropout, and practice phases, which took approximately two hours. All subjects then completed the single test block (three tests). For each test the presentation of items was randomized.

During the practice phase, subsets of the 45 vocabulary items were practiced differently within the different practice conditions. For the vocabulary-practiced items, subjects performed only the vocabulary-retrieval task; for the subtask-practiced items, they practiced only the two subtasks; the no-practice items were not practiced at all during this phase. Thus, 30 items (15 vocabulary-practice items and 15 subtask-practice items) were practiced during the practice phase.

As it was necessary to practice both subtasks for the subtask-practice items, two trials were required for a subtask-practiced item within each practice block. Therefore, the 15 subtask-practice items were blocked and counterbalanced, so that for half the items the keyword subtask occurred first followed by the English subtask, whereas for the other half the English subtask was performed first followed by the keyword subtask. Presentation order of the items within each practice block was randomized for each subject. There were six practice blocks.

The computer display of the vocabulary items was modified as follows. Instead of presenting and testing items in a linear format, as in the first experiment, the three words, and the symbols that later replaced

the various words during the dropout, practice, and test phases, were displayed in a staircase format, as shown in Figure 16.

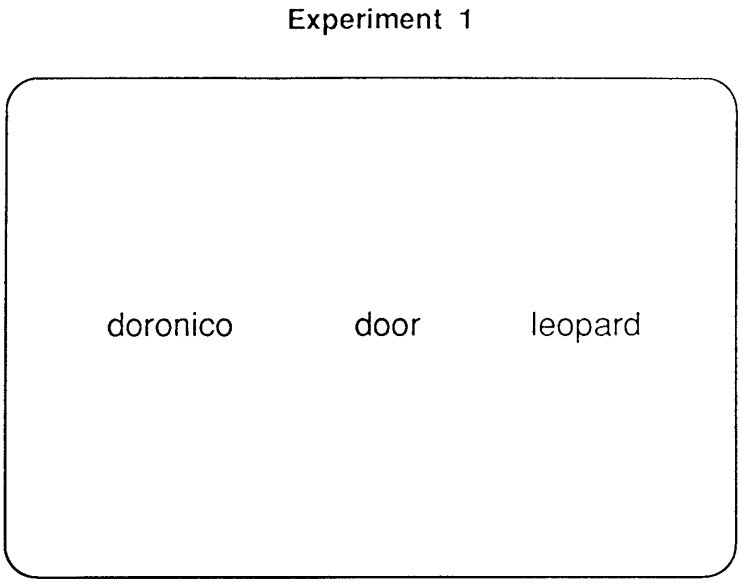
In the current experiment, subjects had 10 instead of 15 seconds to respond on each trial in the dropout, practice, and test phases. Finally, subjects gave verbal reports on all acquisition and test trials; however, no reports were collected during the dropout and practice phases.

Results

Voice key errors occurred on 3.8% of the trials and were excluded from the analyses. As in Experiment 1, subject means were computed for each cell (mean latency and mean proportion recall). There were 3 types of practice and three retrieval tasks, yielding 9 scores for each subject. As all subjects received the same practice and testing, the design was completely within-subjects, with practice condition and retrieval task as the two within-subjects factors. However, in lieu of omnibus tests, specific comparisons were performed to test the various predictions of the direct and mediation models.

Proportion of words recalled. Mean proportion of items correctly recalled for the vocabulary task as a function of the type of practice is displayed in Figure 17. Contrary to the prediction of the direct model, recall performance on the vocabulary task was no better after vocabulary-practice than after subtask-practice, $F(1,17) < 1$. Mean proportion of items recalled for the keyword and English subtasks as a function of the type of practice is displayed in Figure 18. Mean proportion of items recalled for the keyword subtask was no better following vocabulary-practice than following no-practice (.91 versus .89 respectively), $F(1,17) < 1$. However, mean proportion of items recalled for the English subtask was higher after vocabulary-practice than after no-practice (.95 versus .89 of the items respectively), $F(1,17) = 5.98$, $p < .026$, $MSe = .001$, as predicted by the mediation model.

None of the recall results supported the direct model. The only reliable difference, improved recall on the English subtask following vocabulary practice, supported the mediation model.



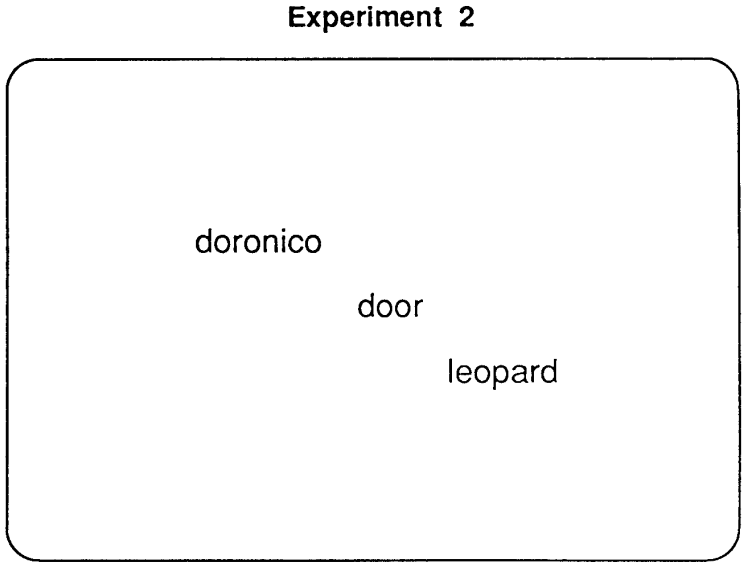


Figure 16. Computer displays for Experiments 1 and 2.

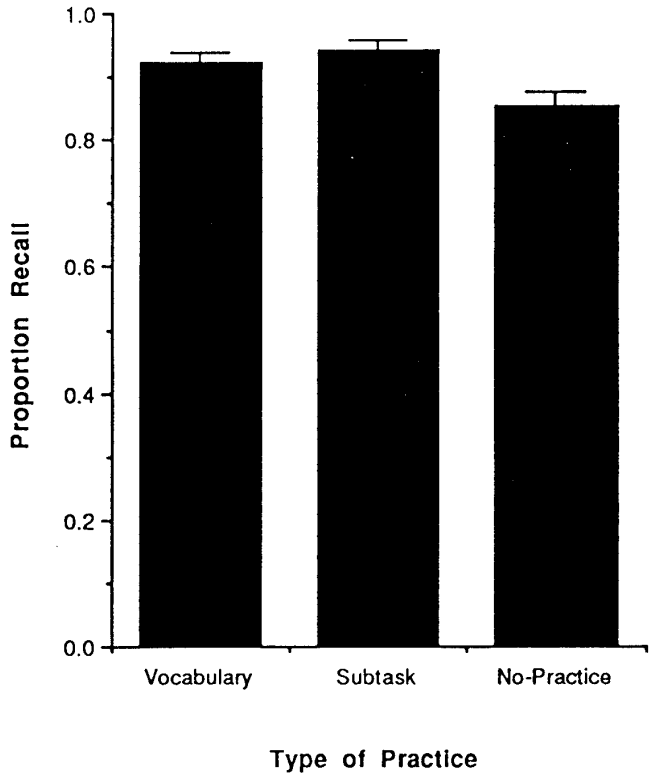


Figure 17. Mean proportion of items recalled for the vocabulary task as a function of practice type (Experiment 2).

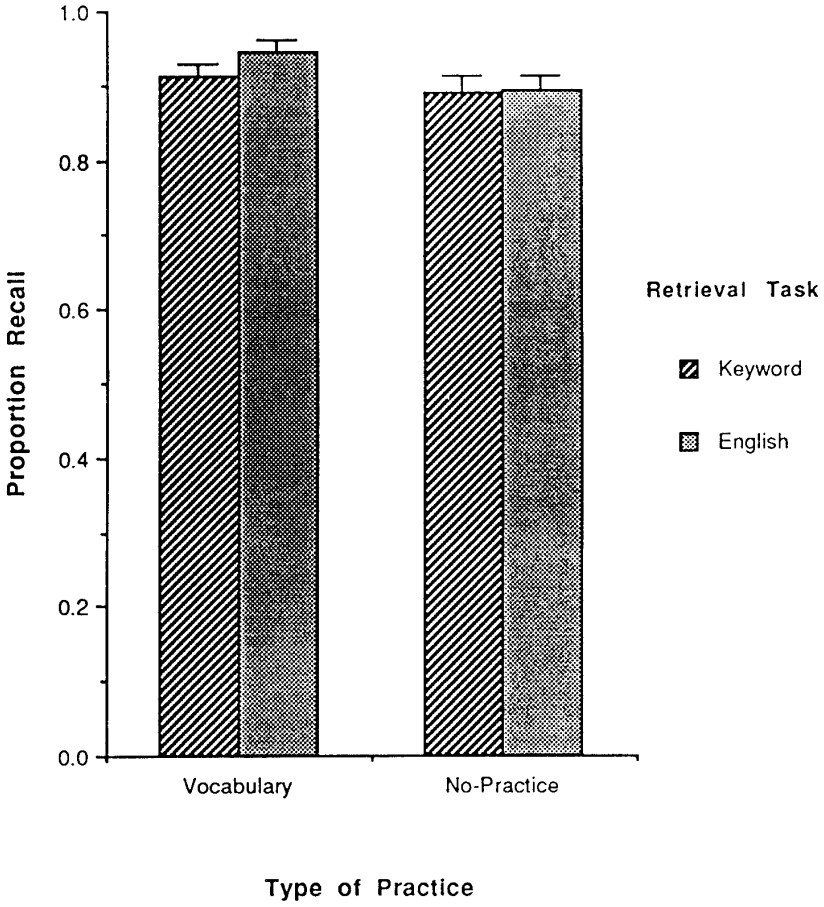


Figure 18. Mean proportion of items recalled for the subtasks as a function of practice type (Experiment 2).

Retrieval Times. Mean retrieval times for the vocabulary retrieval task as a function of practice condition are displayed in Figure 19. As predicted by the direct model, retrieval times for the vocabulary task were faster after vocabulary practice than after practice on the two subtasks (1874 msecs versus 2073 msecs), F (1,17) = 7.98, $p < .012$, MSe= .001. However, retrieval times for the subtasks were reliably faster after vocabulary practice than after no practice (1697 versus 1910 msecs for the keyword subtask; 1750 versus 2176 msecs for the English subtask), $F(1,17) = 17.84$, $p < .001$, MSe = .001 and $F(1,17) = 26.91$, $P < .0001$, MSe = .01, as predicted by the mediation model (see Figure 20).

Verbal Reports. The verbal reports for the vocabulary retrieval task were analyzed using the same encoding scheme described in the first experiment. As before, only the correct retrieval trials were included in the analysis. Less than 1% of the protocols were missing, inaudible, or miscoded in some way. A few items (about 4%) were encoded as ambiguous. The proportions calculated for the different mediation categories included the ambiguous category; however, the ambiguous items were not analyzed further.

Proportion of reported mediation for each of the different types of mediation (keyword+link, keyword, and unmediated) as a function of the practice condition is displayed in Figure 21. Focusing on the vocabulary-practice condition, contrary to the prediction of the direct model, the majority of the reports were not unmediated: only .17 of the reports fell into the unmediated category. Instead, most of the reports (.79) still mentioned accessing the keyword (.55 mentioned both the keyword and image; .24 mentioned just the keyword), supporting the mediation model, as in the first experiment.

Though not supportive of the direct model, the reports were not quite as supportive of the mediation model as in the first experiment, where there were almost no reports of unmediated retrieval. In this respect, a post hoc comparison of the proportion of unmediated reports after vocabulary-practice versus after subtask-practice is interesting. According to the mediation model, these two types of practice should be equivalent, as both require accessing the mediators. We should therefore

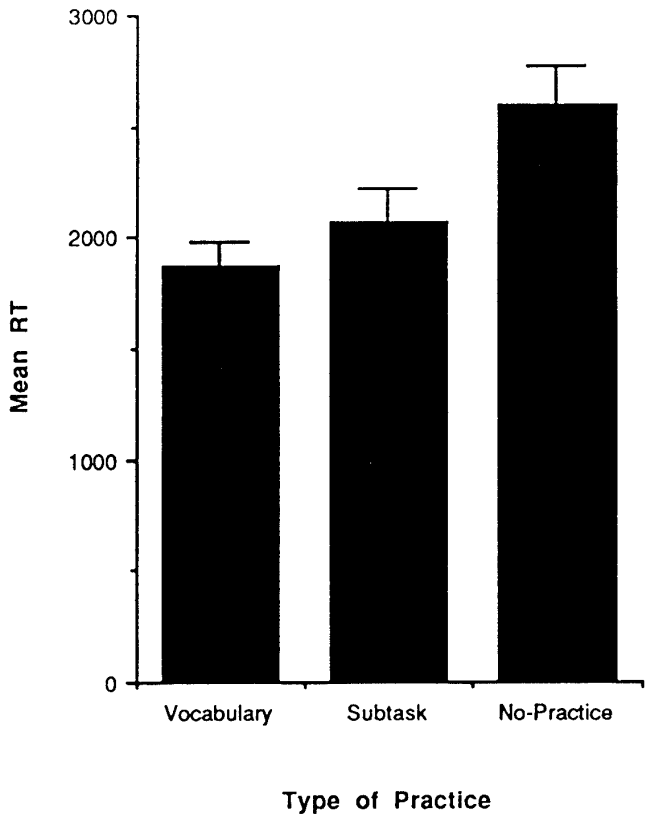


Figure 19. Mean retrieval time for the vocabulary task as a function of practice type (Experiment 2).

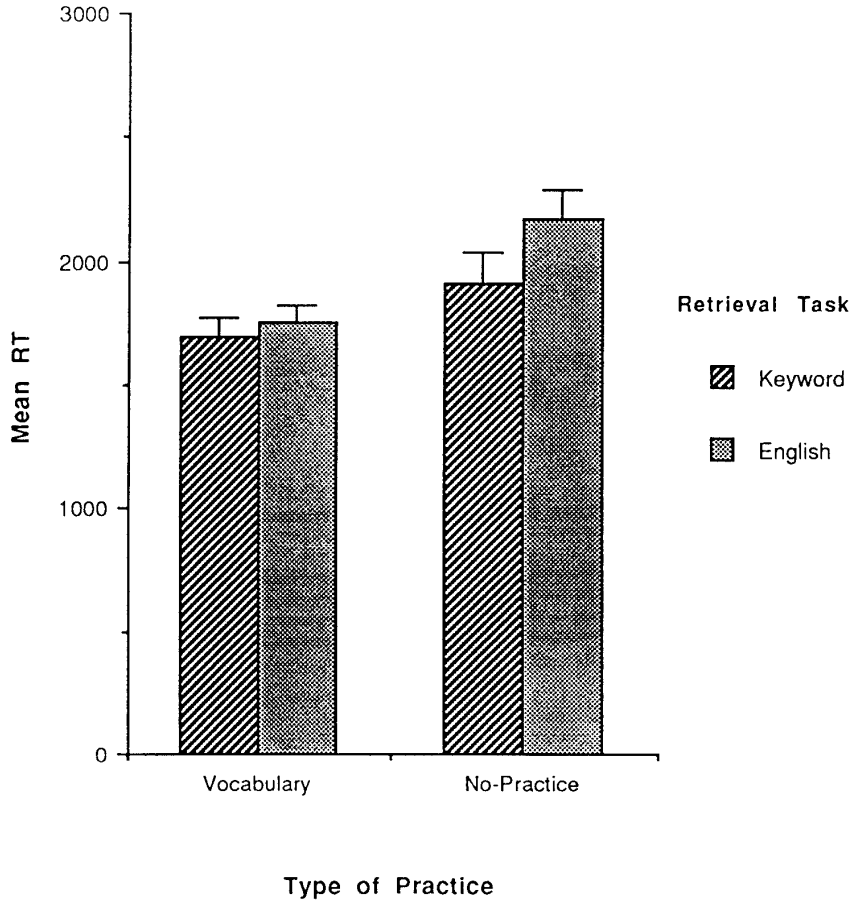


Figure 20. Mean retrieval time for the subtasks as a function of practice type (Experiment 2).

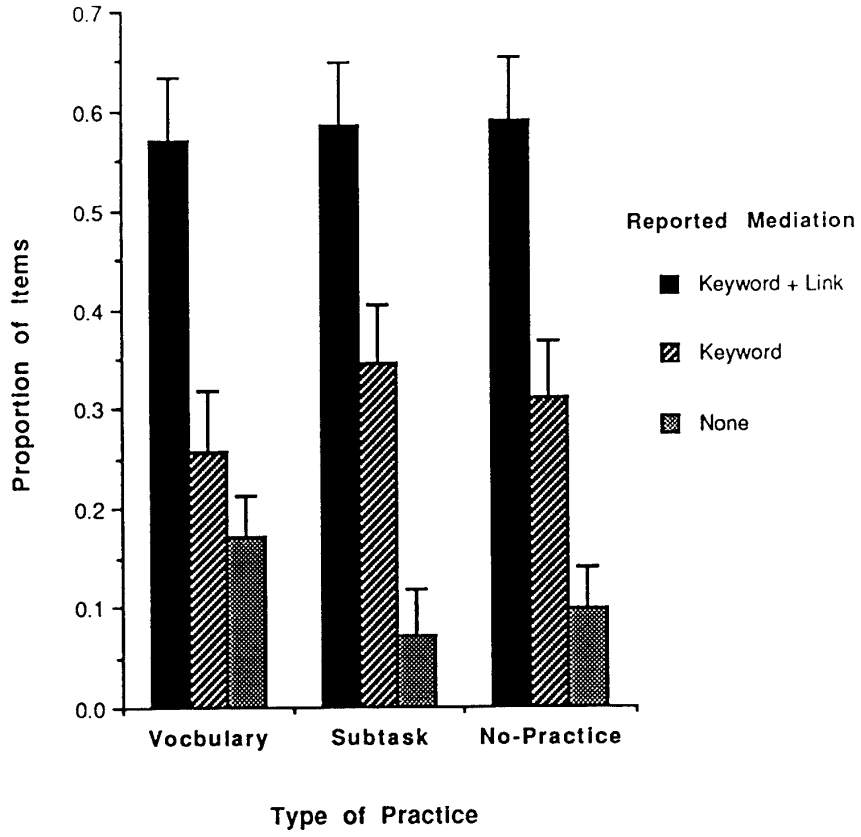


Figure 21. Proportion of vocabulary items for different categories of mediation as a function of practice type (Experiment 2).

expect to see the same pattern of reported mediation for the two types of practice. In fact, though, there were more reports of unmediated retrieval after practicing the vocabulary retrieval task than after practicing the two subtasks (.17 versus .07) and this difference was reliable, F (1,17) = 13.01, $p < .002$, MSe = .00692.

Finally, the retrieval times for the different types of reported mediation are presented in Figure 22 (only the data for the 15 subjects who had RTs for each of the three types of reported mediation are included). Retrieval times were faster for unmediated retrieval (1626 msecs) compared to mediated retrieval (2291 msecs) and the difference was reliable, F (1,14) = 13.67, $p < .002$, MSe = .0349. Retrieval times were also faster when the reported mediation included just the keyword (2195 msecs) than when the report mentioned the keyword and an interactive image (2392 msecs); however, this difference was not statistically reliable, F (1,14) = 2.34, $p < .15$, MSe = .008, though the direction of the difference was consistent with that found in the first experiment.

Discussion

The pattern of results in the second experiment is quite interesting. First, there seems to be little evidence for the direct model. The one result consistent with the direct model was the increased retrieval speed for the vocabulary retrieval task after practicing the vocabulary task itself instead of practicing the subtasks. However, the direct model cannot explain why the subtasks were improved by practicing the vocabulary retrieval task, which according to the direct model is not related to the task of retrieving the keyword or using the keyword to retrieve the English equivalent. Given the results of the first experiment and the results here, the direct model does not appear to be an viable account of how the English definitions of the Spanish words are retrieved from memory.

What of the mediation model? A number of the results supported the mediation model. Practicing the vocabulary retrieval task did improve performance on the subtasks relative to not practicing at all. Both the keyword and English subtasks were faster after vocabulary retrieval practice. In addition, the English subtask showed better recall

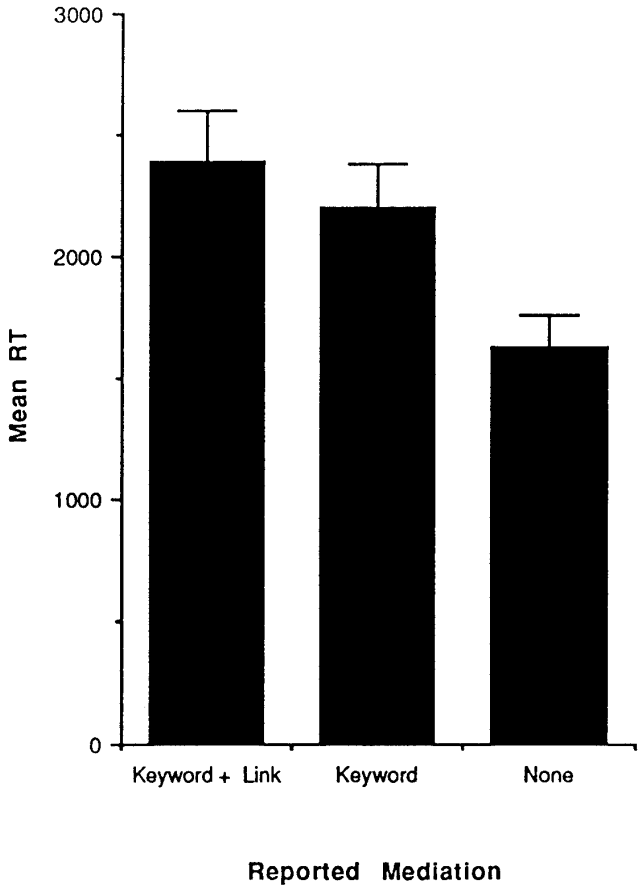


Figure 22. Mean retrieval time as a function of type of reported mediation (Experiment 2).

performance after vocabulary retrieval practice. Of course, any claims based on the recall results must be qualified given the high level of recall overall. However, it is important to note, that given the lack of dramatic recall differences for the different practice conditions, our confidence in the latency differences is strengthened, for we can rule out the possibility of speed/accuracy tradeoff explanations of the differences. In short, the improvements in subtask performance after vocabulary retrieval practice support the mediation model. In addition, the verbal reports strongly suggest mediated retrieval, as even after practicing the vocabulary retrieval task, subjects continued to report the keyword and/or the image on the majority of retrieval trials.

Although the mediation model does a reasonable job of explaining most of the current results, it does not explain all of them. First, the vocabulary retrieval task was faster after vocabulary practice than after subtask practice. According to the mediation model this should not be the case, as either type of practice requires accessing the same mediating cognitive structures. However, our results clearly suggest that practicing the vocabulary retrieval task provided some additional benefit in retrieving the English definitions of the Spanish words. One possible explanation is that for some items, more direct access to the English definition was established with practice, and that as this occured, it was no longer necessary to rely exclusively on retrieving the mediating keyword and image. For the items retrieved via this direct process, we would expect increased retrieval speed, as it was no longer necessary to carry out the mediating steps originally needed to access the English definition.

This account could also explain the other result difficult to account for using the simple mediation model: the increase in reports of unmediated retrieval following vocabulary practice rather than subtask practice. Again, according to the mediation model, these two types of practice are equivalent, so the proportion of unmediated retrievals should have been the same. The increased number of unmediated retrievals reported is consistent with the possibility that for some items in the vocabulary-practice condition the English equivalents were directly accessed.

Of course, the current results do not really provide any direct evidence for this change of mediated to direct retrieval. Nor is it clear exactly what mechanism might underlie such a change. One possibility, as

previous researchers have suggested, would be the addition of a new, direct link between the Spanish word and its English definition. However, another possibility, perhaps even more plausible, would be some sort of composition mechanism (Anderson, 1983; Neves & Anderson, 1981), such that the separate steps of retrieving the keyword, then using the keyword to retrieve the interactive image, and finally retrieving the English definition, are composed into a single production that retrieves the English definition directly from the Spanish word.

In sum, the results of the second experiment support the mediation model more strongly than the direct model. However, it appears that with continued practice retrieving the English definitions, the mediation model itself may not be a complete account. Some modification of the mediation model seems necessary.

CHAPTER IV

GENERAL DISCUSSION

The experiments reported here provide considerable support for the mediation model and little support for the direct model. In the first experiment, each of the predictions of the mediation model was confirmed. Success in recalling the English definitions of the Spanish words depended on whether subjects could successfully perform both of the subtasks. When the two subtasks were successfully performed, the vocabulary definitions were recalled as well; when neither of the subtasks was successfully performed, the English definition of the Spanish word was almost never recalled. These subtasks measured whether subjects could recall the keyword mediatior from the Spanish word and whether the keyword could cue retrieval of the English definition.

These results were strengthened by the latency results for the retrieval tasks. As predicted by the mediation model, the time required to perform the vocabulary task was always greater than that required to perform the subtasks, as one would expect if each subtask corresponds to only a subset of the retrieval steps involved in the vocabulary task itself. More importantly, the vocabulary retrieval task facilitated the English retrieval subtask whereas the keyword subtask did not. This result strongly supported the mediation model because the mediation model assumes that the two subtasks correspond to different steps in the retrieval of the vocabulary items, and therefore, the subtasks were not be expected to help one another, whereas the vocabulary retrieval task was expected to help the English subtask.

Finally, the retrospective verbal reports for the vocabulary retrieval task itself provided additional evidence for the mediation account. Subjects invariably reported keyword mediators in retrieving the English equivalents of the Spanish words, and the content of these reports was supported by the increasingly longer latencies for the reports with more mediating steps.

The second experiment offered additional support for the mediation model: after practice retrieving the vocabulary definitions but not the

mediatiors, subjects continued to report mediatiors for the majority of vocabulary retrievals. In addition, practice retrieving the definitions of the Spanish words improved retrieval of the mediatiors when all three retrieval tasks were subsequently retested. The one result not predicted by the mediation model was that retrieval speed on the vocabulary retrieval task was faster after practice on the vocabulary retrieval task itself than after practice on the two subtasks. This was the only result consistent with the direct model. However, the direct model remains unsatisfying because it cannot explain how practicing the vocabulary retrieval task produces improvement in retrieval of the mediators on the two subtasks.

The results of these two experiments have important implications for research on the keyword method. They provide the first detailed evidence for the retrieval of the keyword in retrieving the English definitions. Previous researchers have asked subjects to report how they used keywords but have provided no independent evidence to validate these reports. The current studies not only provide considerable evidence for the role of keywords in the retrieval process; they provide detailed evidence for the time course of the process. Previous keyword researchers have suggested the need for more direct process measures in studying the keyword method (e.g. Pressley, Levin, & Delaney, 1982). The current studies, for the first time, provide such measures. The use of subtask, latency and verbal report measures to monitor various substeps in the retrieval process should prove extremely useful to future research on the keyword method. For example, using the methodology we have developed here, one could monitor whether subjects use the provided keyword, an alternative mediator (e.g. another keyword), or no keyword at all.

In addition to the implications for research on the keyword method itself, the current results have important implications for research on mediation in general. As outlined earlier, proponents of both mediation and direct retrieval accounts have argued their cases based soley on recall performance or unvalidated verbal reports. As mediation is a covert process, it is difficult to resolve this debate without evidence for intermediate cognitive structures or processing steps. The current results considerably strengthen the mediation hypothesis by providing new evidence for the intermediate steps in the retrieval process.

Previously, proponents of the mediation hypothesis have argued that if recall of the target material is contingent on whether a mediator is recalled or not, this constitutes evidence for mediation. However, proponents of the direct hypothesis have simply argued that recall of mediator and target are correlated by the stimulus to which each is related. By decomposing the vocabulary retrieval task into two different subtasks, one sharing the same stimulus as the vocabulary retrieval task and one not, we have been able to provide stronger evidence for the mediation interpretation. Whereas one might argue that recall of the English equivalent and the keyword are correlated by their relation to the same stimulus (i.e. the Spanish word), this argument will not work for the second subtask. Performance on the vocabulary retrieval task was not only limited by whether subjects could recall the keyword, but by whether given the keyword, they could recall the English definition. In fact, as the keyword subtask was almost always correctly performed, it was failure on the second subtask, the English retrieval subtask, that limited performance on the vocabulary retrieval task. In short, a simple correlational account of mediation is more difficult to maintain given the results we have reported here.

Previous studies of mediation have often employed verbal report measures to obtain evidence for mediation processes, but in such a way that the validity of the reports is questionable. For example, reports have been obtained after the completion of an experiment, with subjects asked to report on their thoughts over many previous trials. As we have pointed out, such procedures are highly undesirable because after many trials with similar stimuli, subjects cannot be expected to remember the details of any given trial, and may therefore generalize and/or infer what their thoughts were instead of simply reporting them. In the current experiments, retrospective verbal reports were gathered immediately after a trial, increasing the probability that whatever information the subject was aware of during retrieval was still available for report. In addition, we were able to relate the content of these verbal reports to our other performance measures: for example, the increasingly longer latencies for vocabulary retrieval trials in which subjects reported some type of mediation.

Few studies of mediation have employed latency measures to constrain the possible interpretations of their other performance measures.

The importance of such measures is crucial in making inferences about intermediate processes. For example, we noted that performance on the English retrieval task is strongly related to performance on the vocabulary retrieval task, though the two tasks share no stimulus. One might argue that in performing the English retrieval task, what subjects actually do is use the keyword to retrieve the Spanish word and then retrieve the English equivalent directly from the Spanish word. In fact, this interpretation has been suggested as a possible account for mediation in other paired associate studies (see Bellezza & Poplawsky, 1974). However, this account can be easily ruled out here by the latency data. If indeed, subjects used such a backward retrieval strategy for the English retrieval task, the latencies for this task would have been longer than those for the vocabulary task itself, which was not the case.

In conclusion, the experiments reported here provided considerable new support for the mediation model. Importantly, this support was not restricted to a single dependent measure. Rather a set of convergent measures related a consistent story. In addition, a number of these measures provided evidence for the intermediate cognitive structures and proceses involved in retrieving Spanish-English vocabulary. The one result inconsistent with the mediation account was the differential increase in retrieval time for the vocabulary retrieval task depending on the type of practice preceding it. Practicing either the subtasks or the vocabulary task itself increased retrieval speed on the vocabulary retrieval task. However, practicing the vocabulary task itself improved performance more than practicing the subtasks alone. This result is inconsistent with a simple mediation model that assumes that practicing the two subtasks is completely equivalent to practicing the vocabulary retrieval task itself. One possibility is that what begins as mediated retrieval is somehow modified with practice: for example, by establishing a direct link between the Spanish word and English word. Alternatively, the separate substeps in the mediated retrieval process might be composed into a single production for retrieving the definition from the Spanish word. In future experiments, we plan to investigate the effect of extended practice on the retrieval of the English equivalent, as well as possible mechanisms that might underlie changes in the nature of the retrieval process.

BIBLIOGRAPHY

Adams, J. A., & McIntyre, J. S. (1967). Natural language mediation and all-or-none learning. Canadian Journal of Psychology, 21, 436-449.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a russian vocabulary. Journal of Experimental Psychology: Human Learning, and Memory, 104, 126-133.

Bellezza, F. S. (1981). Mnemonic devices: classification, characteristics, and criteria. Review of Educational Research, 51, 247-275.

Bellezza, F. S. (1986). Mental cues and verbal reports in learning. In G. H. Bower (Ed.), The psychology of learning and motivation: Vol. 20 (pp. 237-273). New York: Academic Press.

Bellezza, F. S. (1987). Mnemonic devices and memory schemas. In M. A. McDaniel & M. Pressley (Eds.), Imagery and related mnemonic processes: Theories, individual differences, and applications (pp. 34-55). New York: Springer-Verlag.

Bellezza, F. S. & Poplawsky, A. J. (1974). The function of one-word mediators in the recall of word pairs. Memory & Cognition, 2, 447-452.

Bellezza, F. S., Poplawsky, A. J., & Aronovsky, L. A. (1977). The functional role of one-word mediators. Memory & Cognition, 10, 460-462.

Clark, J. M., & Pavio, A. (1987). A dual coding perspective on encoding processes. In M. A. McDaniel & M. Pressley (Eds.), Imagery and related mnemonic processes: Theories, individual differences, and applications (pp. 5-33). New York: Springer-Verlag.

Delaney, H. D. (1978). Interaction of individual differences with visual and verbal elaboration instructions. Journal of Educational Psychology, 70, 300-318.

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.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Hall, J. W., Wilson, K. P., & Patterson, R. J. (1981). Mnemotechnics: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. Journal of Educational Psychology, 73, 345-357.

Jones, B. F., & Hall, J. W. (1982). School applications of the mnemonic keyword method as a study strategy by eighth graders. Journal of Educational Psychology, 74, 230-237.

Klatzky, R. L. (1975). Human memory: structures and processes. San Francisco: W. H. Freeman.

Levin, J. R. (1981). The mnemonic '80s: Keywords in the classroom. Educational Psychologist, 16, 65-82.

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. (1985). Mnemonic Vocabulary Instruction: What's fact, what's fiction. In R. F. Dillon (Ed.) Individual differences in cognition : Vol. 2. (pp. 145-172). Orlando, FL: Academic Press.

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.

Levin, J. R., Shriberg, L. K., Miller, C. E., McCormick, C. B., & Levin, B. B. (1980). The keyword method in the classroom: How to remember the states and their capitols. Elementary School Journal, 80, 185-191.

Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. Journal of Educational Statistics, 5, 269-287.

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

Montague, W. E. (1972). Elaborative strategies in verbal learning and memory. In G. H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory. Vol. VI. New York: Academic Press, 1972.

Montague, W. E., Adams, J. A., & Kiess, H. O. (1966). Forgetting and natural language mediation. Journal of Experimental Psychology, 72, 829-833.

Myers, J. L. (1979). Fundamentals of experimental design (3rd ed.). Boston: Allyn & Bacon, Inc.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ott, E. C., 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, 23, 197-206.

Paivio, A., & Desrochers, A. (1979). Effects of an imagery mnemonic on second language recall and comprehension. Canadian Journal of Psychology, 33, 17-28.

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

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., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. Review of Educational Research, 52, 61-91.

Pressley, M., Levin, J. R., Digdon, N., Bryant, S. L., & Ray, K. (1983). Does method of item presentation affect keyword method effectiveness? Journal of Educational Psychology, 75, 686-691.

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. Journal of Experimental Psychology, 6, 163-173.

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., Nakamura, G. B., Hope, D. J., Bispo, J. G., & Toye, A. R. (1980). The keyword method of foreign vocabulary learning: An investigation of its generalizability. Journal of Applied Psychology, 65, 635-642.

Prytulak, L. S. (1971). Natural language mediation. Cognitive Psychology, 2, pp 1-56.

Ratcliff, R. A. (1978). A theory of memory retrieval. Psychological Review, 85, 59-108.

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

Raugh, M. R., Schupbach, R. D., & Atkinson, R. C. (1977). Teaching a large russian vocabulary by the mnemonic keyword method. Instructional Science, 6, 199-221.

Solso, R. L. (1988). Cognitive psychology (2nd ed.). Boston: Allyn and Bacon, Inc.

Sweeney, C. A. & Bellezza, F. S. (1982). Use of the keyword mnemonic in learning English vocabulary. Human Learning, 1, 155-163.

Turnure, J. E., & Lane, J. F. (1987). Special educational applications of mnemonics. In M. A. McDaniel & M. Pressley (Eds.), Imagery and related mnemonic processes: Theories, individual differences, and applications (pp. 329-357). New York: Springer-Verlag.

Underwood, B. J. (1972). Are we overloading memory? In A. W. Melton & E. Martin (Eds.), Coding processes in human memory (pp. 1-23). Washington, DC: V. H. Winston & Sons.

APPENDIX A

STIMULI FOR EXPERIMENT 1

Spanish word	Keyword	English equivalent
reina	rain	queen
pantano	pants	swamp
tenedor	ten	fork
borrasca	board	storm
bandera	band	flag
cielo	ceiling	sky
balde	ball	bucket
pollo	pole	chicken
crepusculo	crepes	twilight
candada	candle	padlock
maton	mat	bully
fingido	finger	dummy
panuelo	pan	handkerchief
saco	sack	jacket
manchar	man	stain
campana	camp	bell
garra	garage	claw
bombero	bomb	fireman
cocina	coke	cuisine
soplete	soap	blowtorch
forastero	forest	outsider
tarima	tar	platform
bandeja	banjo	tray
tacon	tack	heel
barca	bark	boat
tierra	tea	dirt
codo	code	elbow
pavo	paw	turkey
charco	charcoal	puddle
tonel	tunnel	keg
doronico	door	leopard
pinzas	pizza	tweezers
cabra	cab	goat
sonrisa	sunrise	smile
perro	pear	dog
pena	pen	rock
horno	horn	oven
juguete	jug	toy
beso	baseball	kiss
flecha	flesh	arrow
babucha	baby	slipper
carta	cart	letter

APPENDIX B

STIMULI FOR EXPERIMENT 2

Spanish word	Keyword	English equivalent
reina	rain	queen
pantano	pants	swamp
tenedor	ten	fork
borrasca	board	storm
muneca	money	doll
cielo	ceiling	sky
balde	ball	bucket
pollo	pole	chicken
crepusculo	crepes	twilight
candada	candle	padlock
maton	mat	bully
fingido	finger	dummy
panuelo	pan	handkerchief
saco	sack	jacket
manchar	man	stain
campana	camp	bell
garra	garage	claw
bombero	bomb	fireman
cocina	coke	cuisine
soplete	soap	blowtorch
forastero	forest	outsider
tarima	tar	platform
bandeja	banjo	tray
tacon	tack	heel
barca	bark	boat
tierra	tea	dirt
codo	code	elbow
pavo	paw	turkey
charco	charcoal	puddle
tonel	tunnel	keg
doronico	door	leopard
pinzas	pizza	tweezers
cabra	cab	goat
sonrisa	sun	smile
perro	pear	dog
pena	pen	rock
horno	horn	oven
juguete	jug	toy
bolsa	bowl	purse
flecha	flesh	arrow
babucha	baby	slipper
carta	cart	letter
ropa	rope	clothing
dados	dad	dice
butaca	boot	armchair

APPENDIX C

ENCODING INSTRUCTIONS FOR VERBAL PROTOCOLS

General

1. 5 variables to encode the protocols:

Spanish-word	link	keyword	link	English-word
p1	p2	p3	p4	p5

2. For the time being, the link between the Spanish word and the keyword is coded as a '9' in all cases

3. The 3 primary words (Sp, Kw, Eng) are encoded as a '1' if specifically mentioned. As '0' if not mentioned. As a '9' if there isn't enough info (e.g. 'I went back to the Spanish word', the Spanish word would be encoded as a '9' because the person doesn't explicitly mention the Spanish word or some version of it)

4. For p3 and p5 (kw and Eng) if the person refers to a kw or Eng word that is not correct but is from the list, code a '7.'

5. The cue for a trial (e.g. the kw for a kw-eng trial) is automatically encoded as a '1.'

6. FIRST ENCODE those elements clearly present and then concentrate on the difficult elements. For example, for a Sp-Eng trial:

a. Put a '1' for the cue	(i.e. 1 _ _ _ _)
b. Mentioned kw? Yes.	(i.e. 1 _ 1 _ _)
c. Mentioned Eng? Yes.	(i.e. 1 _ 1 _ 1)
d. '9' for Sp-kw link	(i.e. 1 9 1 _ 1)
e. Now code the kw-Eng link (use the categories to justify your choice)	(e.g. 1 9 1 0 1)

- Keyword to English Link: No Mediation Mentioned -

No Mention of link = '0' The following are rules & examples;

1. Thought or remembered kw and then Sp:

examples: 'Thought paw then turkey.'
'Thought paw and then turkey.'

2. Just remembered x; Just thought of x; or x just popped into my head; I just remembered it; or I just thought of the answer; x triggered y; I just knew it was y

examples: 'Just thought of paw then turkey.'
'Turkey just popped into my head.'
'Just thought of paw then turkey.'
'Paw triggered turkey'
'Saw paw and just knew it was turkey'

3. Use of 'to' or 'then':

examples: 'Paw to turkey.'
'Just paw then turkey.'

note: If some kind of ambiguous phrase is used with 'to' count as a '9' (e.g. 'I thought of paw which brought me to turkey.')

THE ABOVE ARE ALL INTENDED TO CAPTURE WHEN SUBJECT REPORTS GOING DIRECTLY FROM KW TO ENG. If the subject reports anything that suggests he/she thought of something else in between (even if it is irrelevant), DON'T CODE AS A '0'.

4. If either the keyword or the English word are not mentioned, then no link is coded between the keyword and English word (i.e. code as a '0').

5. If another kw or Eng word are mentioned, but the kw and Eng word for the current trial are not mentioned code link as a '0'.

- Keyword to English Link: Link Mentioned -

Code as a '1' or a '3' if there is a specific reference to an image or picture, or the use of a visual word in relating the kw and Eng.

1. To code as a '1' both kw and Eng must be mentioned and they must be related in some specific way. See category 9 below for definition of vagueness.

2. To code a '3' the image must be of the kw and eng word or the visual word must relate the kw and eng word somehow.

e.g. I pictured the paw on the turkey

Note the following are vague and would be '9's because the relationship between kw and Eng is not specified:

I remembered my picture of the paw and the turkey.
I just remembered my visual image.
I saw the word paw and turkey.
I pictured a paw and a turkey.

3. As long as kw and Eng word are mentioned along with specific link, code a '1' or a '3'. Sometimes Ss may use a vague phrase and later specify the link precisely.

4. If subject mentions the kw and a mediating link, then tries to establish what English word fits the link, uses different words and finally says the correct English word, this will be mediation---i.e. must be clear that kw is being specifically related to the Eng word but they do NOT have to appear in same proposition, if the relation established with Kw and link can clearly be used with the English word. Example:

I thought of 'jug' and i/
was trying to think what was in the jug/
I was trying to match them up/
At first I thought it was/
a jacket/
but then/
I remembered it was a toy.

Similarily, if subject inititates an image with a mediating relation and later find the word that fits the relation, it will constitute an example of imagistic mediation. Example:

I remember now picturing the rain falling down/
remember/
that the queen was in the courtyard/
and the rain was falling on the queen/

- Keyword to English link: Not Enough Info ('9') -

Perhaps mediated, perhaps not. Kw and English word ARE mentioned, but the association or mediator between them is vague.

Examples: Phrases such as the following used with the kw and Sp:

1. 'associated with' or 'the association with' (e.g. 'I thought of paw and then the association with turkey').

2. 'correlated with,' 'the correlation of' or 'the correlation between' x and y. (e.g. 'I remembered my correlation between paw and turkey').

3. 'went with,' 'goes with,' 'got me to', 'brought or brings me to', 'brought up or brings up', 'got me to', 'made me think of', 'means,' 'corresponds to,' 'put it together,' 'is,'

e.g. Pizza got me to tweezers
Paw brought up turkey
Door made me think of leopard
The keyword door was correlated with leopard
I just associated paw with turkey
Pavo just went with turkey
Turkey just corresponds to paw
Paw is Turkey
Paw means turkey
Just put together paw and turkey

4. 'And': The use of 'and' without 'then' is ambiguous and should be coded as a '9':

e.g. Remembered paw and turkey.
Just thought paw and turkey

BUT NOTE!! : Thought paw and then turkey = '0'

5. REMEMBER!! BOTH kw and Eng MUST BE MENTIONED to CONSIDIDER a LINK AS A '9.' IF EITHER THE Kw or Eng is NOT MENTIONED, you should code the link as not mentioned (i.e. '0').

e.g. I just remembered my association of a paw that
was attached to the turkey.

ADDITIONAL NOTES:

NOTE: If there is a conflict between '0' and '9', code a '9'

NOTE: Singular/plural forms okay for keyword or English word

NOTE: For the kw, if they use a word that contains part of the keyword (e.g. 'basketball' for 'ball') count it as correct keyword. But no synonyms. 'Bag' for 'sack' is wrong.

NOTE: Synonym for English word would be a '6' (e.g. 'coat' for 'jacket')

NOTE: '7' = a kw or eng from another part of the list (however, if you code a '7' for either the kw or eng word, the link must be coded as a '0' as both the correct kw and eng are not present).

Technical Report 90-10


University of Colorado at Boulder

Institute of Cognitive Science

Muenzinger Psychology Building
Campus Box 344
Boulder, Colorado 80309-0344
(303) 492-5063
FAX: (303) 492-7177
email: cogsci@psych.colorado.edu
URL: http://psychwww.colorado.edu/ics/