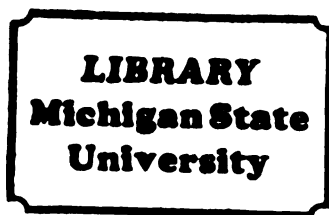24490023

THESIS

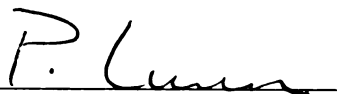This is to certify that the

dissertation entitled

FOREIGN LANGUAGE VOCABULARY ATTRITION:
A STUDY OF COLLEGE-LEVEL SPANISH STUDENTS
OVER AN 18-MONTH PERIOD

presented by

Marilyn Ruth Bierling

has been accepted towards fulfillment
of the requirements for

___Ph. D.___ degree in ___Spanish___

_____
Major professor

Date_31 May '90_

0-12771

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
JUL 0 6 2012		
DEC 2 8 1994 (出 6月7日)		

MSU Is An Affirmative Action/Equal Opportunity Institution

c:\circ\datedue.pm3-p.1

FOREIGN LANGUAGE VOCABULARY ATTRITION:

A STUDY OF COLLEGE-LEVEL SPANISH STUDENTS

OVER AN 18-MONTH PERIOD

By

Marilyn Ruth Bierling

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Romance and Classical Languages

1990

ABSTRACT

FOREIGN LANGUAGE VOCABULARY ATTRITION:
A STUDY OF COLLEGE-LEVEL SPANISH STUDENTS
OVER AN 18-MONTH PERIOD

By

Marilyn Ruth Bierling

This dissertation investigates the attrition of foreign language vocabulary in fifteen students of Spanish over an 18-month period. The study examines characteristics of the students and their learning situation, characteristics of the words and the way they were taught and tested, and the relationship between recognition and production vocabulary.

The 'Natural Approach' was used to teach the first semester of college-level Spanish to these English-speaking students. One group met in a classroom in the United States, and the other group met in Mexico. The first 20 hours of instruction were studied intensively in order to track the students' acquisition of 100 specially-targeted words in Spanish. During the following 18 months the students were retested on these 100 words for both recognition and production in order to study the pattern of attrition. The data were analyzed statistically, and individual cases were studied in depth.

The major findings were the following: 1) there were no differences in retention scores based on motivation, effort, sex, frequency of the words in teacher classroom input, or location of the course in the U.S. or Mexico (when

the effects of post-course contact with the 100 target words were eliminated), 2) there were significant differences in retention scores between words taught and tested for both recognition and production during the course, and words taught and tested for recognition only, 3) there were significant differences in retention scores for words with higher and lower saliency ratings, 4) adjectives tended to be forgotten more quickly than verbs or nouns, 5) retention scores dropped quickly at first and then gradually leveled off, and 6) the best description of recognition and production vocabulary loss was that equal quantities of words were lost over time, rather than that the ratio of production and recognition vocabulary remained constant.

The study concludes that requiring students to study for production of vocabulary is an important factor in long-term retention. Other important factors are the saliency of the vocabulary in teacher input, and the mnemonic devices created and structural analyses performed consciously by the student. Such results support a cognitive theory for foreign language acquisition.

DEDICATED TO MY HUSBAND NEAL,
who through many long hours provided
encouragement and support

ACKNOWLEDGMENTS

This dissertation would not have been possible without the encouragement and assistance of many different people who gave of their time to listen to my ideas and to give their advice. I would especially like to thank Dr. Pat Lunn, the director of my dissertation committee. Pat helped me to shape the idea for my research and spent many long hours correcting the drafts and offering her suggestions. I would also like to thank the other members of my committee, Dr. Natalie Lefkowitz, Dr. Ann White, Dr. Grover Hudson, Dr. Susan Gass, and Dr. George Mansour for the help they gave. Dr. Jim Stapleton and William Noble, also of Michigan State University, gave many helpful suggestions regarding the statistical analysis.

Several colleagues at Calvin College provided valuable assistance. I especially want to thank Dr. Edna Greenway for her wise advice, Dr. Al Reynolds for his statistical expertise, and Diane Lucar for the encouragement she gave me in the office we shared through much of the year. Nick Hengeveld and Jeremy Frens of the Calvin Computer Center helped in the preparation of the graphs for the manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The decade of the 1980s witnessed an explosion in
research on the processes by which adults acquire a second
language. Fueled by new theories and re-worked older ones
and a bewildering array of methodologies and movements, this
research has attempted to sort out the probable from the
improbable and the workable from the unworkable. Closely
related to the field of language acquisition research is the
field of language attrition research, an area of study that
was still declared to be 'antenatal' in the early 1980s
(Berko-Gleason, 1982b:22). A conference on the attrition of
language skills, held at the University of Pennsylvania in
1980, called for extensive research into this new field, a
call which was heeded during the decade. A survey of the
literature in Chapter 2 will demonstrate that most research
in the area of second language attrition is very recent.
And, since any kind of longitudinal study in the area of
language attrition must obviously stretch over a significant
period of time, it is only quite recently that results of
such studies are being published.

Language attrition can refer to at least four distinct
types of loss: 1) loss due to accident, disease, or aging of
an individual, 2) loss of a dying dialect or language by a
whole group of people, 3) loss of a first language in a

second-language environment, and 4) loss of a foreign language in a first-language environment. These different types of attrition have far-reaching ramifications for areas as diverse as psychology, gerontology, dialectology, foreign language education, and U.S. government policy. Yet until recently, categories 2-4 above were relatively unexplored. In the area of foreign language, specifically, attention has been focused on the processes involved in acquiring the foreign language, rather than on how quickly (or slowly) it is forgotten. Yet research in attrition is inevitably tied to research in acquisition and has a good deal to say about acquisition theory and about the ultimate goal of all our well-intentioned efforts as foreign language teachers.

The goal of this study is to focus on one small area of foreign language attrition--the loss of vocabulary that had at one time been either recognized or produced by the foreign language student. In the present chapter I will give a brief overview of second language acquisition theory and methodology in the 1980s and will show the relevance of the present study to this field. I will also give a brief description of the methodology used in this study, state the hypotheses to be investigated, and define terms. Chapter 2 contains an in-depth review of literature in the field of language attrition specifically, as well as recent literature on vocabulary acquisition and human memory. Chapter 3 will explain in detail the methodology used in this research, and Chapter 4 will give an analysis and

evaluation of the data. Finally, in Chapter 5, I will discuss the findings and point out some implications for second language acquisition theory and methodology in the foreign language classroom.

OVERVIEW OF LANGUAGE ACQUISITION THEORY
AND METHODOLOGY IN THE 1980s

One of the most influential models for second language acquisition during the 1980s was Stephen Krashen's Monitor Model. This model has provoked research and produced teaching methodology such as the 'Natural Approach' (Krashen 1982 and 1985, Krashen and Terrell 1983), as well as much criticism. Briefly stated, the Model consists of five hypotheses: 1) the acquisition/learning hypothesis, 2) the natural order hypothesis, 3) the monitor hypothesis, 4) the input hypothesis, and 5) the affective filter hypothesis. Terrell has used these five hypotheses as the theoretical basis for his popular first-year Spanish textbook Dos mundos (1986c), the textbook used in the classroom during the present study. In Dos mundos, vocabulary is taught through a great deal of comprehensible input in the form of teacher talk, utilization of pictures and real objects, and, in the beginning lessons, through Total Physical Response. Because of the emphasis on the importance of comprehensible input, comprehension receives priority over production, at least in the beginning stages of acquisition. Students are encouraged to bind the form of the Spanish word to the

meaning, with the goal being automatic retrieval of the meaning of a word. Later on, students begin to use the new words in their emerging speech, hopefully without reliance on English as an intermediary.

Krashen's Monitor Model has not gone unchallenged. Although all five hypotheses have been criticized, the most controversial seem to be the acquisition/learning distinction and the input hypothesis.

The controversy over the acquisition/learning distinction centers around the emphasis on meaning versus form and subconscious versus conscious learning, and whether learning can turn into acquisition. Sharwood Smith (1981, 1986) has pointed out the value of consciousness raising in the acquisition process. Both VanPatten (1984) and Terrell (1986a,b) minimize the subconscious/conscious dichotomy by saying that binding form to meaning can occur on either a conscious or subconscious level. Some conscious learning activities may be helpful to make certain features more salient and to aid in their comprehension and access. Terrell (1989) states that grammar instruction may be useful as an 'advanced organizer' (giving the learners information that will help them make sense of the input) and will also provide multiple occurrences of a form for acquisition and for practice in monitoring. Byrnes (1984) argues that there is ample evidence that formally learned knowledge can become part of the subconscious automatic processing capability. Gregg (1984) attacks Krashen's terminology, saying that the

acquisition/learning and subconscious/conscious distinctions are ill defined and trivial. Higgs (1985b) asserts that as meaning in the target language becomes obscure, more conscious information becomes important for the student. Long (1983) calls for Krashen to reevaluate the importance of 'learning', and, indeed, in his 1985 book, Krashen admits that acquisition can be aided by learning. McLaughlin (1979, 1987) also criticizes Krashen's acquisition/learning hypothesis and asserts that controlled processing can become automatic. McLaughlin's cognitive theory will be discussed further in Chapter 2 in the section on memory.

The term 'consciousness' itself is ambiguous. In a perceptive paper to be published in **Applied Linguistics**, Richard Schmidt points out that 'consciousness' can be defined as awareness, as intention, or as knowledge. Schmidt concludes that 'subliminal language learning is impossible, and that noticing is the necessary and sufficient condition for converting input to intake' (page 1). He believes that incidental learning is possible, although adults are helped by paying attention and by intending to learn. The issue of implicit learning is difficult to resolve, because there is evidence for implicit learning, as well as for the facilitative effect of conscious knowledge and understanding.

The input hypothesis states that "humans acquire language in only one way -- by understanding messages, or by

receiving 'comprehensible input'"(Krashen 1985:2). Speech is the result of acquisition and not its cause. However, others emphasize the need for output and negotiation of meaning as well. Swain (1985, 1989) points out that after several years of comprehensible input in immersion programs in Canada, children still produce speech with many errors. Input is essential, but not sufficient. When the student is pushed to produce output, he is also forced to produce a more sophisticated analysis of the language, testing the hypotheses that he has formed. What is needed in the immersion programs is greater opportunity for the students to interact with each other in the foreign language, thus producing 'comprehensible output'. VanPatten (1987, 1989) also believes that negotiation and interaction may be crucial. He calls for more sophisticated accounts of the role that consciousness plays in the processing of input.

Other writers have pointed out the difference between input and intake. The term intake was introduced by Corder (1967), who used it to describe 'what goes in and not what is available to go in' (165). According to Hatch (1983), input is what the learner hears and tries to process, and intake is what is successfully processed. Chaudron (1985a,b) has given ideas for testing the way that a learner processes input and turns it into intake--which may be encoded into short-term or long-term memory. Gass (1988) describes a five-step process: a) apperceived input: input which is noticed by the learner, b) comprehended input: that

which is understood by the learner, ranging from semantic comprehension to full structural analysis, c) intake: mental activity mediating between comprehended input and the learner's grammar; process of assimilating linguistic information, d) integration: storage of the information possibly leading to a reanalysis of the learner's grammatical system, and e) output: that which a learner produces.

Also related to this general area of input is the question of the nature of input that learners receive. Lightbown (1983, 1987) has extensively explored the relationship between input in the classroom and the output of students. Gass and Madden have edited a book entitled Input in Second Language Acquisition (1985) containing several articles which 'have a direct link to empirical research' (Scarcella and Perkins 1988:348).

Another large area of research and controversy related to the input issue is the relationship of comprehension (both listening and reading comprehension) to production (both speaking and writing). The skills involving comprehension are also referred to as recognition or the receptive skills, and production is also known as recall. Although there is evidence that production in some cases may precede comprehension (Sheldon and Strange 1982), many writers such as Postovsky (1974), Nord (1980), Winitz (1981), Asher (1981), and Krashen and Terrell (1983) insist on the priority of comprehension in foreign language

instruction. They say that students will produce when they
have received considerable input and are ready to do so.
Ostyn and Godim (1985) believe that a minimum vocabulary in
a foreign language is 5000 words, a goal which can be met in
two years of language study at the college level if more
emphasis is placed on recognition and less on production.
Davies (1976) calls for well-developed receptive skills--a
logical goal for two years of language study in Swedish
classrooms; production skills should be reserved for the
more motivated students. The tests that accompany Terrell's
textbook Dos mundos focus on listening and reading
comprehension and vocabulary and grammar recognition through
a multiple choice format, though suggestions are also given
for speaking and writing tests.

Gass (1988:213) points out that 'there is a lot of
mileage that needs to be travelled between the input to the
learner...and what the learner produces. We cannot assume
that with mere presentation of language information, whether
implicitly or explicitly, students will convert it to
output'. Hatch (1983) describes how vocabulary may be
'penciled in', (for example, a certain word X exists, has
two syllables, and begins with ch), but there is no evidence
to support that 'penciled in' forms later become part of the
learner's system. Rivers (1986) points out that
comprehension and speaking are very different skills--in
comprehension small details may be ignored, but not in
speaking. In other words, ability to recognize words and

forms will not necessarily lead to the ability to produce them.

Krashen's Monitor Model is thus related to a great deal of the current research and controversy in the field of second language acquisition. When this model is criticized, as in Gregg (1984, 1986), Spolsky (1985), and McLaughlin (1979,1987), it is most frequently mentioned that Krashen lacks operational definitions and that his hypotheses are tautological or cannot be falsified.

This brief overview of second language acquisition theory, controversy, and methodology in the 1980s cannot be complete without mentioning the Oral Proficiency Movement. While an approach such as the 'Natural Approach' emphasizes the acquisition of a large recognition vocabulary, textbooks based on the Oral Proficiency Movement, such as Spinelli and Rosso-O'Laughlin's Encuentros (1988), emphasize production. In the Oral Proficiency Movement, vocabulary is still essential for the beginning speaker, as pointed out by Higgs (1982, 1984, 1985a,c). In the relative contribution model presented by Higgs, vocabulary and pronunciation are the most critical factors at the lowest level of proficiency, whereas control of grammer is essential to go from the advanced to the superior level. Since many students seem to 'fossilize' at the advanced (2 or 2+) level, however, it is important to teach grammatical accuracy from the beginning of instruction. Omaggio (1984) stresses concern for development of linguistic accuracy from the beginning. The

proficiency movement has been criticized by several writers, including Lantolf and Frawley (1985), Savignon and Bachman (1985), Schulz (1986), VanPatten (1987), and Kramsch (1986,1987), the chief criticism being that grammatical accuracy has been given too much weight. However, the ACTFL proficiency guidelines are becoming an influential base for organizing foreign language courses.

PURPOSE AND IMPORTANCE OF STUDY

While teaching during the 1987-88 school year with the Natural Approach, an approach that emphasizes large recognition vocabularies at the beginning of second language acquisition as noted above, I became interested in several questions that had direct impact on my teaching. What were the characteristics of the teacher input, of the learning environment, and of the learner that might lead to long-term vocabulary retention? What were the characteristics of the words themselves that led to intake by the student and long-term retention? Through which stages did the students pass in the acquisition and the loss of vocabulary? What is the relationship between amount of recognition and production vocabulary? Does this relationship change over time? Should large vocabularies be taught for recognition only--and be tested for recognition only--during the beginning stages of second language acquisition?

These questions are timely, since they ask about fundamental issues of methodology in current comprehension-based approaches, as well as issues in teaching and testing for recognition and/or production. A basic assumption that was made before beginning this study was that without continued contact with specific vocabulary items, foreign language students over time would lose some of their ability to access these items and that certain patterns in recognition and production would emerge. It is hoped that by studying the attrition of foreign language vocabulary, we can learn something more about human language--how we acquire it and how we lose it.

BRIEF DESCRIPTION OF STUDY

Research hypotheses

The research hypotheses were divided into three groups: 1) hypotheses about students, 2) hypotheses about words, and 3) hypotheses about the relationship of recognition vocabulary to production vocabulary.

Hypotheses about students:

1. Students who study in a target-language (Spanish-speaking) environment retain more vocabulary than those who study in a native-language (English-speaking) environment,

even when the specific vocabulary learned is not rehearsed or used outside of class.

2. Students who have pre-course contact with Spanish (either informally in the culture or formally in the classroom) will retain more vocabulary than those who have no previous contact.

3. Students who have post-course contact with Spanish (either informally in the culture or formally in the classroom) will retain more vocabulary than those who have no further contact, even when the specific vocabulary learned is not rehearsed or used.

4. Vocabulary is retained at different rates according to sex.

5. Students who report more effort while taking a language course retain more vocabulary than those who report less effort.

6. Students who report more intrinsic motivation for the course retain more vocabulary than those who report more extrinsic motivation.

7. Students who receive higher grades for the course retain more vocabulary than students who receive lower grades.

8. Students who receive higher scores on the first retention tests continue to receive higher scores on later retention tests.

Hypotheses about words:

9. Vocabulary that is repeated more frequently in teacher classroom input is retained at higher levels than vocabulary that is repeated less frequently.

10. Verbs, adjectives, and nouns are retained at different rates.

11. Vocabulary learned earlier in the course will be retained at higher rates than vocabulary learned later in the course.

12. Vocabulary with high emotionality or saliency ratings is retained longer than vocabulary with low ratings.

Hypotheses about the relationship of recognition to production:

13. Recognition scores for vocabulary are higher than production scores.

14. Both recognition and production scores for vocabulary drop quickly at first and then gradually level off.

15. The best description of recognition and production vocabulary loss is that equal quantities of words are lost over time rather than that the ratio of production to recognition vocabulary remains constant over time.

16. Smaller percentages of recognition vocabulary are lost than of production vocabulary.

17. Vocabulary taught and tested for both recognition and production is retained at higher rates than vocabulary taught and tested for recognition only.

18. Ability to produce a word initially, rather than ability to recognize a word, is the best predictor of both recognition and production later on.

Subjects

The subjects of this study were college-level students from two beginning Spanish classes that the researcher taught during the summer of 1988.

The first class was an intensive 3 1/2-week course taught at Calvin College in Grand Rapids, Michigan, from May 31 to June 22. The nine English-speaking students were in class for 3 hours each day and used audio tapes and completed homework assignments outside of class. This Spanish class was taught, of course, in an English-speaking environment.

The second class was an intensive 3 1/2-week course taught in Merida, Mexico, from June 26 to July 18. This class was part of a program for English-speaking college-level students who were interested in working with the Presbyterian Church in Mexico and who followed their 3 1/2-week course with field assignments in Central America and Mexico before returning to the U.S. and Canada. Six students participated in this class.

Both classes were taught entirely in Spanish with the 'Natural Approach', using the textbook Dos mundos by Tracy Terrell et al. Only the first 20 hours of instruction were included in the acquisition period for the present project, corresponding to the first five pasos 'steps' in the textbook. Students were followed for eighteen months after the end of their class.

Data collection during class sessions

Students at the beginning of each course were given a questionnaire asking about any former contact with Spanish and motivation and attitudes toward learning Spanish. Students were also aware that vocabulary research was being done (though not the details) because of the consent form they were given the option of signing. The teacher did not see these consent forms, nor the final evaluation of the course, until after the final grades were given.

Vocabulary was introduced in 'teacher talk' with the use of pictures and objects and through Total Physical Response. All new words were written on the board and copied in student notebooks. After the first three hours of instruction, students were encouraged to begin responding with one-word answers in Spanish, and by the end of the first 20 hours students were participating in widely-varying communicative activities. Short grammatical explanations had also begun, and students had begun writing short compositions in their workbooks.

Data collection during the class sessions occurred through normal classroom activities. The following behaviors in class and for tests were assessed:

- motor response to commands
- oral response to pictures
- picture naming and description (both oral and written)
- translation

-guided composition

-matching of words (oral and written) to objects and
 pictures

-written multiple choice items

By using these multiple measures the researcher was able to record which vocabulary items had entered a) the recognition stage and b) the production stage by the end of the first 20 hours.

The only obtrusive instrument used was a tape recorder for recording the input of the instructor during class, as well as oral responses during individual oral testing. The tape recorder was also used at the end of the course, when students were asked to do some introspection about their vocabulary learning strategies. Students were asked why they could recall certain words and not others, what methods they used for studying vocabulary, and how often they mentally translated into English. In the final interview eighteen months later, the tape recorder was also used for oral interviews.

In addition to the regular vocabulary of over 600 words introduced in Pasos A-E in the text, the researcher added 100 words which were not taught in Dos mundos nor in the second-year book at Calvin. Some of the words came from Spanish dialects other than those spoken in Mexico. These 100 words formed the core of the research in determining the attrition rate, since the students were likely to hear these words as input only during the first 20 hours of

instruction. When a student indicated that he or she had further contact with one of these target words, the word was dropped from the data for that particular student.

These 100 words belonged to categories of words already introduced in the lesson, such as family members, clothing, and parts of the body, and were introduced along with the other vocabulary. Examples of these target words are madrastra 'stepmother', cerilla 'match', and barbilla 'chin'. Since students studied all words from their vocabulary lists copied off the board, it was not readily apparent that these words were not included in the vocabulary lists at the end of each paso in the text. By the end of the first 20 hours, these words had been checked for both recognition and production in regular tests in order to form a baseline for later retention testing.

Data collection during retention period

At the end of each 3 1/2 week session, the students were given their first test to measure retention of these 100 words, which were mixed in with other words that had been taught. This test was in written format (see Appendix A) and used pictures that the students had seen before in class. Students were also shown scenes containing several of these 100 vocabulary items during individual oral interviews and were asked to describe these scenes into the tape recorder.

The students in the summer 101 class at Calvin continued immediately with Spanish 102 and then had a break between July 21 and September 12, at which time they continued with second-year Spanish. These students retook the written picture test at the end of 102 and again in April of 1989.

The students in Merida left immediately for field assignments after finishing 101--and went to locations in Central America and Mexico where they needed to use Spanish. They met again in Mexico City on August 8 and then returned to the U.S. and Canada on August 13. During their stay in Mexico City they retook the written picture test, and they repeated the test in May of 1989. These students were asked to indicate which words they had come into contact with in their experiences. Of the 100 target words, only a few were mentioned, and these words were dropped from the data for the student that mentioned them.

All fifteen students were asked to retake the written picture test in November and December of 1989. Oral interviews were also performed with the twelve students who were still in the Michigan area. The entire testing schedule for the two groups is illustrated by Table 1 below:

TABLE 1
TESTING SCHEDULE

CLASS	TIME 0	TIME 1	TIME 2	TIME 3	TIME 4
#1 Grand Ra-pids, MI	during first 20 hours of class	June 22, 1988	July 1988 (at end of Spanish 102)	April 1989	Nov.-Dec. 1989
#2 Merida, Mexico	during first 20 hours of class	July 18, 1988	August 1988 (in Mexico City)	May 1989	Dec. 1989

TIME 0: Baseline testing
TIME 1: at end of Spanish 101
TIME 2: 3-1/2 weeks after
      previous testing

TIME 3: 9 months after
      previous testing
TIME 4: 7 months after
      previous testing

Variables

The dependent variables in this research are the scores on vocabulary recognition and production tests at five different points in time: the baseline tests and retention tests at four different times. Thus there was a set of ten scores per student and also per word, depending on how the data were organized.

Several independent variables were chosen for study because of their possible effects on recognition and production scores. One important variable was the learning environment, whether an English-speaking environment or the target-language environment. Other independent variables related to the learners were their contact with Spanish both before and after the course, their sex, their effort and

motivation, and the grades they received in the course. Variables relating to the vocabulary words themselves were their frequency in classroom input, their part of speech, the order in which they were taught, their emotionality and saliency ratings, and whether they were originally taught for recognition and production or for recognition only.

Potential variables that were controlled in the research were the method ('Natural Approach' used), the level of the instruction (first semester of college-level Spanish), the intensity and length of the courses (3-1/2 weeks), and the instructor (same instructor for both courses).

Analysis of data

The data were analyzed using the SPSS (Statistical Package for the Social Sciences) program at Calvin College. Correlations were performed, as well as several tests to test for significant differences between means (T-test, One-way analysis of variance, and Anova). The level of significance chosen was .05. A multiple analysis of variance was performed at Michigan State University in order to study the relationship between initial recognition and production as predictors of later scores.

Information not analyzed statistically but important for the final discussion included written work produced by the students during the first 20 hours of instruction, final

oral interviews at the end of the course and after the final
vocabulary test, and introspective comments by students on
how they learned and remembered vocabulary items. To
illustrate relationships between recognition and production
vocabulary over time, several graphs have been included in
the text.

Limitations of research

Because of the desire to conduct the research under
natural conditions in actual classrooms, this study has four
major limitations:

1. small sample size--The research was conducted
during intensive summer courses in a small college and a
comparatively small program in Mexico, and thus only fifteen
students participated. This small sample size is partially
offset by the large amount of data resulting from ten
separate tests per student with 100 items per test.
However, care must be taken in applying the results of this
study to the general student population.

2. non-random groups--In real-world educational
settings, classes are not normally formed by random
assignment, but rather by student choice or scheduling
necessities. By choosing a number of independent variables,
the researcher could group the students in different ways,

but again care must be taken in applying the results to other educational situations.

3. post-course contact with vocabulary items--Students were asked to identify which of the vocabulary items on the tests they had read or heard in later contact with Spanish, and these items were eliminated from the data for that individual. However, it is possible that some students missed identifying some of these items.

4. repeated tests--Although the later tests were separated by several months and target vocabulary items were interspersed with different non-target items for each test, scores for some students may have been influenced by seeing the target items on an earlier test.

An additional limitation was caused by the requirement by Michigan State University that students sign consent forms for the research. Thus students knew that they were subjects of research on foreign language vocabulary, although they did not know the exact nature of the research. Such knowledge may have resulted in the 'Hawthorne Effect', in which students perform differently because they are aware that they are participants in an experiment.

DEFINITIONS IN THIS STUDY

Since these terms will be used frequently throughout the study, working definitions are given here in order to clarify their meaning.

Second language acquisition: Internalization of the structure and vocabulary of a second language in either a first language or second language environment, by either formal or informal means. In this definition 'acquisition' is not meant to contrast with 'learning', and 'second language' is not meant to contrast with 'foreign language'.

Second language vocabulary: Spanish words taught with pictures, objects, or actions in context in the classroom.

Vocabulary retention: The ability to access a word through either recognition or production.

Vocabulary attrition: The failure to access a word through either recognition or production.

Vocabulary recognition: The ability to match a vocabulary word to a picture, object, or action or to pick the correct word in a multiple choice item.

Vocabulary production: The ability to write or say a vocabulary word when shown a picture, object, or action (also known as recall).

Input: Teacher talk in the classroom and worksheets containing practice items in Spanish. The Natural Approach calls for methods (such as use of context or pictures) that insure that foreign language input is <u>comprehensible</u> to the students.

Intake: The process during which learners assimilate linguistic information noticed and comprehended in input into an incipient linguistic system.

Frequency in classroom input: Number of times that instructor repeats vocabulary word in input teacher talk.

Saliency of vocabulary word: The ability of a word to attract attention due to its conspicuous, striking, or unusual nature, as indicated by classroom learners on rating scale.

Extrinsic motivation: Motivation resulting from pressure outside the student.

Intrinsic motivation: Motivation resulting from a desire within the student with no external pressure.

Natural Approach: A teaching methodology based on the
following principles: 1) comprehension precedes production,
2) speech emerges in stages, 3) speech emergence is
characterized by grammatical errors, 4) group work
encourages speech, 5) students acquire language only in a
low-anxiety environment, 6) the goal is proficiency in
communication skills (Terrell et al. 1986:7-9 in Dos mundos
Instructor's Manual).

Total Physical Response: A teaching methodology in which
students respond physically to commands in the foreign
language (Asher 1981).

Regression Hypothesis: A hypothesis stating that language
attrition is the mirror image of language acquisition (that
which is last learned is first forgotten).

These definitions provide a useful working vocabulary
for this and following chapters.

CHAPTER 2

REVIEW OF LITERATURE

This chapter will review literature in three areas that are
relevant to this research: 1) language attrition, 2)
foreign language vocabulary, and 3) human memory. At the
end of each section the literature will be summarized, and
at the end of Chapter 2 each research hypothesis for the
present project will be scrutinized in light of what the
literature says about it and what gaps still exist in our
knowledge.

LITERATURE ON LANGUAGE ATTRITION

Different types of language attrition

Although the purpose of the present investigation is to
study the attrition of a foreign language, research in other
types of language attrition should also be mentioned. Such
research can alert us to patterns that may be significant in
foreign language attrition as well. So, before going on to
review literature in the area of foreign language attrition
specifically, I will briefly review recent research in the
following areas: 1) attrition of a first language through
accident, disease, or aging, 2) the attrition of dying
dialects or languages by entire groups of people, and 3)

attrition of a first language in a second language environment.

Aphasia, the loss of language due to some sort of trauma, has been thoroughly researched. "Studies have indicated...that the dissolution of language in aphasia does not mirror acquisition by children. Aphasic subjects may be reduced to a very limited lexicon, but it rarely consists of baby's first words, such as 'mommy' and 'juice', and it is more likely to consist of politeness routines and/or unprintable words" (Berko-Gleason 1982a). Obler (1982) has shown that naming is the most universally impaired skill, both in aphasia and in normal aging, with content words being harder to recall than function words and nouns harder than verbs. Obler and Albert (1984) found that naming ability increases through age 60, then gradually begins to decrease after 60, and steeply declines in the 70s. With disease such as Alzheimer's, naming difficulties occur even in the mild stages (Pan and Berko-Gleason 1986). The regression hypothesis (that which is last learned is first forgotten) has not been substantiated in research on loss of a first language through accident or disease (Caramazza and Zurif 1978:x).

Dorian (1973, 1978, 1981, 1982) is often cited in connection with research in the area of languages in danger of extinction, such as East Sutherland Gaelic. In reference to comments made by the less fluent speakers themselves, she notes that "explicit comment on the decline in the quality

of their Gaelic focuses almost entirely on the lexicon...: the younger speakers feel sure their elders had more 'words for things' than they have themselves" (1978:590-91). Her studies, as well as others on languages as diverse as Navajo, Frisian, and Breton, emphasize change in the language or dialect as a whole across generations, with close attention to sociolinguistic factors such as language prestige and official support for bilingualism.

Recently, research has increased on the attrition of a first language (L1) in a second language (L2) environment. Much of the data has come from migrant workers in Europe and from second and third-generation immigrants in the United States, for example Brewer-Bomar (1981) on the Spanish of two children in the U.S. and Gonzo and Saltarelli (1979, 1983) on emigrant languages. Gonzo and Saltarelli (1979) found that as Spanish-speaking immmigrants gradually lost their first language, there was increased lexical borrowing from English from one generation to the next. Leyen (1986) studied native language vocabulary decline and concluded that 1) extreme L1 decline occurred only among subjects who had left their L1 environment during early childhood, 2) language attrition was mostly in the area of lexicon, and 3) the production of lower frequency L1 lexicon became impaired over time. Lexicon among bilinguals appeared to be stored in a common master lexicon and impairment occurred in the access rather than in the master lexicon itself. In connection with first-language loss research, Jaspaert,

Kroon, and Van Hout (1986:41) make the interesting point
that the relation between language loss and time is probably
non-linear. Many growth processes show an S-like curve, and
this is probably also true in the case of language decline.

In summary, studies in these other areas of language
attrition point to the critical nature of the lexicon. In
the loss of a dialect by a whole group of people or in the
loss of a first language in a second language environment,
the lexicon seems to be the most sensitive to attrition. In
aphasia, diseases such as Alzheimer's, and in normal aging,
naming is also the most impaired skill, with nouns being
harder to recall than verbs. Thus the investigation of the
loss of vocabulary seems to be a potentially important area
of research in foreign language attrition as well.

Early studies on foreign language attrition

Before 1980 there were very few studies on the
retention of a foreign language learned in a first-language
environment. One of the earliest studies is that of
Anderson and Jordan (1928) on the retention of Latin words
and phrases. After a period of eight weeks, only one-half
of the meanings of Latin words was retained (students were
asked to write the English equivalent of the Latin word).
It was also noted that those who learned more, retained
more, although this result showed up only when comparing the
most extreme cases. A few early reports attempted to

measure the effect of summer vacation on the loss of language skills. Kennedy (1932) found that his students showed a 15-35% loss in their Latin syntax after the summer vacation, and he also suggested that there was a positive correlation between language retention and original attainment and desire to continue study of Latin. Geoghegan (1950) attempted to replicate and to extend this research. His study included Latin, French, and Spanish high school students, and he tested vocabulary as well as grammar. The results were mixed, with Latin students showing significant losses in both grammar and vocabulary and Spanish students showing significant losses in vocabulary only. French students even showed gains in vocabulary. Scherer (1957) reported on a series of experiments that he had performed over a period of six years. He had tested college students at the end of the first year of German and again at the beginning of the second year in grammar, reading, and vocabulary. Scores in vocabulary and grammar decreased insignificantly, while reading scores increased slightly. Scherer also concluded that a week of review at the beginning of the new school year was enough to bring students back to their previous level. Pratella (1969) studied the effect of summer vacation in all four skill areas--listening comprehension, speaking, reading, and writing. Significant losses were sustained only in the area of listening comprehension. Pratella also concluded that sex and verbal ability were not important factors in loss

over the summer vacation. In summary, these short-term
studies of loss over three months indicate that vocabulary
may be one of the areas most likely to show attrition.

Smythe, Jutras, Bramwell, and Gardner (1973) conducted
two Canadian studies on the retention of French reading and
listening skills by high school students in London, Ontario.
Over the summer, students showed a small but significant
loss in reading comprehension and a small but significant
gain in listening comprehension. The second study compared
the high school students returning to French class
immediately after the summer semester with those who waited
to begin study of French again until the second semester.
As expected, the group that waited for eight months to
resume study of French scored lower than the group that
waited only three months; yet the total loss of the lower-
scoring group was only 4%. Smythe (1973:405) concluded that
'it has been demonstrated that only relatively small changes
in second language performance were recorded over reasonably
long periods during which students were not receiving
instruction'. The three-month vacation could even have a
facilitating effect on language retention.

Cohen (1974, 1975) studied elementary school children
who had been acquiring Spanish in an immersion setting in
California. In his first project he tested fourteen
children on their speaking abilities before and after summer
vacation (between the first and second grades). After the
summer break the children made more errors overall, and

there were more problems with article and adjective
agreement, as well as errors in person inflection of verbs.
Utterances in general were shorter. The following summer
Cohen studied three children intensively to see if the last
things learned were the first to be forgotten. Two of the
three children provided data that at least partially
supported this hypothesis. For example, one child had
finally shown general command of the ser/estar distinction
just before the end of first grade, but at the end of the
summer recess she was overusing estar. Cohen also suggested
that the forgetting process could produce new incorrect
patterns, but also in some cases 'a pause in the learning
process may actually cause a reduction in certain problem
areas' due to the elimination of incorrect hypotheses about
the language (1975: 137).

Edwards (1976, 1977) studied French-dominant and
English-dominant bilinguals in the Canadian Public Service
who had completed courses of instruction six months, twelve
months, and eighteen months earlier. This study differed
from the ones above in that the subjects were adults and
continued to have contact with the second language after
finishing their course of study. After six months there
were no significant differences between language-dominant
groups, but after twelve months the English-dominant group
showed significant losses in speaking, though they gained in
reading skills. More decline in speaking occurred after
eighteen months. The French-dominant group showed no

decline in their English skills after six, twelve, and eighteen months. In questionnaires this latter group indictated more frequent use of the second language. Edwards concluded that retention was more related to use and a supportive environment than to motivation or attitude.

These studies before 1980 show that much more needed to be done in the area of foreign language attrition research. However, some interesting patterns did emerge: vocabulary usually showed some attrition even over short periods of time, while listening and reading comprehension results varied, and in some cases scores in these skills were even higher after a summer vacation. Cohen (1975) concludes that there is some evidence to support the regression hypothesis (at least in speaking), and Edwards (1976, 1977) believes that a supportive environment which encourages use of the second language after the termination of a course is more important to retention than motivation or attitude.

Research during the 1980s

In May of 1980 a conference on language attrition was held at the University of Pennsylvania, the proceedings of which were published in The loss of language skills (1982), edited by Lambert and Freed. This conference and book provided a major impetus for research in the area of language attrition. The subjects in the book range from the relation of psycholinguistics and sociolinguistics to

language loss (articles by Berko-Gleason, Gardner, Dorian, Obler) to guidelines for conducting research (Oxford). In an article entitled 'The U.S. Government's Foreign Language Attrition and Maintenance Experience', Lowe states that second language atttrition is a serious problem for the U.S. government. The C.I.A. Language School, on the basis of anecdotal information, feels that subjects start losing their language skills as soon as they stop practicing them. Candidates with low-level skills lose them more rapidly than those with higher-level skills. Also Lowe believes that 'Speaking is least stable, perhaps so unstable as to never be constant; Understanding occupies a middle ground, with some languages being harder to retain understanding in than others; and Reading is the most stable of the skill modalities' (1982:181).

In the conclusion to his article on 'Determining the Linguistic Attributes of Language Attrition' in the same book, R. Andersen states his views on vocabulary attrition: one important area of study should be the 'quick retrieval of appropriate vocabulary and idiomatic phrasing in on-going speech production. My own experience as an LA [language acquirer] and especially my observations of my children (as well as other individuals) convinces me that this area far outweighs morphosyntactic and phonological aspects of attrition' (1982:113).

Gardner in his article in this book gives two hypotheses related to second language retention: '1) Since

attitudinal/motivational characteristics are related to the
level of second language proficiency, they will relate to
second language retention, 2) Since attitudinal/motivational
characteristics are related to indices of participation in
language-related situations, they will relate to attempts to
maintain second language skills once training has
terminated' (1982:41). Gardner refers to this post-training
period as the 'incubation period'. In more recent research
on social factors (Gardner, Lalonde, and MacPherson 1985),
attitudinal/motivational factors seem to be more influential
during the acquisition period (since they influence the
level of competence acquired) than during the incubation
period. Attitudes toward the second language were not
directly related to use of the language after the completion
of training. These conclusions are similar to those of
Edwards (1976, 1977).

After the 1980 conference and the 1982 publication of
The loss of language skills, research in the area of second
language attrition drastically increased. In October of
1982 Theo Van Els convened a second conference on language
attrition at the Catholic University of Nijmegen in the
Netherlands, and in March of 1986 a Language Loss Symposium
was held in Kerkrade in the Netherlands. The papers from
this symposium have been published in a volume entitled
Language attrition in progress, edited by Weltens, De Bot
and Van Els. Also in 1986 Applied Psycholinguistics devoted
an entire issue to the theme of language attrition. One of

the most recent testimonies to the importance of this field is that in July of 1989 an entire issue of <u>Studies in Second Language Acquisition</u> was dedicated to discussion of language attrition. Several of the articles coming out of these publications are included as part of the discussion in this chapter.

Out of the 1980 conference grew the Language Skills Attrition Project (LSAP), which is a long-term effort to gather parallel data about language attrition in Arabic, Chinese, and Japanese as second languages. In their article in <u>Language attrition in progress</u>, Lambert and Moore (1986) describe progress that has been made in specifying criterion variables by developing diagnostic tests in listening, speaking, and reading (each with a subtest in lexicon and structure), and in specifying predictor variables. In the same book, Ginsberg (1986:20) writes that the major goal of LSAP is 'to determine the nature and extent of language skill attrition on the part of selected U.S. populations in Arabic, Chinese, and Japanese, both on a global basis and as a function of student-specific, language-specific, initial instruction-specific, and interim exposure/use variables'. It is assumed that 'language competence is a complex, multidimensional phenomenon whose components are not necessarily acquired or lost at the same rates or for the same reasons' (1986:21).

A number of interesting studies have appeared on the loss of a second language by children after returning to

their first language environment. Hansen (1980) wrote a dissertation on the acquisition and attrition of Hindi-Urdu by two English-speaking children. The oldest child (eight years old at the end of the study) spent three periods of time in India, but between periods of exposure apparently totally lost the second language and had to relearn it. The younger child (six years old) spent two periods of time in India, but showed a similar pattern of acquisition and attrition. Even though the children had virtually no conscious memory of Hindi-Urdu, previous experience seemed to facilitate the relearning. Hansen also states that the process of forgetting was the reverse of the acquisitional sequence.

Olshtain has studied the attrition of English in Hebrew-speaking children who had acquired a native-like mastery of English in the U.S. before returning to Israel (Berman and Olshtain 1983, Olshtain 1986, 1989). She shows that the result of attrition is greater variability in the application of marked rules, resulting from both a reversal of the acquisition process and from language transfer from Hebrew, and lower accessibility of specific lexical items. She says that we can assume 'reduced accessibility in vocabulary retrieval in all situations of attrition where there is a reduction of language use over longer periods of time. In cases of lexical retrieval difficulty, we can expect all attriters to develop some kind of compensation strategies' (1989:163). The greatest loss was shown by the

youngest children (ages five through eight), most of whom
had no further contact with English once they returned to
Israel. The older children (through age 14) received
further exposure to English in Israel, and 'although they
lose command of rapid retrieval of certain vocabulary items
and of certain syntactic porperties of English, these
children do retain a fair amount of native-like command of
English' (1983:233).

Cohen (1989) reports on lexical attrition in a third
language, specifically the attrition of Portuguese in his
own two children who were already fluent in English and
Hebrew. Cohen had Daniel (age 9) and Judy (age 13) retell
the same story after one, three, and nine months of
discontinued contact with Portuguese. He found a
significant decrease in lexical production, especially in
the younger child and especially with respect to nouns,
although Daniel could still identify the words in an oral
comprehension task. The children 'used at least six lexical
production strategies in order to compensate for forgotten
words--two of them L1-based (borrowing and foreignizing) and
four of them intralingual (the use of a general word,
approximation, circumlocution, and word abandonment)'
(1989:147). (Foreignizing involves using an L1 root word
with an L2 inflection.)

The interest in the question of whether language
attrition is the reverse of language acquisition is
reflected in the studies of Hansen (1980) and Olshtain

(1986, 1989) above. The most recent article dealing with the regression hypothesis (Jordens, De Bot, and Trapman, 1989) examines attrition in German case markings. For German-speaking immigrants who had spent at least ten years in the Netherlands (L1=German), grammatical case preserves its semantic function. However, for Dutch students who had studied German and then had minimal contact with it later (L2=German), the nominative case was used as the default case, reflecting the reverse of the acquisition process, in which the nominative case was the first learned. These results tend to support the regression hypothesis for L2 speakers of German, but not for L1 speakers of German who are losing their language in a second language environment.

Bahrick (1984a,b,c) carried out some remarkable cross-sectional research at Ohio Wesleyan University by testing 733 alumni, some of whom had never studied Spanish (the control group), some of whom were studying Spanish at the time, and others who had finished studying Spanish one to fifty years before. Independent variables included level of training reached, grades obtained, and opportunities to use Spanish since leaving formal study. Those with more than casual contact with Spanish since finishing their courses were eliminated from the study. Reading comprehension, vocabulary recognition and recall, grammar recognition and recall, idiom recognition and recall, and Spanish word order were all tested.

Bahrick's study produced a wealth of interesting data. Up until six years after taking a language course, retention rates dropped, but then remained level for several decades, finally to fall once more. Thus, after an initial loss, 'the remainder is immune to further losses for at least a quarter of a century, and much of that content survives for fifty years or longer' (1984c:111) in what Bahrick calls 'permastore'. Bahrick also found that 1) rehearsals during the retention interval such as listening to television programs in Spanish or conversing in Spanish were limited and had little if any influence on the retention scores, 2) the original level of training influenced the proportion of knowledge with permastore longevity, 3) production vocabulary stabilized later than recognition vocabulary, 4) the production of grammar declined most rapidly, and 5) reading comprehension was maintained at a level determined by recognition of vocabulary and grammar (1984c:111). Bahrick also notes:

> 1) much content survives more than fifty years; 2) grades received in courses continue to be valid predictors of performance for several decades following training; 3) training of a single course in the language is likely to leave little if any permanent content; 4) the larger the number of courses taken the greater the portion of content with permastore longevity. It appears that the total amount of content to be forgotten during the five years following training is relatively constant for individuals at different levels of training, but this amount becomes a progressively smaller portion of total knowledge with higher levels of training...; 5) attrition affects smaller portions of recognition vocabulary than of recall vocabulary. This difference may reflect the fact that the recognition vocabulary is larger, and that

the absolute amount of attrition is approximately
the same (1984c:116).

Neisser (1984) questions Bahrick's use of the word
'permastore'. Instead 'Bahrick could have made essentially
the same argument simply by postulating that some response
strengths reach a critical threshold during learning; beyond
that threshold, they become immune to interference or decay'
(1984:33). Thus, information that is tied into an extensive
cognitive structure is resistant to forgetting (1984:34).

As an interesting comparison to Bahrick's research, De
Bot and Lintsen (1986) and De Bot and Clyne (1989) have
gathered data on elderly Dutch-English bilinguals living in
Australia. They studied the reversion of these elderly
immigrants to their first language (Dutch) to see if there
was a final decline in L2 ability as observed by Bahrick.
They found language reversion in the immigrants who had
never reached the 'critical threshold' as mentioned by
Neisser above, but other immigrants showed 'surprisingly
little loss of proficiency in both Dutch and English over
the years' (1989:167). These elderly people, in contrast to
the subjects studied by Bahrick, had ample practice in their
second language, and thus were able to break through the
'critical threshold'.

Meara (1989) presents an interesting theory about
vocabulary retention in which students eventually reach an
equilibrium between words forgotten and words spontaneously
regenerated. 'Spontaneous regeneration' refers to

remembering words that the student was previously unable to access. The point of his article is that vocabulary testing techniques should look at the underlying pattern of vocabulary attrition rather than superficial surface phenomena. He gives three suggestions for testing programs: 1) the week should be the basic unit of time, 2) the attrition pattern between week 2 and 3 is the best predictor of the final equilibrium state, and 3) use of a simple, neutral assessment instrument, such as having students cross out foreign vocabulary words that they do not know well enough to say what they mean. Although he includes no data with the article, his equilibrium model is interesting because it predicts that learners will eventually reach a state in which scores will not drop further, and because he suggests a very simple instrument for testing vocabulary attrition. He does not deal with the effects of testing students each week or clarify whether the same words would be tested each time.

Two fairly recent studies have investigated the effect of summer vacation on second language retention. A dissertation by Robison (1985) reported the results of testing 180 junior and senior high Spanish students before and after the summer vacation. Robison concluded that 1) listening and reading comprehension were not affected, 2) less-advanced students showed significant loss in writing while more advanced students did not, 3) vocabulary recognition and recall were significantly affected, though

from a practical point of view loss was minimal, 4) FLES students retained more than non-FLES students, 5) no loss was noted in ability to produce linguistically correct forms and no evidence showed that forgetting is the chronological reverse of learning, 6) a sequence of forgetting was observed, varying according to one's level of achievement, and 7) exposure to Spanish during the summer raised performance. Robison's conclusions that significant loss was noted in vocabulary, but not in listening and reading comprehension, are similar to the conclusions of studies done before 1980 about the effect of summer vacation.

The results of a study by Moorcroft and Gardner (1987) seem to contradict Robison's conclusions about the forgetting sequence. Moorcroft and Gardner tested 114 students of French twice: upon their completion of 9th grade and again as they started the 10th grade. Their conclusion is that 'while grammatical structures in general are more likely to be affected by language loss than vocabulary, it seems to be the most recently acquired structures that are most vulnerable to language loss'. (1987:337) The loss in grammatical accuracy was related to the fact that 'most grammatical structures are incompletely and recently learned' (1987:338).

A number of other recent studies should be mentioned before concluding this section on language attrition research. Godsall-Myers (1981) in a dissertation on the attrition of German in six college students concluded that

1) the attrition proceeds at different rates for different skills, 2) the attrition is inversely related to the cumulative average the student had in the low-intermediate German class, and 3) the attrition of L2 patterns will reflect the learning order and the frequency of occurrence of L2 patterns. She also suggests that vocabulary be taught using semantically paired items and synonyms and antonyms and that students be taught principles of word formation in German.

In 1984, the same year that Bahrick reported on long-term retention of Spanish taught in the U.S., Clark and Jorden reported on retention of Japanese taught in U.S. schools. As in Bahrick's study, participants were asked to fill out retrospective questionnaires. In contrast to Bahrick's study in which participants had little contact with the foreign language after finishing course work, Clark and Jorden found that many of their participants had extensive contact with Japanese after finishing their courses. Clark and Jorden did not include the many college students whose only contact with Japanese had been a single-year part-time course, since 'the language acquisition of such students is so minimal that, except for commonly used ritual expressions and the most basic grammatical patterns, attrition is almost total after a comparatively short period of time away from the classroom' (1984:16-17). No correlation was found between attrition of speaking and listening comprehension skills on the one hand and attrition

of reading skills on the other hand, showing the independence of these skills. Some other interesting conclusions were that "no differences were found between 'attriting' and 'non-attriting' students in the reported ease with which they were able to learn foreign languages in general or Japanese in particular" (1984:52-53) and that 'decreases in language proficiency since the time of maximum performance are directly related to decreases in the extent to which these skills are in fact currently being used' (1984:53). For those students who originally had a good functional control of Japanese and had seemed to reach a critical point in overall proficiency, renewed exposure to real-life settings was a powerful means for re-acquisition (1984:55,58).

Fakhri (1985) examined morphological and syntactic simplification in the interlanguage of American college students of French after periods of three, eight, and ten months. Morphological simplification in the interlanguage occurred in written narratives, and also in written narratives the students increased their use of full noun phrases in order to avoid clitic pronouns. Fakhri called for further study in other areas of language attrition, including the loss of vocabulary in comprehension and production (1985:157).

Weltens, Van Els, and Schils (1989) reported on a four-year study on the attrition of French receptive skills in Dutch students who had finished their study of French four

and six years earlier, but who had had very little contact with French since then. These students had originally been at a very high proficiency level, since they had received 400-600 hours of training in French, and therefore had probably mastered French above Neisser's 'critical threshold'. The researchers report gains in listening and reading proficiency and say that these gains might be attributed to 'general cognitive maturation, further academic training, and continued learning of other foreign languages' (1989:214). Lexical tests and grammar tests showed some attrition. Confirmation of Bahrick's findings were shown in that 1) most loss occurred in the first two years and then leveled off and 2) a fixed <u>amount</u> of material was lost rather than a fixed <u>proportion</u>. However, unlike Bahrick, they found that attrition was independent of training level.

Summary of literature on language attrition

The research during the past few years in second language attrition has been overwhelming. No longer can the field be called 'antenatal'. A summary of the research at this point is difficult, but some general observations can be made:

1. Attrition of the lexicon is an extremely important area for study in all types of language loss. For example,

R. Andersen believes that it 'far outweighs morphosyntactic and phonological aspects of attrition' (1982:113).

2. The level that a student reaches affects subsequent attrition (Kennedy 1932, Godsall-Myers 1981, Lowe 1982, Bahrick 1984, Robison 1985). A single foreign language course is unlikely to leave any content (Bahrick 1984, Clark and Jorden 1984). There may be a 'critical threshold' of language skill that must be reached in order for much to survive (Neisser 1984, Weltens et al 1989, De Bot 1986,1989).

3. Higher grades received in course work are related to higher retention rates (Godsall-Myers 1981, Bahrick 1984). In contrast, Clark and Jorden (1984) report no difference in initial ease in learning the language for 'attriting' and 'non-attriting' students.

4. Motivation and attitude may have little to do with attrition rates (Edwards 1976, 1977, Gardner 1985). Kennedy (1932) did report higher retention rates for students who had a desire to continue their study of Latin.

5. The regression hypothesis (first to be learned is last to be forgotten) does not appear to be true for first language loss (Caramazza and Zurif 1978, Berko-Gleason 1982a), but may apply in cases of second language loss (Cohen 1974, 1975, Hansen 1980, Godsall-Myers 1981, Moorcroft and Gardener 1987, Olshtain 1986, 1989, Jordens et al. 1989). However, more work at this point has been done in morphology and syntax than in vocabulary.

6. Language attrition is most rapid at first, but appears to reach a state where it levels off (Bahrick 1984, Weltens et al. 1989). Anderson and Jordan (1928) found rapid initial attrition of Latin vocabulary. Meara's matrix theory (1989) predicts rapid initial attrition of vocabulary with a leveling off and a final state of equilibrium.

7. A fixed <u>amount</u> of material appears to be lost rather than a fixed <u>proportion</u> (Bahrick 1984, Weltens et al. 1989).

8. Production skills (speaking, writing) usually suffer more attrition than recognition skills (listening, reading) (Scherer 1957, Edwards 1976 and 1977, Lowe 1982, Bahrick 1984, Robison 1985).

These generalizations will be discussed specifically in relation to the research hypotheses for this project at the end of Chapter 2.

LITERATURE ON FOREIGN LANGUAGE VOCABULARY

Importance and organization of the lexicon

With the advent of comprehension-based approaches to foreign language acquisition and the call for communicative competence, there has been renewed interest in the importance of vocabulary. Extensive bibliographies compiled by Meara (1983, 1987) contain entries calling for a massive input of vocabulary from the earliest stages of language learning (such as Judd 1978), an entry insisting on the importance of vocabulary acquisition at the middle stages of language learning (Twaddel 1973), and another stating that vocabulary learning is problem number one at the advanced level (Marton 1977). As noted in Chapter 1, vocabulary and pronunciation are considered the most important components for the novice level in the Oral Proficiency Interview.

Meara (1984) cites an article by Levenston (1979) concerning issues and problems in second language vocabulary acquisition and points out that studies of the lexicon in the interlanguage tradition are not very numerous in comparison with the importance accorded to the lexicon. It has been shown that for native speakers lexical errors are more disruptive than grammatical errors, and lexical errors consistently outnumber grammatical errors by three or four to one (Meara 1984:229). One of the reasons for the neglect of the lexicon by linguists is the belief that it, 'unlike

our knowledge of syntax and phonology, is an inherently
messy part of our linguistic competence, and that it cannot
be handled conveniently using the sorts of rules which have
been used to describe the more manageable parts of a
language' (1984:230).

In response to Meara's assertion that little work has
been done on the lexicon, Sharwood Smith (1984) points out
that such fields as word morphology, semantics, phonology,
and syntactic and pragmatic coding touch on lexical matters.
Even generative grammar has become much more lexically
oriented, and the area of psycholinguistics has much to
offer. Gass (1987) agrees that many areas of research
'impinge on questions of the lexicon', but still 'the
lexicon has been dealt with somewhat tangentially'
(1987:130). Fundamental questions remain as to how and what
is learned, the organization of the lexicon, and access and
retrieval of the lexicon.

Meara (1984) also considers what a typical lexical
entry might look like in a second language learner's memory.
The entry would consist of two parts, a phonological or
orthographical representation (perhaps only partially
spelled out), and a semantic representation specifying
meaning (perhaps as part of a semantic tree or cluster).
Meara mentions various methods that one might use for
ordering, storing, searching for, and accessing these
lexical entries, but there is no principled way of choosing
among them. Meara also compares the lexicons of second

language learners with the lexicons of native speakers and
finds several differences: 1) semantic networks are much
more loosely organized and unpredictable for learners,
especially for material recently learned, and 'semantic
factors are frequently overridden by extraneous phonological
factors' (1984:234); for example, some learners associated
prêtre 'priest' with argent 'money', confusing it with
prêter 'to lend'; 2) phonological representations of words
may be only partially sketched in, and although initial
sounds and consonant clusters tend to be preserved, the
placing and nature of vowels is not so reliable; 3)
different languages may use different word-handling
strategies, for example, syllables may play 'a much more
important role in the representation of words for Spanish
speakers than is the case for English' (1984:235). Henning
(1973) has demonstrated that beginning language learners
tend to encode vocabulary into acoustic clusters (according
to similarities in sound), while higher-proficiency learners
tend to cluster words semantically, relying more on
similarities in meaning than on sound.

Vocabulary teaching and learning

The realization that words come in semantic networks or
sets is extremely important for teachers and learners.
According to Ludwig (1984:559), 'the foreign language
teacher cannot look at words in isolation. In any language,

each word suggests other formally or functionally similar words...' D. Brown (1980) encourages the teaching of collocations (words that often occur together, such as <u>sweet</u> and <u>sugar</u>), clines (gradations of words such as <u>freezing</u>, <u>cold</u>, <u>tepid</u>, <u>warm</u>, <u>hot</u>, and <u>boiling</u> water), and clusters (a group of similar things, such as <u>heat</u>, <u>Fahrenheit</u>, <u>conduct</u>, <u>melt</u>, <u>hot</u>, <u>temperature</u>, <u>sun</u>, and <u>calorie</u>).

In addition to semantic associations, structural and mnemonic associations can be formed (Cohen 1987). Structural associations are made when students analyze the root and affixes of a word to understand the meaning. There are several kinds of mnemonic associations; for example, in a chain-type mnemonic the learner links words through their use in a story or rhyme. An example of a verbal mnemonic is a Spanish speaker's use of the word <u>choca</u> (strikes) to remind him of the English <u>chalk</u>, by forming the sentence <u>La tiza se choca con la pizarra</u> (1987:45).

An example of an imagery mnemonic is the keyword technique, which has been well researched by Cohen and others. In the keyword mnemonic, the learner thinks of a word in his native language that sounds like the foreign language word. For example, <u>pato</u> (duck) in Spanish sounds like the English <u>pot</u>. The next step is to visualize a duck and a pot together, for example, a duck with a pot on his head. The next time that the learner hears the word <u>pato</u>, he will remember the image and the meaning <u>duck</u>. The keyword mnemonic has been tested for effectiveness in many

different research projects. When tested under laboratory
conditions, the results are always positive when compared to
other methods of learning vocabulary, but when tested in the
classroom, results have been mixed. Not all students use
the keyword method when instructed to do so, and some
students conjure up mental images even when not taught to do
so. In the cases of younger children, better results are
obtained when the teacher creates the image for the child,
but adults seem to do just as well whether they create their
own image or use one created for them.

Cohen (1987) also answers criticisms of the keyword
technique. Though creating a keyword image ties the foreign
word to the native language, there is a strong tendency for
students to translate to their native language anyway, even
in a method such as the Natural Approach. The keyword
mnemonic also focuses on just one meaning of the word to be
learned, but in cases of words with multiple meanings,
perhaps this one meaning can trigger the other meanings.
Nouns are definitely the easiest to learn with the keyword
approach, and verbs and adjectives are harder. Even though
it appears that using the keyword mnemonic would take more
time than learning a word through other methods, groups
using keywords do better even when under time constraints.
Some research has shown that associational devices have
lasted up to eight weeks, but there is a need for research
over longer periods of time to see if and when the keyword
drops out.

Although some writers believe that vocabulary can be learned more quickly from word lists than within context (Nation, 1982, cites a number of experiments and concludes that for the initial learning, words in isolation are just as or more effective as words in context), other writers (D. Brown 1980) encourage learners to use the context to find meanings of words. Part of the problem is determining just what 'context' means. Is context the presence of a nondefining sentence, a defining sentence, or the presence of a story? And what can be done to ensure that the learner is really learning from the context in the foreign language, rather than simply forming a covert bond between the foreign language word and a native language word as if he were learning from a word list (Nation 1982:23)? Koster (1987) emphasizes context in the recognition and comprehension of speech; for example, 'He sent me a Christmas guard' was often accepted by Dutch speakers as a correct English sentence, since they interpreted <u>guard</u> as <u>card</u> (1987:25). Schouten-Van Parreren (1989) contends that words should not be taught in isolation since such words present neither a linguistic nor a psychological reality. Isolated words cannot 'evoke emotions or involvement in the learner, a factor which plays an ...important part in long-term acquisition' (1989:77).

Haastrup (1989) provides a cautionary note about context. She believes that if the context is too predictable and the word is too easily guessed,

comprehension is aided, but not long-term learning.
However, if the context does not make the meaning of the
word obvious, and if the learner must do some analysis of
the surface structure of the word, the word will have a more
distinctive representation in memory and will be better
retained.

A word judgment of this type may require more effort,
but it may pay off later in better recognition and recall
because of the depth of analysis (see page 73 of this
dissertation for a description of Craik and Lockhart's
'depth of processing' framework). The keyword mnemonic, as
well as Asher's Total Physical Response, are techniques that
may also increase depth of analysis, owing 'their success to
the dual psychological realities of imagery and concreteness
of individual words' (Ludwig 1984:555). 'Visual information
appears to interact with verbal stimuli, demanding that the
learner make a linguistically sophisticated analysis'.

There is also a good deal of research on the use of
pictures in learning foreign language vocabulary,
substantiating the value of imagery. Deno (1968) observed
that there were problems with using pictures to learn
foreign words, since the meaning provoked by the picture may
not be the one intended. However, Kellogg and Howe in 1971
demonstrated that, at least for children in grades 4-6, the
use of a picture instead of an English word was superior in
a paired-associate learning situation. Some of the
children, though, learned faster with words than with

pictures, showing a variety in learning style. A study by Champagnol in 1972 also showed the superiority of learning with pictures, especially in retentjon of the words twelve weeks later. The use of pictures with words produced better performance than words without pictures, but the best performance of all was produced by the use of real objects.

D. Brown (1980) has compared vocabulary acquisition to a hill that grows taller and taller, but which is constantly threatened by the erosion of forgetting at the bottom. She summarizes the methods for developing vocabulary and preventing forgetting with the formula eight Cs and a G: use of collocations, clines, clusters, cloze procedures, context, consultation (with a book or dictionary), cards (flashcards with native language on one side and target language on the other), creativity, and guessing (with and without context).

Word characteristics: word difficulty and frequency in input

Another interesting area of research is determining the relative difficulty of words. Nation (1982:17) states that words are difficult when they are hard to pronounce or when 'their English translations are adjectives, adverbs, or verbs, rather than nouns'. Cohen (1986:148) makes an interesting point when he states that nouns may be easier to forget because they are easier to learn. Foreign words are easy when they are similar in form and meaning to their

English translations or when they are 'morphologically transparent' and their meanings can be understood from the meanings of their parts (Laufer 1989:11).

Murray (1986) studied the characteristics of words that would make them easier to translate into a foreign language. He obtained ratings for 145 words according to characteristics such as frequency, memorability, intensity, emotionality, goodness, pleasantness, and concreteness, and then asked the students to translate the English words to French. He noted that variables such as emotionality and imagery, which influence long-term episodic memory, had little effect on translation time, but 'translation efficiency was most strongly influenced by the frequency of the word in the language, its familiarity, and the similarity of the French equivalent to the English word' (1986:353). Ludwig's article on word characteristics (1984) points out that nouns and concrete words are easier to learn because they can evoke a visual image in the mind. Also words that are positively loaded are recalled better than words with neutral or negative connotations.

A few writers have reported on the number of times a new word must be repeated before students normally remember it. Nation (1982:17) cites evidence that seven repetitions are sufficient for most students. Carpay (1975, cited in Palmberg 1987:208-209) proposed a 4+1+1+1 formula: 'New words should be used in at least four different contexts in the introductory lesson, followed by (at least) one

occurrence in each of three subsequent lessons'. On the other hand, for indirect (incidental) vocabulary learning in context, Nation (1982:17) says that the average number of encounters needed to recognize a word is approximately sixteen. Nation makes other comments on indirect vocabulary learning, saying that even though the greatest share of vocabulary learning is indirect, direct learning of vocabulary is still an important component that speeds vocabulary development (1982:15-16).

Recognition and production vocabulary

A very important area of study, also mentioned in Chapter 1, is the relationship of recognition vocabulary to production vocabulary. At this point I will review the literature specifically on foreign language vocabulary, but the general concept of recognition and production will again be discussed in the third section of this chapter on studies in human memory. Recognition vocabulary is sometimes known as receptive or passive vocabulary, but the term passive seems to be ill chosen, since it implies that the learner has nothing to do.

Melka Teichroew (1982) has written a perceptive article on recognition (receptive) vocabulary versus productive vocabulary. She first reviews various tests for receptive vocabulary, ranging from the rather arbitrary method of

checking off words in a dictionary to some very commonly-
used methods such as multiple choice and translation tests.
The most common production test is to ask the subject to
complete a sentence in L2 with the aid of a translation in
L1 or a related word in L2. Melka Teichroew goes on to
question the point where receptive knowledge ends and
productive knowledge begins. She notes that in reception we
need to hear or see only enough information to distinguish
one word from another, and she refers to Brown and McNeill's
(1966) tip-of-the-tongue experiment in which subjects had
enough information to recognize a word (and often tell the
number of syllables and the first or final letter) but were
unable to produce it. She points out that there are many
different features indicating familiarity with a word: its
spelling, phonology, morphology, and semantic and syntactic
properties. A student who can produce a word may be
familiar with it to a greater or lesser degree. Melka
Teichroew summarizes various reports that indicate that
receptive vocabulary in a second language is about twice the
amount of productive vocabulary, especially for beginning
students, but at later stages the gap seems to close. She
concludes that recognition and production are points along a
continuum beginning with the first stage of recognition and
ending with various stages of production depending on
familiarity with the word. The form of a test has strong
influence on the results that are produced. She believes
that there is only one lexical store which we use

receptively or productively according to need, and she presents two questions for further research: 1) is there a meaningful gap between comprehension and production, or is this gap more seeming than real? and 2) does the same comprehension/production problem exist in L2 as in L1?

Cohen (1986) also emphasizes that vocabulary acquisition is a continuum. Palmberg (1987:203) asks if all words necessarily pass automatically from recognition to production knowledge, or if some words never reach the production stage and others bypass the recognition stage. He also believes that the relative sizes of the recognition and production vocabulary are 'independent until the productive vocabulary gets exceedingly large'.

Summary of research on foreign language vocabulary

Certain ideas recur in the literature on foreign language vocabulary. These generalizations will be discussed specifically in relation to the research hypotheses for this project at the end of Chapter 2.

1. The lexicon is an extremely important component of foreign language acquisition, although it has often been neglected. Lexical errors are more serious and outnumber grammatical errors (Meara 1984).

2. Words may be only partially 'sketched in', with initial sounds and clusters the most important. Recognition does not require as much information as production (Melka Teichroew 1982, Meara 1984).

3. Beginning learners tend to cluster words according to phonological rather than semantic features (Henning 1973, Meara 1984). It is important to encourage semantic clustering of words and to form semantic networks (D. Brown 1980, Ludwig 1984).

4. Mnemonic devices such as the keyword technique show positive results under laboratory conditions, but have mixed results in the classroom (Cohen 1987).

5. Opinions on the effects of context differ. Some writers say that words can be learned as well in isolation (Nation 1982) or caution that too much context may discourage structural analysis of a word (Haastrup 1989). Others stress the importance of using context to find and retain the meaning of a word (D. Brown 1980, Schouten Van-Parreren 1989).

6. Depth of analysis results in better retention of a word (Craik and Lockhart 1972).

7. Learning words through pictures and real objects generally results in better performance (Kellogg and Howe 1971, Champagnol 1972, Ludwig 1984).

8. Features that make words easier to learn are part of speech (nouns), similarity to L1, morphological transparency, and frequency in the language (Carpay 1975, Nation 1982, Ludwig 1984, Cohen 1986, Murray 1986, Laufer 1989). Emotionality and imagery had little effect on translation time in Murray's research (1986), but other writers (such as Ludwig 1984) assert the importance of emotionality and imagery in vocabulary learning and retention.

9. Recognition and production are points along a continuum (Melka Teichroew 1982, Cohen 1986).

10. Opinion is divided about the relationship of size of recognition and production vocabulary. Melka Teichroew (1982) cites reports indicating that for beginners, recognition vocabulary is about twice the size of production vocabulary, although the gap seems to close for more advanced learners. Palmberg (1987) believes that the sizes of recognition and production vocabularies are relatively independent, except when productive vocabulary is very large.

LITERATURE ON HUMAN MEMORY

As Weltens (1987:23) has noted, research on memory and forgetting has a long tradition, going back to Ebbinghaus in the late nineteenth century. However, much of this research lacks relevance for studies in language attrition, since experiments done under laboratory conditions will not necessarily show what actually happens in real language learning and forgetting situations. Bahrick (1979:298) states:

> Methodologically, most research has been limited to information acquired in the laboratory and tested later for retention. This longitudinal approach offers high control over the conditions of original acquisition, but it limits research to simple material that can be acquired within a few laboratory sessions and to short retention intervals...The point is not that all previous memory research is irrelevant to the acquisition and maintenance of knowledge systems. Past findings dealing with interference, level of processing, distribution of practice, and so on, almost certainly have a bearing on how well knowledge is maintained, but it is impossible to determine the applicability of any findings without a data base relevant to the acquisition and maintenance of knowledge under more ecologically realistic circumstances.

As Bahrick points out, language learning and maintenance are problems that must be studied in a real-world context. However, language attrition is a problem intrinsically connected to human memory, and we cannot fail to note the implications of verbal learning experiments in psychology for language attrition studies. This section will first deal with problems related to short-term and

long-term memory, including problems of recognition versus
recall and intentional versus nonintentional learning. Then
this section will turn to recent discussions of depth of
processing and declarative and procedural memory and will
relate these discussions to the field of language attrition
research.

Short-term and long-term memory

Traditionally, human memory has been divided into three
different stores: 1) sensory memory, which holds visual and
acoustic information for up to two seconds, 2) short-term
memory (also known as primary or working memory), which is
limited and holds information up to thirty seconds, and
which requires rehearsal if the information is to be
retained, and 3) long-term memory (also known as secondary
memory), with no known limit and which may last from minutes
to years. Loss in this memory store is due to interference
from new information. Some psychologists also include
tertiary memory, which is long-term memory immune to
forgetting through interference. Bahrick (1984a,b,c) refers
to this type of memory as 'permastore'.

Short-term memory is thought to have a capacity of
seven 'chunks' of information, plus or minus two. When
subjects are asked to repeat words in a list, they show both
a primacy effect and a recency effect; thus the first few
and last few items are the best remembered, whereas the

words in the middle are more likely to be forgotten. The
resulting memory curve is known as the serial position
curve. Both Meara (1980a) and Spitze and Fischer (1981)
have performed interesting experiments in short-term memory
with language learners. Whereas native speakers are able to
chunk words in a sentence into larger syntactic and semantic
groups, thus being able to hold more words accurately in
short-term memory, non-native speakers show the typical
serial position curve, with the middle items particularly
fuzzy, as if the items in the sentence were unrelated. This
inability to form semantic and syntactic chunks
simultaneously becomes more pronounced the lower the level
of the learner. Data from Henning (1973) support the view
that low-proficiency students attempt to register vocabulary
in memory through sound (acoustic) similarities, whereas
high-proficiency students rely more on semantic
associations. Pimsleur (1967) provides a 'memory schedule'
for learning individual words which is frequently cited:
When a word reaches a probability of 60% of being remembered
(perhaps after only 5 seconds in short-term memory,
especially if there is intervening material), it should be
repeated. Since forgetting is exponential in form, the next
repetition of this word should come 5^2 (or 25) seconds
later, the next repetition 5^3 (125) seconds (or two minutes)
later, and so on. Pimsleur calls this schedule 'graduated
interval recall'.

Many experimenters have found that repeating the presentation of items enhances both recognition and recall of these items. However, giving the learner more time to rehearse items on his own improves recognition memory but has little effect on ability in free recall (Eysenck 1982:221).

Ebbinghaus performed the first experiments on short-term and long-term memory in the second half of the nineteenth century. His attempts to recall lists of nonsense syllables led to the traditional 'forgetting curve', and his work set the pattern for further research in serial learning and paired-associate learning in the twentieth century. Serial learning refers to learning a series of items in order, and paired-associate learning refers to a specific stimulus memorized with a specific response. Retention can be measured by 1) recall (in serial recall the items of a list must be recalled in order, in free recall the items can be recalled in any order, and in cued recall the subject is given a clue and must respond appropriately), 2) savings (difference in time taken to learn information a second time), or 3) recognition (either asking the subject to identify an item as 'old' or 'new' or giving a multiple-choice test).

A great deal has been written in the psychological literature about the difference between recognition and recall. Testing for recognition memory generally produces the higher scores, though scores may also be dependent upon

the subjects' knowing which mode they will be tested in
(Tversky 1973). Wingfield and Byrnes (1981:31) give these
reasons for the superiority of recognition: 1) recognition
tests 'restrict the size of the set of alternatives from
which the correct response must be drawn', 2) correct
judgments can be made on the basis of partial recall, and 3)
if alternatives on the multiple choice test are very
different from each other, the test can be very easy
(however, if the alternatives require discrimination between
fine nuances of meaning, the recognition test may be even
more difficult than the recall test).

J. Brown (1976) has edited a book entitled <u>Recall and
Recognition</u>. In it he states that recognition can be
mediated by recall, especially when the subject must search
through many items (1976:2). Low levels of learning suffice
for recognition, but not for recall, since the amount of
information required for a correct response is lower in
recognition. Brown also speaks about research design:

> In assessing differential effects, an extremely
> tempting within-subject design is available.
> Retention can be tested first by recall and then
> by recognition. This greatly increases the
> precision of the experiment since recall and
> recognition are then tested on the same subjects,
> on the same material and on the same occasions.
> One practical difficulty is that the statistical
> analysis of the results can be tricky. A simple
> solution if recognition and recall are assessed
> using the same measure may be to analyse in terms
> of the difference score (recognition minus recall)
> for each subject: the statistical analysis is
> then conducted on a table of independent
> difference scores (1976:34).

In the second chapter in the same book, Tulving (1976) questions whether recognition and recall are fundamentally two different processes, or if they are basically the same process of using stored information. He comes to the conclusion, as do Lockhart, Craik, and Jacoby in Chapter 3, that the two do not differ in a crucial way: 'they are different only in the sense that in recognition re-presentation of the stimulus provides better information from which the initial encoding can be reconstructed' (1976:85). In Chapter 5 Cooper and Monk deal with details of testing. Testing may cause consolidation of the learned material and stereotyping in recall. The same items tend to be recalled whether they are correct or incorrect (in reference to free recall from a list). On repeated recognition tests, false alarms become consistent. A point made by Murdock in Puff (1982:5) is that recognition memory can be tested for latency (time taken for subject to respond) and subject's confidence level, as well as for accuracy. In the same book Bahrick and Karis (1982:433) make the following point about recall and recognition measures:

> ...with high degrees of original learning and short retention intervals, recognition performance is likely to remain near the ceiling and fail to reveal forgetting, whereas recall measures are more sensitive. With low levels of learning, or very long retention intervals, recall performance may · decline to a point where it is no longer sensitive to further retention losses, whereas recognition performance may continue to reflect changes over time....ecological memory investigations that extend over long time periods are likely to require

both recall and recognition measures to reflect adequately the changes occurring during the entire retention interval.

Research done on the 'tip of the tongue' phenomenon is related to the concept of recognition and recall. Brown and McNeill (1966) wrote the first article on this phenomenon, but other articles have appeared since (for example, Kohn et al. 1987). In the study by Brown and McNeill (1966), definitions were given, and even when subjects were not able to recall the words, they were often able to give some features, such as the beginnings and ends of the words, the number of syllables, and the location of the primary stress. Sometimes this 'tip of the tongue' state ended in recall, and other times only in recognition when shown the correct word. It seemed that some features of the word were 'penciled in' more faintly than others. Brown and McNeill reached the following conclusion:

> It is consequently possible to recognize words when one has not stored the complete letter sequence. The evidence is that we do not store the complete sequence if we do not have to. We begin by attending chiefly to initial and final letters and storing these. The order of attention and of storage favors the ends of words because the ends carry more information than the middles. An incomplete entry will serve for recognition, but if words are to be produced (or recalled), they must be stored in full (1966:335).

Brown and McNeill add that the need to produce encourages the registration of the middle letters of a word. Also the growth of vocabulary should force attention to the

middle of a word, because more detail is needed to distinguish one word from another as vocabulary increases (1966:337).

Retrieval of a word may also be heavily influenced by context or the way that an event was encoded (Craik 1979:88). According to Weltens (1987:32) 'recall should be facilitated by a test situation which more or less reflects the learning situation'. By using think-aloud procedures, Cohen (1986) discovered that in addition to using structural, semantic, and mnemonic associations as retrieval strategies, learners attempted to recall the learning situation, such as the textbook or the situation in which the word came up. Wingfield and Byrnes (1981:44) also assert that recall can improve when a test occurs in the same environment where learning occurred. From his experiments in free recall, Jenkins (1974:788) concludes that 'the most important determiner of recall was the nature of the <u>event</u> the subjects experienced when their task brought them into contact with stimulus words. This finding suggests that subjects recall the quality of the events they have experienced, not stimuli to which they have been exposed'.

Questions have also been asked about the effect of intent to learn. Is incidental learning as effective as intentional learning? According to Eysenck (1982:197, 201-202), most human learning is incidental, since no test of learning is expected. From recent experiments it seems as though the orienting task given before learning and the type

of processing performed by the learner are more influential than intent to learn. 'The most appropriate generalization is probably that intention to learn has little or no effect on recognition memory, whereas its effects on recall are more variable'(1982:206).

Another question frequently asked about long-term memory is what causes forgetting. Do memories spontaneously decay over time, or is access blocked because of interference from new material? Most of the research points to the validity of the interference hypothesis, that what we learn before and after any piece of information will influence how well we will remember that information. Time is a <u>factor</u> in forgetting, but not its <u>cause</u>.

Bahrick's study on long-term retention of Spanish has already been summarized in the first section of this chapter. He has done other cross-sectional studies of long-term ecological memory, that is, memory in real-life social and cultural settings rather than laboratory settings. In a 1979 study, Bahrick used alumni from Ohio Wesleyan University for testing maintenance of knowledge of the city of Delaware, Ohio, and compared this retention to the subjects' opportunities for 'rehearsal' of this knowledge (the frequency, duration, recency, and distribution of their trips back to their alma mater). Bahrick (1982:428) pointed out that we must study variables that are 'ecologically important rather than just those that are easily manageable...It is acceptable to sacrifice some control over

critical variables in order to investigate ecologically important phenomena that would otherwise be neglected'. In connection with longitudinal testing designs, Bahrick (1982:432) states:

> Testing the same individuals repeatedly has the advantage of control over individual differences; however, there is the danger that later test performance will be affected by the practice or interference resulting from the earlier tests and will therefore fail to yield an unbiased estimate of the amount of forgetting during the interval. To avoid this problem, laboratory research in memory has primarily used between-subject comparisons. Semantic memory content is generally retained over much longer time intervals than the intervals involved in typical laboratory research, and the problems created by successive testing are likely to be trivial because of the long intervals between successive tests. For that same reason, however, the longitudinal approach becomes cumbersome.

In this article Bahrick goes on to describe the longitudinal studies done over the summer vacation by Cohen (1975) and Smyth et al. (1973) and gives his own methods for conducting cross-sectional research.

Depth of processing

Craik and Lockhart (1972) proposed an alternative view to the 'duplex' model of short-term and long-term memory. They offered a model based on 'depth of processing', ranging from a very shallow sensory analysis of a stimulus to a deeper analysis based on meaning. They proposed that deeper memory traces or codes are more enduring than shallow ones,

and forgetting is a function of the depth of processing
(Wingfield and Byrnes, 1981:282). Memory is a result of the
cognitive operations that have been performed on the
stimulus. Processing for semantic features yields much
better word retention than processing for phonological
features or for physical structure (for example, if the word
is written in capital letters or not). The depth of
processing concept has resulted in more calls for meaning
and communicative activities in the classroom (for example,
Stevick 1976:43).

Since the appearance of this model, there have been
some amendments to it. Rehearsal does improve recognition,
even if it is on the same and not on a deeper level. Also
some sensory memories, such as the sound of a voice, can be
very enduring. It is difficult to specify the mechanisms
that lead to deep processing or to define it, and it is
impossible to prove that deep processing has <u>not</u> occurred.
Depth of processing can easily be confounded with factors
such as processing time, difficulty of processing, or
distinctiveness of processing. (Depth of processing may also
make a memory more distinctive and thus lessen interference
from other memories.) However, depth of processing remains
a viable model and has been the impulse for a great deal of
research.

The advantages for learning with pictures can perhaps
be explained through depth of processing, because imagery
can lead to a more meaningful (or perhaps a more

distinctive?) analysis. An alternative explanation is that pictures induce a dual code--an imaginal code and a verbal one. In any event, 'pictorial presentation and imagery coding substantially facilitate both recall and recognition memory' (Paivio 1976:120).

Another interesting item to note is the effect of personal significance on memory. Stevick (1976: 38-39) reports on several experiments showing that memory is better when a subject's ego is involved (for example, when the subject reads a passage supposedly derived from his own personal inventory) or when a word is emotionally loaded (such as money, slut). Schouten-Van Parreren (1989:77) concurs that words are better learned when they evoke emotions or involvement in the learner.

There are other terms for talking about memory and dividing it up into different types. Common terms are semantic and episodic memory. Semantic memory refers to general rules that we know, but that are removed from the context in which we learned them--such as 2 X 2 = 4 or that an apple is fruit (Bahrick 1979:297). Episodic memory refers to memory of particular events related to time and place, such as yesterday's menu or an accident we had. Variables such as word frequency have a greater effect on semantic memory and variables such as emotionality and imagery affect retrieval from episodic memory to a greater extent (Murray 1986).

Cognitive theory

In their article applying cognitive theory to second
language acquisition, O'Malley, Chamot, and Walker (1987)
discuss the role of memory. According to them, language is
not learned separately from other cognitive skills but is
stored and retrieved from memory like other complex
cognitive skills. They believe that the short-term, long-
term memory model is inadequate for explaining language
acquisition, and they prefer the theoretical model developed
by John Anderson (1983, 1985). Things we know about (static
information) constitute declarative knowledge, and things we
know how to do (dynamic information) constitute procedural
knowledge. Declarative knowledge is represented in memory
by 'nodes that are associated with other nodes through
connecting associations or links' (1987:290). Schema are
interconnected networks of nodes depicting complex concepts.
These schema allow us to organize and understand new
information.

Procedural knowledge is used over and over again to
solve problems and to understand and generate language.
Unlike declarative knowledge which may be acquired rapidly,
procedural knowledge is acquired gradually and with
extensive opportunities for practice. Instructional methods
for second language acquisition need to focus on
communicative activities which give practice for language as
a skill rather than as an object of study. The language

learner must procede from rule-bound declarative behavior to an automatic procedural stage. During the first or cognitive stage, the knowledge is declarative and the learner tends to be consciously aware of learning. During the second or associative stage the learner eliminates errors in the information and strengthens the skills, resulting in an autonomous stage in which the skill becomes virtually automatic and unconscious (see also McLaughlin 1987). O'Malley et al. propose that vocabulary is declarative knowledge and predict fairly rapid attrition:

> ...aspects of the language that are at the first or cognitive stage of acquisition and are therefore represented by declarative knowledge would be forgotten first, whereas those aspects of the language that have become automatic or proceduralized would be retained...Vocabulary is identified with concepts, and thus can be represented as declarative knowledge (1987:303).

Jenkins (1974) cautions that complex human behavior is qualitatively different than the simple behaviors often measured by experiments in the verbal-learning tradition. What memory is depends on context. The contextualist takes the uncomfortable position that a "'final' analysis is a myth, that analyses mean something only in terms of their utilities for some purposes" (1974:787). 'What is remembered in a given situation depends on the physical and psychological context in which the event was experienced, the knowledge and skills that the subject brings to the context, the situation in which we ask for evidence for

remembering, and the relation of what the subject remembers to what the experimenter demands' (1974:793). The contextualist must work from the top down and remember that human memory is much more than simply a sum of its parts.

Summary of literature on human memory

The literature on human memory is vast, and a brief but accurate summary is difficult. However, the following generalizations can be made that are relevant to second language learning.

1. It is important to conduct memory research in ecologically realistic circumstances (real-life settings) as well as in the laboratory (Bahrick 1979, 1982, 1984).

2. Short-term memory holds a capacity of about seven chunks for up to thirty seconds. When repeating sentences, native speakers can form larger semantic and syntactic chunks, but non-native speakers show a typical serial position curve, as if the items in the sentence were unrelated (Meara 1980, Spitze and Fischer 1981).

3. Forgetting is exponential in form (Pimsleur 1967).

4. Long-term memory can last for years, but is subject to interference from new material. Bahrick (1984) writes about another stage of memory ('permastore') which is immune to forgetting through interference.

5. Recognition and production vocabulary differ in the following ways: a) rehearsal improves recognition but not production; repeated presentation improves both scores, b) recognition usually produces higher scores, since the amount of information required for a correct response is lower, c) lower levels of learning suffice for recognition, d) with high levels of learning, production is more sensitive, and with low levels of learning, recognition is more sensitive, e) intention to learn has more effect on production than on recognition (Tversky 1973, J. Brown 1976, Wingfield and. Byrnes 1981, Eysenck 1982).

6. Beginnings and ends of words are easier to recall than middles (Brown and McNeill 1966, Kohn et al. 1987).

7. The learning situation in which the word was encoded affects the retrieval of the word (Jenkins 1974, Craik 1979, Cohen 1986, Weltens 1987). Memory is better when ego and emotions are involved (Stevick 1976).

8. Depth of analysis seems to be an important factor. (Craik and Lockhart 1972, Wingfield and Byrnes 1981). Imagery and

pictures can lead to a more meaningful analysis of vocabulary (Paivio 1976).

9. Word frequency has a greater effect on semantic memory, and emotionality and imagery have a greater effect on episodic memory (Murray 1986).

10. Cognitive theory predicts fairly rapid attrition of foreign language vocabulary, since it is declarative knowledge (O'Malley et al. 1987).

THE HYPOTHESES

In this section, the hypotheses for the present research are examined in the light of the literature reviewed in Chapter 2. Several of the hypotheses are supported by the literature, but other hypotheses are more controversial. Gaps in the literature are noted--gaps which the current research seeks to begin to fill.

Hypotheses about students

1. Students who study in a target-language environment retain more vocabulary than those who study in a native-language environment, even when the specific vocabulary learned is not rehearsed or used outside of class.

   None of the literature reviewed in this chapter deals with this question.

2. Students who have pre-course contact with Spanish will retain more vocabulary than those who have no previous contact.

   None of the literature reviewed in this chapter deals specifically with this question.

3. Students who have post-course contact with Spanish will retain more vocabulary than those who have no further

contact, even when the specific vocabulary learned is not rehearsed or used.

Several studies cited in this chapter concluded that post-course contact with the foreign language was important for retention (Edwards 1976, 1977, Lowe 1982, Gardner 1982, Robison 1985, Clark and Jorden 1984). However, all of these studies assumed that students had the chance to hear and practice many of the items for which they were being tested, amd thus the items were reinforced through time. The present research seeks to eliminate target vocabulary with which the student has contact after the termination of the course.

4. Vocabulary is retained at different rates according to sex.

Pratella (1969) concluded that sex was not an important factor in loss over the summer vacation.

5. Students who report more effort while taking a language course retain more vocabulary than those who report less effort.

None of the literature cited in this chapter deals with this question.

6. Students who report more intrinsic motivation for the course retain more vocabulary than students who report more extrinsic motivation.

Kennedy (1932) reported higher retention rates for students who had a desire to continue their study of Latin. Edwards (1976, 1977) and Gardner (1985) indicate that attitudinal/motivational factors were not related to retention. This question has not yet been resolved.

7. <u>Students who receive higher grades for the course retain more vocabulary than students who receive lower grades</u>.

Bahrick (1984) and Godsall-Myers (1981) specifically note that higher grades were related to higher retention rates. However, Clark and Jorden (1984) did not find differences between 'attriting' and 'non-attriting' students according to the ease with which they originally learned Japanese.

8. <u>Students who receive higher scores on the first retention tests continue to receive higher scores on later retention tests</u>.

The only researcher who retested the same subjects was Meara (1989), but he gave no data on the relationship of the first retention scores to the later retention scores. He suggests that the words retained from Week 2 to Week 3 are a better predictor of future scores and the final equilibrium state than words retained from Week 1 to Week 2.

Hypotheses about words

9. <u>Vocabulary that is repeated more frequently in teacher classroom input is retained at higher levels than vocabulary that is repeated less frequently</u>.

Godsall-Myers (1981) concluded that attrition reflects the frequency of occurrence of L2 patterns. Murray (1986) noted that translation efficiency from English to French was influenced by word frequency. Nation (1982) believes that words should be encountered seven times in order to be remembered (and sixteen times for incidental learning). Word frequency seems to have a greater effect on semantic memory than on episodic memory, in which a single vivid encounter with a word may fix it in memory (Schouten-Van Parreren 1989:82).

10. <u>Verbs, adjectives, and nouns are retained at different rates</u>.

Obler (1982) has shown that in aphasia and in normal aging, nouns are harder to recall than verbs. Nation (1982) states that nouns are easier to learn than adjectives, adverbs or verbs, and Cohen (1986) makes the point that nouns may be easier to forget because they are easier to learn, since they have not passed through as much processing. The relationship of part of speech to retention merits further investigation.

11. <u>Vocabulary learned earlier in the course will be retained at higher rates than vocabulary learned later in the course</u>.

Although the regression hypothesis (first-learned, last forgotten) does not appear to be true in the loss of a first language (Caramazza and Zurif 1978, Berko-Gleason 1982), several researchers report finding cases of regression in the loss of a second language (Cohen 1974 and 1975, Hansen 1980, Godsall-Myers 1981, Moorcroft and Gardner 1987, Olshtain 1983, 1986, 1989, Jordens et al. 1989). However, more work has been done in morphology and syntax than in vocabulary.

12. <u>Vocabulary with high emotionality or saliency ratings is retained longer than vocabulary with low ratings</u>.

Emotionality is most important for retention in episodic memory (Murray 1986). Several writers believe that emotionally loaded words are easier to remember than neutral words (Stevick 1976, Ludwig 1984, Schouten-Van Parreren 1989). None of the literature reviewed tested for effects of saliency.

Hypotheses about the relationship between recognition and
production vocabulary

13. Recognition scores for vocabulary are higher than
production scores.

In general, production skills suffer more attrition than
recognition skills (Scherer 1957, Edwards 1976 and 1977,
Lowe 1983, Bahrick 1984, Robison 1985). Bahrick (1984)
states that attrition affects smaller portions of
recognition vocabulary than recall vocabulary. Several
writers on human memory mention that recognition produces
higher scores (Tversky 1973, J. Brown 1976, Wingfield and
Byrnes 1981).

14. Both recognition and production scores for vocabulary
drop quickly at first and then gradually level off.

Pimsleur (1967) states that forgetting is exponential in
form. Bahrick (1984a,b,c) concludes that retention rates
drop for up to six years and then level off. For Weltens et
al. (1989), most loss occurred during the first two years
and then leveled off. Meara's matrix model of acquisition
(1989) predicts that remembering and forgetting will
eventually reach an equilibrium. None of the vocabulary
studies presents data from measuring the same learners
several times.

15. <u>The best description of recognition and production</u>
<u>vocabulary loss is that equal quantities of words are lost</u>
<u>over time rather than that the ratio of production to</u>
<u>recognition vocabulary remains constant over time</u>.

Bahrick (1984) believes that the absolute amount of
attrition in production and recognition vocabulary is
approximately the same. Weltens et al. (1989) also believe
that a fixed amount of material is lost rather than a fixed
proportion, although they do not refer specifically to
vocabulary. Other writers mention size of these recognition
and production vocabularies in terms of proportions; for
example, Melka Teicroew (1982) cites reports indicating that
for beginning students, recognition vocabulary is about
twice the size of production vocabulary. This question has
not yet been resolved.

16. <u>Smaller percentages of recognition vocabulary are lost</u>
<u>than of production vocabulary</u>.

If recognition vocabulary is larger than production
vocabulary and equal amounts are lost from both, it follows
that the percentage loss from recognition will be smaller
(see Bahrick 1984c).

17. <u>Vocabulary taught and tested for both recognition and</u>
<u>production is retained at higher rates than vocabulary</u>
<u>taught and tested for recognition only</u>.

Brown and Mc Neill (1966) believe that the need to produce a word promotes storage of the word in full. Tversky (1973) states that scores may be dependent upon the subjects' knowing which mode they will be tested in. None of the literature reviewed in this chapter specifically tested these two different modes for teaching and learning vocabulary.

18. <u>Ability to produce a word initially, rather than ability to recognize a word, is the best predictor of both recognition and production later on</u>.

None of the literature reviewed in Chapter 2 specifically tested for this hypothesis.

In conclusion, most literature on foreign language attrition is very recent, and the study of foreign language vocabulary is only beginning to attract the attention it deserves. Both of these areas are related to the study of human memory, including ideas on levels of processing and cognitive theory. There are still many questions about the role of vocabulary in language attrition, questions which the present research can play a part in answering.

CHAPTER 3

METHODOLOGY

The purpose of this study is to investigate the retention of foreign language vocabulary in college-level students over varying intervals of time. Several characteristics of the students and the learning environment are studied, as well as characteristics of the words themselves. Special attention is paid to the relationship of recognition vocabulary to production vocabulary over the retention period. A brief overview of the methodology has already been given in Chapter 1, and the present chapter will describe in greater detail the 1) subjects, 2) materials and procedures, and 3) methods of analysis.

In this study a balance was sought between degree of control (in order to increase internal validity) and degree of approximation to the real-world classroom (in order to increase external validity and generalizability to other situations). If similar patterns can be found in differing learners and differing situations, then the results may be applicable to other learners in other real-world situations. In the previous chapter it was noted that writers such as Bahrick (1979) have called for research in ecologically realistic circumstances.

SUBJECTS

The fifteen subjects were students enrolled in two intensive introductory Spanish courses. The first course was taught at Calvin College in Grand Rapids, Michigan, from May 31 to June 22, 1988, and the second course, composed of different students, was taught in Merida in the Yucatan Peninsula of Mexico from June 26 to July 18 of the same summer.

The nine students who enrolled in the Calvin course did so for very different reasons, as indicated by a questionnaire that they were given the first day of class. Six students were in the class because of graduation requirements at Calvin, although two of these six also indicated a special interest in the Spanish language. A seventh student was hoping to go to Spain (though his plans changed midway through the course), an eighth student believed that Spanish would help her in her nursing career, and the ninth student (a 16-year old high school student) mentioned his interest in Latin America. Except for this 16-year-old student, the subjects in this class ranged in age from 19 to 22.

These nine students also varied widely in their backgrounds of foreign language study. Latin and French were two of the languages that some of these students had studied in high school. Three of the students had already attempted Spanish at Calvin College, had dropped out because

of low grades, and were repeating the course. Final grades
in the summer course at Calvin ranged from A to D.

The six students who enrolled in the course in Merida
were involved in the STS (Summer Training Session in
Missions) program in Mexico, conducted for twenty years by
the Reformed Bible College in Grand Rapids and presently
associated with IDEA ministries in Grand Rapids. In terms
of motivation, this was a much more homogeneous group. All
of the students mentioned an interest in missions, as well
as an interest in Latin America and its people. However,
their foreign language background was widely varied: two had
never studied a foreign language while others had studied
German, French, or Dutch. Only one subject had studied
Spanish previously, and for only one semester. One subject
had already graduated from college and was teaching in a
school in south Texas with many Spanish-speaking children.
The ages of these subjects ranged from 20 to 25. Final
grades in this class were high, ranging from A to B-.

Since this was a retention study, follow-up
questionnaires asked the students to evaluate the course and
their effort in it and to detail further contact with
Spanish, either formally in the classroom or informally in
the culture. Two of the students in the Calvin class had no
further contact with Spanish, although most went on to take
the second semester of first year (another intensive summer
course) and then on to second-year Spanish during the
regular school year. The students in the Merida class lived

with Mexican families in pairs while taking the class and
then moved on in pairs to 3-week field assignments in
Mexican or Central American villages and heard and used
varying degrees of Spanish, depending on the principal
language spoken in the village, whether Spanish or an
indigenous language (such as Maya or Tzeltal). Five of
these six students returned to the U.S. or Canada at the end
of August and took no further classes in Spanish, but one
remained in Mexico City to continue her study of Spanish and
eventually went to work in Honduras. Although it would have
been possible to eliminate this student from the study, in
the end she was included, in view of the fact that the
attrition of the 100 target words showed the same patterns
in her case as in the other students.

Of the total number, eight subjects were male and seven
were female. Tables 2 and 3 give a summary of the
characteristics of these subjects. The data file in Table
40 in Appendix B indicates how these characteristics were
coded for the statistical analysis and also includes the
scores for the subjects on each of the recognition and
production tests. Although not all of the subjects
responded to each of the four retestings, all participated
in the final retest; thus, there were no drop-outs.

TABLE 2
SUBJECT CHARACTERISTICS

SUBJECT	SEX	AGE	LOCATION OF CLASS	MOTIVATION	EFFORT	COURSE GRADE
01	F	20	Mexico	interest in missions	much	B+
02	M	21	Mexico	interest in missions, Latin America	some	A
03	F	24	Mexico	interest in missions, Latin America	much	A
04	F	21	Mexico	interest in missions, Latin America	much	B–
05	F	25	Mexico	interest in missions, Latin America	much	A
06	M	24	Mexico	interest in missions, Latin America	some	B+
07	M	22	U.S.	graduation requirement, career in communications	some	D
08	M	21	U.S.	graduation requirement	much	B–
09	M	20	U.S.	graduation requirement	much	B–
10	M	19	U.S.	graduation requirement, interest in Latin America	some	C
11	F	20	U.S.	graduation requirement	much	C
12	F	22	U.S.	graduation requirement	some	C
13	M	22	U.S.	plans for travel to Spain, though later the plans fell through	some	B+
14	F	20	U.S.	career in nursing	much	A–
15	M	16	U.S.	interest in Latin America	much	A

TABLE 3
SUBJECT CHARACTERISTICS: PREVIOUS AND LATER CONTACT

Subject	Previous language study	Later contact with Spanish
01	2 years junior high French	some in Mexico, though stayed in village where Tzeltal was spoken; no further formal study
02	2 years high school German, 1 semester college German, 2 years college Greek	4 weeks with Spanish-speaking family; no further formal study
03	1 semester college Spanish, 2 years college Dutch, 10-week visit to Peru	3 weeks in Spanish-speaking village; no further formal study
04	2 years high school Latin	further language study and stay in Honduras until present
05	short trips to Mexico, teacher in bilingual school	3 weeks in Spanish-speaking village; no further formal study; teacher in bilingual school
06	none	3 weeks in Mexico, though stayed in village where Maya was spoken
07	1 semester high school Spanish	3 semesters college Spanish
08	1 year college Spanish	3 semesters college Spanish
09	1 semester college Spanish	3 semesters college Spanish
10	2 years high school Latin	2 semesters college Spanish
11	4 years high school Latin	none
12	2 years high school Latin, 1 year college Spanish	2 semesters college Spanish
13	1 year college Latin	none
14	3 years high school French	1 semester college Spanish
15	4 hours Spanish (Total Physical Response)	3rd year high school Spanish; talked with migrant workers

As a control to see what kind of scores would result from a class of Spanish students who had not studied this specific material, the same recognition and production test was given to a class of 24 students at Calvin College. These 102-level students had been taught Spanish by an entirely different approach with a different text (Communicating in Spanish by Lamadrid el al. 1984) and had encountered only two of the target words (nido and cartel) in their course work. Their mean score on recognition was 1.6 (out of 100 words) and their mean score on production was .2 (also out of 100), scores far below those of the students in the study, even 18 months after the course had ended.

MATERIALS AND PROCEDURES

Both the course taught in Grand Rapids and the course taught in Merida used exactly the same methods, text, and additional materials. The courses were taught using the 'Natural Approach' with the text Dos mundos (1986) by Terrell et al. Classes met for three hours each morning for a period of seventeen days in Grand Rapids and for sixteen days in Mexico. Both classes took the same period of time (twenty hours) to cover Pasos A-E (the period intensively studied in this investigation), although the Grand Rapids class was able to go through Capítulo 5 in the text because

of the extra day, while the Merida class completed only the material through Capítulo 4.

The researcher was the instructor for both classes and used the same lesson plans for both classes, thus insuring uniformity in classroom procedure. The language of instruction for both classes was Spanish.

On the first day of class all students were given a questionnaire asking for name, age, class rank, address, previous contact with Spanish or other foreign languages, motivation for studying Spanish, and what the student hoped to learn in the class. This questionnaire was handed in to the instructor and the form is included in Appendix A. Students also filled out a consent form which asked for permission to use classroom data for a study on vocabulary retention, although the exact nature of the study was not specified. This consent form is also included in Appendix A. Students gave these consent forms to a third party, and the instructor did not see them until after she had given the final course grades. All of the students gave their consent as it turned out, and thus the researcher was able to include the data from all fifteen students.

New material was first presented in class orally by the teacher through teacher talk with pictures, objects, and commands. Students were asked to indicate comprehension with actions, by answering with sí or no, or by answering with single words or short sentences. After a new group of vocabulary words was introduced, the instructor wrote the

new words on the board and students copied them in their notebooks, often including their own English translation. Students were asked to study the text outside of class after the presentation of new material and to fill out their workbooks and worksheets. Homework assignments (including grammatical explanations) were discussed briefly in class the following day and sometimes collected. By the end of the first twenty hours, students were participating in widely-varying communicative activities and writing short compositions.

The material in Dos mundos is organized around language functions and communicative situations. Paso A contains classroom commands, names, adjectives to describe people, colors, clothing, numbers, parts of the body, and expressions for greeting and leave-taking. Paso B teaches informal and formal distinctions, classroom objects, more commands, and more adjectives for describing people. Paso C includes family, expressions for possession, telling time and telling age, and more numbers. Paso D teaches some weather expressions, expressions for origin and nationality, class subjects, and expressions for location. Paso E contains more numbers, addresses and phone numbers, dates and birthdays, and expressions for prices. All together these five pasos teach a total of about 650 new words or expressions. After finishing these preliminary pasos, the students begin the regular capítulos, which are set up in the same basic format as the preliminary pasos, but longer.

In addition to these 650 new words, the instructor chose another 100 words which formed the core of the investigation on retention. These 100 words were chosen because of their concreteness and their ability to be represented by pictures (all were nouns, verbs, or adjectives), their possibility of inclusion in the categories already being taught (in order to form networks of meanings and to make them less obvious), and their non-appearance in later chapters in Dos mundos or the second-year book used at Calvin, Charlemos by Jarest and Robinson (1986). Thus these 100 words would not be reviewed after the first twenty hours of class and could be used to measure the attrition rate. On subsequent retention tests students were asked to mark words with which they had had contact after the course, and when any of these 100 words were mentioned, they were eliminated from that student's data. Table 4 gives a list of these 100 words, the paso in which they were taught, their part of speech, their frequency in teacher talk, and whether they were originally taught and tested for production and recognition or for recognition only.

TABLE 4
WORD CHARACTERISTICS

WORD NO.	WORD	ENGLISH TRANSLATION	PASO	PART OF SPEECH	FRE-QUENCY IN INPUT	TAUGHT AND TESTED FOR REC. AND PROD.?
001	¡pellizque!	pinch!	A	verb		yes
002	¡ande de puntillas!	tiptoe!	A	verb		yes
003	¡rece!	pray!	A	verb		yes
004	¡pliegue!	fold!	A	verb	low	yes
005	¡tache!	cross out!	A	verb	low	yes
006	¡amarre!	tie!	A	verb	low	yes
007	cojo	lame	A	adj.	low	yes
008	tuerto	one-eyed	A	adj.	medium	yes
009	manso	gentle	A	adj.	high	yes
010	huraño	timid	A	adj.	high	yes
011	ocioso	lazy	A	adj.	high	yes
012	agudo	sharp,smart	A	adj.	high	yes
013	torpe	stupid	A	adj.	low	yes
014	delantal	apron	A	noun	medium	yes
015	faja	belt	A	noun	medium	yes
016	cremallera	zipper	A	noun	medium	yes
017	aretes	earrings	A	noun	medium	yes
018	tacón	heel	A	noun	low	yes
019	chaleco	vest	A	noun	medium	yes
020	alhajas	jewels	A	noun	medium	yes
021	meñique	little finger	A	noun		yes
022	tez	face	A	noun		yes
023	puño	fist	A	noun		yes
024	pulgar	thumb	A	noun		yes
025	panza	stomach	A	noun		yes
026	barbilla	chin	A	noun		yes
027	quijada	jaw	A	noun		yes
028	sien	temple	A	noun		yes
029	hocico	snout	A	noun	medium	yes
030	espinazo	back (of animal)	A	noun	low	yes
031	garra	claw	A	noun	low	yes
032	pescuezo	neck (of animal)	A	noun	medium	yes
033	mariposa	butterfly	A	noun	low	yes
034	ciervo	deer	A	noun	low	yes
035	venado	deer	A	noun	low	yes
036	ballena	whale	A	noun	low	yes
037	tiburón	shark	A	noun	medium	yes
038	¡exprima!	squeeze, wring out!	B	verb	very low	yes
039	¡guiñe!	wink!	B	verb	medium	yes
040	¡ronque!	snore!	B	verb	low	yes

TABLE 4 (cont'd)

041	¡silbe!	whistle!	B	verb	medium	yes
042	¡bostece!	yawn!	B	verb	low	yes
043	descalzo	barefoot	B	adj.	low	yes
044	calvo	bald	B	adj.	very low	yes
045	cano	white-haired	B	adj.	very low	yes
046	mohoso	moldy	B	adj.	low	yes
047	harapiento	ragged	B	adj.	very low	yes
048	arrugado	wrinkled	B	adj.	very low	yes
049	empapado	soaked	B	adj.	low	yes
050	maceta	flower pot	B	noun	low	yes
051	tabla	board	B	noun	low	yes
052	colcha	blanket	B	noun	low	yes
053	cuenco	bowl	B	noun	medium	yes
054	rótulo	sign	B	noun	low	yes
055	cartel	poster	B	noun	medium	yes
056	pasillo	hallway	B	noun	low	yes
057	huella	footprint	B	noun	medium	yes
058	fango	mud	B	noun	low	yes
059	ala	wing	B	noun	low	yes
060	rama	branch	B	noun	low	yes
061	mirlo	blackbird	B	noun	high	yes
062	nido	nest	B	noun	low	yes
063	antepasado	ancestor	C	noun	low	yes
064	vástago	descendant	C	noun	low	yes
065	bisabuelo	great-grand-father	C	noun	medium	yes
066	tatarabuelo	great-great-grandfather	C	noun	very low	yes
067	padrastro	stepfather	C	noun	very low	yes
068	madrastra	stepmother	C	noun	very low	yes
069	padrino	godfather	C	noun	very low	yes
070	madrina	godmother	C	noun	very low	yes
071	mellizo	twin	C	noun	low	yes
072	lágrima	tear	C	noun		yes
073	beca	scholarship	C	noun		yes
074	cesto	basket	C	noun	very low	yes
075	incendio	fire	C	noun		yes
076	llama	flame	C	noun		yes
077	cerilla	match	C	noun		yes
078	aguanieve	sleet	D	noun		no
079	estanque	pool	D	noun	medium	no
080	dehesa	meadow	D	noun	low	no
081	loma	hill	D	noun	low	no
082	cerro	hill	D	noun	low	no
083	vereda	path	D	noun	very low	no
084	sendero	path	D	noun	low	no
085	agujero	hole	D	noun	low	no
086	carpa	tent	D	noun	low	no
087	aldea	village	D	noun	very low	no
088	verja	window grating	D	noun	very low	no

TABLE 4 (cont'd)

089	peña	rock	D	noun	very low	no
090	bolívar	monetary unit (Venezuela)	E	noun	very low	no
091	rastrillo	rake	E	noun	very low	no
092	enano	dwarf	E	noun	very low	no
093	azulejo	tile	E	noun	low	no
094	navaja	blade	E	noun	low	no
095	cascote	garbage	E	noun	medium	no
096	hilo	thread	E	noun	low	no
097	semilla	seed	E	noun	very low	no
098	frasco	bottle	E	noun	medium	no
099	globo	balloon	E	noun	medium	no
100	aguja	needle	E	noun	low	no

KEY FOR FREQUENCY IN CLASSROOM INPUT

very low: 1-4 repetitions
low: 5-9 repetitions
medium: 10-19 repetitions
high: 20+ repetitions

These 100 words were taught just like the other vocabulary words in that they were introduced through teacher talk and the use of commands, pictures, and realia. Like the other vocabulary words, they were written on the board and copied by students into vocabulary notebooks, and since students were asked to study vocabulary from their notebooks, it was not readily apparent that these words were not included in the text. Since students received additional practice with the regular vocabulary words in their workbooks, extra worksheets were assigned to give students written practice with these 100 words, which were mixed in with the regular vocabulary words.

Tests were given after Paso A, Pasos B-C, and Pasos D-E. The greater part of each test came directly from the testing manual provided with <u>Dos mundos</u>, but items testing the extra 100 vocabulary words were mixed in with the other items, and students were called in individually for oral interviews after each of these tests. In addition, two quizzes (not in the <u>Dos mundos</u> testing program) were written to test production of vocabulary during the first 20 hours of the course. The two production quizzes and sample items from the tests are included in Appendix A.

The tests provided with <u>Dos mundos</u> are mainly multiple choice tests of vocabulary and grammar, thus asking the student to recognize but not to produce the words and forms. Because of the emphasis on recognition, the authors found it possible to include the large number of 650 vocabulary words

in the first five pasos of the text. Suggestions are also given in the testing manual for compositions and oral interviews in which the students are asked to produce words and forms, though the students always have a choice about what to say. When the researcher added additional parts to the tests and wrote the two extra quizzes, she insured that each of the 100 additional words in Pasos A-E would be tested at least once for recognition during the first twenty hours. After students had been tested for recognition on a test or quiz, the next quiz or test would ask them for the production of this same vocabulary, and students were asked to study with production in mind. Because production tests lagged behind the recognition tests, students were not tested for production of the extra vocabulary in Pasos D and E and thus did not study for the production of this vocabulary either. This difference in testing during the first twenty hours shows up in Table 4 in the column headed 'Taught and tested for recognition and production', where some words were learned for both recognition and production and other words for recognition only.

The data collected during the first twenty hours of class during these tests and quizzes provided the information for setting a baseline for each student: whether the student at one point in time could recognize and produce a word. After the first twenty hours of class, the students did not see or hear these 100 words again in class, and they were told to study only the vocabulary lists in

their books (which did not contain the 100 target words) for the final exam. However, at the end of the course, the students unexpectedly received their first retention test. The first part of the retention test consisted of pictures of the 100 words followed by blanks for writing the corresponding words. The pictures had been used many times in the past, but the teacher was also available to identify in English any picture that a student could not interpret. Pictures of other familiar words that the students had learned were also included so that it was impossible for the subjects to determine exactly which words the researcher was interested in. When students had completed as much of this part as possible, they turned it in to the instructor and took the second part, which was identical to the first part, except that a long list of Spanish words was provided for the students to choose from. Students needed to match the letter of the word to the correct picture, thus performing a recognition test. All 100 target words were listed and pictured, but several other pictures and words were also included as distractors. Although the researcher recorded the amount of time each student took to complete each part, no time limit was imposed, since the goal was to learn if the student had retained each word and not to force a quick guess (for a precedent, see Bahrick 1984b:3). On the matching test students were also asked to indicate when they were guessing by starring the word, thus indicating a lack of confidence in their decision. Although this information

about confidence level was not analyzed statistically, it will be mentioned again in Chapter 4.

Pictures, instead of English translations, were used on the retention test because of the nature of the classroom experience in which pictures, realia, and commands had been used to teach the words. Students were also familiar with these pictures because of the worksheets and tests they had completed earlier. Also, the literature reviewed in Chapter 2 of this work indicates that words are more easily learned and retained when accompanied by pictures (Kellogg and Howe 1971, Champagnol 1972, Paivio 1976:120).

The research design in which each subject is tested first for recall and then immediately for recognition of the same vocabulary is mentioned in Chapter 2 of this dissertation. J. Brown (1976:34) has stated that 'This greatly increases the precision of the experiment since recall and recognition are then tested on the same subjects, on the same material and on the same occasions'.

The second retention test was given to the subjects approximately three-and-a-half weeks later. The form for the test was the same: first a production test and then a recognition test over the same words. The test in Grand Rapids was administered by a different instructor at the end of Spanish 102 (second semester of first year), and the test in Mexico was administered by the researcher in Mexico City when the students met again after their field assignments in Mexican and Central American villages.

The third retention test was taken by the subjects in April and May of 1989, approximately nine months after the previous test. The format was the same, except that several different words and pictures had been mixed in with the 100 target words. The students were widely dispersed geographically, and for this reason the tests were mailed to the students with explicit instructions about how to complete them. By this time the subjects were quite familiar with the protocol for taking the tests, but in case a student did not recognize a picture, a separate sheet of English words for the pictures was included.

The fourth and final retention test was given in November and December of 1989, approximately eight months after the previous test. The format of the test was the same as in the spring, but this time the researcher was able to personally test twelve of the fifteen subjects. She still needed to rely on mailed forms for the three subjects living in Texas, Minnesota, and Illinois. A copy of the retention test (both production and recognition parts) is included in Appendix A.

Although testing all 100 words (along with other words) in each of the four retention tests leaves open the possibility that students retain more <u>because</u> of the tests, Berko Gleason (1982:22) sets a precedent by suggesting that longitudinal studies be conducted using the same measure. Because of the small number of subjects in this study, the researcher felt that more accurate data could be obtained by

using approximately the same test each time. However, the possibility remains that some subjects obtained higher scores because of the effect of previous tests, in spite of the relatively long periods of time between tests.

All retention tests were scored using the same criteria. On the recognition test an item was either correct or not correct, and one point was awarded for a correct answer. On the production test it was felt that partial answers were important in tracing the attrition pattern, so partial credit was given for some answers. The precedent for giving partial credit is found in Bahrick (1984b:5), and Ginsberg (1986:23,183). On the production test, the researcher gave one point for a totally correct word or for a word that was correct except for an accent mark or for a spelling that would not change the pronunciation of the word (at least in Latin American pronunciation). One-half point was given for a word that could not meet these criteria, but that had at least half of the correct letters. For example, if the correct word was azulejo (tile), asulejo received one point, azuljes received one-half point, and azor received no credit. Raw scores were converted to percentages, thus allowing the researcher to obtain a score even after having deleted those words for which the subject indicated subsequent contact. If a subject indicated that he had had contact with a specific word but still got it wrong on the test, the word was not deleted from the data.

In addition to the regular tests and quizzes during the first twenty hours of the course and the later retention tests, the researcher obtained data using other methods. During class a tape recorder was left on so that the researcher could record her input to the students. These recordings were used later to determine the number of times certain words were repeated in the input. The tape recorder was left on even after completion of the first twenty hours of the course to minimize clues about the research design. Written compositions collected by the teacher were scrutinized for any of the 100 target words that the subjects had used spontaneously. Also oral interviews were conducted at the end of 101 and after the final retention tests. During both interviews subjects were asked to describe a picture that contained several of the target words (see Figure 22, Appendix A), and they were also asked to do some introspection about their vocabulary learning and retention. The precedent for think-aloud procedures is found in Hosenfeld (1977), Cohen and Hosenfeld (1981), and in Cohen (1984). Students were asked how they studied vocabulary, if they mentally translated to English, and how they remembered words. After the final retention test they were asked which devices they used to remember specific words.

The 100 target words were also rated for emotionality and saliency to see if there was any relationship between these qualities and retention. Brown and Ure (1969)

published a list of 650 English words with ratings on a seven-point scale for goodness, pleasantness, emotionality, concreteness, and associative difficulty. This same idea was used to create a five-point scale in the present study for rating words on emotionality and saliency. Classes of Spanish students who were not involved in the present study were asked to rate the words.

The rating for emotionality was done by a class of 30 students with the target words in English, since the target words in Spanish were unknown to the students. The following instructions come from Brown and Ure (1969:234): 'Decide just how strongly emotional the meaning is. Notice that the kind of emotion is not relevant; only the intensity of emotion matters.' A rating of 1 indicated high intensity and a rating of 5 indicated low intensity. A different class of 26 students rated these same English words on saliency, with a definition of saliency coming from Webster's dictionary. The specific instructions read: 'Decide just how salient (noticeable, conspicuous) the meanings of these words are. How easy would they be to remember if you had to learn them in Spanish?' Students again rated these words from 1 to 5.

The third class of 25 students received a list of the Spanish words with the English translations in parentheses. They were again asked to rate the words on saliency, but this time the instructions were slightly different: 'Decide just how salient (noticeable, conspicuous) the forms of the

Spanish words are. How easy would they be to remember if you had to learn them?' Ratings again were from 1 to 5.

The results of the one rating for emotionality and the two ratings for saliency are found in Table 41 in Appendix B, along with the numbers for recoding words according to frequency in input, paso in which they were learned, part of speech, and whether they were taught and tested for both recognition and production. Table 41 also contains the percentage scores for each of the words on each of the recognition and production tests over time.

ANALYSIS

As can be seen from Appendix B (Tables 40 and 41), the variables in this study can be organized in two ways, by student and by word. When the data are organized by student, n equals 15, and when the data are organized by word, n is equal to 100. Both data files deal with recognition and production, and this distinction can be set apart as a third area of analysis. The research hypotheses are also organized around these three areas (dealing with students, with words, and with recognition/production), and these three areas thus form the basis for the organization of results in Chapter 4. The research hypotheses are listed below and the method of investigating each one is described.

The recognition and production tests were scored by hand, and data and control files were created using SPSS (Statistical Package for the Social Sciences), available at Calvin College. Where statistical analysis (such as a T-test, One-way analysis of variance, or Anova) seems appropriate, the hypotheses are restated in null form and a probability level of .05 or lower is cause for rejecting a null hypothesis. T-tests are independent rather than dependent, since the subjects are in only one group at a time. One-tailed T-tests are used when the hypotheses are directional, and two-tailed T-tests are used when the hypotheses are nondirectional.

For some hypotheses, a correlation or a graph seems to be the most effective way of studying the problem. In addition, individual case studies make important points that are missed when all the data are lumped together for statistical purposes. Thus, in Chapter 4 on the findings, the questions are looked at from varying perspectives to get a more rounded and more complete picture of the data. One must also remember that patterns are important. If one test is significant at the .05 level, but other similar ones are not, the overall pattern clearly overrides the individual test. This is especially important in the cases where a series of statistical tests are used to analyze recognition and production tests over time. A conservative and careful approach to results should be taken, and results should not be overinterpreted.

Hypotheses about students

independent variables: classroom location (Mexico or U.S.),
pre- and post-course contact with Spanish, sex, effort,
motivation, final grade in course

dependent variables: ten scores per student on a series of
recognition and production tests over time

other data: initial and final questionnaires (written),
interviews at end of 101 and at end of 18-month retention
period

Hypothesis 1: Students who study in a target-language
environment retain more vocabulary than those who study in a
native-language environment, even when the specific
vocabulary is not rehearsed or used. The corresponding null
hypothesis is that there is no difference between the two
groups. In order to test this null hypothesis, an
independent one-tailed T-test was used to test for
significant differences between the Mexico group and the
U.S. group on their recognition and production scores at the
different points in time.

Hypothesis 2: Students who have pre-course contact with
Spanish (either informally in the culture or formally in the
classroom) will retain more vocabulary than those who have

no previous contact. The corresponding null hypothesis is
that there is no difference between the students who have
pre-course contact and the students who have no pre-course
contact. In order to test this null hypothesis, an
independent one-tailed T-test was used to test for
significant differences between these two groups based on
their test scores.

Hypothesis 3: Students who have post-course contact with
Spanish (either informally in the culture or formally in the
classroom) will retain more vocabulary than those who have
no further contact, even when the specific vocabulary
learned is not rehearsed or used. The corresponding null
hypothesis is that there is no difference between the
students who do and do not have post-course contact. For
this analysis the students were divided into four small
groups depending on type of post-course contact and a One-
way analysis of variance was run on the test scores.

Hypothesis 4: Vocabulary is retained at different rates
according to sex. The corresponding null hypothesis is that
there is no difference between the sexes in retention of
vocabulary. Since this hypothesis is nondirectional, an
independent two-tailed T-test was used to test for
significant differences in the test scores between groups
based on sex.

Hypothesis 5: Students who report more effort while taking a language course retain more vocabulary than those who report less effort. The corresponding null hypothesis is that there is no difference based on effort. In order to test the null hypothesis, the students were divided into two groups based on the amount of effort they reported and an independent one-tailed T-test was run on scores of these two groups.

Hypothesis 6: Students who report more intrinsic motivation for the course retain more vocabulary than those who report more extrinsic motivation. The corresponding null hypothesis is that there is no difference based on motivation. In order to test the null hypothesis, the students were divided into three groups based on the type of motivation they reported (1 for intrinsic motivation, 2 for mixed motivation, and 3 for extrinsic motivation), and a One-way analysis of variance was performed on test scores to look for significant differences between the groups. An independent one-tailed T-test was also performed to look for significant differences between the most extreme groups (groups 1 and 3). In order to examine the role of effort and motivation from a slightly different perspective, the rating for effort and the rating for motivation were added for each student to obtain ordinal data, and a nonparametric correlation was performed with test scores to determine Spearman's rank correlation coefficient.

Hypothesis 7: Students who receive higher grades for the course retain more vocabulary than students who receive lower grades. The corresponding null hypothesis is that there is no difference in scores based on final course grade. In order to test the null hypothesis, the students were divided into two groups (those with grades ranging from A to B+ and those with grades ranging from B- to D) and an independent one-tailed T-test was performed on the scores in order to test for differences between the groups. The relationship between the two groups was also graphed.

Hypothesis 8: Students who receive higher scores on the first retention tests continue to receive higher scores on later retention tests. The corresponding null hypothesis is that there is no systematic relationship between retention scores on earlier tests and scores on later tests. This hypothesis was examined by performing a Pearson product-moment correlation with test scores and looking for strong correlations among earlier and later recognition and production scores.

In addition to performing these statistical tests, graphs of the scores for each of the fifteen students were constructed, as well as a composite graph. Individual cases were discussed in relation to the composite graph.

Hypotheses about words

independent variables: frequency of words in classroom
input, part of speech, words taught earlier or later,
emotionality and saliency of words

dependent variables: ten scores for each of 100 words

other data: interviews by students at end of 101 and at end
of 18-month retention period

Hypothesis 9: Vocabulary that is repeated more frequently
in teacher classroom input is retained at higher levels than
vocabulary that is repeated less frequently. The
corresponding null hypothesis is that there is no difference
in scores between vocabulary that is repeated more
frequently and vocabulary that is repeated less frequently
in classroom input. In order to test the null hypothesis,
the words were divided into two groups according to
frequency in input, and an independent one-tailed T-test was
run on the scores to test for significant differences. In
addition, since the frequencies could also be represented by
ordinal data, a nonparametric correlation was run between
frequencies recoded on a scale of 2 to 8 and test scores.

Hypothesis 10: Verbs, adjectives, and nouns are retained at
different rates. The corresponding null hypothesis is that

there is no difference in the retention rates of verbs, adjectives, and nouns. In order to test the null hypothesis, a One-way analysis of variance was conducted on all the words in Pasos A-E by part of speech (with the Tukey option to tell where any significant differences occurred). Next, the One-way analysis of variance was repeated on just the words in Pasos A and B, since these were the only pasos in which all three parts of speech were taught. A graph was also constructed to illustrate these relationships.

In order to study the combined effects of frequency in input, part of speech, and whether a word was taught and tested for both recognition and production or for recognition only, an Anova was run on the scores by these three factors. The size of the F ratio (the ratio of 'between group variance' to 'within group variance') is given, as well as the probability of obtaining these results only by chance. The Anova also gives an idea of the relative contribution of each of these factors to the scores.

Hypothesis 11: Vocabulary learned earlier will be retained at higher rates. The corresponding null hypothesis is that there is no difference in scores between vocabulary learned earlier and vocabulary learned later. This null hypothesis was tested by doing a One-way analysis of variance by paso and including the Tukey option in order to discover where significant differences occurred.

Hypothesis 12: Vocabulary with high emotionality or saliency ratings is retained longer than vocabulary with low ratings. The corresponding null hypothesis is that there is no difference in scores between vocabulary with high emotionality or saliency ratings and vocabulary with low ratings.

For a preliminary analysis, Pearson product-moment correlations were determined between the scores and the one rating for emotionality and each of the two ratings for saliency. The Pearson correlation was chosen because all of these data were interval scaled. The rating for emotionality and the rating for the saliency of the words in English showed practically no correlation with the scores, but the rating for saliency of the forms of the Spanish words showed a weak correlation with test scores.

Based on this preliminary analysis, the words were divided into groups of high and low saliency according to the third scale, and an independent one-tailed T-test was run on the scores to test for significant differences between the two groups. The 21 best-retained words and 24 worst-retained words also formed two separate groups, and an independent one-tailed T-test was performed on their saliency ratings in order to look for significant differences between these two groups.

As mentioned in connection with Hypothesis 12 above, the best-retained words and the worst-retained words were

identified (according to their scores on the final retention tests). These words were further examined by summarizing comments made about them by the students after their final retention tests. Also a brief discussion of networks of meanings is included.

Hypotheses about the relationship between recognition and production vocabulary

independent variable: whether vocabulary was taught and tested for both recognition and production or for recognition only

dependent variables: ten scores for each of 15 students and for each of 100 words

other data: written answers on test forms for each of 15 students

Hypothesis 13: Recognition scores for vocabulary are higher than production scores.

Hypothesis 14: Both recognition and production scores for vocabulary drop quickly at first and then gradually level off.

Hypothesis 15: The best description of recognition and production vocabulary loss is that equal quantities of words are lost over time rather than that the ratio of production to recognition vocabulary remains constant over time.

Hypothesis 16: Smaller percentages of recognition vocabulary are lost than of production vocabulary.

Hypotheses 13-16 are best examined by constructing and studying a graph of the means for recognition and production vocabulary over time.

Hypothesis 17: Vocabulary taught and tested for both recognition and production is retained at a higher rate than vocabulary taught and tested for recognition only. The corresponding null hypothesis is that there is no difference in the scores of vocabulary taught and tested in the two different ways. The null hypothesis was tested by running an independent one-tailed T-test for a significant difference in the scores between words from Pasos A-C (words taught and tested for both recognition and production) and words from Pasos D-E (words taught and tested for recognition only).

The Anova test included under the hypotheses about words is also reexamined, and another Anova is run to test the combined effects of saliency and whether a word is taught and tested for both recognition and production.

Hypothesis 18: Ability to produce a word initially, rather than ability to recognize a word, is the best predictor of both recognition and production later on.

This hypothesis was examined from three different perspectives. In the first test a null hypothesis was formulated stating that words that are recognized and produced at high levels initially will not have higher scores on later retention tests than words that are initially recognized at high levels but produced at lower levels. Words were identified that would fit into these two initial groups, and an independent one-tailed T-test was run on their scores for later retention tests.

The second perspective included a Pearson product-moment correlation for test scores to look for strong correlations between recognition and production scores.

The third perspective included a multivariate analysis run at Michigan State University to determine the best predictors at Time 1 for retention at Time 4. Multiple correlation coefficients are also given.

Finally the data from one student are examined closely to follow the attrition process in detail. Interesting features such as the process of forgetting, confidence level, repetition of errors, and the 'tip-of-the-tongue' phenomenon are briefly discussed.

CHAPTER 4

ANALYSIS AND FINDINGS

INTRODUCTION

This chapter presents the findings of the research
described in the previous chapter. Results are shown for
statistical tests, but individual cases are also described
in order to give a more complete perspective on the data.

The chapter is divided into three sections based on the
division of the research hypotheses into three groups:
analysis of the students, analysis of the words, and
analysis of recognition and production of vocabulary words.
The conclusions for the hypotheses are given within each
section, along with some discussion. An over-all perspective
on the meaning of the results is saved for the summary and
conclusion in Chapter 5.

ANALYSIS BY STUDENT

Although the students were grouped in several different
ways in order to study the effect of various independent
variables on retention rates over time, few important
overall patterns emerged. The small sample size of only

fifteen students may have impeded the emergence of more overall patterns, and indeed differences between groups would need to have been quite large in order to reveal statistical significance at the .05 level. Among the fifteen students involved in this study, there was great variation in total amount of material retained, though some similarities in the pattern of forgetting were observed. The statistical results for the analysis by student are given in this section, but it seems that the most interesting results come from studying individual cases.

Statistical study

First the students were separated into the two groups representing the two different classroom locations, one in Mexico and the other in the United States. Since the researcher eliminated words with which students had contact outside the classroom, she did not expect to see large differences between the groups. However, the fact that there was no difference between the two groups over time merits further discussion. It seems that in the group in the United States there were more students who had difficulties with the class, but there were also some very good students who offset the lower scores (notice in Table 5 that the standard deviation of the U.S. group is higher than that of the Mexico group). Also it seems that the students in the United States had more time to concentrate on

specific course content, while the students in Merida living with Mexican families had a broader communicative goal in mind than simply studying specific course vocabulary. Students in Merida also reported that it was hard to find a time and place for concentrated study. In Table 5 note that R stands for a recognition test and P stands for a production test. The number 0 represents the original baseline tests during the first twenty hours of class, and the number 1 represents the first retention tests given at the end of the three-and-a-half week course. Time 2 is approximately a month after Time 1, Time 3 is approximately 9 months after the previous test, and the final test at Time 4 is another eight months later (see testing schedule in Table 1 on page 20). Thus the ten different tests for recognition and production are labeled R0, P0, R1, P1, R2, P2, R3, P3, R4, and P4.

TABLE 5
T-TESTS FOR GROUPS IN MEXICO AND U.S.

TEST	N	MEAN	SD	T VALUE	1-TAIL PROB
RO MEXICO	6	80.3	11.5	.11	.46
US	9	79.7	14.1		
PO MEXICO	6	56.0	15.4	-.19	.43
US	9	58.2	25.3		
R1 MEXICO	6	69.5	18.3	.40	.35
US	8	65.0	22.4		
P1 MEXICO	6	32.0	21.3	-.81	.22
US	8	42.3	24.8		
R2 MEXICO	6	56.3	25.5	-.30	.38
US	7	60.3	21.9		
P2 MEXICO	6	26.7	17.0	-.46	.33
US	7	30.9	22.6		
R3 MEXICO	5	34.6	25.2	-.37	.18
US	7	40.4	27.5		
P3 MEXICO	5	7.4	7.3	-.94	.37
US	7	12.4	10.2		
R4 MEXICO	6	29.5	21.3	-.02	.49
US	9	29.8	24.2		
P4 MEXICO	6	7.0	6.8	-.59	.28
US	9	9.8	10.0		

Judging from the mean scores on the later tests and their proportionately high standard deviations, one can assume that these scores are positively skewed. Such a distribution results from several of the students scoring near 0 on these tests; they have 'bottomed out' and show little measurable retention on these tests.

Similar T-tests were conducted to see if pre-course contact with Spanish was significant (see Appendix B, Table 43). The first group consisted of seven students who had had some previous contact with Spanish, and the second group consisted of eight students who had not. Although the means for the first group were slightly higher on each test, there were no significant differences on any of the tests. To see if post-course contact was significant, students were

divided into four small groups: those who had later contact in both a Spanish course and in the culture (2 students), those who had later contact only through the culture (5 students), those who had later contact only in the classroom (6 students), and those who had no further contact with Spanish (2 students). A One-way analysis of variance showed no differences between any of the groups on vocabulary retention scores (Appendix B, Table 44). This result was not surprising since the researcher eliminated the vocabulary items from tests if students reported having further contact with them.

Students were also divided into groups of females (7 students) and males (8 students). T-tests showed that there were no differences between females and males (Appendix B, Table 45).

In their final evaluation of the course, students indicated how much effort they had put into the course, and accordingly were given a 1 (much effort) or a 2 (some effort). Nine students were in the first group and six in the second group. T-tests showed no differences between the groups (Appendix B, Table 46). On the initial questionnaire students were asked to give their reasons for taking the course and were grouped the following three ways: 1) took the course because of interest (8 students), 2) took the course because it was a requirement, but also because of interest (3 students), and 3) took the course because it was a requirement for graduation (4 students). On a One-way

analysis of variance no differences showed. When T-tests
compared the most extreme groups (group 1 and group 3),
there was a significant difference only at R1 (Appendix B,
Tables 47 and 48).

A slightly different way of looking at this problem was
to combine the scores for effort and motivation to arrive at
a four-level scale (2 for a combination of high effort and
motivation through 5 for low motivation and effort). Since
this scale is ordinal in nature rather than interval, a
nonparametric correlation was performed to calculate
Spearman's Rho. We would expect a negative correlation with
test scores since a 2 represents the most combined
motivation and effort and a 5 represents the least. The
results are shown in Table 6.

TABLE 6
CORRELATIONS OF TEST SCORES WITH COMBINED
MOTIVATION AND EFFORT

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
MOT.+ EFF.	-.42	-.45	-.44	-.15	-.18	.09	.11	.29	-.07	.09
PROB.	.06	.05	.06	.31	.28	.39	.37	.18	.40	.37

It appears that if there is any connection at all
between motivation and effort and test scores, it is only at
the very beginning. As time goes on, any kind of a
relationship disappears.

The next question was whether students who did better
in class (earned the higher grades) also did better in
retaining their knowledge. Again the students were divided
into two groups, those who earned an A, A-, or B+ in the

course (8 students) and those who earned a B- or lower (7

students). Results of the T-tests are shown in Table 7.

TABLE 7
T-TESTS FOR GROUPS BASED ON COURSE GRADE

TEST		N	MEAN	SD	T VALUE	1-TAIL PROB.
R0	HIGHER	8	88.4	8.7	4.00	.00
	LOWER	7	70.1	8.9		
P0	HIGHER	8	69.8	21.1	3.05	.00
	LOWER	7	43.1	9.8		
R1	HIGHER	8	79.0	14.1	3.58	.00
	LOWER	6	50.8	15.3		
P1	HIGHER	8	45.5	26.7	1.50	.08
	LOWER	6	27.7	13.0		
R2	HIGHER	8	65.5	20.6	1.48	.08
	LOWER	5	47.2	23.3		
P2	HIGHER	8	32.3	22.9	.87	.20
	LOWER	5	22.4	12.6		
R3	HIGHER	7	42.0	27.5	.62	.27
	LOWER	5	32.4	24.4		
P3	HIGHER	7	11.1	9.9	.35	.37
	LOWER	5	9.2	8.8		
R4	HIGHER	8	35.4	24.3	1.06	.15
	LOWER	7	23.1	19.5		
P4	HIGHER	8	11.3	10.1	1.25	.12
	LOWER	7	5.7	6.2		

It is interesting to note that the only significant

differences found between the two groups are on the

beginning tests, but this is to be expected since course

grades were based in part on knowledge of vocabulary words

at Time 0. Even though later times do not show significant

differences, Group 1 maintains a higher mean each time than

Group 2. Note that at the end of the retention period, the

mean for production vocabulary for the higher group stays

about the same (has possibly leveled off), whereas the mean

score for the lower group continues to decline. The

relationship between these two groups is illustrated by the

graphs in Figures 1 and 2.
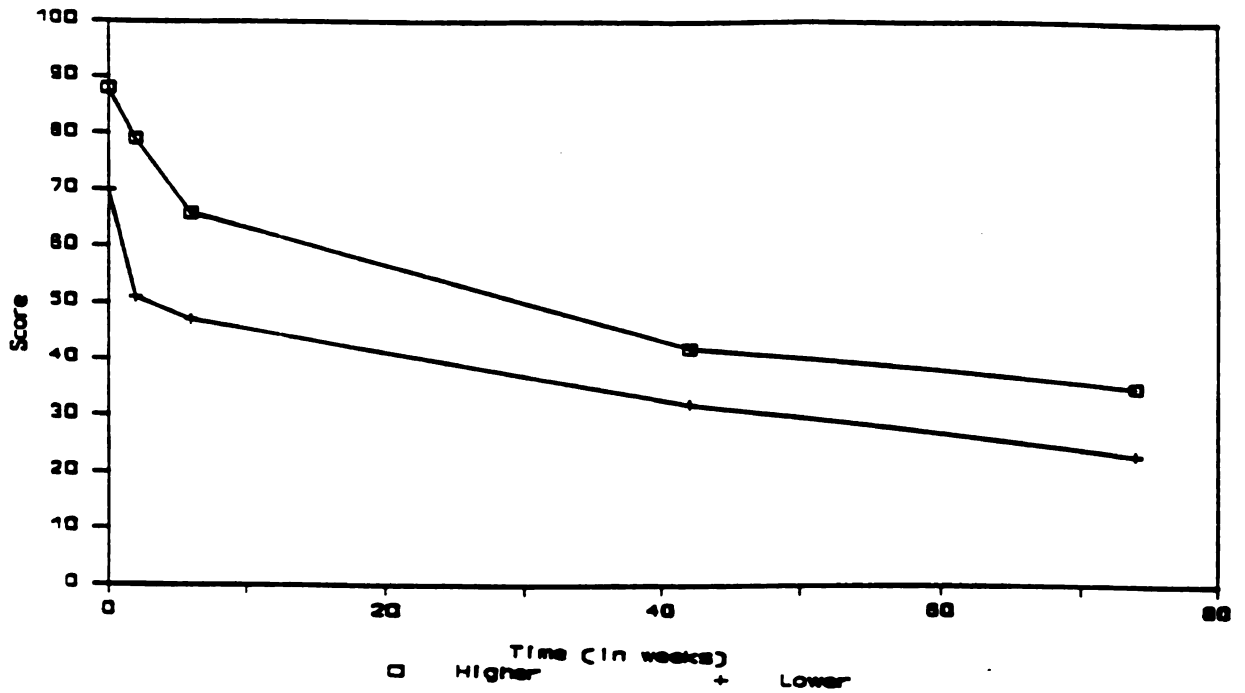
FIGURE 1
RECOGNITION SCORES FOR HIGHER AND LOWER GRADES
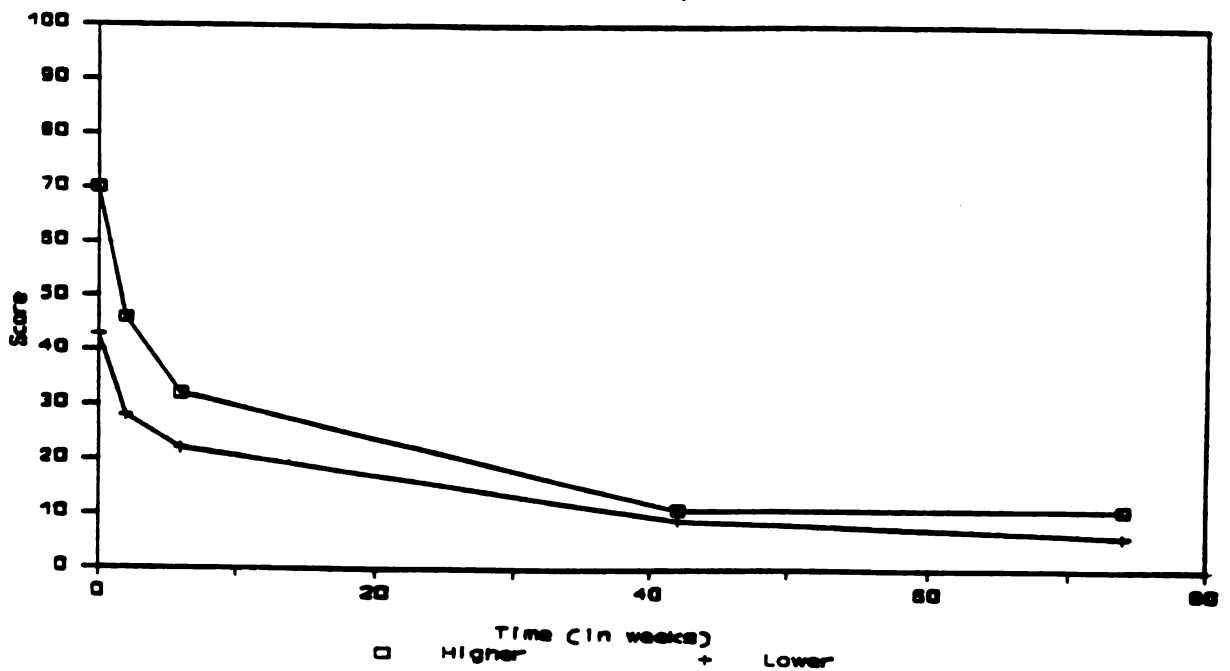
FIGURE 2
PRODUCTION SCORES FOR HIGHER AND LOWER GRADES

Another way to look at the data is to correlate the test scores for each time. Do students who receive high recognition and production scores at Times 0 and 1 continue to receive high scores as time goes on? Table 8 gives the Pearson correlation coefficients for the fifteen students on their ten tests over time.

TABLE 8
CORRELATIONS OF TEST SCORES FOR 15 STUDENTS OVER TIME

	R1	R2	R3	R4	P0	P1	P2	P3	P4
R0	.76	.61	.35	.47	.87	.67	.43	.33	.55
R1		.89	.65	.70	.73	.81	.70	.65	.59
R2	.89		.90	.83	.55	.74	.83	.82	.65
R3	.65	.90		.92	.28	.65	.88	.89	.76
R4	.70	.83	.92		.50	.81	.82	.97	.88
P0	.73	.55	.28	.50		.77	.47	.40	.58
P1	.81	.74	.65	.81	.77		.77	.77	.72
P2	.70	.83	.88	.82	.47	.77		.83	.67
P3	.65	.82	.89	.97	.40	.77	.83		.93
P4	.59	.65	.76	.88	.58	.72	.67	.93	

The strong positive correlations show that students who score high on the original and first tests continue to score high on later tests, although the correlations generally grow weaker over time.

Is early production or recognition a better predictor of final scores when we look at the data organized by student? Consider Table 9:

TABLE 9
CORRELATIONS WITH SCORES AT TIME 4

R0 and R4 .47	R1 and R4 .70	R2 and R4 .83	R3 and R4 .92
R0 and P4 .55	R1 and P4 .59	R2 and P4 .65	R3 and P4 .76
P0 and R4 .50	P1 and R4 .81	P2 and R4 .82	P3 and R4 .97
P0 and P4 .58	P1 and P4 .72	P2 and P4 .67	P3 and P4 .93

Although production scores seem to have slightly higher correlations than recognition scores with both R4 and P4, the numbers are too close to draw any conclusions, especially since only fifteen students were involved. This same question will be raised again in Section 3 of this chapter when we specifically discuss the relationship of recognition and production vocabulary.

Case study

Since variations among individual students were so great, it seems best to study some of the individual cases in depth. Figures 3-17 show the graphs for each student's recognition and production scores over time. All student names have been changed to protect their anonymity. Note that Neal was available for testing only at the beginning and end of the 18-month period, and therefore his graph appears as a straight line. The data on which these graphs are based are found in Table 40 in Appendix B.
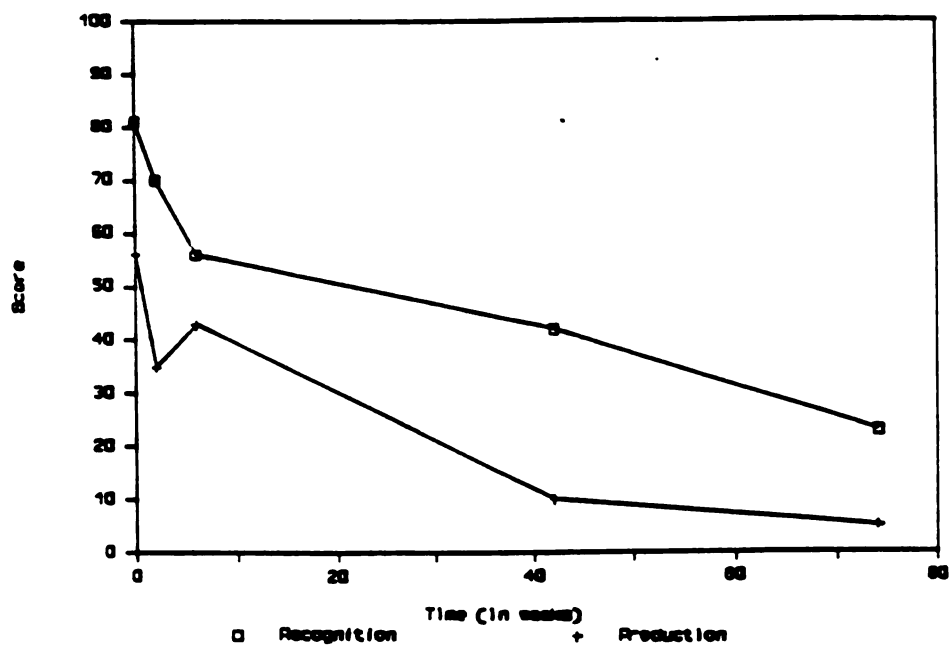
FIGURE 3
RETENTION SCORES FOR SARA (STUDENT 1)
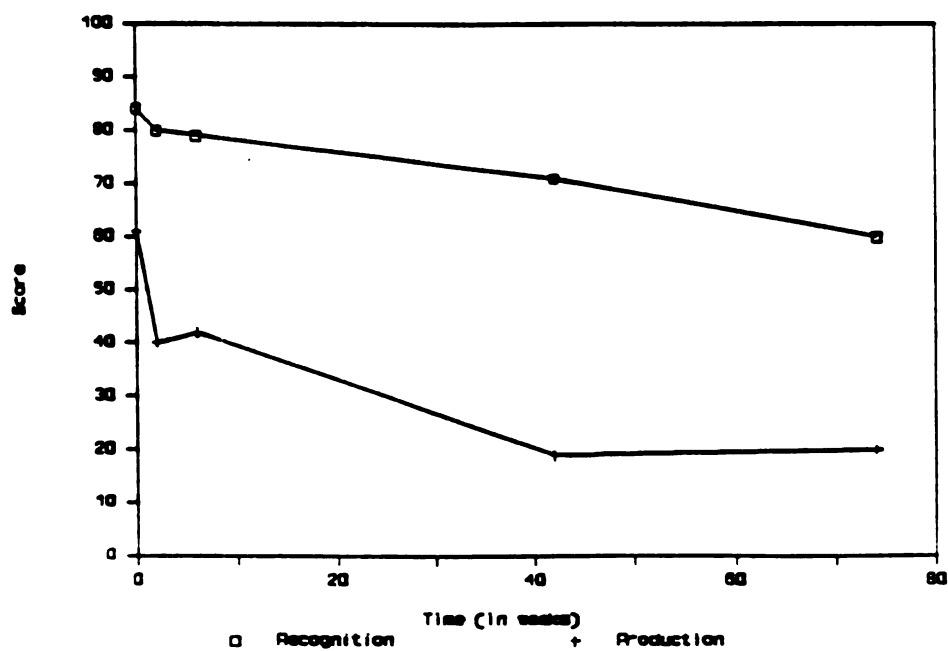
FIGURE 4
RETENTION SCORES FOR DAVID (STUDENT 2)
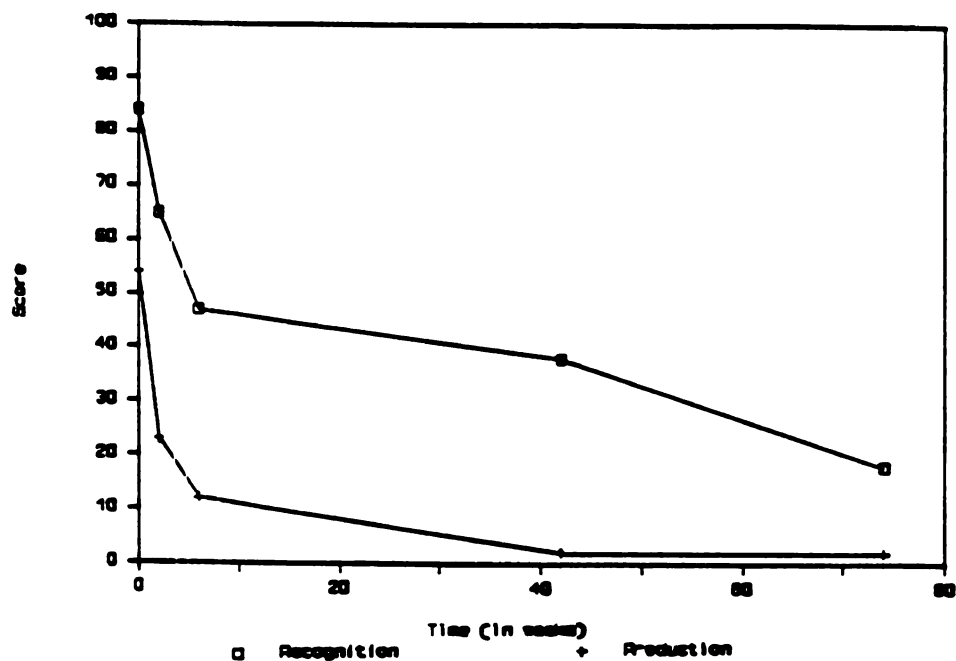
FIGURE 5
RETENTION SCORES FOR JUDY (STUDENT 3)
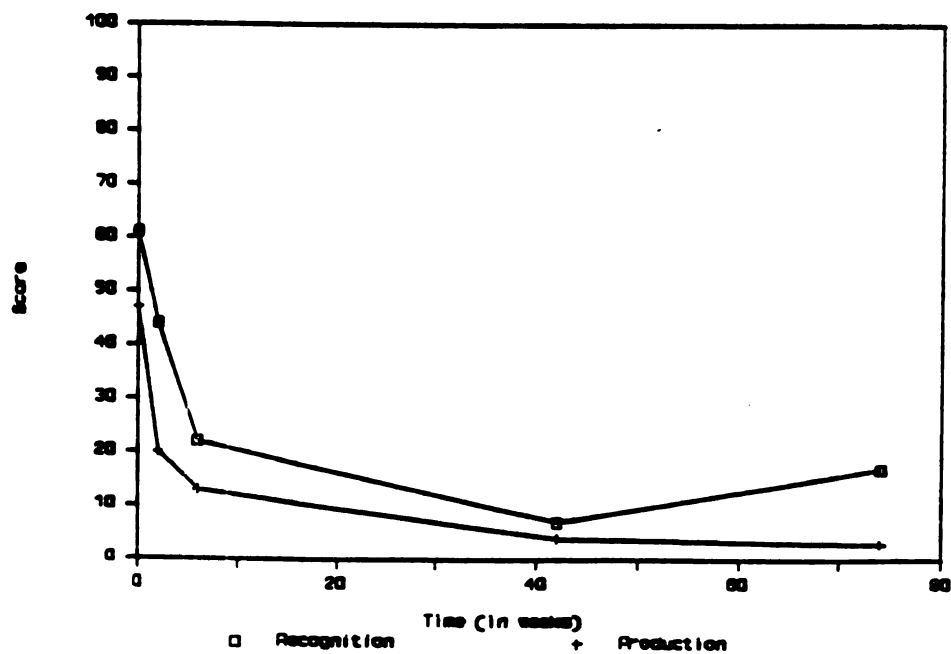
FIGURE 6
RETENTION SCORES FOR LAURA (STUDENT 4)
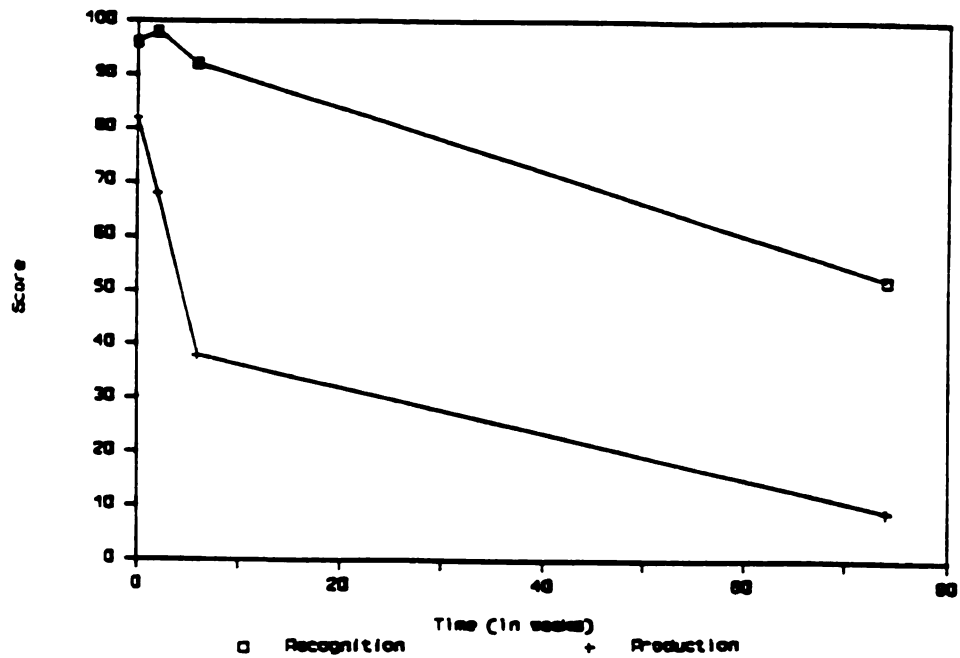
FIGURE 7
RETENTION SCORES FOR RACHEL (STUDENT 5)
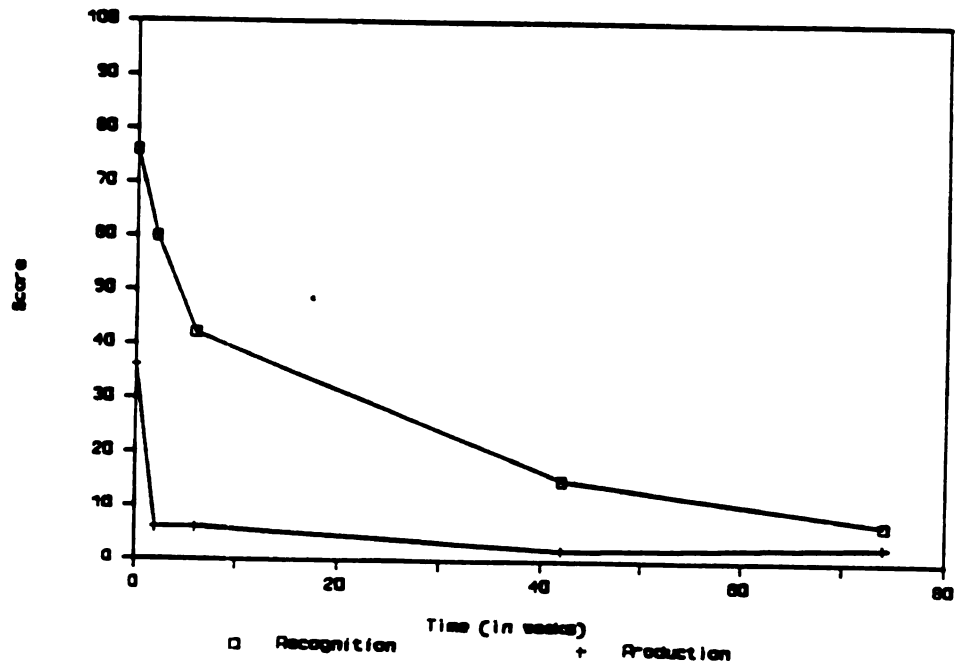
FIGURE 8
RETENTION SCORES FOR GARY (STUDENT 6)
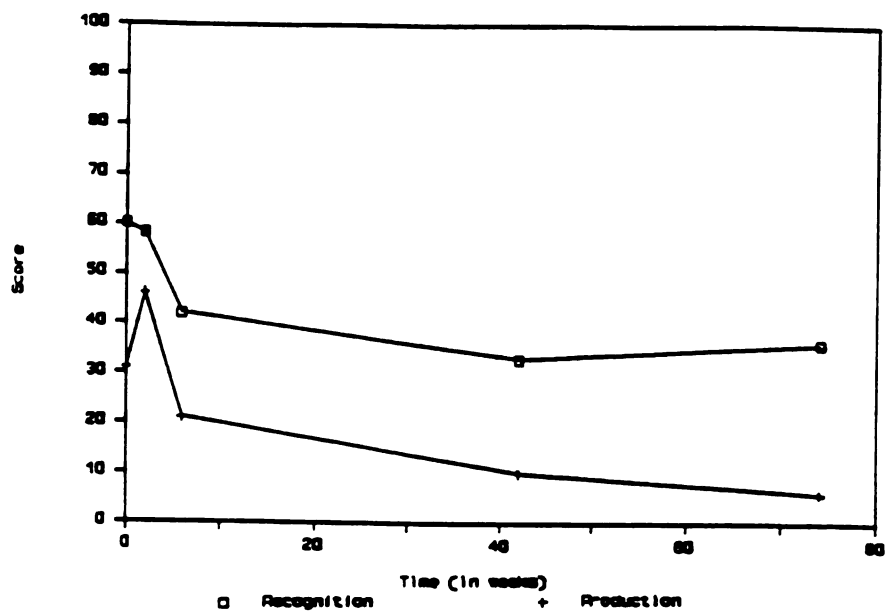
FIGURE 9
RETENTION SCORES FOR ED (STUDENT 7)
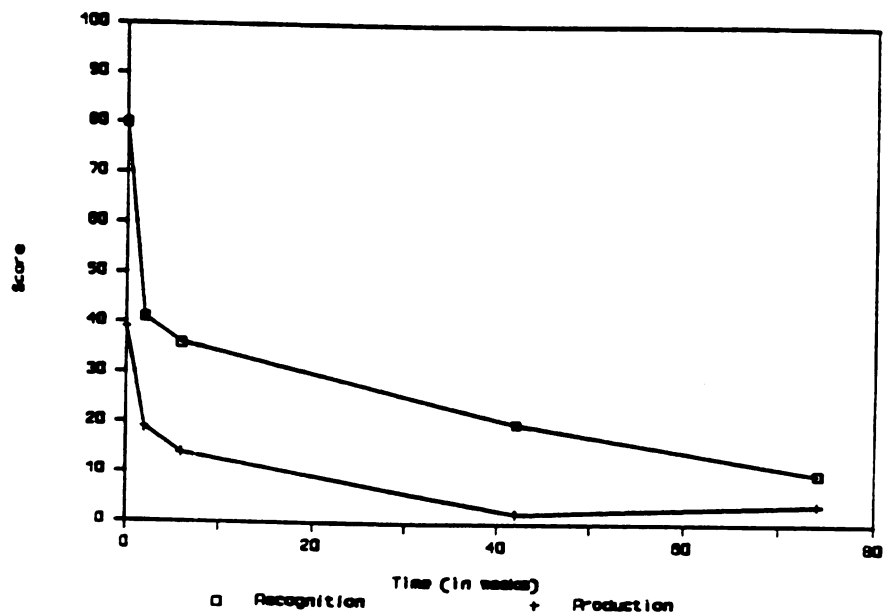
FIGURE 10
RETENTION SCORES FOR TODD (STUDENT 8)
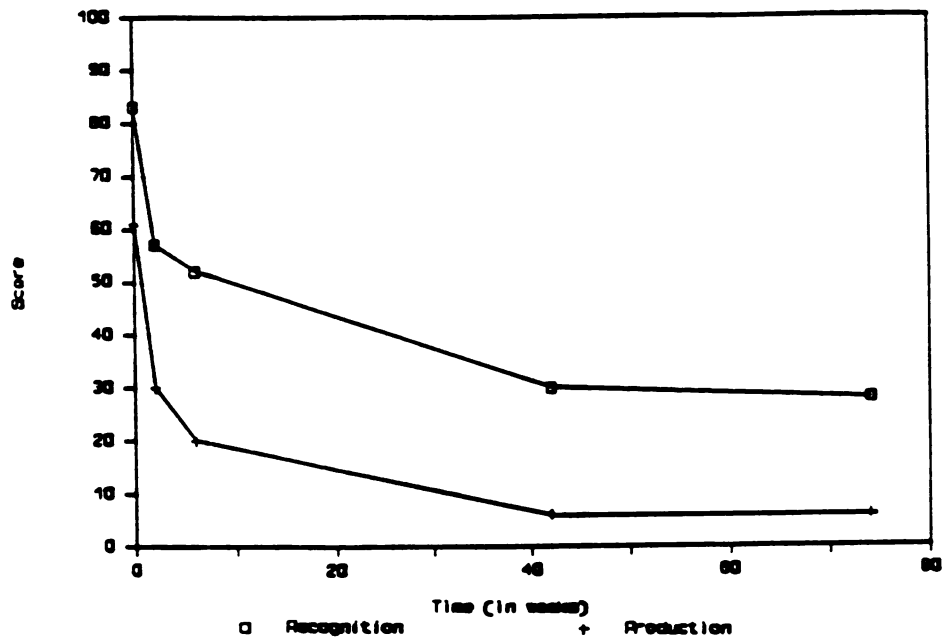
FIGURE 11
RETENTION SCORES FOR JEFF (STUDENT 9)
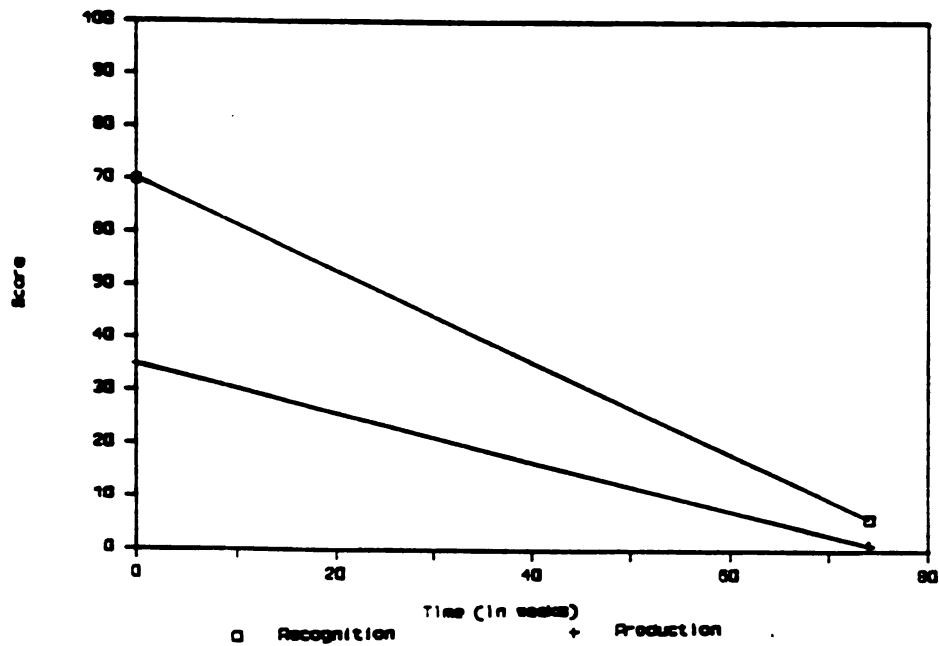
FIGURE 12
RETENTION SCORES FOR NEAL (STUDENT 10)
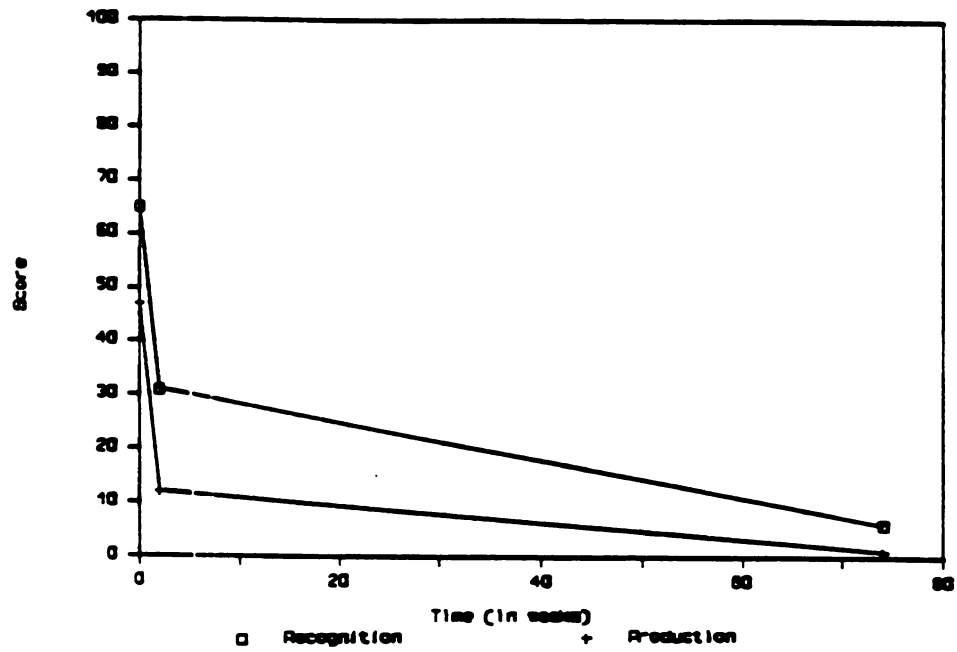
FIGURE 13
RETENTION SCORES FOR KIMBERLY (STUDENT 11)
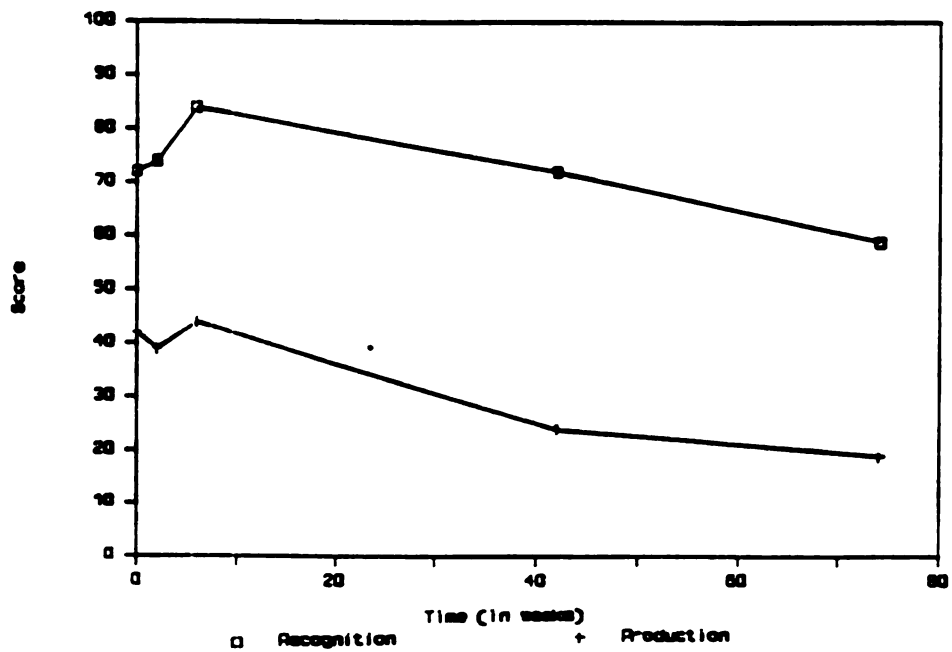
FIGURE 14
RETENTION SCORES FOR DIANE (STUDENT 12)
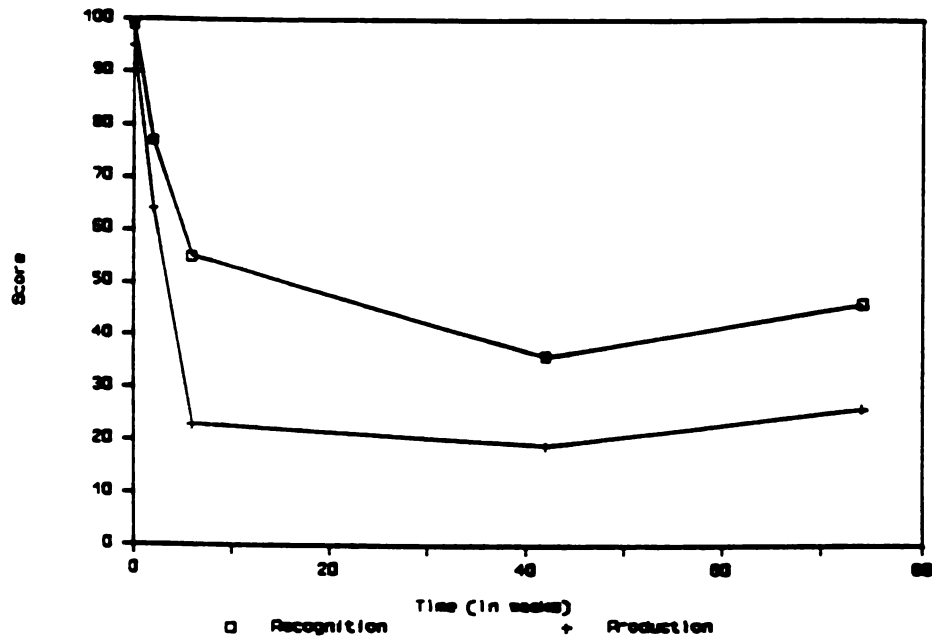
FIGURE 15
RETENTION SCORES FOR JOEL (STUDENT 13)
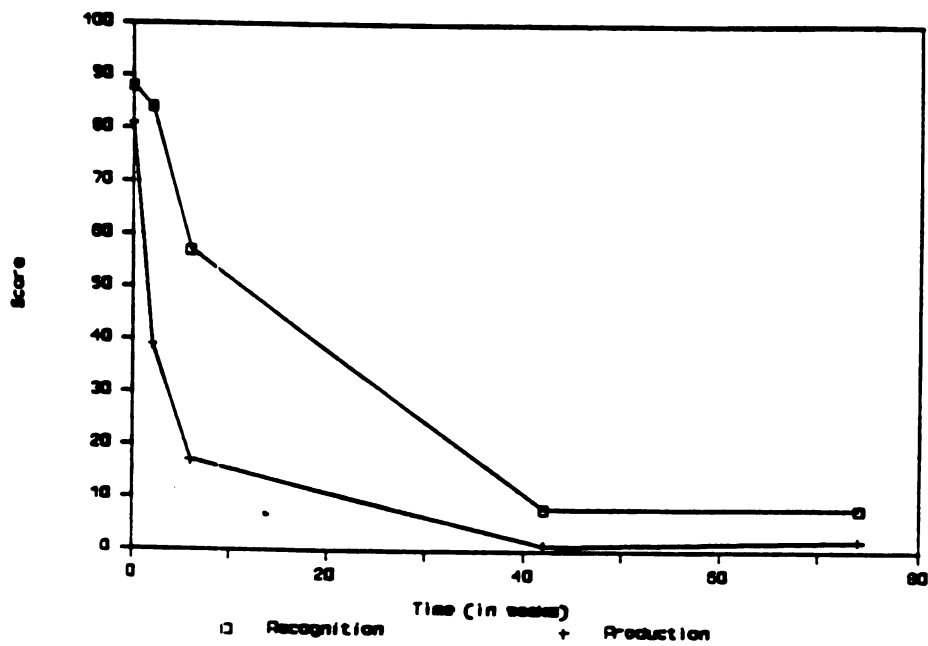
FIGURE 16
RETENTION SCORES FOR KAREN (STUDENT 14)
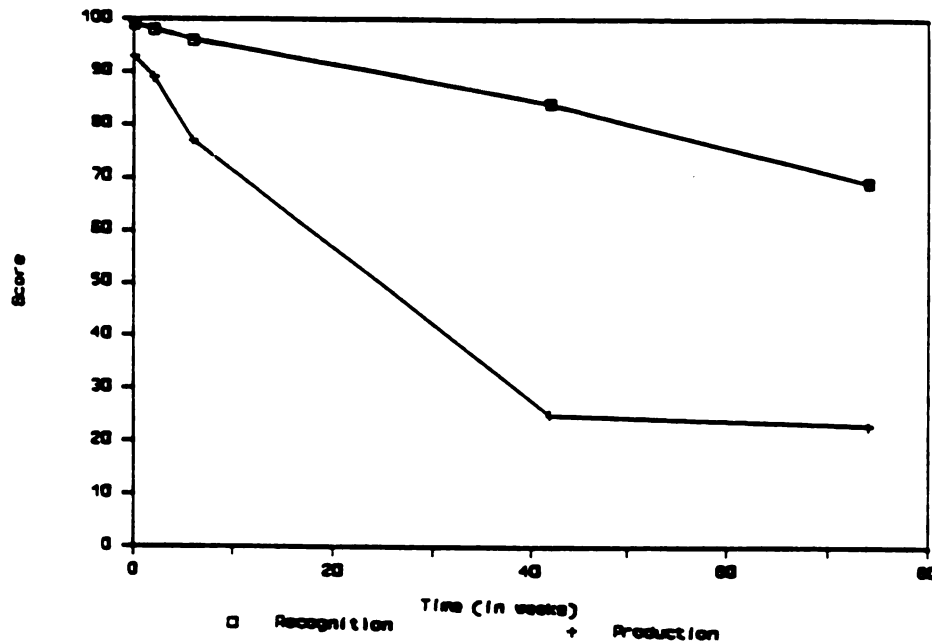
FIGURE 17
RETENTION SCORES FOR GLEN (STUDENT 15)

Figure 18 is a composite of all fifteen graphs. At this point we can make two important observations about the mean curve: 1) retention drops rapidly at first and then levels off and 2) recognition scores are consistently higher than production scores. Later in Section 3 of this chapter we will specifically discuss the relationship of recognition and production vocabulary, while at this point we will focus on individual students and why they might have deviated from the mean curve.
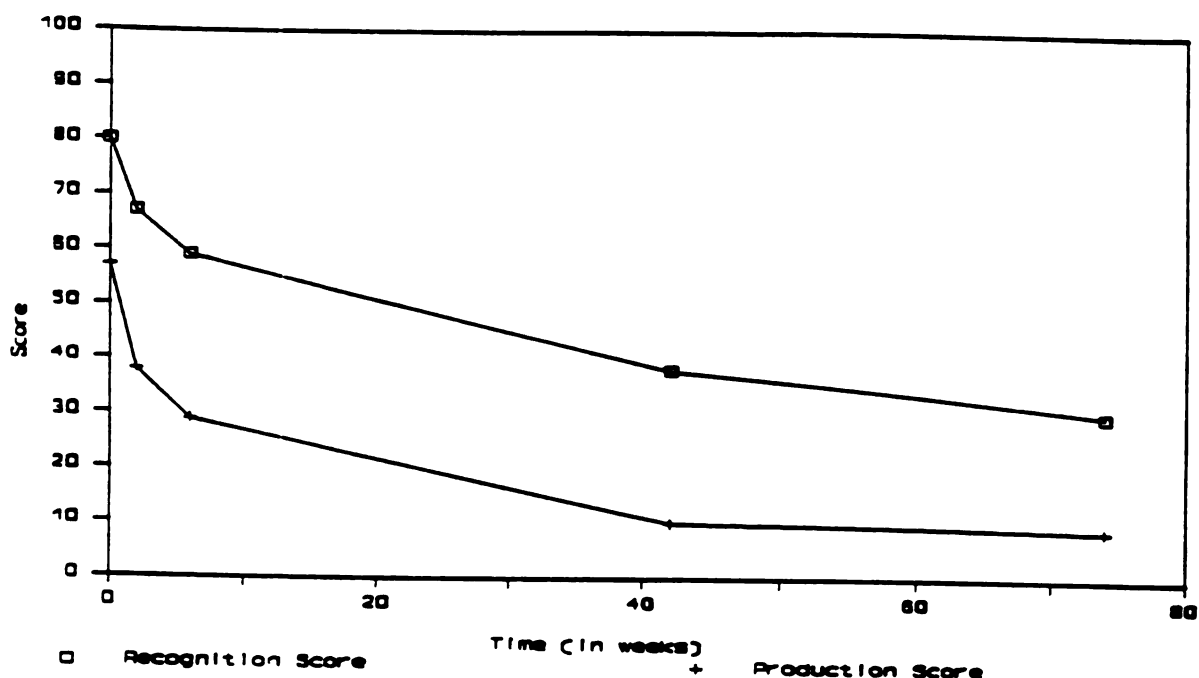
FIGURE 18
COMPOSITE OF RETENTION SCORES (RECOGNITION AND PRODUCTION)

Jeff and Sara are the first students that we will discuss since their scores are quite close to the composite. Jeff was a student in the class in Grand Rapids. He had dropped out of a first semester Spanish class a year and a half earlier at Calvin (with a different text), and was now beginning over with 101. He was taking the course as a college requirement and said that he put a great deal of effort into the class. His final grade for this course was a B-, and he went on to finish three more semesters of Spanish at Calvin.

Jeff liked to study vocabulary by associating a Spanish word directly with its picture and trying not to think in

English. He used a mnemonic device from time to time (for example, Mastercard to remember masticar 'to bite'), but he preferred to access the words directly in Spanish. His grade for the course was about average for Calvin students, and his scores on the retention tests were also near the mean. His vocabulary dropped off rapidly at first; in fact his production score was halved by the end of the three-and-a-half-week course. However, his scores on tests at Times 3 and 4 remained nearly the same.

Sara was another student whose scores were close to the mean. Sara participated in the course in Mexico because of her interest in missions. She had no background in Spanish, but she came from Canada and had studied two years of French in junior high. Sara's grade in the Spanish course was a B+, and she also said that she put a great deal of effort into the course. After finishing the course in Merida, she worked for three weeks in a Tzeltal-speaking village and used Spanish only with her host family. Her production score (but not her recognition score) jumped somewhat after this three-week experience. At the end of the total 18-month period, though, both of her scores were slightly below the mean. After her experience in Mexico, she did not have any further experience with Spanish.

Sara preferred to study vocabulary from English-Spanish word lists, first covering up one side and then the other. She used few devices to aid memory, but one example was English tarp to remember Spanish carpa 'tent'. When she was

asked to describe the picture of the classroom (see Appendix A, Figure 22) at the end of 101, she was able to say the following: 'Un profesora --how do you say teacher?--enseña un estudiante. Ocho y mediodía...puerta, libro, cuaderno, tiza, borrador.' After eighteen months all she could say about the same picture was 'Well, son las ocho. I can't remember anything.'

In contrast, let us consider three students who consistently maintained scores above the mean. David participated in the course in Merida and was interested in missions and in Latin America. He had already studied German and Greek, but had no previous contact with Spanish. He received an A in the course, even though he admitted that he could have spent more time studying. During his field experience he did not have many opportunities to speak Spanish (he stayed in a Maya-speaking area and used English with his friends); however, he returned to Mexico the following summer and had a few more opportunities to speak Spanish. All together his Spanish-speaking experience totaled about four weeks, and he received no further formal study in Spanish.

David studied vocabulary from flashcards he had made. He said that he liked to look for relationships between words, either cognates in English and Spanish or relationships within Spanish. For example, on his final production test he remembered both aguja 'needle' and agudo 'smart, sharp' because a needle is sharp. Barbilla 'chin'

reminded him of <u>barber</u>, and words like <u>padrino</u> and <u>bisabuelo</u> were closely related to other words he knew (<u>padre</u> and <u>abuelo</u>).

Another student who performed consistently well was Joel, a participant in the Grand Rapids class. Joel had studied Latin, but had no previous contact with Spanish. He originally joined the class because he was planning a trip to Spain, but his plans fell through midway in the course, and he admitted that he studied far less toward the end. Joel received a B+ in the course and had no further contact with Spanish.

Joel's scores showed a precipitous decline at first, but they leveled off sooner than the mean and even showed a slight rise at the end. Joel explained that he worked harder on the final test at Time 4, spending more time trying to access the vocabulary. When he studied vocabulary, he reported using as many devices as he could think of to help memorize it, and this showed up on his final test in which he said he remembered the word <u>torpe</u> 'stupid' because another student used it to joke with him, the word <u>lágrima</u> sounded like <u>tear</u> in Latin, and <u>fango</u> (mud) reminded him of fangs. In his original picture description task at the end of 101 he described the village (see Appendix A, Figure 22) with the following words: 'La pictura tiene mucho cosas. Hay mariposa y calle, hace sol, hay montaña. Hay perro, un gato. Hay hombre, la máquina, hay loma, hay tree.' His picture description eighteen months

later was not too different: 'Es la sol. Es la mariposa.
Es la iglesias. Es la tractora. Es la ganada. Es la loma.
Es la casa. What are dog and cat? I can't remember.' Note
that he was still able to produce the two target words
mariposa and _loma_.

A third student who did very well throughout was Glen,
the sixteen-year-old student in the Grand Rapids class. A
few years earlier he had received 4 hours of contact with
Spanish through TPR (Total Physical Response), but he had no
other foreign language background. He joined the class
because of an interest in Latin America and said that he
studied a great deal. His final grade was an A, and he went
on to take more Spanish in high school and also had contact
with Spanish-speaking migrants during the summer.

Glen said that he tried to think the words in Spanish
without using reference to English as a crutch. He learned
various pairs of words together, like _cano_ 'grey-haired' and
calvo 'bald', or _huella_ 'footprint' and _fango_ 'mud'.
Another thing that helped him was that he went back
immediately to study any word that he had missed on a test
during 101. If he missed a word, he said that it was a sure
sign that he would not miss it a second time.

Looking at these three students with consistently high
scores, certain patterns emerge. All three of them had some
intrinsic motivation for taking the course, and all three
had shown previous interest in foreign language. Even
though amount of effort was not uniform, they all received

high grades in the class. All three learned vocabulary through relationships and associations with other words.

On the other hand, let us look at two students whose scores were less than the mean. Gary started out with a fairly high recognition score, but his production score was less than half of the recognition. By the end of the eighteen months Gary retained very little, either in recognition or in production vocabulary. Gary had no previous foreign language experience before he joined the class in Merida, but he did have an interest in missions and in Latin America. His efforts to study outside of class were sporadic because of the conditions under which he was living, but in class he paid attention and ended up with a course grade of B+. His field assignment was in a Mayan village where little Spanish was spoken, and after leaving Mexico he had no further contact with Spanish. His time in Mexico was difficult because of problems his wife and two small children had in adjusting to a foreign culture, and by the end of the summer he had made a decision that his family would be happiest staying in the United States.

Gary admitted that he had no special method for studying vocabulary--he just tried to remember the word and that was all. He also said that after the course he did not have any special reason for trying to remember the words. A few came to mind because of vivid experiences--seeing the globos 'balloons' in Mexico or remembering the teacher's

story of fishermen catching <u>tiburones</u> 'sharks' where she had just taken a swim.

Kimberly was another student whose scores were below the mean, both in recognition and production. Kimberly had had four years of high school Latin but no previous contact with Spanish. She took the summer language course in Grand Rapids in order to fill a single course requirement in foreign language at another college, and thus had no further contact with Spanish after finishing 101. She reported putting a great deal of effort into the course at Calvin. Her course grade was a C.

Kimberly said that she was more acoustically than visually oriented and that she would read words to herself to learn them by sound. She remembered a word like <u>pupitre</u> 'desk' by knowing that she was 'pooped' when she sat in it. Although her Latin helped her to remember <u>lágrima</u>, some French she had learned in France the summer after 101 got in her way when she took the final oral interview (<u>fleurs</u> instead of <u>flores</u>, <u>garçon</u> instead of <u>muchacho</u>.) Kimberly was also extremely nervous about tests, and this most likely impeded her ability to do well.

Looking at both cases, we see that neither student ended up with a desire to go on in Spanish. Gary did not have an effective way of studying vocabulary. Kimberly tried to learn through association with sounds, but her hold on Spanish was too nebulous after just one course to last for much time, especially after some contact with French.

Probably the most instructive cases are those in which
the student performed entirely differently than what could
be expected from class performance and original test scores.
One of these students was Karen, a student who took the
class in Grand Rapids because she was interested in nursing
and thought she might use Spanish in her career. Karen came
to the course with a background of three years of high
school French, though she had no previous contact with
Spanish. She worked hard in the course and received an A-.
She went on to take Spanish 102 at Calvin during the summer
but had no further contact with Spanish after that.

Karen's original test scores were originally very high,
88% for recognition and 81% for production. Note that the
production score is nearly as high as the recognition score.
On her original picture description test of the village, she
was able to say the following: 'Es una mañana y es un
ciudad avec--con una iglesia y las casas. Tengo un montaña
y los plantos y arbols. Hay una mariposa.'

Even though Karen's recognition score remained high on
the first retention test, her production score was only 39%,
which was right at the mean. On the final tests her
recognition score was only 8% and her production score 2%,
far below the mean. A look at her final production test
(written) shows more French than Spanish: <u>piscine</u>, <u>nez</u>, and
<u>oiseaux</u>, for example. Her final oral picture description
showed the same problem; all she could say was 'un mariposa,
un eglise, montagne, les casas', very little for an

investment of seven intensive weeks and several hundred
dollars.

Karen seemed to use effective methods for studying
vocabulary during the class. She read and pronounced the
words and tried to form pictures in her mind. She related
the words to French that she knew, and just before a test
she studied the word pairs in English and Spanish, covering
up the Spanish and giving the words by memory. One
interesting comment that she made was that she knew the
beginnings of the words much better than the endings.

It is difficult to explain why Karen retained so few
words, when it seemed like her French should have helped her
Spanish. Instead the three years spent studying high school
French seemed to remain much stronger in her memory than the
seven intensive, but quick weeks of Spanish. It seems that
she never reached a 'critical threshhold' in Spanish over
the seven weeks in spite of her excellent grades, and that
what was quickly learned was quickly forgotten.

A student in whom the opposite effect was observed was
Diane. Although Diane came to Calvin with two years of high
school Latin, she had difficulty with Spanish and had
already dropped out of one Spanish course. She took the
summer 101 course because of the language requirement and
received a C, but might have done better if she had not
worked at another job for so many hours that at times she
could barely stay awake in class. During the 102 course
that followed, Diane again dropped out. Diane finally

finished 102 during the following school year, but dropped out once more during second-year Spanish.

Diane's original scores were below the mean, 72% for recognition and 42% for production. However, on the first retention test her recognition score was slightly higher than her first score (74%), and her production score had hardly dropped at all (39%). At the final test, her scores were far above the average: 59% for recognition and 19% for production. These are not the results we would have expected when looking at her original scores.

After the final test, Diane explained how she remembered so many words--the letters of the words themselves visually represented their meaning. For example, for Diane the word calvo actually looked like a scalp and the ñ in peña looked like a rock. To remember ala (wing) she visualized a bird covered with a's instead of feathers. The two r's in garra looked like claws, while the two r's in arrugado looked like wrinkles. The two o's in globo looked like two balloons, and the two ll's in llama 'flame' reminded her of two matches. Cuenco 'bowl' was a bowl with the number cinco 'five' written on it. Other Spanish words had links to English; for example, manso 'gentle' reminded Diane of gentleman. Pasillo 'hall' reminded Diane of a pass, which she needed to go down the hall in high school. Lágrima 'tear' had the word grim in it and dehesa 'meadow' had the word hay in it. The word mariposa reminded her of the Virgen Mary, someone quite beautiful. Diane's rich

visual and acoustic imagery undoubtedly helped her to retain words that other students soon forgot.

The final case that we will look at is that of Laura. She is the only one of the Mexico group that stayed in Latin America after the termination of the program. Laura went to Mexico City for further language training and then on to Honduras where she worked in a mission program. She took her final retention tests while home for Christmas in December 1989. At first appearance it seemed that it might be necessary to drop the data from Laura because her exposure to Spanish was obviously far greater than with any of the other students, and by December of 1989 her Spanish was by far the most fluent. However, the attrition of target words closely paralleled that of the other students. For example, in her first rather painful oral description of the picture of the classroom at the end of 101, she mentioned several objects, including <u>maceta</u> 'flowerpot', one of the target words. In her final oral description of the same scene in December of 1989 her Spanish was very fluent and she told an interesting story about the picture, but all of the original target words (including <u>maceta</u>) were absent from her speech.

Laura's original scores were lower than the mean, 61% for recognition and 47% for production, and once the words with which she had contact were eliminated from the final data, her final scores were also low (17% and 3%). Laura's written tests were far from blank, though. She had

substituted words from the Spanish dialect where she was living for the target words she had been taught; for example, _mojado_ for _empapado_ 'wet', _fósforo_ for _cerilla_ 'match', _basura_ for _cascote_ 'trash', _cubierta_ for _colcha_ 'blanket', and _canasta_ for _cesto_ 'basket'. Laura remarked that. even a North American woman living and speaking Spanish in Honduras for seven years did not know the names for all the pictures, and Laura felt that it was quite impossible to expect her to know them all.

The case of Laura seems to be one of forgetting through interference, in which the original vocabulary words are supplanted by words that the subject hears continuously. Although her Spanish had improved a great deal, there was little measurable retention left of the original words.

These case studies point out the variation that exists among individual students. A certain pattern seems to be true for all students: rapid attrition at first with a gradual leveling off and recognition scores that are consistently higher than production scores. However the ratio of recognition vocabulary to production vocabulary varies with each student. Although most students who perform well at the beginning also perform well at the end, and vice versa, there are a couple of students who do not follow this general pattern, and they either retain or forget an unexpectedly high amount. Visual and acoustic mnemonic devices relating words to an image or to the sound of an English word seem to be important in the retention of

the words in this study, as well as the chance to study
Spanish long enough to build up a 'critical threshold' from
which knowledge is not so readily forgotten.

Results for the hypotheses about students

Hypothesis 1: **Students who study in a target-language
environment retain more vocabulary than those who study in a
native-language environment, even when the specific
vocabulary is not rehearsed or used outside of class.**

The corresponding null hypothesis <u>cannot</u> be rejected at
the .05 probability level. (Table 5)

Hypothesis 2: **Students who have pre-course contact with
Spanish (either informally in the culture or formally in the
classroom) retain more vocabulary than those who have
no previous contact.**

The corresponding null hypothesis <u>cannot</u> be rejected at
the .05 probability level. (Page 125)

Hypothesis 3: **Students who have post-course contact with
Spanish (either informally in the culture or formally in the
classroom) retain more vocabulary than those who have
no further contact, even when the specific vocabulary
learned is not rehearsed or used.**

The corresponding null hypothesis <u>cannot</u> be rejected at
the .05 probability level. (Page 126)

Hypothesis 4: **Vocabulary is retained at different rates according to sex.**

The corresponding null hypothesis cannot be rejected at the .05 probability level. (Page 126)

Hypothesis 5: **Students who report more effort while taking a language course retain more vocabulary than those who report less effort.**

The corresponding null hypothesis cannot be rejected at the .05 probability level. (Pages 126-127, Table 6)

Hypothesis 6: **Students who report more intrinsic motivation for the course retain more vocabulary than those who report more extrinsic motivation.**

The corresponding null hypothesis cannot be rejected at the .05 probability level. (Pages 126-127, Table 6)

Hypothesis 7: **Students who receive higher grades for the course retain more vocabulary than students who receive lower grades.**

The corresponding null hypothesis cannot be rejected at the .05 probability level (Table 7). However, Figures 1 and 2 show a pattern of consistently higher scores for students who received higher grades and consistently lower scores for students who received lower grades.

<u>Hypothesis 8</u>:   <u>Students who receive higher scores on the</u>
<u>first retention tests continue to receive higher scores on</u>
<u>later retention tests.</u>

Strong correlations were noted between scores on early tests with scores on later tests. (Tables 8 and 9)

Retention scores for individual students were examined by means of graphs and discussion. Although most students who perform well at the beginning also perform well at the end, and vice versa, there are a few students who do not conform to the overall pattern. Further discussion of the results above and of individual variation will be found in Chapter 5.

ANALYSIS BY WORD

A second way to look at the data is to analyze it by word. What percentage of the students could recognize or produce a word at each point in time? What changes occur in a word over time in its recognition and production strength, and what characteristics of the word might affect the likelihood of its long-term retention?

Statistical study

The independent variables for each word are given in Table 3 in Chapter 3, where each word is listed in Spanish with its English translation, the paso in which it was taught (A,B,C,D,E), its part of speech (noun, verb, or adjective), its frequency in teacher input (very low to high), and whether it was taught and originally tested for both recognition and production or for recognition only. Table 41 in Appendix B shows the recoding of these variables for statistical analysis, as well as each word's rating on one scale of emotionality and two scales of saliency. Also in Table 41 the percentage scores for each word are given for each recognition and production test. Notice that there are no scores for words in Pasos D and E for baseline production, since part of the design was to originally teach and test these particular words for recognition only. Other missing data occur because of problems in carrying out the

project—for example, for parts of two lessons in Merida the
tape recorder did not function, thereby making it impossible
to collect data on input for several words. Whenever data
were left out of the tables, they were also left out of the
calculations.

In order to analyze the role of frequency in input on
the retention scores of words, the categories very low and
low (fewer than 10 times) were recoded as 'low,' and the
categories medium and high (10 times or more) were recoded
as 'high'. As a result, 49 words were included in the low
category and 25 words in the high. Independent one-tailed
T-tests were then run on the two groups of words, with the
results shown in Table 10:

TABLE 10
T-TESTS FOR FREQUENCY IN INPUT

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
T VALUE	-.72	-.15	-.25	-.59	.38	-.36	1.22	-.40	.51	1.00
PROB.	.24	.44	.40	.28	.35	.36	.11	.35	.31	.16

The results of this test for significant difference
show that retention of words in this project was <u>not</u>
dependent on the number of times the instructor repeated the
words in the input. These results are reinforced by a
nonparametric correlation test in which the frequency was
recoded on a scale of 2 to 8 and correlated with test
scores.

TABLE 11
CORRELATIONS OF FREQUENCY IN INPUT WITH TEST SCORES

TEST	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
Spearman										
rho	.07	.07	.03	.14	-.06	.10	-.11	.07	-.14	-.09
PROB	.26	.28	.41	.11	.30	.18	.16	.28	.11	.22

Again, there is no pattern of correlation between the number of times words were repeated in class and the test scores. The difference between repeating a word 5 or 20 times may not have been great enough to make a difference in subsequent test scores, especially since none of these words was repeated again in input after the first twenty hours of class.

Another research question was whether part of speech influenced the retention of vocabulary. It must be noted that because of the nature of the vocabulary taught in the beginning pasos of Dos mundos, verbs and adjectives could be inserted only into Pasos A and B. Thus the number of verbs (11) and adjectives (14) is much lower than the number of nouns (75) taught. Moreover, all words, whether verb, adjective, or noun, were concrete and were taught through actions, pictures, or real objects. Since verbs were taught through Total Physical Response and tested in command form only, they showed no morphological complexity during the first twenty hours of class. Adjectives showed some complexity in that students heard them inflected for gender

and number. However, they were tested only in the masculine singular form.

The first test performed to find a significant difference between parts of speech was a One-way analysis of variance, with verbs, adjectives, and nouns each forming a group. The results of these tests are shown in Table 12.

TABLE 12:
ONE-WAY ANALYSES OF VARIANCE FOR PART OF SPEECH

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
F	3.45	.12	.16	1.88	.51	.62	1.93	.90	.62	1.39
F PROB	.04	.89	.86	.16	.60	.54	.15	.41	.54	.28

The test at R0 is the only one that shows a significant difference, and the accompanying Tukey test showed that the difference was between adjectives (with a mean of 89.57) and nouns (with a mean of 78.17). Thus it appears that only on the baseline recognition test was there a significant difference related to part of speech, a difference that does not exist in any of the further testings.

Another way to look at the data is to limit the analysis to the words taught in Pasos A and B only, since these were the only pasos to include all three parts of speech. If this limitation is made we are dealing with only 37 nouns rather than with 75. The number of verbs remains at 11 and the number of adjectives at 14. Another One-way analysis of variance provides us with the results in Table

13. The set-up of Table 13 is similar to that of Table 12, but the data include only the words from Pasos A and B.

TABLE 13
ONE-WAY ANALYSES OF VARIANCE FOR PART OF SPEECH
(PASOS A AND B ONLY)

TEST	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
F	2.26	.20	1.89	.28	2.75	.80	3.09	.50	2.47	.99
F PROB	.11	.82	.16	.76	.07	.46	.05	.61	.09	.38

Although only test R3 reaches significance at the .05 level, an interesting pattern is established in that probabilities on the recognition tests are lower than on production tests. This seems to result from the adjective scores--which were high on the original tests but are lower on later tests than the scores for verbs and nouns, especially on the recognition tests. Table 14 gives the means and standard deviations for verbs, adjectives, and nouns at the different test times. Note that there are two columns of scores for nouns: the first column includes nouns from all pasos and the second includes nouns from Pasos A and B only.

TABLE 14
MEANS AND STANDARD DEVIATIONS FOR PART OF SPEECH

PART OF SPEECH	VERB	ADJECTIVE	NOUN (A-E)	NOUN (A-B)
R0 MEAN	78.6	89.6	78.2	81.7
SD	13.6	11.3	15.7	14.8
P0 MEAN	55.5	58.6	57.6	59.0
SD	10.9	15.6	17.1	17.6
R1 MEAN	68.6	68.7	66.0	76.4
SD	18.3	11.2	21.7	15.6
P1 MEAN	42.0	46.2	35.9	46.6
SD	23.4	15.6	19.7	17.1
R2 MEAN	64.3	55.7	58.4	68.3
SD	20.5	8.8	23.1	18.4
P2 MEAN	30.5	33.2	27.7	37.2
SD	15.9	12.7	18.9	18.3
R3 MEAN	48.3	31.2	38.2	45.2
SD	20.0	15.5	22.7	21.1
P3 MEAN	14.4	10.1	9.8	13.4
SD	14.5	8.9	10.3	11.6
R4 MEAN	31.0	24.4	30.7	36.2
SD	10.8	11.1	21.8	20.1
P4 MEAN	12.8	7.5	8.4	10.4
SD	11.6	7.9	8.8	9.5

Figures 19 and 20 on the following pages show the graphs for part of speech for recognition and for production tests. Notice the pattern that adjectives follow, particularly on the recognition tests. This pattern will be discussed further in Chapter 5.
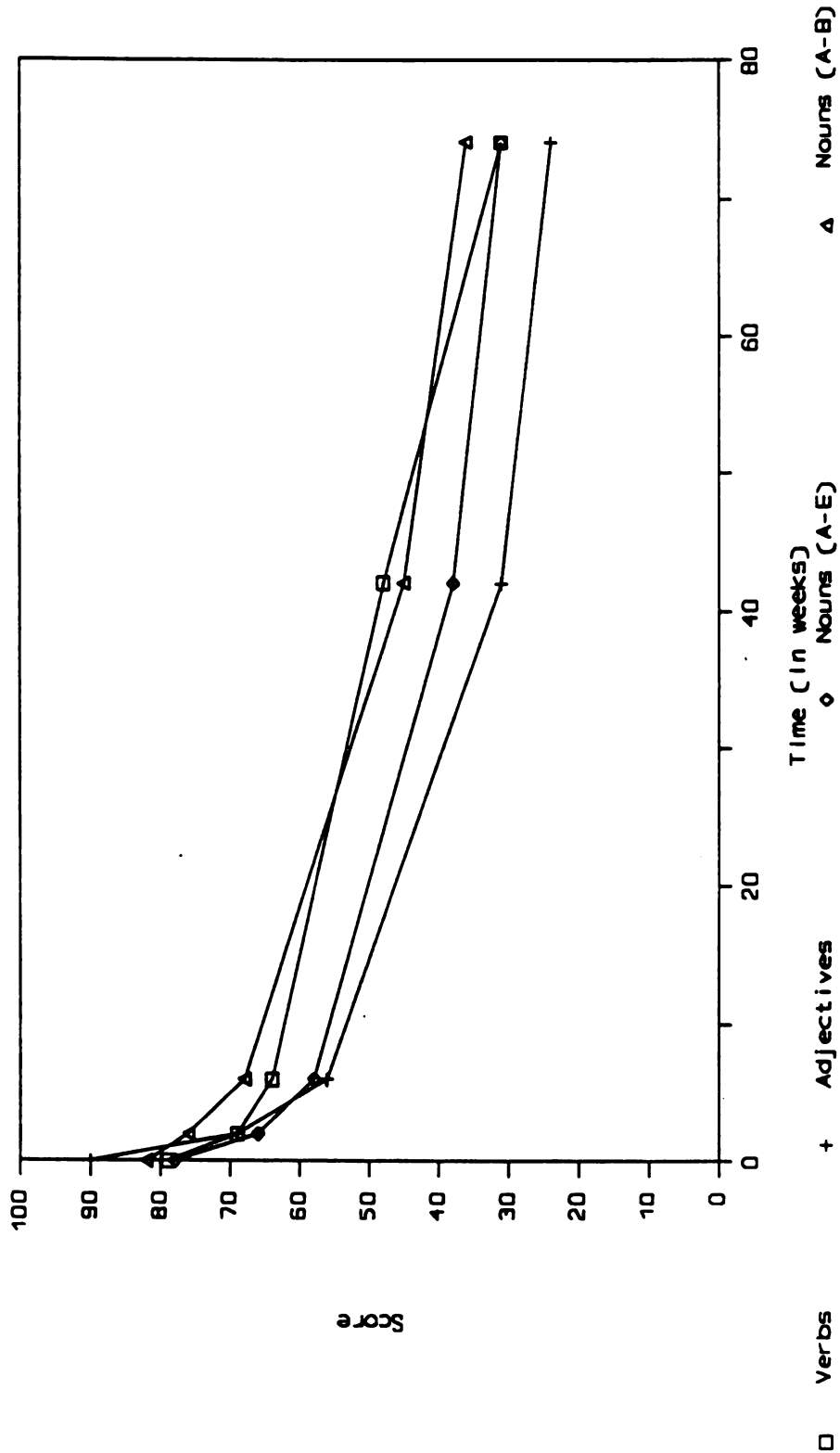
FIGURE 19

RECOGNITION SCORES FOR PART OF SPEECH
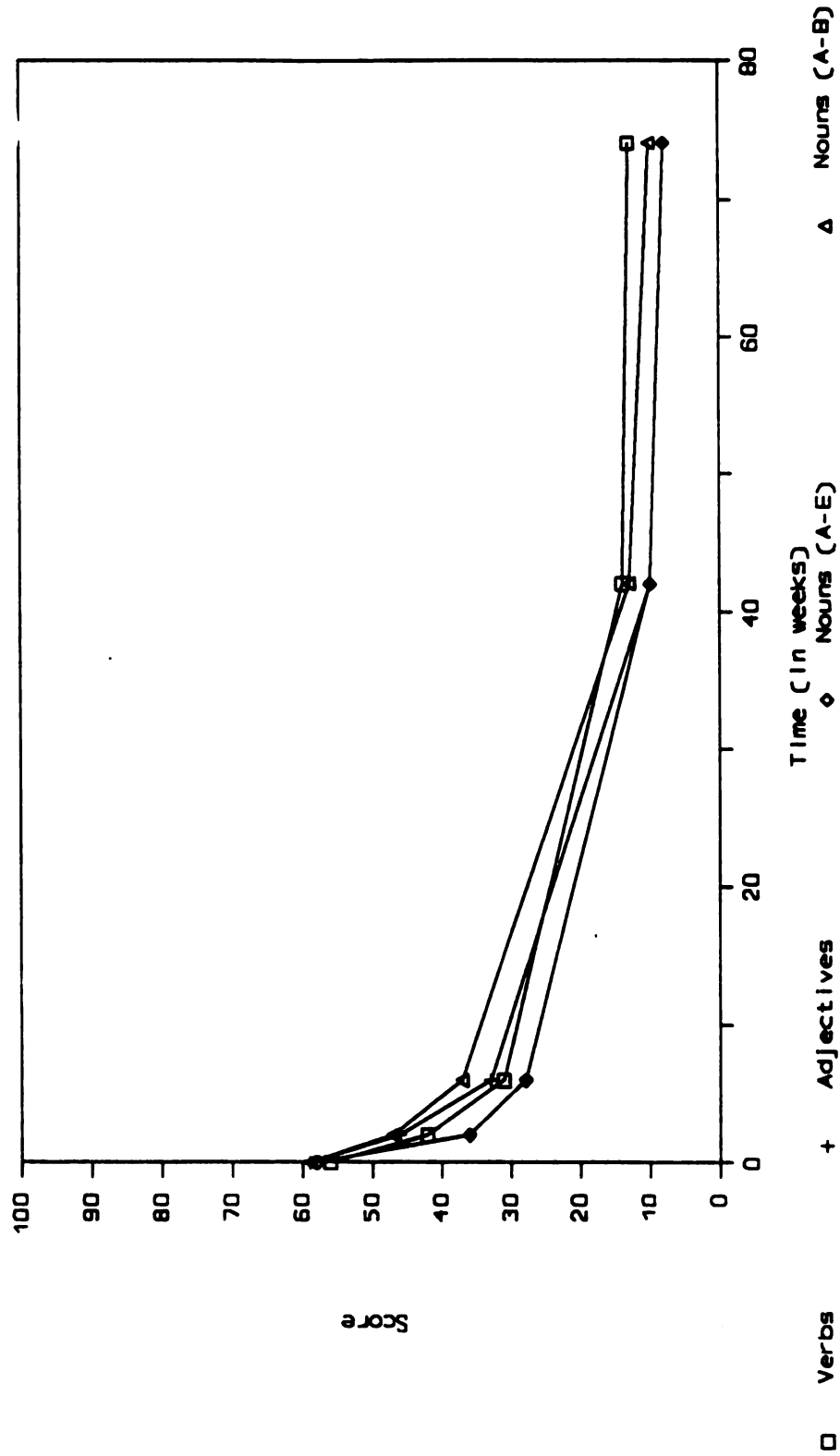
FIGURE 20

PRODUCTION SCORES FOR PART OF SPEECH

Anova tests were also performed to study the interactions between the different independent variables, and we find the same pattern repeated in that part of speech seems to make some difference on recognition, but not on production. The following tables show the Anova tests for the independent variables of 1) part of speech, 2) frequency in input (rated on a scale of 2-8), and 3) whether a word was taught and tested originally for both recognition and production or only for recognition. Table 15 shows the size of the F ratio (the ratio of 'between group variance' to 'within group variance'), and Table 16 shows the probability of obtaining an F value at least that large if all means were equal. The symbols used are S (part of speech), I (frequency in input), and R (taught and tested for both recognition and production or for recognition only).

TABLE 15
ANOVAS FOR PART OF SPEECH, FREQUENCY IN INPUT,
AND RECOGNITION/PRODUCTION: F RATIO

	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
S	.97	–	2.47	.05	1.49	.43	4.41	.71	4.52	1.31
I	1.21	–	.79	1.25	1.17	.47	1.77	1.13	.94	.66
R	9.81	–	39.30	22.24	27.10	17.73	25.02	10.72	25.37	10.50
SOURCE OF VARIATION EXPLAINED										
	2.74	–	5.22	3.97	3.78	2.71	4.27	2.10	3.78	2.19

TABLE 16
ANOVAS FOR PART OF SPEECH, FREQUENCY IN INPUT,
AND RECOGNITION/PRODUCTION: SIG. OF F

	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
S	.39	–	.09	.95	.23	.65	.02	.49	.01	.28
I	.32	–	.56	.29	.33	.80	.13	.35	.46	.65
R	.00	–	.00	.00	.00	.00	.00	.00	.00	.00

SOURCE OF VARIATION EXPLAINED

	.01	–	.00	.00	.00	.01	.00	.05	.00	.04

As can easily be seen, whether a word has been taught and tested originally for both recognition and production explains by far the greatest share of variation in test scores. This factor will be examined in the third section of this chapter on recognition vs. production. Frequency in input again appears to play no role. However, the pattern of part of speech affecting recognition tests, but not production tests, can again be seen.

Tables 17 and 18 illustrate a similar Anova for the following variables: 1) part of speech, 2) frequency in input, and 3) paso (A, B, and C only). Notice that by eliminating Pasos D and E, we have also eliminated the effect of whether a word was taught for both recognition and production or for recognition only, since all words in Pasos A, B , and C were taught and tested originally for both recognition and production. The symbols used are S (part of speech), I (frequency in input), and P (paso).

TABLE 17
ANOVAS FOR PART OF SPEECH, FREQUENCY IN
INPUT, AND PASO (A-C): F RATIO

	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
S	.03	.43	4.50	1.79	4.04	1.56	3.90	1.15	3.82	.55
I	.83	.90	1.25	2.85	2.01	.62	1.67	1.40	.74	1.20
P	3.22	2.21	1.13	5.78	1.34	2.50	.19	.78	.06	.10
SOURCE OF VARIATION EXPLAINED	1.62	1.34	1.40	2.87	1.69	.88	1.62	1.00	1.23	1.00

TABLE 18
ANOVAS FOR PART OF SPEECH, FREQUENCY IN
INPUT, AND PASO (A-C): SIG. OF F

	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
S	.97	.66	.02	.18	.02	.22	.03	.33	.03	.58
I	.54	.49	.30	.02	.09	.68	.16	.24	.60	.32
P	.05	.12	.33	.01	.27	.09	.83	.47	.94	.91
SOURCE OF VARIATION EXPLAINED	.14	.24	.22	.01	.12	.55	.14	.45	.30	.45

Although the probability that the source of variation has been explained is much less in this test, the pattern of the effect of part of speech still holds: significant for the recognition tests but not for the production tests.

Why should part of speech have somewhat more effect on recognition than production, and why should adjectives be the least-well remembered? These questions will be brought up again in Chapter 5.

The previous test brings up the question of whether a word taught earlier or later in the 20 hours under study has a better probability of being retained. This question is clouded by the research design in which words in the first

three pasos were taught and tested for both recognition and production and words taught in the last two pasos were taught and tested only for recognition. As mentioned above and as discussed in the third section of this chapter, the goals for teaching and testing are definitely important. The One-way analysis of variance in Table 19 also clearly shows the difference in scores between Pasos A-C and D-E.

TABLE 19
ONE-WAY ANALYSES OF VARIANCE BY PASO (A-E)

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
F RATIO	4.44	4.63	9.45	9.68	6.12	7.08	3.47	3.23	2.53	2.22
F PROB.	.00	.01	.00	.00	.00	.00	.01	.02	.05	.07
GROUPS SIG. DIFFERENT AT .05 LEVEL (TUKEY)	A/B B/D	A/B	A/D A/E B/D B/E C/D	A/D A/E B/C B/D B/E	A/D B/D	A/D A/E B/D B/E	--	A/E	--	--

It can be noted that most of the significant differences between groups occur between Pasos A, B, and C (which were taught and tested for both recognition and production) and Pasos D and E (taught and tested for recognition only). In the three places where there are significant differences within A, B, and C, the data show that B has the larger mean, and thus no conclusion can be drawn about words learned earlier or later (Appendix B, Table 49). There are no differences either between Pasos D and E. Recall that data for PO are missing in this test for

Pasos D and E, since these words were not tested for baseline production.

Thus there are no data here that would suggest that first learned is first forgotten or vice versa. What these data do suggest is that as time goes on, the difference in how the words were taught becomes less important. Notice that the one-way test at P4 no longer shows significance at the .05 level.

Another factor that was studied was the role of emotionality or saliency of a word in its retention. The scales were obtained from second-year college Spanish students as explained in Chapter 3. Two of the scales, the ones for emotionality and for saliency of meaning, were given in English only, and the third scale for the saliency of the Spanish word was given in Spanish followed by an English translation in parentheses. It should be remembered that the students who performed these ratings were from different classes and had not studied these particular words in Spanish.

A rating of 1 represents highest emotionality and a rating of 5 represents lowest emotionality. Among the highest rated words were pray! (1.83) and tear (1.33—associated with crying), and among the lowest were apron (4.38), board and bowl (both 4.17). On saliency (defined as noticeability, conspicuousness), smart and stupid both rated high (1.81) and balloon rated a 1.70, while apron (4.08), animal back (4.10), and board (4.04) rated low. As might be

suspected, these two ratings correlated fairly strongly with each other (.65); however, the negative correlation that we might expect with test scores (negative because the lower the number of the rating, the higher the emotion or saliency) is extremely low or nonexistent (Appendix B, Tables 50 and 51).

The students who rated the Spanish words as to saliency of form gave ratings such as 1.75 to _tabla_ 'board', 1.71 to _bisabuelo_ 'great-grandfather', and 1.68 to _globo_ 'balloon'. No words were rated a 4, but ratings such as 3.92 were given to _hocico_ 'animal snout' and 3.80 to _dehesa_ 'meadow'. When the Pearson correlation coefficients were determined for these saliency ratings and test scores, the results in Table 20 were noted.

TABLE 20
CORRELATIONS OF TEST SCORES WITH SALIENCY OF FORMS
OF SPANISH WORDS

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
r	-.04	-.28	-.29	-.13	-.29	-.20	-.24	-.09	-.34	-.21
PROB.	.33	.01	.00	.10	.00	.02	.01	.20	.00	.02

The correlation coefficient of this last rating with emotionality is only .11 and with saliency of the English word is only .02. It seems that the form of the Spanish word is somewhat more important than the emotional connotation in determining retention of the word.

It is interesting to compare the means for the saliency of Spanish words by paso to see if any one paso contained words that were more salient than another. Table 21 shows that the means for saliency were quite close in all five Pasos. The most salient words were in Pasos C and E.

TABLE 21
SALIENCY OF SPANISH WORDS BY PASO

PASO	A	B	C	D	E
MEAN	3.06	3.00	2.70	3.08	2.72

The Spanish words were divided into two groups according to whether they were more salient than the mean on this scale (2.96) or less salient, and independent one-tailed T-tests were performed on the scores to see if there was a significant difference between the two groups.

TABLE 22
T-TESTS FOR GROUPS BASED ON SALIENCY OF SPANISH WORD

TEST		N	MEAN	SD	T VALUE	1-TAIL PROB
R0	MORE	38	81.29	15.92	.84	.20
	LESS	57	78.56	15.10		
P0	MORE	30	61.33	17.23	1.84	.04
	LESS	45	54.49	14.77		
R1	MORE	39	72.67	18.47	2.48	.01
	LESS	56	62.46	20.54		
P1	MORE	41	40.07	19.09	.94	.18
	LESS	57	36.25	20.56		
R2	MORE	41	64.20	20.46	2.23	.02
	LESS	57	54.54	21.64		
P2	MORE	41	31.39	18.56	1.25	.11
	LESS	57	26.79	17.56		
R3	MORE	41	41.59	22.36	1.35	.09
	LESS	57	35.53	21.48		
P3	MORE	41	9.76	9.67	-.41	.34
	LESS	57	10.67	11.44		
R4	MORE	41	36.12	18.72	2.67	.00
	LESS	57	25.60	19.59		
P4	MORE	41	9.49	8.72	.62	.17
	LESS	57	8.33	9.44		

The degree of saliency reaches significance on three of the recognition tests but on only one of the production tests. It appears that saliency is more important for recognition than for production.

The list in Table 23 contains the words that were best retained at the last testing. They had recognition scores of at least 50% (at least half the students recognized the word) or production scores of over 15% (at least 2 students could produce the word).

TABLE 23
WORDS BEST RETAINED AT TIME 4

RECOGNITION	PRODUCTION
meñique (50%)	¡rece! (21%)
pulgar (50%)	¡tache!(29%)
mariposa (89%)	mariposa (44%)
ballena (75%)	¡ronque! (25%)
tiburón (85%)	tiburón (32%)
¡exprima! (50%)	¡exprima! (32%)
mohoso (50%)	mohoso (25%)
cuenco (60%)	cuenco (30%)
rama (62%)	rama (19%)
mirlo (67%)	mirlo (15%)
antepasado (64%)	antepasado (17%)
padrastro (53%)	bisabuelo (17%)
madrastra (53%)	madrina (19%)
padrino (67%)	padrino (19%)
aguanieve (87%)	aguanieve (23%)
	llama (30%)

In contrast, the words in Table 24 are remarkable in that <u>no</u> students retained them at Time 4 in either their recognition or production vocabulary:

TABLE 24
WORDS WITH A SCORE OF 0 AT TIME 4

vástago	'descendant'	estanque	'pool'
cerro	'hill'	agujero	'hole'
verja	'window grating'	navaja	'knife blade'

The additional words found in Table 25 had scores of under 20% (remembered by two students or fewer) in both recognition and production at Time 4:

TABLE 25
OTHER POORLY RETAINED WORDS

faja 'belt'	quijada 'jaw'
cerilla 'match'	cascote 'garbage'
hilo 'thread'	aguja 'needle'
huraño 'timid'	delantal 'apron'
sien 'temple' (of head)	manso 'gentle'
agudo 'sharp, smart'	lágrima 'tear' (crying)
beca 'scholarship'	sendero 'path'
vereda 'path'	aldea 'village'
enano 'dwarf'	frasco 'bottle'

When an independent one-tailed T-test is performed on the best-retained words (Table 23) and the worst-retained words (Tables 24 and 25) to learn if their saliency ratings are significantly different, we obtain the results shown in Table 26.

TABLE 26
T-TESTS FOR SALIENCY OF BEST-RETAINED
AND WORST-RETAINED WORDS

GROUPS	N	MEAN	SD	T VALUE	1-TAIL PROB
BEST-RETAINED	21	2.80	.53	-2.88	.00
WORST-RETAINED	24	3.17	.32		

Again the retention rate of the Spanish word seems to be related to.the saliency of its form.

Descriptive word study

When students were interviewed after taking the test at Time 4 about why they were able to remember certain words, these were the responses given for the words in the best-retained list (Table 23):

meñique 'little finger' -- it reminds me of mini; it reminds me of miniature

pulgar 'thumb' -- I remember from class and class tests; reminds me of puny (a thumb is short and stubby); reminds me of pulling (on a thumb)

mariposa 'butterfly' -- it's a pretty word (2 students); reminds me of the Virgin Mary (someone beautiful); beautiful word; posa reminds me of purple (color of butterfly); the m sounds like a flower; I remember the joke you told about the the word butterfly in German and in Spanish (2 students)

ballena 'whale' -- opposite of a small ballerina; it's like a balloon that someone blew up; a balloon as big as a whale; a mellow word; like baleen in English

tiburón 'shark' -- I learned by rote memorization; sounds mean (2 students); reminds me of jaws; reminds me of tiger shark; a different word, not at all typical

¡exprima! 'wring out!' -- I remember doing this in class (4 students); I think of strangling the cloth

mohoso 'moldy' -- I remember picture; reminds me of mold (3 students); simply sticks in my head

cuenco 'bowl' -- I remember bowl in class; this word was drilled in; I see a 5 (cinco) written over the bowl; I had a hard time with this word so I worked especially hard at it; I missed this word once so studied extra hard; I like the word and associate it with cesto 'basket'; I like the sound of the word

rama 'branch' -- reminds me of bracchia; I see a tree branch; it popped into my head

mirlo 'blackbird' -- I remember the picture; you look at a blackbird like you look in a mirror; interesting word; rote memory

antepasado 'ancestor' -- popped into head; made up of two words, ante and pasado; after + passed; ante means before

padrastro 'stepfather' -- reminds me of padre (3 students); must be distinguished from padrino (godfather)

madrastra 'stepmother' -- like madre (2 students); must be distinguished from madrina (2 students)

padrino 'godfather' -- like padre (5 students)

aguanieve 'sleet' -- agua + nieve; I know two parts of the word; agua is water (2 students); like water-rain; water + snow; contains the word snow

¡rece! 'pray!' -- the r reminds me of the pr in pray; I can visualize the word

¡tache! 'cross out!' -- I remember doing in class (3 students)

¡ronque! 'snore!' -- sounds like rude; you caught me doing this during class; I remember from class; I like the word

bisabuelo 'great-grandfather' -- contains abuelo; bi means two

madrina 'godmother' -- like madre (3 students)

llama 'flame' -- like the word 'call' in Spanish (llamar); the ll looks like two matches; looks like flame; struggled with this word

Reasons that students gave for remembering words can be summarized by the following: 1) mnemonic devices relating the Spanish word to the sound of the English word (meñique and mini) or to another Spanish word (cinco written on bowl), 2) linking the form of the word to a physical object

it represents (ll looks like two matches), 3) structural
analysis (bi means two, madrina looks like madre), 4)
recognizing cognates in English (ballena and baleen), 5)
remembering vivid experiences related to learning the word
(actions, jokes), 6) having struggled extra hard with the
word in the past.

It is interesting to note that all but one of these
words (aguanieve) comes from the pasos taught for both
recognition and production. There seems to be no
relationship between the frequency that these words were
repeated in classroom input, but some (such as mariposa,
¡exprima!, ¡ronque!) are definitely linked to vivid
classroom experiences as recalled by the students. Only one
of these well-remembered words is an adjective.

When we look at the lists of the worst-retained words
(Tables 24 and 25), we note that none of these words is a
verb and that slightly more than half (13 out of 24) come
from Pasos D and E, the lessons that were taught for
recognition only (only 23% of total words come from Pasos D
and E).

Another question can be asked of the words that were
well retained. Did words tend to form networks of meanings
so that if one word was retained, another closely-related
word was usually retained?

One closely-related set of words was mirlo 'blackbird',
rama 'branch', ala 'wing', and nido 'nest'. None of the
students in the final interview could remember how they had

been taught these words, but in fact they had been taught together in a story told with pictures in class. Using the pictures from Are you my mother?, a children's storybook in English, the instructor retold the story in Spanish, repeating each of the target words several times. In the final recognition test, there were four students who knew none of these four words, and the data from another two students were eliminated because they had further contact with these words. Of the remaining students, two recognized only mirlo. But seven of the students recognized at least two of these words, and four students recognized all four words. In fact, mirlo and rama ended up in the list of best-retained words, both in production and recognition.

Huella 'footprint' and fango 'mud' are two more words that were closely related. They were taught with a picture of footprints tracking through mud. There were four students who recognized both huella and fango, but no students who recognized only one of the words. One word seemed to bring the other one to mind in close association.

Results for the hypotheses about words

Hypothesis 9: Vocabulary that is repeated more frequently in teacher classroom input is retained at higher levels than vocabulary that is repeated less frequently.

The corresponding null hypothesis cannot be rejected at the .05 probability level. (Tables 10 and 11)

Hypothesis 10: <u>Verbs, adjectives, and nouns are retained at different rates.</u>

The corresponding null hypothesis <u>cannot</u> be rejected at the .05 probability level. (Tables 12 and 13)

There is a tendency for adjectives to be forgotten more quickly, a tendency which shows up more on the recognition tests. (Tables 14-18, Figures 19 and 20)

Hypothesis 11: <u>Vocabulary learned earlier will be retained at higher rates.</u>

The corresponding null hypothesis <u>cannot</u> be rejected at the .05 probability level when applied to data from Pasos A, B, and C. (Table 19)

Hypothesis 12: <u>Vocabulary with high emotionality or saliency ratings is retained longer than vocabulary with low ratings.</u>

High emotionality or high saliency of English meanings is not correlated with high retention scores. (Page 168)

High saliency of the Spanish form is weakly correlated with high retention scores. (Table 20)

Spanish words that have higher saliency ratings also have significantly higher scores on recognition tests. (Table 22)

The best-retained Spanish words have significantly higher saliency ratings than the worst-retained Spanish words. (Table 26)

THE RELATIONSHIP BETWEEN RECOGNITION AND
PRODUCTION VOCABULARY

Statistical study

In order to analyze the relationship of recognition and production vocabulary, the data files organized by both student and word were utilized. Mean scores and standard deviations for times 0, 1, 2, 3, and 4 were calculated for both recognition and production, and the results are displayed in Tables 27 and 28 below:

TABLE 27
DATA ORGANIZED BY STUDENT

TIME	0		1		2		3		4	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
REC	80	13	67	20	59	23	38	26	30	22
PROD	57	21	38	23	29	20	10	9	9	9

TABLE 28
DATA ORGANIZED BY WORD

TIME	0		1		2		3		4	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
REC	80	15	67	20	59	21	38	22	30	20
PROD	57	16	38	20	29	18	10	11	9	9

As can be seen from these tables, the mean scores are the same, whether organized by student over time or by a particular word over time. The standard deviations are somewhat different and can be explained by the two different ways of grouping the data.

It is readily apparent from these tables that the size of both recognition and production vocabulary decreases over time, and that at any one time, recognition vocabulary is greater in size than production vocabulary.

What is the relationship in size between recognition and production vocabulary? Recall that Melka Teichroew (1982) cites reports indicating that for beginners, recognition vocabulary is about twice the size of production vocabulary, but for more advanced learners the gap seems to close. Her statement permits the inference that production vocabulary will be a certain percentage of recognition vocabulary. Bahrick (1982a:116) states that 'attrition affects smaller portions of recognition vocabulary than of recall vocabulary. This difference may reflect the fact that the recognition vocabulary is larger, and that the absolute amount of attrition is approximately the same'. Thus Bahrick presents the view that the absolute numbers of words lost from recognition and production vocabulary are about equal.

In order to study the numerical relationship in loss of recognition and production vocabulary over time, the quotients and differences in Tables 29 and 30 were calculated. The same results were obtained from using either Table 27 or Table 28.

TABLE 29
RATIO OF PROD. SCORES TO REC. SCORES: P÷R

TIME	0	1	2	3	4
RATIO	71%	57%	49%	26%	30%

TABLE 30
DIFFERENCE BETWEEN PROD. AND REC. SCORES: P-R

TIME	0	1	2	3	4
DIFFERENCE	23	29	30	28	21

From Table 29 it appears that the ratio of production to recognition vocabulary does not remain constant during the attrition process. Rather, the ratio of production vocabulary to recognition vocabulary becomes progressively smaller until production vocabulary begins to level off between Times 3 and 4. Instead of a constant ratio, Table 30 gives a different way of looking at the relationship. The difference between the amount of recognition and production vocabulary remains between 21 and 30 words, with the lowest difference being at Time 4 when production vocabulary has nearly leveled off. In other words, the students lost an average of 13 words (out of 100) between times 0 and 1 in their recognition vocabulary, but they also lost 19 words (out of 100) in their production vocabulary. Between times 1 and 2, they lost an average of 8 words in recognition and 9 words in production. The averages for between times 2 and 3 are 20 and 18 words, and for between times 3 and 4 the averages are 8 and 2 words (when production has leveled off). Thus the data support the hypothesis that the same amount of recognition and

production vocabulary is lost over time, rather than that
the ratio of production to recognition vocabulary remains
constant. This relationship is graphed below in Figure 21
(the same as Figure 18 in the first section of this
chapter).
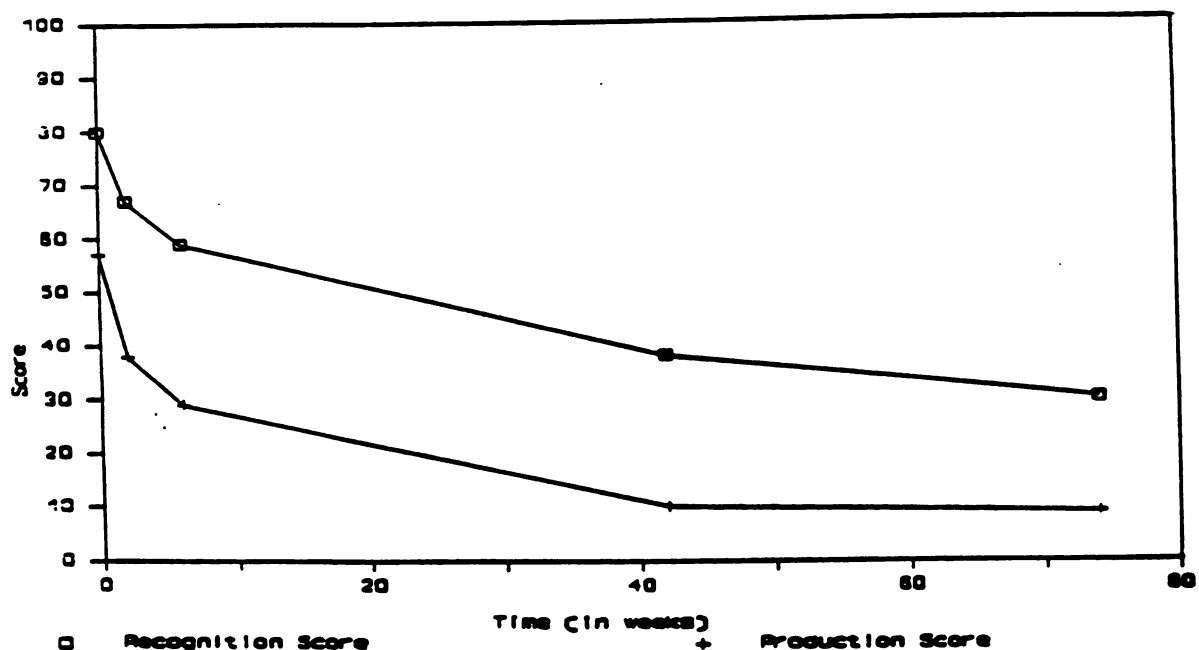
FIGURE 21
COMPOSITE OF RETENTION SCORES (RECOGNITION AND PRODUCTION)

These results confirm Bahrick's view that the absolute
amount of attrition is approximately the same for
recognition and production vocabulary. Attrition affects
smaller portions of recognition vocabulary than of
production vocabulary in the present study, illustrated by
the fact that after 18 months, 37% of original recognition

vocabulary (from Time 0) remains, whereas only 16% of production vocabulary remains. This relationship is illustrated in Table 31.

TABLE 31
PERCENTAGE OF ORIGINAL VOCABULARY (TIME 0)
RETAINED AT LATER TIMES

TIME	0	1	2	3	4
RECOGNITION	100%	84%	74%	48%	37%
PRODUCTION	100%	67%	51%	18%	16%

The graph in Figure 21 also confirms Hypothesis 14 concerning the rate of attrition. The drop in both recognition and production vocabulary is steepest between Times 0 and 1. The slope is not so steep between Times 1 and 2 and begins to level off between Times 2 and 3 and 3 and 4. Production vocabulary seems to be bottoming out by Time 4, although it would be unwise to make predictions about what would happen if students were tested again several months later.

Another important question is one of teaching methodology, stated in Hypothesis 17. Is vocabulary taught and tested for both recognition and production retained at higher rates than vocabulary taught and tested for recognition only? This hypothesis was tested by using independent one-tailed T-tests to find differences in scores for words taught and tested for both recognition and production (Group 1 = Pasos A, B, and C) and for words taught and tested for recognition only (Group 2 = Pasos D

and E). The results from these T-tests are summarized in Table 32 below. Note that Time 0 is not included in the table since Pasos D and E were not originally tested for production.

TABLE 32
T-TESTS FOR WORDS TAUGHT AND TESTED FOR
RECOGNITION AND PRODUCTION AND WORDS
TAUGHT AND TESTED FOR RECOGNITION ONLY

TEST	GROUP	MEAN	SD	T VALUE	DF	1-TAIL PROB.
R1	1	72.12	15.44	5.70	95	.00
	2	48.09	23.02			
P1	1	42.83	18.08	4.94	98	.00
	2	21.91	16.87			
R2	1	63.25	17.50	4.23	98	.00
	2	43.39	26.15			
P2	1	32.73	17.25	4.40	98	.00
	2	15.57	13.03			
R3	1	42.52	20.47	3.77	98	.00
	2	24.17	20.42			
P3	1	12.13	11.07	3.25	98	.00
	2	4.30	5.83			
R4	1	33.10	17.84	3.14	98	.00
	2	19.04	21.90			
P4	1	10.19	9.39	2.99	98	.00
	2	4.00	5.92			

As shown by Table 32, all of the differences in the recognition and production tests are significant at the .01 level, showing that words taught and tested for both recognition and production are retained at higher rates than words taught and tested for recognition only. The results of the Anova shown in Tables 15 and 16 in the second section

of this chapter also support this finding, showing that teaching and testing for both recognition and production carries far more weight in retention than part of speech or frequency in input.

An alternate explanation for the lower retention of words from Pasos D and E is that these words were taught later than the words in Pasos A, B, and C, and that the last learned is the first to be forgotten. Although it is not possible to reject this explanation entirely, it should be noted that for the One-way analysis of variance in Section 2 of this chapter (Table 19), there was no evidence that words taught earlier in Paso A were remembered better than words from Pasos B or C; in fact, the retention scores of words from B had a higher mean than the scores of words from A.

Because saliency of the form of the Spanish word was also a significant factor in retention (see section 2 of this chapter), Anova tests were run on the test scores to see if teaching and testing the word for both recognition and production, in combination with saliency, could account for the variation. In this test, words were coded a 1 if they were above the mean score on this scale, or a 2 if they were below it. The symbols used are S (saliency) and R (taught and tested for both recognition and production or for recognition only).

TABLE 33
ANOVAS FOR SALIENCY OF SPANISH WORD AND TEACHING AND
TESTING FOR RECOGNITION AND PRODUCTION: F RATIO

	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
S	2.12	–	9.95	1.88	6.23	2.76	1.39	.05	5.89	.71
R	8.66	–	35.86	24.48	18.33	20.01	11.81	10.00	9.88	9.31
SOURCE OF VARIATION EXPLAINED	4.69	–	14.55	8.61	7.83	7.32	4.23	3.52	4.92	3.35

TABLE 34
ANOVAS FOR SALIENCY OF SPANISH WORD AND TEACHING AND
TESTING FOR RECOGNITION AND PRODUCTION: SIG. OF F

	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
S	.15	–	.00	.17	.01	.10	.24	.83	.02	.40
R	.00	–	.00	.00	.00	.00	.00	.00	.00	.00
SOURCE OF VARIATION EXPLAINED	.00	–	.00	.00	.00	.00	.01	.02	.00	.02

Together these two factors seem to explain quite well the source of variation in the scores. Notice that the F ratios in the bottom row in Table 31 are higher than those in Table 13, which studied a different combination of factors. Notice also that saliency again seems to be more important for recognition that for production.

A different, but related, question, is whether ability to produce a word initially, rather than ability to recognize a word, is the best predictor of both recognition and production later on. Is a word at Time 0 or at Time 1 more likely to be retained if the student can recognize and produce it at the beginning than if he can only recognize it but not produce it? Remember that it is necessary to consider Time 1 as well as Time 0, since not all words were tested for production at Time 0.

Three different tests were devised for examining Hypothesis 18 statistically from slightly different perspectives. The first test divided words into two groups so that a T-test could be run. In the first group were words that were recognized at a high level (mean of 80% or higher) and also produced at a high level (mean of over 50%) at either Time 0 or Time 1 or both. In the second group were words that were recognized at a high level (mean of 80% or higher), but produced at a lower level (mean of under 50%) at either Time 0 or Time 1 or both. Some words fell into both groups because of widely varying scores at Times 0 and 1, and these words were eliminated from this test. Words that had low production and low recognition scores were also eliminated. Group 1 contained many more words (46 words) than the Group 2 (12 words), since high recognition and high production are so closely correlated. The results of the T-tests are shown below in Table 35.

TABLE 35
T-TESTS FOR WORDS AT HIGH LEVELS OF RECOGNITION
AND PRODUCTION AND WORDS AT HIGH LEVELS
OF RECOGNITION BUT LOWER PRODUCTION

TIME	GROUP	MEAN	SD	T VALUE	DF	1-TAIL PROB
R2	1	66.80	17.70	.74	56	.23
	2	62.33	21.91			
P2	1	39.07	18.50	2.78	56	.00
	2	23.17	13.58			
R3	1	44.59	21.68	.47	56	.32
	2	41.33	19.35			
P3	1	14.37	12.25	1.57	56	.06
	2	8.58	6.96			
R4	1	35.65	19.16	1.20	56	.12
	2	28.33	17.65			
P4	1	12.52	9.89	2.92	56	.00
	2	3.83	5.44			

The results of these T-tests show that there are significant differences (at the .01 level) between these groups of words on the later production tests at P2 and P4, which is not unexpected, since there were large differences between these groups at the beginning in production. But words that had high recognition rates at the beginning, even though they were low in production, continue to hold their own in recognition scores. Although their means are slightly below those of the high-recognition, high-production group, there are no significant differences between groups on recognition tests, even at Time 4. It seems that high initial recognition is the best predictor for high recognition at Time 4, and high initial production is the best predictor for high production at Time 4.

Another way to look at initial production and recognition as predictors of final production and recognition is to study the Pearson correlation coefficients for the scores over time:

TABLE 36
CORRELATIONS OF SCORES FOR 100 WORDS OVER TIME

	R1	R2	R3	R4	P0	P1	P2	P3	P4
RO	.41	.27	.32	.36	.49	.54	.40	.21	.27
R1		.88	.67	.69	.52	.76	.73	.52	.53
R2	.88		.68	.72	.35	.67	.69	.51	.52
R3	.67	.68		.84	.27	.56	.57	.57	.64
R4	.69	.72	.84		.26	.56	.64	.53	.64
P0	.52	.35	.27	.26		.65	.56	.29	.36
P1	.76	.67	.56	.56	.65		.85	.62	.62
P2	.73	.69	.57	.64	.56	.85		.66	.69
P3	.52	.51	.57	.53	.29	.62	.66		.76
P4	.53	.52	.64	.64	.36	.61	.69	.76	

All of these correlations have a probability of .01 of occurring by chance. The strongest correlations are between the same type of test (recognition and recognition, or production and production) on two consecutive testings; for example, R1 and R2 (.88), R3 and R4 (.84), or P1 and P2 (.85) and P3 and P4 (.76). Correlations between recognition tests and production tests given at the same time are also fairly strong: R1 and P1 (.76), R2 and P2 (.69), R3 and P3 (.57), and R4 and P4 (.64).

As time passes, correlations between the same tests weaken; for example, the correlation between R1 and R2 is .88, between R1 and R3 is .67, and between R1 and R4 is .69. The correlation between P1 and P2 is .85, between P1 and P3 is .62, and between P1 and P4 is .61. Although correlations exist with tests at R0 and P0, these relationships are the weakest. The greatest attrition occurred between Time 0 and Time 1, and a word correct at Time 1 had a greater chance of being retained over the 18-month period than a word correct at Time 0.

When we look at the predictability of a later score, we again notice that recognition seems to be a slightly better predictor of recognition, and production seems to be a slightly better predictor of production. For example, the correlation between R1 and R4 is .69, whereas the correlation between P1 and R4 is lower--.56. The correlation between R1 and P4 is .53, and the correlation between P1 and P4 is higher--.61. Note the following pairs of correlations with tests at Time 4:

TABLE 37
CORRELATIONS WITH SCORES AT TIME 4

R1 and R4	.69	R2 and R4	.72	R3 and R4	.84
R1 and P4	.53	R2 and P4	.52	R3 and P4	.64
P1 and R4	.56	P2 and R4	.64	P3 and R4	.53
P1 and P4	.61	P2 and P4	.69	P3 and P4	.76

Thus, ability to recognize a word seems to be a slightly better predictor of later recognition than ability

to produce that word initially. Likewise, ability to produce a word is a slightly better predictor of later production. The correlations between tests at Time 1 and Time 4 are illustrated by the scattergrams in Appendix B (Figures 23-26).

The third perspective on the relationship of recognition and production is given by a multivariate analysis of scores at Time 1 and Time 4, showing how both recognition and production scores at Time 1 can be used to predict recognition and production scores at Time 4. Table 38 gives the multiple regression equations that best predict these final scores.

TABLE 38
BEST PREDICTORS AT TIME 1 FOR SCORES AT TIME 4

FOR FINAL REC.: $\hat{R4} = \overline{R4} + .612\ (R1-\overline{R1}) + .0918\ (P1-\overline{P1})$

FOR FINAL PROD.: $\hat{P4} = \overline{P4} + .0674\ (R1-\overline{R1}) + .229\ (P1-\overline{P1})$

The symbols $\hat{R4}$ and $\hat{P4}$ stand for the predicted values of a certain word on the final recognition and production tests. The symbols $\overline{R1}$, $\overline{R4}$, $\overline{P1}$, and $\overline{P4}$ stand for the means of all words on each of those tests. For example, if we wish to find the predicted recognition score for <u>delantal</u> at Time 4, we need the following information:

```
mean of recognition scores for R4:    30
scores for delantal at R1:            36 (recognition)
                                      25 (production)
mean scores at R1:                    67 (recognition)
                                      38 (production)
```

Substituting these values in the formula for $\widehat{R4}$, we get

$$
\begin{aligned}
\widehat{R4} &= 30 + .612 \ (36-67) + .0918 \ (25-38) \\
&= 30 + .612 \ (-31) \quad + .0918 \ (-13) \\
&= 30 - 19 \qquad\qquad\quad - 1 \\
&= 10
\end{aligned}
$$

Our projected recognition score for delantal at Time 4 is
10, and the actual recognition score for delantal at Time 4
is 13.   The full table for predicted scores and actual
scores is found in Table 42 in Appendix B.

The formulas show that in order to predict final
recognition, the recognition scores carry more weight than
the production scores at Time 1 (.612 vs .0918) and to
predict final production, the production scores carry more
weight than the recognition scores at Time 1 (.229 vs.
.0674).   This result is consistent with the earlier tests
showing that recognition is the best predictor of
recognition and production is the best predictor of
production.

The multiple correlation coefficient of R4 with R1 and
P1 is .693, and the multiple correlation coefficient of P4
with R1 and P1 is .624.   When these quantities are squared,
we see that scores of R1 and P1 explain about 48% of the
variance of R4 and about 39% of the variance at P4.

The results showing the relative importance of recognition and production as predictors when we analyze by word in Tables 36 and 37 are slightly different than the results shown in Tables 8 and 9 when we analyze by student. These results will be taken up again in the conclusions in the final chapter.

Case study

The final table in this chapter is a case study that shows what happens to the vocabulary of one student over time. The student chosen is Jeff, since his attrition curve is so close to the mean curve. An X in the appropriate space shows that Jeff could recognize the word on a certain test (a star indicates that he guessed at the answer). A word that caused confusion on the recognition test is enclosed in parentheses. The Spanish word in the space below the X shows what Jeff actually produced on a specific production test. A horizontal dash indicates that data were not available for that specific slot.

TABLE 39
THE ATTRITION PROCESS

TIME	0	1	2	3	4
pellizque					
ande de puntillas				X	X
rece	rece	X rece	X rece	X rece	X respen
pliegue	piejan				
tache					
amarre	X amarre	X amarre			
cojo	X				
tuerto					
manso	X				
huraño	X				
ocioso	X			*X	*X
agudo	X agudo				
torpe	X				
delantal	X delanta				
faja	X faja	X faja	X faja		
cremallera	X cremallera	X cremallero	X cremalera		
aretes	X	(alhajas)	(alhajas) arrejas		
tacón	X	X	X toquen		
chaleco	X	X			X
alhajas	X alhajas	X	X arreja	*X	
meñique	mennique	X meññique	X mennique	X	X mennique
tez	tez	X tez	X	X teza	
puño	X				
pulgar	X pulgar	X pulgar	(partido) pulgar	X	*X
panza	X				

TABLE 39 (cont'd.)

TIME	0	1	2	3	4
barbilla	X brarillo	X	X		
quijada	X	X	X		
sien	X	X	X		
hocico	X hocio	(horario) hocico	(horario)		
espinazo	espinazo	X espinaro	X espinaza		
garra	X garra	X garra	X garra	*X gantes	*X
pescuezo	X	X	X		
mariposa	X maripasa	X mariposa	X mariposa	X mariposa	X mariposa
ciervo	X cervo			X	
venado	vierbo	X	X		X
ballena	X ballena	X ballena	X ballena	X	X ballena
tiburón	X tiburón	X tiburon	X tiburon	X tarzó	X
exprima	X	exprima	X espiren		
guiñe	X quiñe	quinen			
ronque	X ronque	X ronque	X ronque		
silbe	X silbe				
bostece	X bostece	X	X		
descalzo	X descalzo				
calvo	X calvo	X calvo	X colva	X	
cano	X cano				
mohoso	X mohoso	(manso) mohoso	(manso) mohoso	X mo	*X monado
harapiento	X harapiento	X	X	*X	*X
arrugado	X arrugado	X arragudo	X		
empapado	X empapado				
maceta	X mesata	X meceta	X meseta	X mesa	X

TABLE 39 (cont'd.)

TIME	0	1	2	3	4
tabla	X tabla	X tabla	X tablo	X	X tabla
colcha	X colcha	X colcha	X colcha	X	X
cuenco	X cuenco	X cuenco	X	X	X
rótulo	X	X	X	X	
cartel			X		
pasillo	X pasillo	X			*X
huella	X huella	X huela	X		
fango	X falgo	(falda) falgo	X fuejo	*X	
ala	X ala	X ala	X ala		
rama	X .rama	X	X rama	X rama	X
mirlo	X mirlo	X	X	X	X
nido	X nido	X nido	X		
antepasado	antepasado	X	X	X	X
vástago	X véstado				
bisabuelo	bisabuelo	X	X biabuelos		
tatarabuelo	X tatarabuelo	X tetrabuelo	X tetrabuelos		
padrastro	-- padrastro	padrestro			
madrastra	-- madrastro	madrestro			
padrino	X	X	X	X padrismo	*X
madrina	madrina	X	X	X madrismo	*X
mellizo	X mellizo	X mezenas	X		
lágrima	lagrima	X	X		
beca	beca				
cesto	X ceta	X	X		
incendio	X incendio	X incendio	X	X	X

TABLE 39 (cont'd.)

TIME	0	1	2	3	4
llama	X llama	X			
cerilla	X				
aguanie- ve	X --	X	X	X	X aguanieve
estanque	X --	X	X		
dehesa	X --				
loma	X --				
cerro	X --				
vereda	X --				
sendero	X --				
agujero	 --				
carpa	X --	X carpa	X	X carpa	X
aldea	X --				
verja	X --				
peña	X --				
bolívar	X --	X	X		X
rastrillo	X --				
enano	X --				
azulejo	X --	X azulijas	X azulita		
navaja	X --				
cascote	X --	X			
hilo	X --				
semilla	X --	X silla			
frasco	X --	X	X		
globo	X --	X globo	X globa	X	X
aguja	X --				

Note some interesting features of the data in Table 39:
the general progression of attrition is from 1) recognition
and full production, to 2) recognition and partial
production, to 3) recognition only, to 4) no observable
retention. Jeff shows that his confidence level is not very
high on the recognition of some of the final words by
indicating guesses with a star. Words that are good
examples of this typical attrition pattern are garra,
tiburón, calvo, maceta, and fango.

When Jeff produces part of a word incorrectly, he
always gets the first part right. Notice that even when
words are grossly misspelled, the first letters, and to some
extent the final letters, are preserved, while the middle of
the word suffers the most transformations: garra > gantes,
tiburón > tarzó, mohoso > monado, maceta > mesa, fango >
falgo > fuejo. Mistakes are sometimes repeated, as in the
case of falgo, which was repeated on the next test before
becoming fuejo on a later test. When mistakes occur in
recognition, a word is often mistaken for another that has
the same beginning letters; for example, pulgar was replaced
by partido on the matching test, hocico was replaced by
horario, and mohoso was replaced by manso. Examples such as
these support the belief that words may be 'penciled in'
incompletely in memory, with the first part of the word the
'darkest', and the middle part the 'lightest'. This

phenomenon is similar to the 'tip-of-the-tongue' phenomenon discussed in Chapter 2.

Jeff gives some examples of missing a word on one test and then remembering the word on a subsequent test (for example, see <u>chaleco</u>, <u>venado</u>, <u>ciervo</u>, <u>pasillo</u>, and others). This phenomenon, though not the norm, should indicate that we need to use caution when we say that a student has 'forgotten' a word. Although the student may not be able to access a word for a particular test at a particular time, we cannot prove that the word has disappeared entirely from memory and that it cannot be accessed at a future time.

Results for the hypotheses about recognition and production

Hypothesis 13: Recognition scores for vocabulary are higher than production scores.

The graph in Figure 21 shows this hypothesis to be true for the current research.

Hypothesis 14: Both recognition and production scores for vocabulary drop quickly at first and then gradually level off.

The graph in Figure 21 shows this hypothesis to be true for the current research.

Hypothesis 15: The best description of recognition and production vocabulary loss is that equal quantities of words

are lost over time rather than that the ratio of production to recognition vocabulary remains constant over time.

The graph in Figure 21 shows this hypothesis to be true for the current research. This relationship is also illustrated by Tables 29 and 30.

Hypothesis 16: Smaller percentages of recognition vocabulary are lost than of production vocabulary.

Table 31 shows this hypothesis to be true for the current research.

Hypothesis 17: Vocabulary taught and tested for both recognition and production is retained at a higher rate than vocabulary taught and tested for recognition only.

The corresponding null hypothesis can be rejected at the .01 probability level (Table 32).

Hypothesis 18: Ability to produce a word initially, rather than ability to recognize a word, is the best predictor of both recognition and production later on.

A corresponding null hypothesis states that words that are recognized and produced at high levels initially will not have higher scores on later retention tests than words that are initially recognized at high levels but produced at lower levels. This null hypothesis cannot be rejected, since there were no differences at the .05 probability level

between the two groups of words on later <u>recognition</u> tests (Table 35).

Correlations between scores and multivariate analysis also showed that early recognition seems to be the best predictor of later recognition, and early production of later production (Tables 36-38).

CHAPTER 5

SUMMARY AND CONCLUSION

.

The main focus of this study has been to investigate
the attrition of foreign language vocabulary by examining
factors such as student characteristics, lexical
characteristics, and the relationship between recognition
and production vocabulary over time. This final chapter
will provide a summary of the research, discuss the findings
and their implications for second language acquisition
theory and foreign language teaching, and make
recommendations for further research in the area of foreign
language attrition.

SUMMARY OF RESEARCH

Fifteen college-level students of first-semester
Spanish were studied intensively during their first 20 hours
of the course in order to track their acquisition of 100
specially-targeted words in Spanish. During the following
18 months these students were retested on these 100 words
for both recognition and production in order to study the
pattern of foreign language vocabulary attrition.

Research hypotheses were grouped into three categories:
1) hypotheses about students, 2) hypotheses about words, and

3) hypotheses about the relationship of recognition and production vocabulary. Most of these hypotheses were tested through statistical procedures such as T-tests, One-way analyses of variance, Anovas, and correlations. Some of the hypotheses, however, were studied by means of descriptive data, and individual cases were also considered.

An important component of the research was to gather data from students taught in actual classrooms under ecologically realistic conditions rather than in a laboratory situation. Although conducting a longitudinal study of this scope involved some problems of control, it was felt that this type of study could provide important data about what happens to real foreign language learners in their very real battle with vocabulary attrition.

FINDINGS AND DISCUSSION

Before discussing the findings, the limitations of the research should be reviewed. As mentioned in Chapter 1, the sample size of only 15 students is small, which must be taken into account especially when discussing the first eight hypotheses which are based on the data file organized by student. The results for Hypotheses 9-18 are based on the data file organized by word containing 100 items, and since n is larger, statistical significance is easier to achieve.

Three factors may have influenced the scores on the
retention tests: 1) possible post-course contact with
vocabulary items, even though the researcher attempted to
eliminate the influence of all such contact, 2) repeated
tests over the same vocabulary items, even though these
tests were spread far apart in time, and 3) the knowledge of
the students that they were involved in some kind of
vocabulary research. Any of these three factors could have
raised retention scores and made the attrition appear less
severe.

Student characteristics

Hypotheses 1-8 dealt with student characteristics.
Students were grouped as to whether they took the class in
the United States or Mexico, how much pre-course and post-
course contact they had with Spanish, sex, type of
motivation and effort, and grades they received for the
course. No differences at the .05 level were found for
these different groups with regard to their retention
scores. Graphs showed, however, that students who finished
the course with higher grades did retain slightly more
vocabulary than students who finished the course with lower
grades. Also, strong correlations were found among
retention tests over time, showing that students who scored
higher on the first retention tests also scored higher on
the final retention tests. Production scores correlated

slighty higher than recognition scores with both final
production and recognition.

Individual case studies of students showed that the
students who retained the most vocabulary were ones who had
rich capabilities for imagery, who analyzed words
structurally, and who looked for relationships with other
words. Those who did poorly on retention tests did not make
as much use of these memory aids. Also, interference from
lexical items from other languages or from another dialect
of Spanish seemed to play a role in reducing the retention
rate of the target words.

When evaluating these results, it was somewhat
surprising to find that there were no differences between
the students who were taught in the United States or Mexico,
or between those who had pre-course or post-course contact
with Spanish and those who had none. However, the set-up of
the research required that the 100 target words be heard and
learned only during the first 20 hours of the course, and
that students have no further contact with these words.
Thus, the research conditions prevented reinforcement of
vocabulary caused by previous or later contact with Spanish
either in Mexico or in subsequent Spanish classes.

It would be very hard to deny that learning Spanish in
a native-language environment and having subsequent contact
with Spanish language and culture are very valuable
influences on language retention. But the present research

seems to indicate that rather than influencing vocabulary retention in a general way, these environmental factors have a more specific result--to reinforce specific vocabulary items. When these vocabulary items are missing from input outside of the course, there can be no reinforcement from environment, and the vocabulary items will attrite quite rapidly. In fact, a target-language environment may even increase the attrition rate of non-rehearsed vocabulary by adding interference from other words, as in the case of Laura.

It was also surprising to note that there was no difference in retention scores based on grouping by type of motivation (extrinsic or intrinsic) or amount of effort reported by the students. This finding on motivation, though, is consistent with data from Edwards (1976, 1977) and Gardner (1985). Gardner states that attitudinal/ motivational factors seem to be more influential during the acquisition period than during the retention period. The case of Karen in the present study is representative of the good student who works hard in class and hopes to use Spanish in a future career, yet who is able to access only a very few words after 18 months.

The data in the present study confirm the report by Pratella (1969) that sex seems to have no effect on retention rates.

The question of whether course grades can predict retention rates remains open. Course grades are based on

many different factors, one of which is initial acquisition
of vocabulary. Thus, the T-tests in this research at Time 0
show a significant difference between students who received
higher grades (A, A-, B+) and students with lower grades (B-
or lower). However, this significant difference disappears
by Times 2, 3, and 4, even though the graphs continue to
show a small difference between groups, especially on the
recognition tests. Godsall-Myers (1981) and Bahrick (1984)
found that higher grades received in course work were
related to higher retention rates, but Clark and Jorden
(1984) reported that there was no difference in initial ease
in learning Japanese for 'attriting' and 'non-attriting'
students. In the present research, course grades gave some
indication of future retention rates, but there was also a
great deal of individual variation.

Retention scores at Time 1 were a better indicator of
future retention rates than were course grades. It is
interesting that Meara (1989) suggests that researchers use
retention rates between Weeks 2 and 3 of the retention
period as the best predictors of future retention. The
strong correlations among scores in the present research
show that once students have shown themselves to be good
vocabulary 'retainers' at Time 1 (about 2 weeks after their
last contact with the target vocabulary), they continue to
get high scores on later retention tests. The highest
correlations occur between subsequent tests in the same mode
toward the end of the retention period (for example, the

correlation coefficient between R3 and R4 is .92, and between P3 and P4 is .93), but the correlation coefficients between R1 and R4 (.70) and between P1 and P4 (.72) show that the scores at Time 1 can be good early predictors of future retention.

Whether production or recognition scores are the best predictors for future retention is really too close to call from the data organized according to student. Although the correlation coefficients slightly favor production, it must be remembered that this data file is organized around only 15 students, and the correlation coefficients are extremely close. This question will be discussed later in this chapter under the section on recognition and production.

Word characteristics

Hypotheses 9-12 dealt with characteristics of the 100 target words. The data showed no differences in retention rates for words the instructor repeated with higher or lower frequency during class time or for words that were presented earlier or later. Differences in retention rate for part of speech were noted, caused by the more rapid loss of adjectives than of nouns or verbs.

The 100 target words had been given ratings in the following three categories: 1) emotionality of the English translation, 2) saliency of the meaning of the English . translation, and 3) saliency of the Spanish form. Saliency

is defined as the ability of a word to attract attention due to its conspicuous, striking, or unusual nature. The first two ratings failed to correlate with retention scores; however, the third rating showed weak correlations with these scores. Moreover, when lists were compiled of the best-retained and the worst-retained Spanish words, the words on the best-retained list had significantly higher ratings in the third category above than the words on the worst-retained list.

Recognition tests rather than production tests were more sensitive in detecting differences based on part of speech or saliency rating. It is also interesting to note that on the graphs illustrating the retention rates for students who received higher grades and those who received lower grades, the recognition tests also showed greater differences between groups.

Students were asked individually to describe the devices they had used to remember the Spanish words on the best-retained list. Mnemonic devices had frequently been used, such as relating the Spanish word to the sound of an English word or linking the form of the word to a physical object. Students also mentioned using structural analysis, struggling extra hard with a word they had missed in the past, and remembering vivid experiences related to the learning of the word in the classroom. Also, students tended to remember certain words (such as <u>mirlo</u>, <u>nido</u>, <u>ala</u> and <u>rama</u>) as part of a network.

Several interesting observations can be made about these findings on word characteristics and the way in which the words were taught. In the present study, the most influential factor on retention was whether the word was taught and tested for both recognition and production (and this factor will be discussed further in the next section), although the saliency of the Spanish word also appeared to play a role. Judgments about greater saliency were made by Spanish students outside of the main study and seemed to be based on shortness of the word, similarity of the form to English (even when not a cognate), lack of diphthongs and accent marks, and ease of structural analysis. Within the main study, subjects made comments related to saliency during their individual interviews: the Spanish word reminded them of the sound of an English word or phrase or of the actual physical object that the word represented, or the structure of the word was easy to analyze as in aguanieve and madrina. And, the instructor played an important role in making words more salient by associating the words with vivid images or actions during class; the students commented that they remembered some words because of an action, story, or joke they had experienced in class.

Such experiences are related to episodic rather than semantic memory. According to Bahrick (1979), episodic memory refers to memory of particular events related to time and place, whereas semantic memory refers to general rules. Schouten-Van Parreren (1989:82) reports that even weak

pupils could remember a difficult word that they had encountered only once because they had enjoyed the story in which it had occurred.

Murray (1986) states that variables such as imagery and emotionality affect retrieval from episodic memory. Therefore, it was a surprise to find that the ratings for the emotionality of the English translations of the target words had extremely low or nonexistent correlations with retention scores (Appendix B, Table 50). For the present research, emotionality seemed to be based on what the instructor and the student did with the word when it was presented in class, rather than on characteristics of the English translation.

Murray (1986) also states that variables such as word frequency have the greatest effect on semantic memory. The frequency of repetition of words during the first 20 hours of the course may simply have been too limited to show any correlation with retention scores, or to influence semantic memory to a great degree. Words were remembered because of their saliency in input and because students had to study them for both recognition and production, not because of the number of times students heard them in input during the first 20 hours of class. When we consider the information from Nation (1982) about the number of times a word should be repeated in input (seven for intentional learning and sixteen for incidental learning), we must keep in mind that Nation is not referring to retention over long periods of

time. It appears that for frequency of repetition to be an important factor in retention of vocabulary, words must be repeated far more than sixteen times, a feat that becomes increasingly difficult for an instructor to perform as total amount of vocabulary increases.

Another factor which was not significant in the present research was whether a word was learned earlier or later in Paso A, B, or C. Since the words in these three pasos were taught in a matter of a few days, there was not much spread in time between words taught earlier or words taught later. Thus, there is insufficient evidence to make a judgment on the regression hypothesis for vocabulary retention.

A factor which did seem to play a small role was the part of speech of the word, a factor which showed up on recognition tests but not on the less sensitive production tests. Although adjectives scored better than nouns and verbs on the original tests, their scores trailed on later tests. This result was unexpected because research in language loss in aphasia (Obler 1982) and in loss of a second language (Cohen 1986) indicates that nouns are more easily lost than other parts of speech. Cohen comments that because nouns are easily learned, they are also easily forgotten. Possible explanations for the results in the present research are that 1) adjectives were the only words during the first 20 hours of the course to show any morphological variation and 2) nouns and verbs are the major constituents of the Spanish sentence and therefore are

essential to comprehension and expression; adjectives are somewhat peripheral because they are modifiers.

Sapir (1921:117) pointed out the primacy of noun and verb: 'There must be something to talk about and something must be said about this subject of discourse once it is selected.' More recently, categorial grammar has presented the concept of syntactic rank; in this analysis of part of speech, nouns are of a higher rank than both verbs and adjectives, since both verbs and adjectives say something about nouns (Lyons 1977:439). It would be interesting to apply this concept to further research on the role of part of speech in second language retention.

Recognition and production

Hypotheses 13-18 dealt with the relationship of recognition and production in vocabulary retention. Graphs showed that recognition scores were higher than production scores at all points in time. Both recognition and production scores dropped the most quickly at the beginning and then gradually leveled off over time. By the end of the 18-month period, production scores had nearly leveled off and were close to 0 for many of the weaker students, while recognition scores were still continuing to decline.

The best description of the relationship of recognition and production vocabulary loss is that equal numbers of words are lost from each type of vocabulary over time.

Because production vocabulary represents a smaller absolute number of words, the percentage loss from production vocabulary is greater than the percentage loss from recognition vocabulary. Thus the ratio of production vocabulary to recognition vocabulary is constantly decreasing, until production vocabulary 'bottoms out' while recognition vocabulary continues to decline.

There was a significant difference in retention scores for vocabulary that was taught and tested for both recognition and production, and vocabulary that was taught and tested for recognition only. However, for predicting later recognition and production scores, early recognition scores were slightly better for predicting recognition, and early production scores were slightly better for predicting production.

Data from individual students showed that a typical attrition order was a) recognition and full production, b) recognition and partial production--with the first letters of the word being the easiest to produce, c) recognition only, and d) no observable retention. Occasionally, words that a student failed to access on a retention test reappeared on a later test, making it impossible to judge at any point whether a certain word was completely erased from memory, or whether the difficulty was in accessing the word.

There are many observations about the relationship of recognition and production that can be made from the data.

Recognition was the easier skill, and scores on recognition were always higher than scores on production. In order to score well on production, a student needed more information than simply the first and last sounds in a word. The descriptive data from the production tests showed that students may 'pencil in' words in memory, with the first letters being the 'darkest' or easiest to remember. The middle parts of words were the first to blur over time. When words were confused on the recognition (matching) tests, the two words mixed up were often words that had the same initial sounds, but different middle parts (such as hocico and horario or mohoso and manso).

By far the most influential factor in determining retention scores was whether a word had been taught and tested for both recognition and production or for recognition only. Requiring students to study for a production test (after a recognition test) during the first 20 hours of the course most likely forced students to pay attention to the middles of words, as well as to review the words again from a different perspective. This different perspective can be related to Craik and Lockhart's depth of processing theory (1972), in which learners process vocabulary on different levels, thereby increasing their retention. Recognition and production skills seem to be different enough that they require processing a word at different levels, and when students are required to go the

extra step to produce a word, they add another level to their analysis of a word.

Recognition testing was more sensitive for judging lower levels of learning. When some of the students scored close to 0 on their production tests, their recognition tests still showed evidence of some retention. In the present research, recognition tests were also more sensitive than production tests for showing differences for students with higher and lower course grades, for part of speech, and for saliency rating of the Spanish word.

In terms of absolute numbers, roughly as many words were lost in recognition vocabulary as in production vocabulary. This result is difficult to explain, but Bahrick (1984c) also reports the same phenomenon. Thus, it does not appear that production vocabulary remains a constant ratio of recognition vocabulary, but that attrition affects larger percentages of production vocabulary than of recognition vocabulary. Production vocabulary that is 'forgotten' is usually not lost entirely, but rather becomes recognition vocabulary. Near the end of the attrition process, production vocabulary appears to 'bottom out' first while recognition vocabulary continues to decrease. The data from this research show that at the end of the 18-month period the ratio of production to recognition vocabulary increased slightly.

Retention rates for both recognition and production vocabulary fell rapidly at first and then gradually leveled

off. These results are consistent with both Bahrick (1984c) and Weltens et al. (1989), though vocabulary decline in the present study was more drastic. Bahrick's data show that attrition continues for up to 6 years and then what is retained at that point continues to be remembered in 'permastore' for at least 25 years.

In contrast to the present research, Bahrick's data base (and Weltens' as well) included groups of individuals who had reached a high level of proficiency. Bahrick remarks that content from a single course is not likely to endure at all, a point also made by Clark and Jordan (1984). In his 'critical threshold' theory, Neisser (1984) states that information must be tied into an extensive cognitive structure if it is to be resistant to forgetting.

This kind of critical threshold cannot be supplied by a single language course or by a few repetitions of vocabulary. The data from the present study show that after 18 months, eleven of the fifteen students were producing less than ten percent of the vocabulary that they had been taught, and five students were recognizing ten percent or less. Only four students could produce more than twenty percent of the words and recognize more than half. It is difficult to predict what would have happened if the testing had continued longer than eighteen months, but for most students it is probable that little vocabulary would remain over a longer period of time or ever reach Bahrick's 'permastore' state.

In the earlier section on student characteristics, the correlation between earlier and later retention scores was discussed and the question was raised of whether early recognition or production scores are the best predictor for future retention. This same question can be examined from a different perspective by studying the data file organized by word. Again the correlations are quite close, but this time the data file is organized around 100 items instead of just 15. The correlations are not as high overall as in the data file organized by student (for example, the correlation coefficient between R3 and R4 is .84 and P3 and P4 is .76, rather than .92 and .93), showing that a strong word is not quite as likely to continue with a high score as a strong student is to continue with a high score. The correlations from this data file slightly favor recognition scores as a predictor of later recognition, and production scores as a predictor of later production. Thus there is no clear-cut indication of the superiority of production over recognition for predicting future retention scores.

That production is not necessarily superior to recognition for predicting future scores is borne out by the other tests (see Tables 34 and 37). It seems that production and recognition skills, although very related, still tap different learner capabilities.

IMPLICATIONS OF RESEARCH

For second language acquisition theory

An overview of current second language acquisition
theory was presented in Chapter 1, with discussion centering
around Krashen's acquisition/learning distinction and the
relationship between comprehension and production and the
value of output. In Chapter 2, cognitive theory was
discussed with its emphasis on how the language learner
proceeds from rule-bound declarative behavior to an
automatic procedural stage.

According to Krashen's Monitor Model, acquisition is a
subconscious process that results from meaningful
interaction in a communicative setting. Learning, on the
other hand, is a conscious process that comes about when the
focus is on the form that the language takes.

Cognitive theory, on the other hand, sees the
acquisition/learning distinction as unnecessary. Even
automatic processes can be brought to consciousness.
O'Malley et al. (1987:302) state that 'a cognitive model of
language learning sees active conscious processes involved
in all language settings, at least in the initial stages of
learning. These processes can be described and used to
assist learning instead of relegated to inaccessible
unconscious mechanisms'.

The data in the present study provide evidence that adult second language learners are highly aware of the process of vocabulary acquisition, even when the Natural Approach is used in the classroom. Although all words were presented in Spanish during the research through a comprehension approach, students wrote down accompanying English translations in their notebooks without prompting from the instructor. Several students, on their own initiative, made flashcards with English translations on the back, even when the instructor urged them to avoid English or to use pictures. Also, the formation of mnemonic devices by the students was a highly intentional activity. The descriptive data from the research show that some students were still using these same devices eighteen months later.

In terms of cognitive theory, the fairly rapid attrition of vocabulary in this study showed that students had not yet turned their declarative knowledge of words into procedural form. This failure can be attributed to the short duration of the learning period, in which the students did not have adequate time to strengthen their vocabulary skills. Although both the Monitor Model and cognitive theory strongly encourage classroom practice with communicative activities, only the cognitive theory includes an explanation for the rapid attrition of foreign language vocabulary.

The data collected in this research have a good deal to say about what constitutes high quality teacher input and

consequent learner intake. Whether or not a student retains a word from a beginning language class does not seem to depend on if the teacher repeats the word 5 or 20 times, but rather on the saliency of the foreign language word. This saliency can depend on the form of the Spanish word itself (its concreteness, its length, its morphological transparency, or its similarity to an English sound), or saliency can depend on what the teacher does to make the word conspicuous by presenting a picture or object, by telling a story or joke, by having a student perform an action, or by tying the word into a network of meaning. The students can also make the word salient for themselves by constructing memory devices such as keywords, by going back to study a word that was especially difficult before, or by analyzing the structure of a word.

The data contain evidence that recognition and production, although closely related, are somewhat different skills in vocabulary acquisition and retention. High recognition predicts high recognition later on and high production predicts high production later on. There is no evidence from the data that recognition turns into production over time (without further input), whereas production can disappear and recognition remain as the only proof of retention.

The data also show that whether or not a word has been taught and tested for recognition and production is an influential factor in determining the retention of the word

over time. The mean scores for such words were significantly different at the .01 level from the mean scores for words taught and tested for recognition only. Forcing students to study for production on tests seems to push them to analyze words at a deeper level, thus aiding in the general retention of these words.

Thus, while the importance of comprehensible input is not challenged, the data point to the additional importance of output in the long-term retention of vocabulary. The data also point to the role of consciousness in initial vocabulary learning—even when the words are presented through the 'Natural Approach'—and emphasize the need to convert declarative knowledge to procedural knowledge if long-term retention is to be achieved.

For foreign language teaching

At the conclusion to a dissertation in the field of second language acquisition, readers have the right to ask 'so what'? Although there are implications for theory, there should also be implications for pedagogy. As cautioned by Hatch (1978), results of research must be applied with care to the classroom. Nevertheless, the following recommendations for the classroom seem to be supported by the data in this research:

1. The saliency of a word in input is an important factor for later retention. The instructor can increase saliency of vocabulary by using imagery, vivid experiences, structural analysis, and by encouraging students to use their own mnemonic devices to remember words. This implies teacher training in methodology, but also in language itself, so that the instructor will have the language skills adequate to the use of imagery, narration, and structural analysis in the classroom.

2. Vocabulary from a single course is rapidly forgotten. Students should be encouraged to take a series of courses in the same foreign language, rather than to 'sample' a foreign language for a short period of time. Teachers and curriculum planners should remember that students need to form networks of meanings tied into an extensive cognitive structure and that isolated words studied for a short period of time are excellent candidates for attrition.

3. Although this research in no way discredits a comprehension-first approach to language learning, it indicates that vocabulary production is also important if students are to retain words over a long period of time. Production forces the students to pay attention to all parts of a word and to analyze the vocabulary at a different level, thus increasing the probability of long-term retention.

4. If the lexicon plays a crucial role in foreign language development, then attention must be focused on vocabulary acquisition in the classroom. A large vocabulary for comprehension is important, but if long-term retention is desired, some vocabulary size may have to be sacrificed in order to allow time for making vocabulary sufficiently salient, for practicing the vocabulary that has already been introduced, and for encouraging production of vocabulary that students can already recognize.

5. Vocabulary testing should reflect course goals for both recognition and production. Recognition testing requires lower levels of learning, but seems to be the most sensitive measure of retention, especially for beginning students. Production testing encourages students to process words at a deeper level and aids long-term retention.

RECOMMENDATIONS FOR FURTHER RESEARCH

Although research in second language attrition exploded during the decade of the 1980s, much more needs to be done. The following areas of research are suggested specifically by questions raised in this dissertation.

1. What is the best predictor for future language attrition--early recognition or early production scores? Although the data from the present research suggest that

recognition is the best predictor for recognition and production for production, the research should be repeated with a larger data base.

2. When is the optimal time for predicting future attrition? The correlations in this research suggest that scores taken 2 weeks after the end of the contact period are good early predictors, and Meara (1989) suggests using the retention rate between Week 2 and Week 3 of the retention period. Are there better points in time for making predictions?

3. Can the ratio of production vocabulary to recognition vocabulary predict the state of acquisition or attrition for an individual?

4. Why are recognition and production vocabulary lost in roughly equal amounts over time?

5. Are certain parts of speech more likely candidates for attrition than other parts of speech? Why in the present study were adjectives more quickly forgotten than nouns or verbs? Can this finding be related to categorial grammar?

6. Is the regression hypothesis true for vocabulary acquisition and attrition?

7. How can the acquisition process be tied to post-classroom maintenance?

8. Is forgetting equivalent to an inability to access information because of interference, or to the complete extinction of a memory trace for a word?

9. How does vocabulary attrition in a foreign language fit into a general theory of language attrition?

In the answers to questions like these lies the hope of understanding more completely the process of language attrition, which is a small step toward understanding human language in general and a smaller step toward understanding ourselves. The challenge for the foreign language profession today is not only to understand how best our students can acquire a foreign language, but also how best they can retain it.

APPENDIX A
STUDENT MATERIALS

STUDENT QUESTIONNAIRE

1. Name _____

2. Age _____   3. Class rank: F   Soph   J   Sen   Grad

4. Current address   Phone _____
   _____
   _____

5. Permanent address   Phone _____
   _____
   _____

6. Have you spoken Spanish at home with relatives? Explain.

7. Have you studied Spanish before? How long and where?

8. Have you studied other foreign languages? Which ones, how long, and where?

9. What contact have you had with Spanish outside of the classroom? Explain please.

   -travel to Spanish-speaking countries (when, where, how long?)

   -contact with Spanish-speaking people in U.S. (describe)

   -listening to radio, TV in Spanish or reading books, newspapers in Spanish (how often?)

10. Why do you want to study Spanish?
    ___ Graduation requirement from college
    ___ I like the way it sounds
    ___ Interest in Latin America
    ___ May be useful in my future career in _____

    Write more about why you are in this class:

11. What specific things do you want to do and to learn in this class?

CONSENT FORM

The research being done by Professor Marilyn Bierling is part of a project for a doctoral dissertation on vocabulary acquisition and retention. There are important questions in the area of second language acquisition that affect all language students, questions such as:

-What is the effect of an intensive course on vocabulary development?
-What is the difference between a Spanish class taught in an English-speaking environment and in a Spanish-speaking environment?
-What is the relationship between the teacher's output of Spanish in class and what the student actually acquires? What stages do students pass through in acquiring vocabulary? ·
-What is the relationship between being able to recognize a word in a foreign language and being able to produce it?
-At what rate is foreign language vocabulary forgotten?

Data for this research will come from your normal written assignments in the workbook, your written tests, and your oral interviews. All of this material is required in the regular 101 courses that the professor teaches. By giving your consent, you allow the professor to use this data for research after the 101 course has been finished and your final grade has been given. If you have given your consent, she may later ask if you would like to participate in a retention study which would involve short testing of vocabulary retention through the fall of 1989.

Data for the research will also come from tape recordings of class meetings. These tape recordings will not be available to the professor until after the final grades have been given, and only the recorded utterances of students who have given their consent will be used in the study.

The research has been explained to me and I understand the basic purposes and procedures. I freely consent to participate and understand that at any time I may withdraw my consent without penalty. I understand that the Spanish 101 professor will not know if I have consented or not until after she has given the final grades for 101, and only after that time will she analyze any of the data.

I understand that all results will be treated with strict confidence and that all subjects will remain anonymous; upon request, results will be made available to me as long as the anonymity of other subjects can be maintained.

Name_____

Date_____

SAMPLE ITEMS FROM RECOGNITION TESTS (FIRST 20 HOURS)

<u>Multiple choice</u>
Escoja la respuesta más lógica.

1. Un hombre con un ojo es
   a. rubio.   b. grande.   c. tuerto.   d. ocioso.

2. Un insecto muy bonito es
   a. la faja.   b. el gato.   c. el chico.   d. la mariposa.

3. _____ es parte de la mano.
   a. El pulgar   b. La sien   c. La oreja   d. El chaleco

4. _____ es parte del zapato.
   a. El traje   b. La tez   c. La cremallera   d. El tacón

5. Un hombre con una pierna es
   a. cojo.   b. torpe.   c. verde.   d. joven.

6. Un hombre calvo no tiene
   a. cuaderno.   b. pelo.   c. secretaria.   d. huella.

7. ¿Dónde está el cartel?
   a. en la pared   b. en la tiza   c. en el ejemplo   d. en el piso

8. La planta está en
   a. la pizarra.   b. el nido.   c. la maceta.   d. la rama.

9. <u>Gemelo</u> es un sinónimo de
   a. mellizo.   b. casado.   c. apellido.   d. primo.

10. Cuando hay un incendio, hay muchas
    a. solteras.   b. llamas.   c. becas.   d. primas.

11. ¡Camine en
    a. el sendero!   b. la piscina!   c. el estanque!   d. la química!

12. Hay líquido en
    a. enero.   b. el frasco.   c. la cédula.   d. el hilo.

13. Un _____ es un hombre corto.
    a. mundo   b. viudo   c. cascote   d. enano

14. La planta tiene varias
    a. camisetas.   b. fechas.   c. semillas.   d. ciudadanías.

15. El padre de mi abuelo es mi
    a. padrino.   b. bisabuelo.   c. tatarabuelo.   d. padrastro.

SAMPLE ITEMS FROM RECOGNITION TESTS (FIRST 20 HOURS) (cont'd.)

Matching

___ 1. el estanque
___ 2. el cascote
___ 3. la calle
___ 4. la aldea
___ 5. el globo
___ 6. el rastrillo
___ 7. la piscina
___ 8. el agujero
___ 9. la peña
___ 10. los piojos
___ 11. la aguanieve
___ 12. la dehesa
___ 13. la carpa
___ 14. la aguja
___ 15. la vereda
___ 16. el estacionamiento
___ 17. el hilo
___ 18. el edificio
___ 19. la navaja
___ 20. la verja
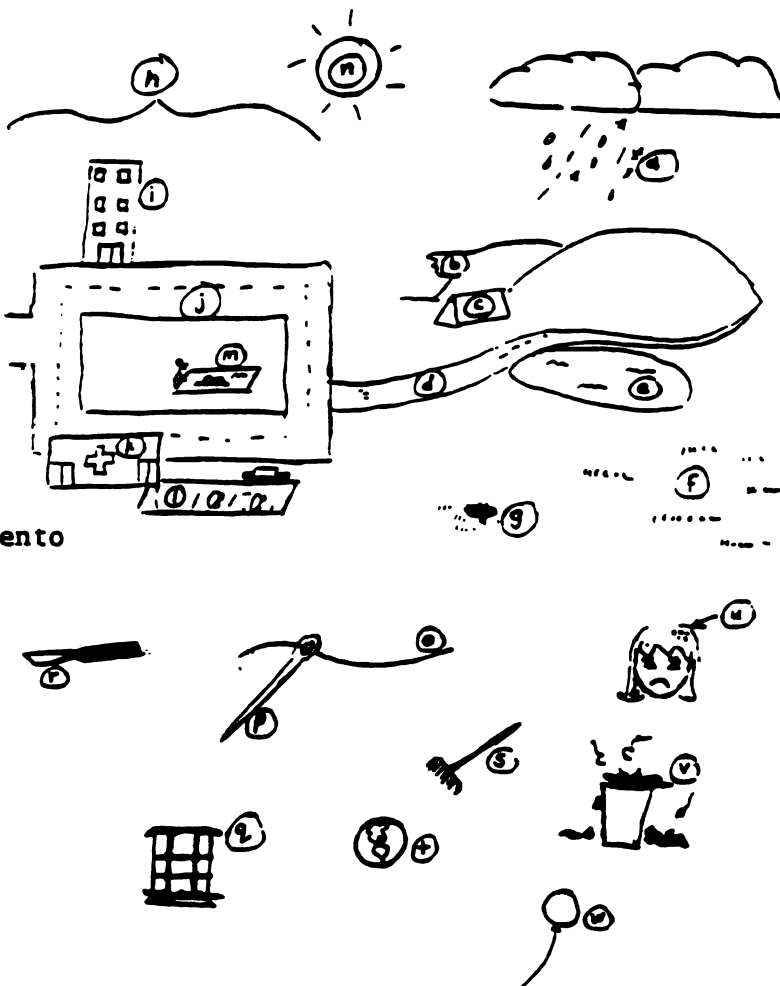
Oral (TPR)

1. Tóquese...
   a. la boca.
   b. la quijada.
   c. el puño.
   d. la nariz.
   e. la sien.
   f. el pie.
   g. la espalda.
   h. el meñique.

2. Apunte las partes del animal.

   a. la garra
   b. el espinazo
   c. el pescuezo

3. Silbe.
4. Guiñe el ojo derecho.
5. Levante la mano izquierda.
6. Ronque.
7. Ande de puntillas.
8. Pellizque al chico.
9. Levántese y dé una vuelta a la derecha.
10. Bostece.

PRODUCTION QUIZZES (FIRST 20 HOURS)

¿Cómo se llama usted?   (EXAMENCITO #1)

_____

I. __Ser__: Write in Spanish, including the pronoun. (8)

I am   __yo soy__   we are _____

you are
(informal) _____

you are   you are
(formal) _____   (plural) _____

II. __Tener__: Write in Spanish, including the pronoun. (10)

I have   _____   we have _____

you have
(informal) _____

he has   _____   they have _____

III. Escriba lo contrario (the opposite). (16)

alto _____   a la derecha _____

guapo _____   escuchen _____

gordo _____   levántense _____

joven _____   miren abajo __miren_____

grande _____   corto _____

agudo _____   lacio _____

el hombre __la_____   negro _____

Buenos días. _____   no _____

IV. Escriba las partes del animal. (4)

① el _____   ② el _____

③ el _____

④ la _____
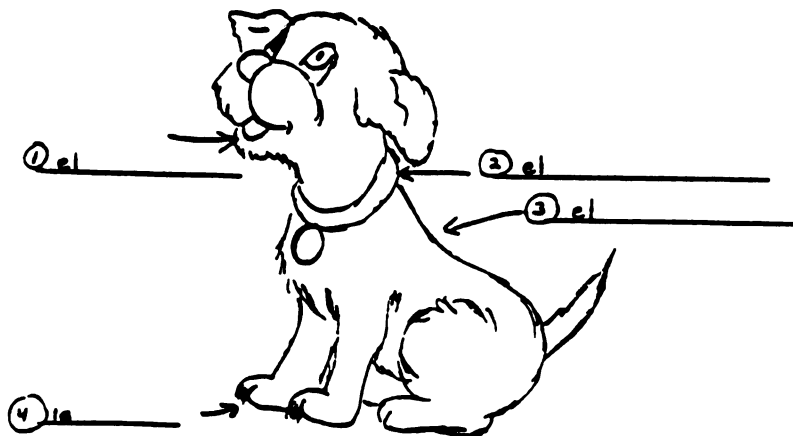
PRODUCTION QUIZZES (FIRST 20 HOURS) (cont'd.)

- 2 -

V. Escriba las partes del cuerpo humano.
(12)

VI. Escriba en español.
(15)

1. the page   la _____
2. lame _____
3. shy _____
4. lazy _____
5. sharp, clever

   _____
6. one-eyed _____
7. gentle _____
8. also _____
9. this _____
10. ..ith _____
11. very _____
12. walk! ¡_____
13. Good morning.

   _____
14. How are you?

   _____
15. 7 _____
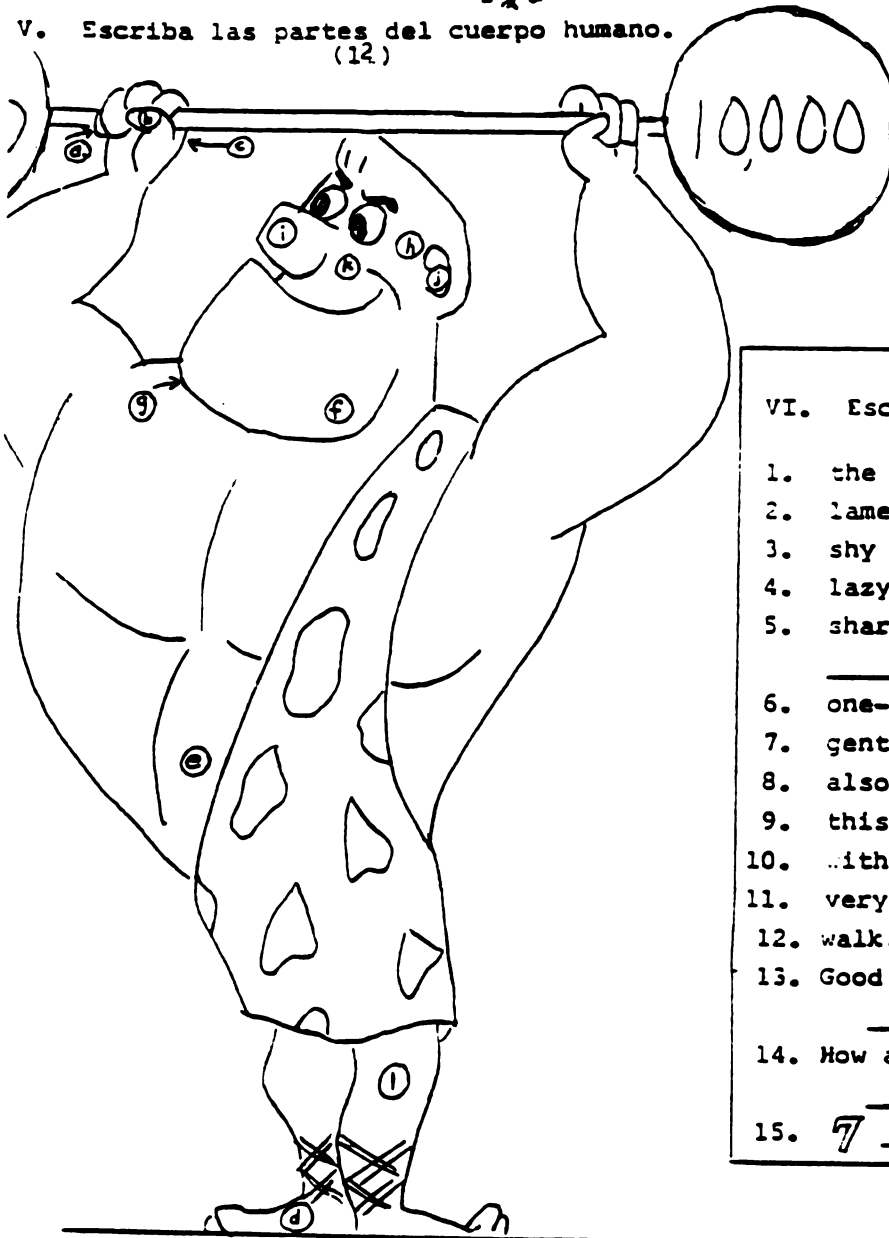
a. el _____
b. el _____
c. el _____ (fist)
d. el _____
e. la _____ (estómago)
f. la _____
g. la _____

h. la _____
i. la _____
j. la _____
k. la _____ (skin, complexion)
l. la _____

PRODUCTION QUIZZES (FIRST 20 HOURS) (cont'd.)

¿Cómo se llama usted?   (EXAMENCITO #2)   ¿Cuál es la fecha?

_____   _____

I.   El verbo <u>estar</u> (include the pronouns) (8)

   I am   _____   we are   <u>nosotros estamos</u>
   you are
   (informal) _____
   she is   _____   they are _____

II. Escriba los mandatos. (8)

   Squeeze! ¡_____!   Answer! ¡_____!
   Wink!! ¡_____!   Whistle! ¡_____!
   Lower your hand!   Yawn! ¡_____!
      ¡_____ la mano!   Describe! ¡_____!
   Snore! ¡_____!

III. Write the adjective to describe the picture. (12)
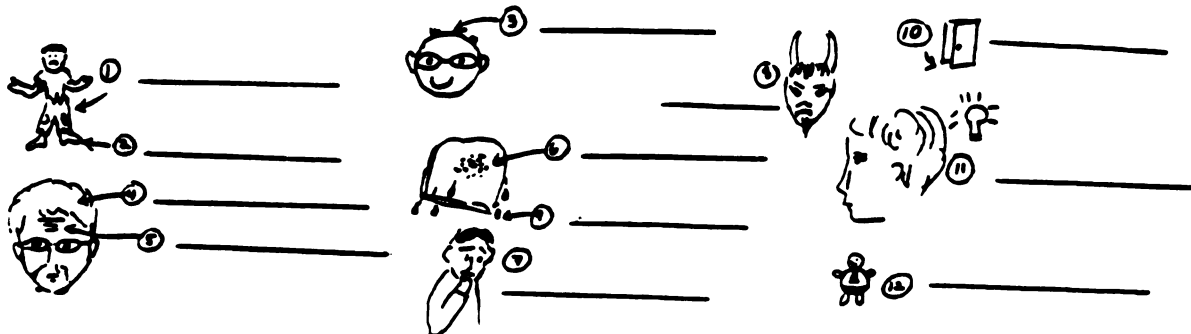
IV. <u>Escriba los sinónimos.</u>   (7)   <u>Escriba lo contrario (opposite).</u>
   huraño _____   adelantado _____
   el gemelo   el _____   el mediodía   la _____
   papá   el _____   mayor _____
   mamá   la _____

V.   La familia (fill in the blanks) (8)   <u>otra vez</u> = again

   1. Anita es mi madre. Ella es divorciada, pero se casa (marries)
       otra vez con Victor. Victor es mi _____, y yo soy
       el (la) _____ de Victor.
   2. Carlos es mi padre. El es viudo (widower), pero se casa otra
       vez con María. María es mi _____.
   3. El padre de mi abuelo es mi _____, y el abuelo de mi
       abuelo es mi _____. Todos (All) estos hombres son
       mis _____, y yo soy el _____ de ellos.
   4. Los hijos de mis padres son mis _____.

PRODUCTION QUIZZES (FIRST 20 HOURS) (cont'd.)

-2-

VI. Escriba en español. (23)

1. the godfather   el _____
2. the godmother   la _____
3. the question   la _____
4. the year   el _____
5. the boyfriend   el _____
6. late   _____
7. You're welcome. _____
8. our _____
9. his/her _____
10. 40 _____
11. 60 _____
12. 80 _____
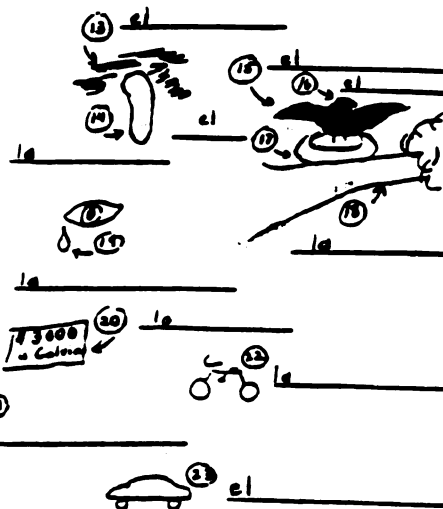
VII. Mire a la profesora. (15)

1. la _____   9. el _____
2. el _____   10. la _____
3. el _____   11. la _____
4. el _____   12. la _____
5. la _____   13. el _____
6. la _____   14. el _____
7. el _____   15. el _____
8. la _____
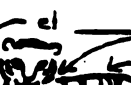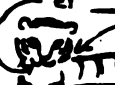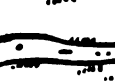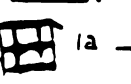
RETENTION TEST (FOR PRODUCTION)

NAME _____   VOCABULARY RECALL   DATE _____
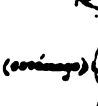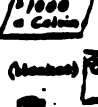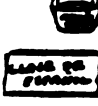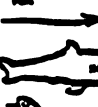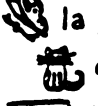
ADJECTIVES

1. _____
2. lazy _____
3. gentle _____
4. shy (tímido) _____
5. pelo _____
6. (inteligente) _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. el _____
18. el _____
19. el _____
20. la _____
21. el _____
22. el _____
23. el _____
24. la _____
25. la _____
26. el _____
27. la _____
28. el _____
29. la _____
30. la _____

31. el _____
32. la _____
33. la _____
34. el _____
35. la _____
36. el _____
37. el _____
38. la _____
39. la _____
40. la _____
41. el _____
42. la _____
43. (_____) el _____
44. el _____
45. la _____
46. el _____
47. el _____
48. la _____
49. la _____
50. el _____
51. la _____
52. el _____
53. el _____
54. el _____
55. la _____
56. el _____ alto
57. la _____
58. RUSSIAN el _____
59. GERMAN el _____

RETENTION TEST (FOR PRODUCTION) (cont'd.)

60. (tios) el _____

61. la _____

62. la _____

63. la _____

64. (sobrinos) la _____

65. el _____

66. el _____

67. la _____

68. el _____

69. las _____

70. face, complexion   la _____

71. la _____

72. (bloches)   la _____

73. el _____

74. el _____

75. el _____

76. el _____

77. el _____

78. el _____

79. el _____

80. el _____

81. la _____

82. el _____

83. la _____

84. el _____

85. la _____

86. la _____

87. el _____

88. godfather   el _____

89. godmother   la _____

90. el _____

91. el _____

92. (gemelos)   las _____

93. la _____

94. el _____

95. el _____

96. ancestor   el _____

97. descendant   el _____

98. el _____

99. el _____

100. la _____

101. la _____

102. el _____

103. la _____

104. el _____

105. la _____

COMMANDS

106. ¡_____!

107. ¡_____!

108. ¡_____!

109. ¡_____!

110. ¡_____!

111. ¡_____!

112. ¡_____!

113. ¡_____!

114. ¡_____!

115. ¡_____!

116. ¡_____!

117. ¡_____!

118. ¡_____!

WORDS IN ENGLISH

LOOK HERE if you don't know what the pictures are.

1. fat
2. lazy
3. gentle
4. shy
5. curly (hair)
6. smart, sharp, clever
7. stupid
8. one-eyed
9. ragged
10. barefoot
11. moldy
12. wet, soaked
13. lame
14. white-haired
15. wrinkled
16. bald
17. the blackbird
18. the wing
19. the nest
20. the branch
21. the (animal) snout
22. the (animal) neck
23. the (animal) back
24. the (animal) claw
25. the rock
26. the path (2 ways)
27.   "
28. the hole
29. the swimming pool
30. the window grating
31. the hallway
32. the flame
33. the match
34. the balloon
35. the seed
36. the poster
37. the rake
38. the board
39. the tear
40. the needle
41. the thread

42. the flower pot
43. the Venezuelan monetary unit
44. the clock
45. the footprint
46. the mud
47. the rubbish
48. the sleet
49. the jewel
50. the apron
51. the belt
52. the vest
53. the dress
54. the earring
55. the zipper
56. the heel (high)
57. the shirt
58. Russian
59. German
60. the fist
61. the temple
62. the chin
63. the jaw
64. the stomach
65. the little finger
66. the thumb
67. the nose
68. the moustache
69. the contact lens
70. face, complexion
71. the scholarship
72. the blanket
73. the bottle
74. the bowl
75. the dwarf
76. the fire
77. the basket
78. the sign
79. the deer (2 ways)
80.   "
81. the whale
82. the shark
83. the butterfly
84. the cat

85. the window
86. the blade (razor)
87. the tile
88. the godfather
89. the godmother
90. the son
91. the grandson
92. the twins
93. the stepmother
94. the stepfather
95. the stepson
96. the ancestor
97. the descendant
98. the great-great-grandfather
99. the great-grand-father
100. the tent
101. the meadow
102. the hill (2 ways)
103.   "
104. the pond
105. the village
106. Tie!
107. Squeeze/Wring!
108. Wink!
109. Yawn!
110. Pinch!
111. Fold!
112. Snore!
113. Walk!
114. Open!
115. Whistle!
116. Walk on tiptoe!
117. Cross out!
118. Pray!

238

RETENTION TEST (FOR RECOGNITION)

NAME   **VOCABULARY RECOGNITION**   **DATE**

MATCHING

___1.

___2. lazy

___3. gentle

___4. shy (tímido)

___5. pelo —

___6. (inteligente)

___7.

___8.

___9.

___10.

___11.

___12.

___13.

___14.

___15.

___16.

___17.

___18.

___19.

___20.

___21.

___22.

___23.

___24.

___25.

___26.

___27.

___28.

___29.

___30.

a. la comida
b. el sendero
c. fácil
d. ocioso
e. rizado
f. la piscina
g. cano
h. el ala
i. harapiento
j. empapado
k. lacio
l. calvo
m. delgado
n. la calle
o. la lectura
p. el mirlo
q. cojo
r. el nido
s. huraño
t. mohoso
u. el agujero
v. agudo
w. la fecha
x. gordo
y. mejor
z. la peña
aa. el viudo
bb. el pescuezo
cc. descompuesto
dd. descalzo
ee. la garra
ff. el horario
gg. la verja
hh. abierto
ii. caro
jj. arrugado
kk. la rama
ll. el espinazo
mm. la cédula
nn. la vereda
oo. el otoño
pp. el hocico
qq. barato
rr. tuerto
ss. la firma
tt. torpe
uu. fiel
vv. manso

___31.

___32.

___33.

___34.

___35.

___36.

___37.

___38.

___39.

___40.

___41.

___42.

___43. (Venezuela)

___44.

___45.

___46.

___47.

___48.

___49.

___50.

___51.

___52.

___53.

___54.

___55.

___56.

___57.

___58. RUSSIAN

___59. GERMAN

a. la faja
b. el cuaderno
c. la corbata
d. el chaleco
e. el rastrillo
f. el ruso
g. el reloj
h. el libro
i. el lápiz
j. la cerilla
k. la huella
l. la pared
m. la semilla
n. el piso
o. el casco
p. el tacón
q. la pizarra
r. el delantal
s. la aguamiel
t. el pupitre
u. la cremallera
v. el papel
w. el vestido
x. el globo
y. el saco
z. el fango
aa. el teano
bb. el bolívar
cc. el alemán
dd. la tiza
ee. la llama
ff. la luz
gg. el arete
hh. la falda
ii. la camisa
jj. la tabla
kk. la pluma
ll. la alhaja
mm. el pasillo
nn. la silla
oo. la aguja
pp. el borrador
qq. la maceta
rr. el abrigo
ss. la lágrima
tt. el hilo
uu. el bolígrafo
vv. el cárcel

RETENTION TEST (FOR RECOGNITION) (cont'd.)

-2-

___60. (ties)
___61.
___62.
___63.
___64. (entonces)
___65.
___66.
___67.
___68.
___69.
___70. face, complexion
___71.
___72. (Número)
___73.
___74.
___75.
___76.
___77.
___78.
___79.
___80.
___81.
___82.
___83.
___84.
___85.
___86.
___87.
___88. godfather

a. la tez
b. el azulejo
c. el bigote
d. el frasco
e. la espalda
f. la colcha
g. la cinta
h. la beca
i. el ciervo
j. el incendio
k. la barbilla
l. el brazo
m. la navaja
n. la pierna
o. la barba
p. la mujer
q. la sien
r. el hombro
s. el meñique
t. la página
u. la boca
v. la tarea
w. el cuenco
x. el gato
y. la quijada
z. la novia
aa. el perro
bb. el cesto
cc. la nariz
dd. el rótulo
ee. la pregunta
ff. el pulgar
gg. el viento
hh. el enano
ii. la ciudad
jj. el puño
kk. el venado
ll. el padrino
mm. el país
nn. la lentilla
oo. el tiburón
pp. el pie
qq. la panza
rr. la vecina
ss. la ballena
tt. la mesa
uu. la ventana
vv. la mariposa

___89. godmother
___90.
___91.
___92. (gemelo.)
___93.
___94.
___95.
___96. ancestor
___97. descendant
___98.
___99.
___100.
___101.
___102.
___103.
___104.
___105.
___106.
___107.
___108.
___109.
___110.
___111.
___112.
___113.
___114.
___115.
___116.
___117.
___118.

a. el primo
b. el antepasado
c. ¡Lea!
d. el hijo
e. la medusa
f. ¡Baile!
g. ¡Bostece!
h. el tatarabuelo
i. el padrastro
j. el estacionamiento
k. el estanque
l. el vástago
m. ¡Amarre!
n. la madrastra
o. ¡Camine!
p. la loma
q. la madrina
r. el nieto
s. el almuerzo
t. ¡Rasca!
u. la dehesa
v. el aula
w. ¡Guiña!
x. ¡Estire!
y. ¡Pellizque!
z. ¡Exprima!
aa. la rectoría
bb. ¡Abra!
cc. ¡Diga!
dd. el bisabuelo
ee. la biblioteca
ff. ¡Ronque!
gg. el apellido
hh. ¡Ande de puntillas!
ii. la parada
jj. ¡Tache!
kk. ¡Toque!
ll. la carpa
mm. el ahijado
nn. la aldea
oo. la librería
pp. ¡Pliegue!
qq. ¡Escuche!
rr. ¡Silba!
ss. ¡Baja!
tt. el cerro
uu. ¡Muestre!
vv. ¡Salte!
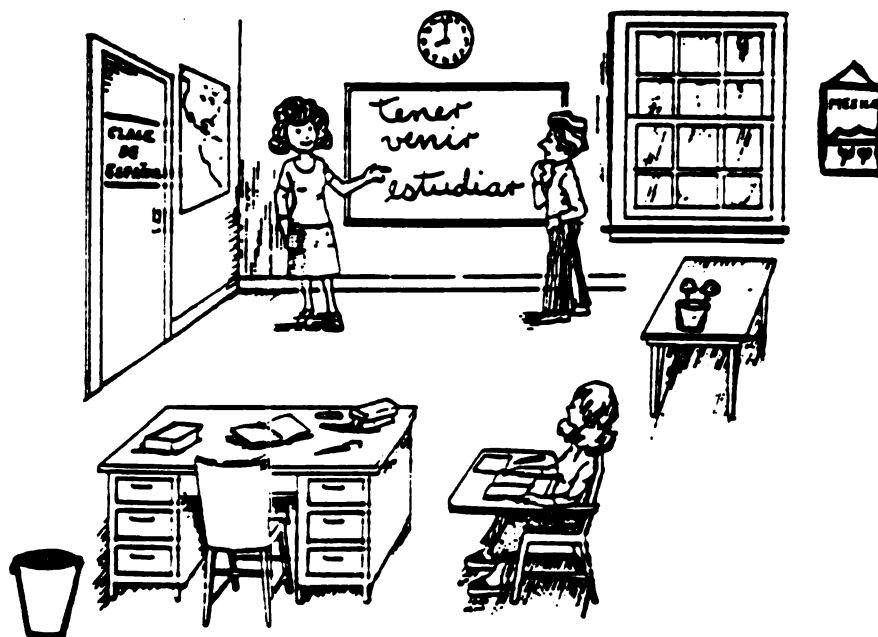
FIGURE 22
PICTURES FOR FINAL ORAL INTERVIEW

EVALUATION OF SPANISH 101

You have been taught with what can best be described as a modified
'Natural Approach.' Emphasis during your first semester of Spanish has
been on <u>understanding</u> the spoken language, though we have also done some
speaking, reading, writing, and grammar study.

1. Have you studied a foreign language before?
   If so, for how long?
   How does this class compare to other foreign language study you have
done? (more positive experience, about the same, more negative
experience, etc.)

2. How do you feel about the amount of Spanish you have learned this
semester? (more than you expected, less than you expected, etc.>)

3. What are your feelings during the class period? (bored, scared,
challenged, interested, etc.)

4. How does the workload in this course compare to other introductory
courses?

   the difficulty?   the grading?

5. How would you describe your own effort in this class? (both in
class and outside of class)

6. Please comment on the following components of the course as learning
tools:

-text (vocab lists, readings, blue pages, actividades)

-lab and workbook

-quizzes and tests

-other classroom activities (singing, bringing in native speakers,
cultural points, etc.)

7. The one thing I appreciate most about Spanish 101 is...

   The one thing I would change in Spanish 101 is...

8. Any other comments would be very much appreciated!

APPENDIX B
ADDITIONAL DATA AND STATISTICAL RESULTS

ADDITIONAL DATA AND STATISTICAL RESULTS

TABLE 40
DATA BY STUDENT

A	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4	B	C	D	E	F	G	H	I	J
01	81	56	70	35	56	43	42	10	23	05	1	03	1	1	1	2	2	2	1
02	84	61	80	40	79	42	71	19	60	20	1	01	1	2	1	3	2	2	2
03	84	54	65	23	47	12	38	02	18	02	1	01	1	1	1	2	1	2	1
04	61	47	44	20	22	13	07	04	17	03	1	05	2	1	1	2	2	1	1
05	96	82	98	68	92	38			52	09	1	01	1	1	1	2	1	2	1
06	76	36	60	06	42	06	15	02	07	03	1	03	1	2	1	3	2	2	2
07	60	31	58	46	42	21	33	10	36	06	2	10	2	2	2	4	1	3	2
08	80	39	41	19	36	14	20	02	10	04	2	05	2	1	3	4	1	3	2
09	83	61	57	30	52	20	30	06	28	06	2	05	2	1	3	4	1	3	2
10	70	35							06	01	2	07	2	2	2	4	2	3	2
11	65	47	31	12					06	01	2	07	2	1	3	4	2	4	1
12	72	42	74	39	84	44	72	24	59	19	2	07	2	2	3	5	1	3	1
13	99	95	77	64	55	23	36	19	46	26	2	03	1	2	2	4	2	4	2
14	88	81	84	39	57	17	08	01	08	02	2	02	1	1	1	2	2	3	1
15	99	93	98	89	96	77	84	25	69	23	2	01	1	1	1	2	1	1	2

KEY

A   student number
RO baseline recognition score
PO baseline production score
R1 first recognition test score
P1 first production test score
R2 second recognition test score
P2 second production test score
R3 third recognition test score
P3 third production test score
R4 fourth recognition test score
P4 fourth production test score
B   location of class: 1) Mexico, 2) U.S.
C   course grade: 1) A, 2) A-, 3) B+, 4) B, 5) B-, 6) C+, 7) C, 8) C-,
                 9) D+, 10) D
D   grade (grouped): 1) A, A-, and B+, 2) B- and lower
E   effort: 1) much, 2) some
F   motivation: 1) enrolled because of interest, 2) enrolled because of
             ·interest and requirement, 3) enrolled to meet requirement
G   effort + motivation
H   previous contact with Spanish: 1) yes, 2) no
I   later contact with Spanish: 1) yes, in class and culture, 2) yes, in
                             culture, 3) yes, in class, 4) no
J   sex: 1) female, 2) male

TABLE 41
DATA BY WORD

A	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4	B	C	D	E	F	G	H	I	J
001	064	043	071	037	050	023	050	013	027	007	1	1	1			2.47	2.80	2.46	1
002	064	057	086	018	092	015	073	005	043	000	1	1	1			3.20	3.00	2.96	2
003	060	063	093	054	091	058	060	040	033	021	1	1	1			1.83	2.00	3.12	2
004	073	037	057	032	067	027	033	000	020	003	1	1	1	4	1	4.10	3.68	3.56	2
005	073	060	086	071	077	054	073	032	029	029	1	1	1	4	1	3.57	3.65	3.20	2
006	085	060		036	038	017	025	000	020	007	1	1	1	4	1	3.70	3.39	3.24	2
007	080	063	057	039	050	042	033	013	020	010	1	1	2	4	1	2.72	3.04	3.32	2
008	060	043	064	025	054	015	025	004	020	000	1	1	2	6	2	2.80	2.35	3.58	2
009	100	067	064	057	055	036	027	008	015	014	1	1	2	8	2	2.03	2.58	3.04	2
010	093	040	064	043	046	035	033	000	013	000	1	1	2	8	2	2.52	2.77	3.48	2
011	093	047	064	032	054	031	042	021	043	004	1	1	2	8	2	2.60	2.04	3.20	2
012	087	077	071	054	042	033	009	018	014	000	1	1	2	8	2	2.23	1.81	3.12	2
013	087	060	086	036	054	038	045	018	029	018	1	1	2	4	1	2.00	1.81	2.28	1
014	067	037	036	025	023	023	017	008	013	000	1	1	3	6	2	4.38	4.08	3.32	2
015	087	088	086	046	075	021	017	004	007	007	1	1	3	6	2	4.00	3.50	2.72	1
016	067	053	093	043	092	046	033	017	033	003	1	1	3	6	2	3.72	2.92	3.48	2
017	073	053	071	039	050	029	033	018	031	011	1	1	3	6	2	3.63	3.35	2.96	2
018	093	053	086	054	077	050	050	025	046	019	1	1	3	4	1	4.03	3.89	3.56	2
019	067	023	057	021	046	008	025	000	020	007	1	1	3	6	2	4.14	3.44	3.20	2
020	073	053	071	032	054	031	033	004	020	007	1	1	3	6	2	3.00	1.96	3.12	2
021	060	047	079	043	077	027	064	005	050	003	1	1	3			3.47	3.35	2.88	1
022	080	063	079	057	067	027	058	033	029	014	1	1	3			2.80	3.42	2.12	1
023	047	053		039	100	046	040	030	046	017	1	1	3			2.07	2.85	2.32	1
024	093	073	071	057	062	046	073	009	050	013	1	1	3			3.35	2.69	3.17	2
025	093	053	079	039	067	038	020	010	015	015	1	1	3			3.00	2.50	2.80	1
026	093	073	079	050	085	046	042	008	027	010	1	1	3			3.45	3.77	2.71	1
027	080	033	057	014	046	008	025	000	007	000	1	1	3			3.43	3.39	3.20	2
028	060	053	086	046	085	039	033	017	013	003	1	1	3			3.20	3.23	3.68	2
029	093	060	050	039	042	012	033	004	027	000	1	1	3	6	2	3.03	2.58	3.92	2
030	073	050	093	043	092	027	058	008	033	007	1	1	3	4	1	4.10	4.23	3.08	2
031	093	067	086	061	085	054	033	008	040	007	1	1	3	4	1	2.57	3.00	2.60	1
032	080	017	050	014	058	019	033	008	033	003	1	1	3	6	2	3.13	3.08	3.60	2
033	100	067	093	079	100	083	100	039	089	044	1	1	3	4	1	3.17	2.58	2.32	1
034	073	043	057	021	054	012	045	021	021	007	1	1	3	4	1	3.23	2.97	3.04	2
035	053	037	064	036	064	032	018	013	031	008	1	1	3	4	1	3.23	2.97	2.80	1
036	093	057	100	079	092	067	080	009	075	013	1	1	3	4	1	3.24	2.80	2.56	1
037	100	067	100	079	092	079	080	041	085	032	1	1	3	6	2	2.27	2.50	3.56	2
038	093	057	086	096	092	046	075	033	050	032	2	1	1	3	1	3.21	3.50	3.52	2
039	093	057	050	029	046	015	050	009	020	007	2	1	1	6	2	2.69	2.08	3.40	2
040	100	067	050	032	054	035	042	018	039	025	2	1	1	4	1	2.77	2.15	3.12	2
041	087	070	043	039	046	031	025	008	040	007	2	1	1	6	2	2.69	2.05	2.76	1
042	073	040	064	018	054	015	025	000	020	003	2	1	1	5		2.76	2.05	3.26	2
043	100	050	057	043	062	031	017	004	020	000	2	1	2	4	1	2.76	2.00	2.96	2
044	100	080	086	075	067	058	027	000	036	004	2	1	2	2	1	2.80	1.95	2.58	1
045	100	067	071	057	069	042	025	013	020	013	2	1	2	2	1	3.11	2.81	3.26	2

TABLE 41 (cont'd.)

A	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4	B	C	D	E	F	G	H	I	J
046	100	083	086	071	069	046	075	029	050	025	2	1	2	4	1	2.55	2.35	2.38	1
047	087	030	050	025	046	008	027	009	021	000	2	1	2	2	1	2.77	2.63	2.83	1
048	080	057	071	054	062	027	027	005	021	007	2	1	2	2	1	3.00	2.40	3.00	2
049	087	057	071	036	050	023	025	000	020	010	2	1	2	4	1	2.82	2.38	3.26	1
050	100	057	086	046	069	027	055	000	036	004	2	1	3	4	1	3.50	3.43	3.29	1
051	087	090	079	050	064	029	064	000	038	013	2	1	3	5		4.17	4.04	1.75	1
052	093	070	086	050	075	042	033	000	033	007	2	1	3	4	1	3.70	2.92	2.80	1
053	093	063	086	064	069	054	067	025	060	030	2	1	3	6	2	4.17	3.12	3.24	2
054	073	070	064	036	042	021	042	000	020	003	2	1	3	4	1	3.77	3.73	3.52	2
055	087	060	071	054	058	033	050	008	020	000	2	1	3	6	2	3.27	3.15		
056	060	053	079	039	058	033	027	013	036	014	2	1	3	4	1	3.83	3.35	2.20	1
057	067	050	064	039	042	023	009	000	029	000	2	1	3	6	2	3.03	3.92	3.42	2
058	087	060	071	043	067	046	045	021	029	010	2	1	3	4	1	2.90	2.69	3.32	2
059	100	093	093	061	075	050	050	017	042	008	2	1	3	5		3.70	3.54	2.68	1
060	093	080	086	064	064	064	055	018	062	019	2	1	3	5		3.90	3.46	3.28	2
061	100	093	100	082	091	055	082	036	067	015	2	1	3	8	2	3.47	3.27	3.20	2
062	093	073	064	039	069	031	050	017	027	013	2	1	3	5		3.43	3.69		
063	047	037	079	021	083	019	058	008	064	017	3	1	3	4	1	2.97	3.73	2.00	1
064	067	033	043	014	031	000	000	000	000	000	3	1	3	4	1	3.13	3.58	3.20	2
065	080	080	086	050	073	055	027	014	036	017	3	1	3	6	2	2.54	3.28	1.71	1
066	093	083	079	039	075	046	042	013	043	013	3	1	3	2	1	2.66	3.39	2.04	1
067		043	079	029	062	015	067	008	053	000	3	1	3	2	1	3.31	3.31	2.44	1
068		043	064	025	062	019	067	008	053	007	3	1	3	2	1	2.93	3.46	2.24	1
069		053	086	032	062	031	080	010	067	019	3	1	3	2	1	3.03	3.15	2.40	1
070	060	067	079	043	069	042	060	015	038	019	3	1	3	2	1	3.00	3.15	2.21	1
071	093	060	057	021	054	000	042	004	033	003	3	1	3	4	1	1.97	2.12	3.04	2
072	087	060	079	050	060	018	030	027	018	014	3	1	3			1.33	2.42	3.36	2
073	080	070	071	057	062	027	027	000	014	007	3	1	3			2.00	3.04	3.40	2
074	087	043	043	014	039	015	042	008	027	007	3	1	3	2	1	3.61	3.15	3.16	2
075	073	043	057	021	054	019	042	000	040	000	3	1	3			2.00	2.04	2.76	1
076	073	053	057	036	033	031	033	025	033	030	3	1	3			2.20	2.58	3.12	2
077	093	040	050	014	036	004	017	000	007	000	3	1	3			2.93	3.08	3.48	2
078	100		093	036	100	027	083	013	087	023	4	2	3			2.86	3.50	3.04	2
079	093		043	014	018	000	009	000	000	000	4	2	3	6	2	3.37	3.12	3.56	2
080	053		043	014	039	008	033	008	020	003	4	2	3	4	1	3.07	2.89	3.80	2
081	067		036	025	025	025	050	008	021	007	4	2	3	4	1	3.50	2.80	3.00	2
082	067		021	011	023	008	017	000	000	000	4	2	3	4	1	3.50	2.80	2.72	1
083	047		014	007	023	012	017	000	000	003	4	2	3	3	1	3.59	3.19	3.04	2
084	080		036	007	023	008	008	004	013	000	4	2	3	5		3.59	3.19	2.80	1
085	060		050	007	025	008	017	000	000	000	4	2	3	4	1	3.66	3.35	3.64	2
086	093		071	054	069	031	042	017	047	007	4	2	3	4	1	3.60	3.54	2.16	1
087	060		021	007	023	012	008	000	013	000	4	2	3	2	1	3.40	2.46	3.36	2
088	047		029	004	023	004	000	000	000	000	4	2	3	2	1	3.41	3.92	3.32	2
089	080			039	042	013	018	009	023	000	4	2	3	3	1	3.43	3.15	2.56	1
090	053		057	007	062	008	033	000	033	013	5	2	3	2	1	2.97	2.69	2.68	1
091	093		071	025	085	015	050	000	047	000	5	2	3	3	1	3.25	2.50	2.26	1
092	093		050	018	039	012	008	000	013	000	5	2	3	3	1	3.08	1.85	2.84	1
093	073		064	036	062	023	017	000	020	003	5	2	3	5		3.56	2.50	3.17	2
094	073		036	000	008	004	009	000	000	000	5	2	3	5		2.31	2.37	2.48	1
095	060		043	011	039	012	017	000	007	000	5	2	3	6	2	2.70	2.00	3.13	2

TABLE 41 (cont'd.)

A	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4	B	C	D	E	F	G	H	I	J
096	073		029	036	033	023	009	008	000	007	5	2	3	4	1	3.53	2.93	3.17	2
097	093		086	050	067	038	050	017	036	009	5	2	3	3	1	3.36	2.88	2.56	1
098	073		043	021	039	000	008	000	013	000	5	2	3	6	2	2.92	2.00	2.68	1
099	087		093	057	100	055	044	011	045	013	5	2	3	6	2	2.57	1.70	1.68	1
100	053		029	018	031	012	009	004	000	004	5	2	3	5		2.44	2.44	3.27	2

KEY

A word number (see Table 4 for Spanish words with English translations)
R0 baseline recognition score
P0 baseline production score
R1 first recognition test score
P1 first production test score
R2 second recognition test score
P2 second production test score
R3 third recognition test score
P3 third production test score
R4 fourth recognition test score
P4 fourth production test score
B paso: 1) Paso A, 2) Paso B, 3) Paso C, 4) Paso D, 5) Paso E
C recognition and production: 1) taught and tested for both recognition
   production, 2) taught and tested for recognition only
D part of speech: 1) verb, 2) adjective, 3) noun
E frequency in input: 2) very low (1-4 times), 4) low (5-9 times),
   6) medium (10-19 times), 8) high (20+ times)
F frequency in input (grouped): 1) low (1-9 times), 2) high (10+ times)
G emotionality of English word (5--lowest emotionality, 1--highest
   emotionality)
H saliency of English word (5--lowest saliency, 1--highest saliency)
I saliency of Spanish form (5--lowest saliency, 1--highest saliency)
J saliency of Spanish form (grouped): 1) high saliency, 2) low saliency

CORRELATIONS OF SCORES AT TIMES 1 AND 4
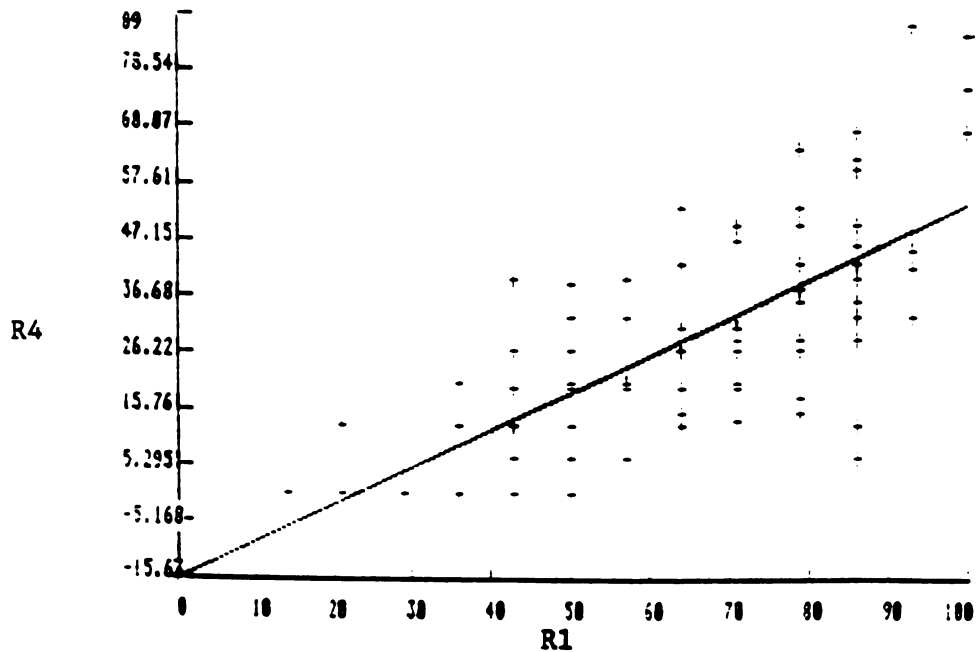
FIGURE 23
SCATTERGRAM FOR R1 AND R4
$\widehat{R4}$ = -15.631 + .683(R1) / r = .689
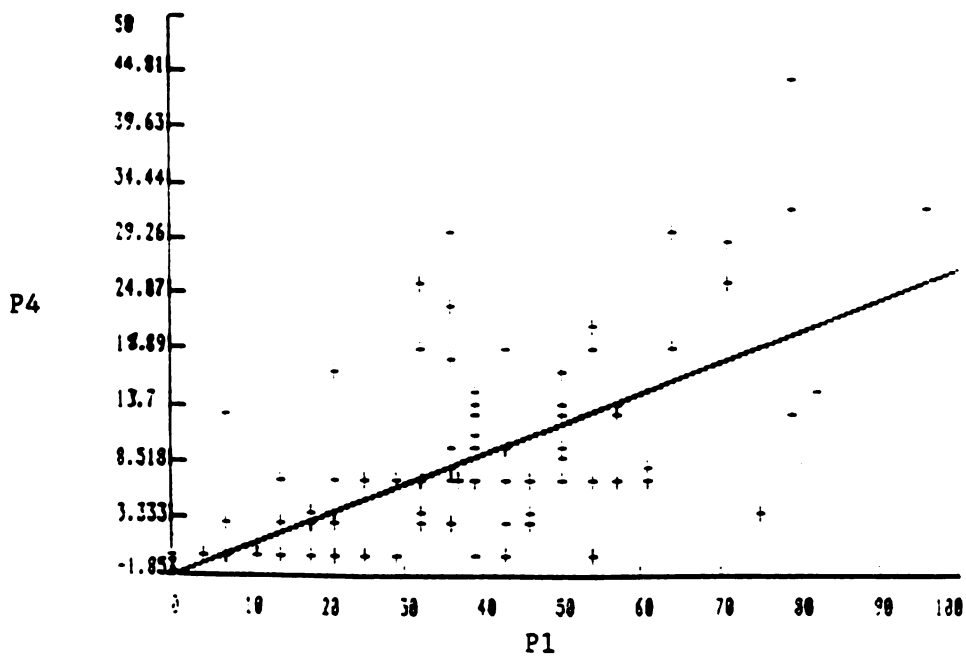
FIGURE 24
SCATTERGRAM FOR P1 AND P4
$\widehat{P4}$ = -1.853 + .280(P1) / r = .617

CORRELATIONS OF SCORES AT TIMES 1 AND 4 (cont'd.)
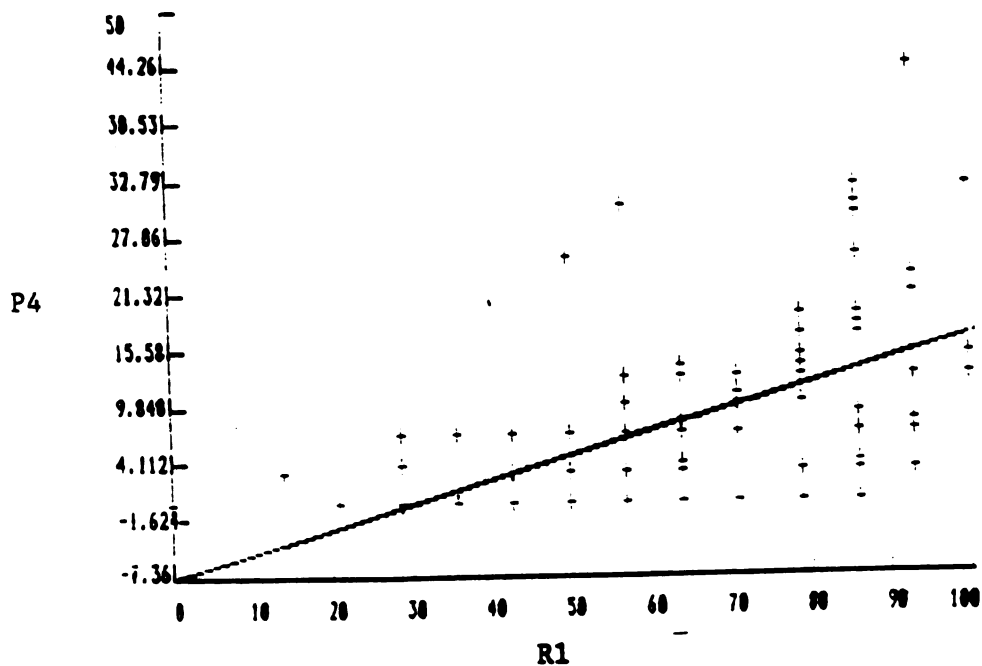
FIGURE 25
SCATTERGRAM FOR R1 AND P4
$\hat{P4}$ = -7.36 + .242(R1) / r = .532
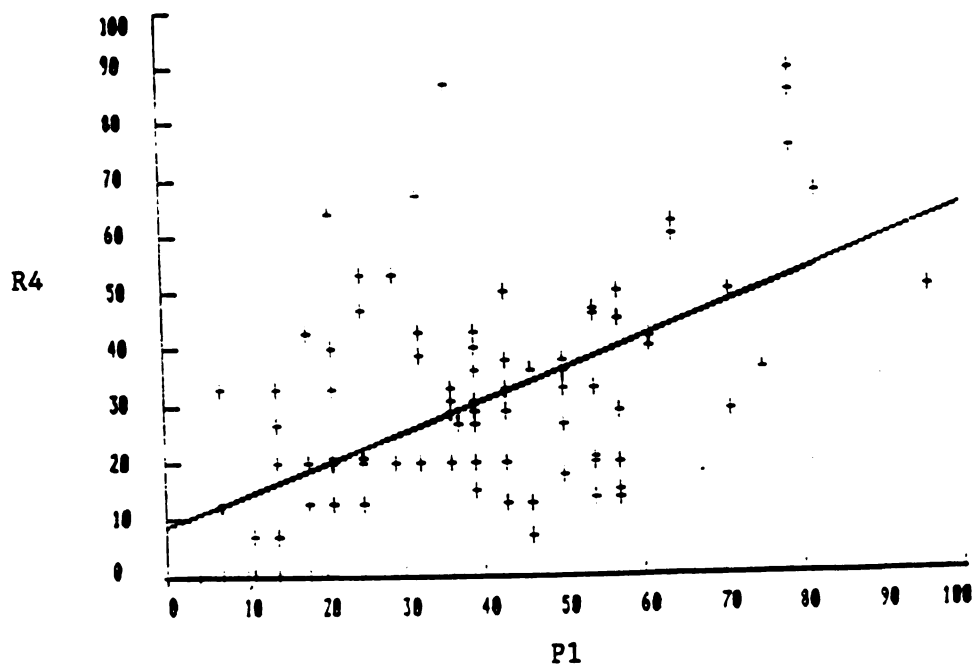
FIGURE 26
SCATTERGRAM FOR P1 AND R4
$\hat{R4}$ = 8.73 + .556(P1) / r = .564

TABLE 42
PREDICTIONS FOR R4 AND P4 (BASED ON R1 AND P1)

A	B	R1	P1	R4	P4	$\widehat{R4}$	C	$\widehat{P4}$	D
001	pellizque	071	037	027	007	32.4	-05.4	08.9	-01.9
002	ande de puntillas	086	018	043	000	39.9	03.1	05.5	-05.5
003	rece	093	054	033	021	47.5	-14.5	14.2	06.8
004	pliegue	057	032	020	003	23.4	-03.4	06.8	-03.8
005	tache	086	071	029	029	44.7	-15.7	17.6	11.4
007	cojo	057	039	020	010	24.1	-04.1	08.4	01.6
008	tuerto	064	025	020	000	27.1	-07.1	05.6	-05.6
009	manso	064	057	015	014	30.0	-15.0	13.0	01.0
010	huraño	064	043	013	000	28.7	-15.7	09.8	-09.8
011	ocioso	064	032	043	004	27.7	15.3	07.2	-03.2
012	agudo	071	054	014	000	34.0	-20.0	12.7	-12.7
013	torpe	086	036	029	018	41.5	-12.5	09.6	08.4
014	delantal	036	025	013	000	09.9	03.1	03.7	-03.7
015	faja	086	046	007	007	42.4	-35.4	11.9	-04.9
016	cremellera	093	043	033	003	46.5	-13.5	11.7	-08.7
017	aretes	071	039	031	011	32.6	-01.6	09.3	01.7
018	tacón	086	054	046	019	43.2	02.8	13.8	05.2
019	chaleco	057	021	020	007	22.4	-02.4	04.2	02.8
020	alhajas	071	032	020	007	32.0	-12.0	07.7	-00.7
021	meñique	079	043	050	003	37.9	12.1	10.8	-07.8
022	tez	079	057	029	014	39.2	-10.2	14.0	00.0
024	pulgar	071	057	050	013	34.3	15.7	13.4	-00.4
025	panza	079	039	015	015	37.5	-22.5	09.9	05.1
026	barbilla	079	050	027	010	38.5	-11.5	12.4	-02.4
027	quijada	057	014	007	000	21.8	-14.8	02.6	-02.6
028	sien	086	046	013	003	42.4	-29.4	11.9	-08.9
029	hocico	050	039	027	000	19.8	07.2	07.9	-07.9
030	espinazo	093	043	033	007	46.5	-13.5	11.7	-04.7
031	garra	086	061	040	007	43.8	-03.8	15.4	-08.4
032	pescuezo	050	014	033	003	17.5	15.5	02.2	00.8
033	mariposa	093	079	089	044	49.8	39.2	20.0	24.0
034	ciervo	057	021	021	007	22.4	-01.4	04.2	02.8
035	venado	064	036	031	008	28.1	02.9	08.2	-00.2
036	ballena	100	079	075	013	54.0	21.0	20.4	-07.4
037	tiburón	100	079	085	032	54.0	31.0	20.4	11.6
038	exprima	086	096	050	032	47.0	03.0	23.4	08.6
039	guiñe	050	029	020	007	18.9	01.2	05.6	01.4
040	ronque	050	032	039	025	19.1	19.9	06.3	18.7
041	silbe	043	039	040	007	15.5	24.5	07.4	-00.4
042	bostece	064	018	020	003	26.4	-06.4	04.0	-01.0
043	descalzo	057	043	020	000	24.4	-04.4	09.3	-09.3
044	calvo	086	075	036	004	45.1	-09.1	18.6	-14.6
045	cano	071	057	020	013	34.3	-14.3	13.4	-00.4
046	mohoso	086	071	050	025	44.7	05.3	17.6	07.4
047	harapiento	050	025	021	000	18.5	02.5	04.7	-04.7

TABLE 42 (cont'd.)

A	B	R1	P1	R4	P4	$\widehat{R4}$	C	$\widehat{P4}$	D
048	arrugado	071	054	021	007	34.0	-13.0	12.7	-05.7
049	empapado	071	036	020	010	32.3	-12.3	08.6	01.4
050	maceta	086	046	036	004	42.4	-06.4	11.9	-07.9
051	tabla	079	050	038	013	38.5	-00.5	12.4	00.6
052	colcha	086	050	033	007	42.8	-09.8	12.8	-05.8
053	cuenco	086	064	060	030	44.1	15.9	16.0	14.0
054	rótulo	064	036	020	003	28.1	-08.1	08.2	-05.2
055	cartel	071	054	020	000	34.0	-14.0	12.7	-12.7
056	pasillo	079	039	036	014	37.5	-01.5	09.9	04.1
057	huella	064	039	029	000	28.3	00.7	08.8	-08.8
058	fango	071	043	029	010	33.0	-04.0	10.2	-00.2
059	ala	093	061	042	008	48.1	-06.1	15.8	-07.8
060	rama	086	064	062	019	44.1	17.9	16.0	03.0
061	mirlo	100	082	067	015	54.3	12.7	21.1	-06.1
062	nido	064	039	027	013	28.3	-01.3	08.8	04.2
063	antepasado	079	021	064	017	35.9	28.1	05.7	11.3
064	vástago	043	014	000	000	13.2	-13.2	01.7	-01.7
065	bisabuelo	086	050	036	017	42.8	-06.8	12.8	04.2
066	tatarabuelo	079	039	043	013	37.5	05.5	09.9	03.1
067	padrastro	079	029	053	000	36.6	16.4	07.6	-07.6
068	madrastra	064	025	053	007	27.1	25.9	05.6	01.4
069.	padrino	086	032	067	019	41.2	25.8	08.7	10.3
070	madrina	079	043	038	019	37.9	00.1	10.8	08.2
071	mellizo	057	021	033	003	22.4	10.6	04.2	-01.2
072	lágrima	079	050	018	014	38.5	-20.5	12.4	01.6
073	beca	071	057	014	007	34.3	-20.3	13.4	-06.4
074	cesto	043	014	027	007	13.2	13.8	01.7	05.3
075	incendio	057	021	040	000	22.4	17.6	04.2	-04.2
076	llama	057	036	033	030	23.8	09.2	07.7	22.3
077	cerilla	050	014	007	000	17.5	-10.5	02.2	-02.2
078	aguanieve	093	036	087	023	45.8	41.2	10.1	12.9
079	estanque	043	014	000	000	13.2	-13.2	01.7	-01.7
080	dehesa	043	014	020	003	13.2	06.8	01.7	01.3
081	loma	036	025	021	007	09.9	11.1	03.7	03.3
082	cerro	021	011	000	000	-00.5	00.5	-00.5	00.5
083	vereda	014	007	000	003	-05.2	05.2	-01.9	04.9
084	sendero	036	007	013	000	08.3	04.7	-00.4	00.4
085	agujero	050	007	000	000	16.8	-16.8	00.6	-00.6
086	carpa	071	054	047	007	34.0	13.0	12.7	-05.7
087	aldea	021	007	013	000	-00.9	13.9	-01.4	01.4
088	verja	029	004	000	000	03.7	-03.7	-01.5	01.5
090	bolívar	057	007	033	013	21.1	11.9	01.0	12.0
091	rastrillo	071	025	047	000	31.3	15.7	06.1 .	-06.1
092	enano	050	018	013	000	17.8	-04.8	03.1	-03.1
093	azulejo	064	036	020	003	28.1	-08.1	08.2	-05.2
094	navaja	036	000	000	000	07.6	-07.6	-02.0	02.0
095	cascote	043	011	007	000	12.9	-05.9	01.0	-01.0

TABLE 42 (cont'd.)

A	B	R1	P1	R4	P4	$\widehat{R4}$	C	$\widehat{P4}$	D
096	hilo	029	036	000	007	06.6	-06.6	05.8	01.2
097	semilla	086	050	036	009	42.8	-06.8	12.8	-03.8
098	frasco	043	021	013	000	13.8	-00.8	03.3	-03.3
099	globo	093	057	045	013	47.7	-02.7	14.9	-01.9
100	aguja	029	018	000	004	05.0	-05.0	01.7	02.3

KEY

A   word number

B   Spanish word

R1 first recognition test score

P1 first production test score

R4 fourth recognition test score

P4 fourth production test score

$\widehat{R4}$ predicted score for fourth recognition test

C   error in prediction for R4

$\widehat{P4}$ predicted score for fourth production test

D   error in prediction for P4

TABLE 43
T-TESTS FOR PRE-COURSE CONTACT WITH SPANISH

TEST	N	MEAN	SD	T VALUE	1-TAIL PROB
R0 SOME CONTACT	7	82.0	13.4	.60	.28
NO CONTACT	8	78.0	12.6		
P0 SOME CONTACT	7	57.4	23.0	.02	.49
NO CONTACT	8	57.3	21.3		
R1 SOME CONTACT	7	70.1	21.5	.58	.29
NO CONTACT	7	63.7	20.9		
P1 SOME CONTACT	7	44.9	25.5	1.15	.14
NO CONTACT	7	30.9	19.8		
R2 SOME CONTACT	7	64.1	25.5	.97	.18
NO CONTACT	6	51.8	18.9		
P2 SOME CONTACT	7	32.3	23.1	.75	.24
NO CONTACT	6	24.0	15.4		
R3 SOME CONTACT	6	46.2	25.6	1.12	.14
NO CONTACT	6	29.8	24.9		
P3 SOME CONTACT	6	11.5	10.5	.43	.34
NO CONTACT	6	9.2	8.2		
R4 SOME CONTACT	7	38.9	21.9	1.57	.07
NO CONTACT	8	21.6	20.6		
P4 SOME CONTACT	7	9.9	8.0	.48	.32
NO CONTACT	8	7.6	9.7		

TABLE 44
ONE-WAY ANALYSES OF VARIANCE FOR POST-COURSE CONTACT WITH SPANISH

TEST	R0	P0	R1	P1	R2	P2	R3	P3	R4	P4
F	.39	.86	.55	.36	.11	.55	.12	.50	.33	.47
F PROB	.76	.49	.66	.79	.95	.66	.95	.70	.81	.71

TABLE 45
T-TESTS FOR DIFFERENCES BASED ON SEX

TEST		N	MEAN	SD	T VALUE	2-TAIL PROB
R0	FEMALE	7	78.1	12.7	-.48	.64
	MALE	8	81.4	13.4		
P0	FEMALE	7	58.4	16.4	.18	.86
	MALE	8	56.4	25.9		
R1	FEMALE	7	66.6	22.9	-.06	.95
	MALE	7	67.3	18.9		
P1	FEMALE	7	33.7	18.3	-.66	.52
	MALE	7	42.0	27.9		
R2	FEMALE	6	59.7	25.4	.17	.87
	MALE	7	57.4	22.1		
P2	FEMALE	6	27.8	15.4	-.10	.92
	MALE	7	29.0	23.8		
R3	FEMALE	5	33.4	27.1	-.51	.62
	MALE	7	41.3	26.1		
P3	FEMALE	5	8.2	9.5	-.67	.52
	MALE	7	11.9	9.2		
R4	FEMALE	7	26.1	20.9	-.56	.59
	MALE	8	32.8	24.4		
P4	FEMALE	7	5.9	6.4	-1.19	.26
	MALE	8	11.1	10.1		

TABLE 46
T-TESTS FOR DIFFERENCES BASED ON EFFORT

TEST		N	MEAN	SD	T VALUE	1-TAIL PROB
RO	MUCH EFFORT	9	81.9	12.6	.74	.24
	SOME EFFORT	6	76.8	13.4		
PO	MUCH EFFORT	9	62.2	18.7	1.10	.15
	SOME EFFORT	6	50.0	24.5		
R1	MUCH EFFORT	9	65.3	24.5	-.38	.35
	SOME EFFORT	5	69.8	10.1		
P1	MUCH EFFORT	9	37.2	25.4	-.13	.45
	SOME EFFORT	5	39.0	21.0		
R2	MUCH EFFORT	8	57.3	25.5	-.23	.41
	SOME EFFORT	5	60.4	20.1		
P2	MUCH EFFORT	8	29.3	22.6	.18	.43
	SOME EFFORT	5	27.2	15.9		
R3	MUCH EFFORT	7	32.7	26.4	-.84	.21
	SOME EFFORT	5	45.4	25.1		
P3	MUCH EFFORT	7	7.1	8.5	-1.52	.08
	SOME EFFORT	5	14.8	8.8		
R4	MUCH EFFORT	9	25.7	21.4	-.84	.21
	SOME EFFORT	6	35.7	24.3		
P4	MUCH EFFORT	9	6.1	6.8	-1.45	.09
	SOME EFFORT	6	12.5	10.4		

TABLE 47
ONE-WAY ANALYSES OF VARIANCE FOR MOTIVATION

TEST	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
F	.73	.84	2.30	1.25	.23	.17	.03	.25	.08	.13
F PROB	.50	.46	.15	.32	.80	.84	.97	.78	.92	.88

TABLE 48
T-TESTS FOR DIFFERENCES BASED ON MOTIVATION

TEST	N	MEAN	SD	T VALUE	1-TAIL PROB
RO MOST	8	83.6	11.9	1.29	.11
LEAST	4	75.0	8.1		
PO MOST	8	63.8	19.6	1.56	.08
LEAST	4	47.3	9.7		
R1 MOST	8	74.9	18.8	2.09	.03
LEAST	4	50.8	18.8		
P1 MOST	8	40.0	26.9	1.05	.16
LEAST	4	25.0	11.9		
R2 MOST	8	61.4	25.7	.23	.41
LEAST	3	57.3	24.4		
P2 MOST	8	31.0	23.7	.33	.37
LEAST	3	26.0	15.9		
R3 MOST	7	37.9	30.6	-.14	.45
LEAST	3	40.7	27.6		
P3 MOST	7	9.0	9.5	-.24	.41
LEAST	3	10.7	11.7		
R4 MOST	8	31.8	24.7	.40	.35
LEAST	4	25.8	24.1		
P4 MOST	8	8.4	8.5	.17	.43
LEAST	4	7.5	7.9		

TABLE 49
MEANS FOR SCORES BY PASO

PASO	A	B	C	D	E
RO	78.6	89.3	77.8	70.6	74.9
PO	53.8	65.1	53.9	--	--
R1	73.9	72.6	67.3	41.5	54.6
P1	43.0	49.7	31.1	18.8	25.4
R2	66.4	62.4	57.0	36.1	51.4
P2	35.0	35.4	22.7	13.0	18.4
R3	42.5	42.8	42.3	25.2	23.1
P3	13.8	11.3	9.3	4.9	3.6
R4	31.5	34.2	35.1	18.7	19.5
P4	9.8	10.8	10.2	3.6	4.5

TABLE 50
CORRELATIONS OF TEST SCORES WITH EMOTIONALITY
OF ENGLISH TRANSLATIONS

TEST	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
r	-.06	-.03	-.03	-.03	-.04	-.05	.03	-.20	-.02	-.11
PROB	.27	.39	.40	.37	.36	.32	.37	.02	.44	.13

TABLE 51
CORRELATIONS OF TEST SCORES WITH SALIENCY
OF ENGLISH TRANSLATIONS

TEST	RO	PO	R1	P1	R2	P2	R3	P3	R4	P4
r	-.15	-.03	.12	.05	.10	-.01	.17	.00	.08	.05
PROB	.08	.39	.11	.31	.15	.46	.05	.48	.21	.33

LIST OF REFERENCES

LIST OF REFERENCES

Andersen, Roger W. 1982. Determining the linguistic attributes of
    language attrition. The loss of language skills, ed. by Richard D.
    Lambert and Barbara F. Freed, 83-118. Rowley, MA: Newbury House.

----- (ed.) 1983. Pidginization and creolization as language
    acquisition. Rowley, MA: Newbury House.

Anderson, J.P., and A.M. Jordan. 1928. Learning and retention of Latin
    words and phrases. Journal of Educational Psychology 19.485-96.

Anderson, John R. 1983. The architecture of cognition. Cambridge:
    Harvard University Press.

-----. 1985. Cognitive psychology and its implications. 2nd edn. New
    York: Freeman.

Asher, James J. 1981. Learning another language through actions: The
    complete teacher's guidebook. Los Gatos, CA: Sky Oak Productions.

Bahrick, Harry P. 1979. Maintenance of knowledge: Questions about
    memory we forgot to ask. Journal of Experimental Psychology:
    General 108.296-308.

-----. 1984a. Semantic memory content in permastore: fifty years of
    memory for Spanish learned in school. Journal of Experimental
    Psychology: General 113.1-29.

-----. 1984b. Associations and organization in cognitive psychology:
    A reply to Neisser. Journal of Experimental Psychology: General.
    113.36-7.

-----. 1984c. Fifty years of second language attrition: Implications
    for programatic research. Modern Language Journal 68.105-18.

-----, and Demetrios Karis. 1982. Long-term ecological memory.
    Handbook of research methods in human memory and cognition, ed.
    by C. Richard Puff, 427-465. New York: Academic Press.

Berko-Gleason, Jean. 1982a. Converging evidence for linuistic theory
    from the study of aphasia and child language. Exceptional language
    and linguistics, ed. by L.K. Obler and L. Menn, 347-56. New York:
    Academic Press.

------. 1982b. Insights from child language acquisition for second-language loss. The loss of language skills, ed. by Richard D. Lambert and Barbara F. Freed, 13-23. Rowley, MA: Newbury House.

Berman, Ruth A., and Elite Olshtain. 1983. Features of first language transfer in second language attrition. Applied linguistics 4. 222-34.

Brewer-Bomar, K. 1981. Second language lexical and syntactical inference in the first language of two four-year-old Spanish speakers. DAI 42.5105A.

Brown, Dorothy F. 1980. Eight Cs and a G. Guidelines for vocabulary teaching, ed. by Rosemary Khoo, 1-17. (RELC Journal Supplement 3.)

Brown, John (ed.) 1976. Recall and recognition. New York: John Wiley.

Brown, Roger, and David McNeill. 1966. The 'Tip of the Tongue' phenomenon. Journal of Verbal Leaning and Verbal Behavior 5. 325-37.

Brown, W.P., and Dorothy M.J. Ure. 1969. Five rated characterstics of 650 word association stimuli. British Journal of Psychology 60. 233-49.

Byrnes, Heidi. 1984. The role of listening comprehension: A theoretical base. Foreign Language Annals 17.317-29.

Caramazza, Alfonso, and Edgar B. Zurif (eds.) 1978. Language acquisition and language breakdown: parallels and divergencies. Baltimore: John Hopkins University Press.

Carpay, J.A.M. 1975. Onderwijsleerpsychologie en leerangontwikkeling in het moderne vreemdetalenoundervijs. Groningen: Wolters-Noordhoff.

Champagnol, R. 1972. Conditions d'apprentissage du vocabulaire. Les languages modernes 66.413-26.

Chaudron, Craig. 1985a. Intake: on models and methods for discovering learner's processing of input. Studies in Second Language Acquisition 7.1-14.

------. 1985b. A method for examining the input/intake distinction. Input in second language acquisition, ed. by Susan M. Gass and Carolyn G. Madden, 285-300. (Issues in second language research.) Rowley, MA: Newbury House.

Clark, J.L.D., and E.H. Jorden. 1984. A study of language attrition in former U.S. students of Japanese and implications for design of curriculum and teaching materials. Washington, DC: Center for Applied Linguistics. ED 243 317.

Cohen, Andrev D. 1974. The Culver City Spanish Immersion Program: How does summer recess affect Spanish speaking ability? Language Learning 24.55-68.

------. 1975. Forgetting a second language. Language Learning 25. 127-38.

------. 1984. Studying second-language learning strategies: How do we get the information? Applied Linguistics 5.101-12.

------. 1986. Forgetting foreign-language vocabulary. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els. Dordrecht: Foris.143-58.

------. 1987. The use of verbal and imagery mnemonics in second-language vocabulary learning. Studies in Second Language Acquisition 9.43-61.

------. 1989. Attrition in the productive lexicon of two Portuguese third language speakers. Studies in Second Language Acquisition 11.135-49.

------, and Carol Hosenfeld. 1981. Some uses of mentalistic data in second language research. Language Learning 31.285-313.

Cooper, Aldwyn, J.R., and Andrev Monk. 1976. Learning for recall and learning for recognition. Recall and recognition, ed. by John Brown, 131-56. New York: John Wiley.

Corder, S.P. 1967. The significance of learners' errors. IRAL 5.161-70.

Craik, Fergus I.M. 1979. Human memory. Annual Review of Psychology 30. 63-102.

Craik, Fergus I.M., and Robert S. Lockhart. 1972. Levels of processing: A Framework for memory research. Journal of Verbal Learning and Verbal Behavior 11.671-84.

Davies, Alan; C. Criper; and A.P.R. Howatt (eds.) 1984. Interlanguage. Edinburgh: University of Edinburgh Press.

Davies, Norman F. 1976. Receptive versus productive skills in foreign language learning. Modern Language Journal 60.440-3.

De Bot, Kees, and Toke Lintsen. 1986. Foreign-language proficiency in the elderly. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els, 131-41. Dordrecht: Foris.

------, and Michael Clyne. 1989. Language reversion revisited. Studies in Second Language Acquisition 11.167-77.

Deno, Stanley L. 1968. Effects of vords and pictures as stiumuli in learning language equivalents. Journal of Educational Psychology 59.202-6.

Dorian, Nancy. 1973. Grammatical change in a dying dialect. Language 49.413-438.

------. 1978. The fate of morphological complexity in language death: evidence from East Sutherland Gaelic. Language 54.590-609.

------. 1981. Language death: the life cycle of a Scottish Gaelic dialect. Philadelphia: University of Pennsylvania Press.

------. 1982. Language loss and maintenance in language contact situations. The loss of language skills, ed. by Richard D. Lambert and Barbara F. Freed, 44-59. Rowley, MA: Newbury House.

Ebbinghaus, H. 1885. Ueber das Gedächtnis. Untersuchungen zur experimentellen Psychologie. Leipzig: Duncker & Humbolt.

Edwards, G. 1976. Second-language retention in the Canadian Public Service. Canadian Modern Language Review 32.305-8.

------. 1977. Second-language retention in the Public Service of Canada. Ottawa Public Service Commission of Canada.

Eysenck, Michael V. 1982. Incidental learning and orienting task. Handbook of research methods in human memory and cognition, ed. by C. Richard Puff, 197-228. New York: Academic Press.

Fakhri, Ahmed. 1985. Linguistic outcomes of foreign language loss: short-term changes in the interlanguage of American learners of French. DAI 46.964A.

Gardner, Robert. 1982. Social factors in language retention. The loss of language skills, ed. by Richard D. Lambert and Barbara F. Freed, 24-43. Rowley, MA: Newbury House.

------; R.N. Lalonde; and J. MacPherson. 1985. Social factors in second language attrition. Language Learning 35.519-40.

Gass, Susan M. 1987. Introduction. Studies in Second Language Acquisition 9.129-32.

------. 1988. Integrating research areas: A framework for second language studies. Applied Linguistics 9.198-217.

------, and Carolyn Madden (eds.) 1985. Input in second language acquisition. (Issues in second language research.) Rowley, MA: Newbury House.

Geoghegan, B. 1950. The retention of certain secondary school subject matter over the period of the summer vacation. Ph.D dissertation. Fordham University.

Ginsberg, Ralph B. 1986. Issues in the analysis of language loss: methodology of the language skills attrition project. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els, 19-36. Dordrecht: Foris.

Godsall-Myers, Jean Elizabeth. 1981. The attrition of language skills in German-classroom bilinguals: A case study. DAI 43.157A.

Gonzo, Susan, and Mario Saltarelli. 1979. Monitoring, pidginization, and immigrant languages. The acquisition and use of Spanish and English as first and second languages, ed. by Roger Andersen, 153-163. Washington, DC: TESOL.

-----,-----. 1983. Pidginization and linguistic change in emigrant languages. Pidginization and creolization as language acquisition, ed. by Roger Andersen, 181-197. Rowley, MA: Newbury House.

Gregg, Kevin R. 1984. Krashen's monitor and Occam's razor. Applied Linguistics 5.79-100.

-----. 1986. Review of The input hypothesis, issues and implications, by Stephen D. Krashen. TESOL Quarterly 20.116-22.

Haastrup, Kirsten. 1989. The learner as word processor. Vocabulary acquisition, ed. by Paul Nation and Ron Carter, 34-46. (AILA Review 6.)

Hansen, Lynne. 1980. Learning and forgetting a second language: The acquisition, loss and re-acquisition of Hindi-Urdu negative structures by English-speaking children. DAI 42.193A.

Hatch, Evelyn. 1978. Apply with caution. Studies in Second Language Acquisition 2.123-43.

-----. 1983. Simplified input and second language acquisition. Pidginization and creolization as language acquisition, ed. by Roger Andersen, 64-86. Rowley, MA: Newbury House.

Henning, Grant H. 1973. Remembering foreign language vocabulary: acoustic and semantic parameters. Language Learning 23.185-196.

Higgs, Theodore (ed.) 1982. Curriculum, competence, and the foreign language teacher. (ACTFL foreign language education series 13.) Skokie, IL: National Textbook Company.

-----. 1984. Teaching for proficiency, the organizing principle. (ACTFL foreign language education series 15.) Lincolnwood, IL: National Textbook Company.

-----. 1985a. The input hypothesis: an inside look. Foreign Language Annals 18.197-203.

-----. 1985b. Language acquisition and language learning: A plea for syncretism. Modern Language Journal 69.8-14.

-----. 1985c. Teaching grammar for proficiency. Foreign Language Annals 18.289-96.

Hosenfeld, Carol. 1977. A learning-teaching view of second-language instruction: the learning strategies of second-language learners with reading-grammar tasks. DAI 38.6581A.

Jarest, Jackie, and Marsha Robinson. 1986. Charlemos un poco. Boston: Heinle and Heinle.

Jaspaert, Koen; Sjaak Kroon; and Roeland Van Hout. 1986. Points of reference in first-language loss research. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els, 37-49. Dordrecht: Foris.

Jenkins, James J. 1974. Remember that old theory of memory? Well, forget it! American Psychologist 29.785-95.

Jordens, Peter; Kees De Bot; and Henk Trapman. 1989. Linguistic aspects of regression in German case marking. Studies in Second Language Acquisition 11.179-203.

Judd, E.L. 1978. Vocabulary teaching and TESOL: A need for reevaluation of existing assumptions. TESOL Quarterly 12.71-6.

Kellogg, Gillian S., and Michael J.A. Howe. 1971. Using words and pictures in foreign language learning. The Alberta Journal of Educational Research 17.89-94.

Kennedy, L.R. 1932. The retention of certain Latin syntactical principles by first and second year Latin students after various time intervals. Journal of Educational Psychology 23.132-46.

Kohn, Susan E.; Arthur Wingfield; Lise Menn; Harold Goodglass; Jean Berko-Gleason; and Mary Hyde. 1987. Lexical retrieval: The tip-of-the-tongue phenomenon. Applied Psycholinguistics 8.245-266.

Koster, Cor J. 1987. Word recognition in foreign and native language. Dordrecht, Netherlands: Foris.

Kramsch, Claire. 1986. From language proficiency to interactional competence. Modern Language Journal 70:366-70.

-----. 1987. The proficiency movement: Second language acquisition perspectives. (Review article.) Studies in Second Language Acquisition 9.355-62.

Krashen, Stephen D. 1982. Principles and practice in second language acquisition. Oxford: Pergamon.

-----. 1985. The Input Hypothesis: Issues and implications. New York: Longman.

-----, and Tracy D. Terrell. 1983. The Natural Approach: Language acquisition in the classroom. Oxford: Pergamon.

Lamadrid, Enrique E.; Laurel A. Briscoe; Robert Phillips; Enrique R. Lamadrid. 1984. Communicating in Spanish: A first course. 2nd edn. Boston: Houghton Mifflin.

Lambert, Richard D., and Barbara F. Freed (eds.) 1982. The loss of language skills. Rowley, MA: Newbury House.

-----, and Sarah J. Moore. 1984. Recent research on language skill attrition. ERIC/CLL News Bulletin 8.1-8.

-----,-----. 1986. Problem areas in the study of language attrition. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els, 177-86. Dordrecht: Foris.

Lantolf, James P., and William Frawley. 1985. Oral proficiency testing: A critical analysis. Modern Language Journal 69.337-45.

Laufer, Batia. 1989. A factor of difficulty in vocabulary learning: Deceptive transparency. Vocabulary acquisition, ed. by Paul Nation and Ron Carter, 10-20. (AILA Review 6.)

Levenston, E.A. 1979. Second language vocabulary acquisition: Issues and problems. The Interlanguage Studies Bulletin 4.147-160.

Leyen, Ina Angelika. 1986. Native language attrition: A study of vocabulary decline. DAI 46.2602A.

Lightbown, Patsy M. 1983. Exploring relationships between developmental and instructional sequences in L2 acquisition. Classroom oriented research in second language acquisition, ed. by Herbert W. Seliger and Michael H. Long, 217-45. Rowley, MA: Newbury House.

-----. 1984. The relationship between theory and method in second language acquisition research. Interlanguage, ed. by Alan Davies, C. Criper, and A.P.R. Howatt, 241-52. Edinburgh: Edinburgh University Press.

-----. 1987. Classroom language as input to second language acquisition. First and second language acquisition processes, ed. by Carol Wollman Pfaff, 169-87. Cambridge: Newbury House.

Lockhart, Robert S.; Fergus I.M. Craik; and Larry Jacoby. 1976. Depth of processing, recognition and recall. Recall and recognition, ed. by John Brown, 75-102. New York: John Wiley.

Long, Michael H. 1983. Does second language instruction make a difference? A review of research. TESOL Quarterly 17.359-82.

Lowe, Pardee, Jr. 1982. The U.S. government's foreign language attrition and maintenance experience. The loss of language skills, ed. by Richard D. Lambert and Barbara F. Freed, 176-90. Rowley, MA: Newbury House.

Ludwig, Jeannette. 1984. Vocabulary acquisition as a function of word characteristics. Canadian Modern Language Review 40.552-62.

Lyons, John. 1977. Semantics, vol. 2. Cambridge: Cambridge University Press.

McLaughlin, Barry. 1979. The Monitor Model: Some methodological considerations. Language Learning 28.309-32.

-----. 1987. Theories of second-language learning. Baltimore: Edward Arnold.

Marton, W. 1977. Foreign language vocabulary learning as problem number one of language teaching at the advanced level. The Interlanguage Studies Bulletin 2.33-57.

Meara, Paul. 1980a. Syntactic structure and memory span in second language learners. Interlanguage Studies Bulletin 5.31-54.

-----. 1980b. Vocabulary acquisition: A neglected area of language learning. Language Teaching and Linguistics: Abstracts 13.221-46.

----- (ed.) 1983. Vocabulary in a second language, vol. 1. London: Centre for Information on Language Teaching and Research.

-----. 1984. The study of lexis in interlanguage. Interlanguage, ed. by Alan Davies, C. Criper, and A.P.R. Howatt, 225-35. Edinburgh: Edinburgh University Press.

----- (ed.) 1987. Vocabulary in a second language, vol. 2. London: Centre for Information on Language Teaching and Research.

-----. 1989. Matrix models of vocabulary acquisition. Vocabulary acquisition, ed. by Paul Nation and Ron Carter, 66-74. (AILA Review 6.)

Melka Teichroew, Francine J. 1982. Receptive vs. productive vocabulary: A survey. Interlanguage Studies Bulletin 6.5-32.

Moorcroft, R., and Robert C. Gardner. 1987. Linguistic factors in second-language loss. Language Learning 37.327-40.

Murdock, Bennet B., Jr. 1982. Recognition memory. Handbook of research methods in human memory and cognition, ed. by C. Richard Puff, 1-26. New York: Academic Press.

Murray, David J. 1986. Characteristics of words determining how easily they will be translated into a second language. Applied Psycholinguistics 7.353-72.

Nation, I.S.P. 1982. Beginning to learn foreign vocabulary: A review of the research. RELC Journal 13.14-36.

Neisser, Ulric. 1984. Interpreting Harry Bahrick's discovery: What confers immunity against forgetting? Journal of experimental psychology: General 113.32-5.

Obler, Loraine K. 1982. Neurolinguistic aspects of language loss as they pertain to second language attrition. The loss of language skills, ed. by Richard D. Lambert and Barbara F. Freed, 60-79. Rowley, MA: Newbury House.

-----, and M.L. Albert. 1984. Language in aging. Clinical neurology of aging, ed. by M.L. Albert, 245-53. New York: Oxford University Press.

Olshtain, Elite. 1986. The attrition of English as a second language with speakers of Hebrew. Language attrition in progress, ed. by Bert Weltens, Kees De Bot, and Theo Van Els, 187-204. Dordrecht: Foris.

-----. 1989. Is second language attrition the reversal of second language acquisition? Studies in Second Language Acquisition 11. 151-65.

Omaggio, Alice C. 1984. The proficiency-oriented classroom. Teaching for proficiency, the organizing principle, ed. by Theodore V. Higgs, 43-84. (ACTFL foreign language education series.) Lincolnwood, IL: National Textbook Company.

O'Malley, J. Michael; Anna Uhl Chamot; and Carol Walker. 1987. Some implications of cognitive theory to second language acquisition. Studies in Second Language Acquisition 9.287-306.

Ostyn, Paul, and Pierre Godim. 1985. RALEX: An alternative approach to language teaching. Modern Language Journal 69.346-53.

Oxford, Rebecca L. 1982a. Research on language loss: A review with implications for foreign language teaching. Modern Language Journal 66.160-9.

-----. 1982b. Technical issues in designing and conducting research on language skill attrition. The loss of language skills, ed. by Richard D. Lambert and Barbara Freed, 119-37. Rowley, MA: Newbury House.

Paivio, Allan. 1976. Imagery in recall and recognition. Recall and recognition, ed. by John Brown, 103-130. New York: John Wiley.

Palmberg, Rolf. 1987. Patterns of vocabulary development in foreign-language learners. Studies in Second Language Acquisition 9.201-19.

Pan, Barbara Alexander, and Jean Berko-Gleason. 1986. The study of language loss: models and hypotheses for an emerging discipline. Applied Psycholinguistics 7.193-206.

Pimsleur, P. 1967. A memory schedule. Modern Language Journal 51.73-5.

Postovsky, Valerian A. 1974. Effects of delay in oral practice at the beginning of second-language learning. Modern Language Journal 58.229-39.

Pratella, W.C. 1969. The retention of first and second year Spanish over the period of the summer vacation. Ph.D. dissertation. Fordham University.

Puff, C. Richard. 1982. Handbook of research methods in human memory and cognition. New York: Harcourt, Brace, Jovanovich.

Rivers, Wilga M. 1986. Comprehension and production in interactive language teaching. Modern language journal 70.1-7.

Robison, Robert Eugene. 1985. The effect of the summer vacation on language attrition in secondary school students of first-year Spanish. DAI 46.636A.

Sapir, Edward. 1921. Language. New York: Harcourt, Brace & World.

Savignon, Sandra. 1985. Evaluation of communicative competence: the ACTFL Provisional Guidelines. Modern Language Journal 69.129-34.

-----, and Lyle F. Bachman. 1985. The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. Modern Language Journal 70.380-9.

Scarcella, Robin, and Leroy Perkins. 1987. Shifting gears and Krashen's input hypothesis. (Review article.) Studies in Second Language Acquisition 9.347-53.

Scherer, George A.C. 1957. The forgetting rate in learning German. The German Quarterly 30.275-7.

Schmidt, Richard V. The role of consciousness in second language learning. (to be published in Applied Linguistics)

Schouten-Van Parreren, Caroline. 1989. Vocabulary learning through reading: Which conditions should be met when presenting words in texts? Vocabulary acquisition, ed. by Paul Nation and Ron Carter, 75-85. (AILA Review 6.)

Schulz, Renate. 1986. From achievement to proficiency: some caveats. Modern Language Journal 70.373-9.

Seliger, Herbert V., and Michael H. Long (eds.) 1983. Classroom oriented research in second language acquisition. Rowley, MA: Newbury House.

Sharwood Smith, Michael. 1981. Consciousness raising and the second language learner. Applied Linguistics 11.159-168.

———. 1984. Discussion of 'The study of lexis in interlanguage.' Interlanguage, ed. by Alan Davies, C. Criper, and A.P.R. Howatt, 236-9. Edinburgh: University of Edinburgh Press.

———. 1986. Comprehension vs. acquisition: The ways of processing input. Applied Linguistics 7.239-56.

Sheldon, Amy, and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. Applied Psycholinguistics 3.243-61.

Smythe, P.C., G.C. Jutras, J.R. Bramwell, and Robert C. Gardner. 1973. Second language retention over varying intervals. Modern Language Journal 57.400-5.

Spinelli, Emily, and Marta Rosso-O'Laughlin. 1988. Encuentros. New York: Holt, Rinehart, Winston.

Spitze, K., and Susan D. Fischer. 1981. Short-term memory as a test of language proficiency. TESL Talk 12.32-41.

Spolsky, Bernard. 1985. Formulating a theory of second language learning. Studies in Second Language Acquisition 7.269-88.

Stevick, Earl W. 1976. Memory, meaning, and method: Some psychological perspectives on language learning. Rowley, MA: Newbury House.

Swain, Merrill. 1985. Communicative competence: some roles of comprehensible output in its development. Input in second language acquisition, ed. by Susan Gass and Carolyn Madden, 236-56. Rowley, MA: Newbury House.

———, and Sharon Lapkin. 1989. Canadian immersion and adult second language teaching: What's the connection? Modern Language Journal 73.150-9.

Terrell, Tracy David. 1986a. Acquisition in the Natural Approach: The binding/access framework. Modern Language Journal 70.211-27.

———. 1986b. Recent trends in research and practice: Teaching Spanish. Hispania 69.193-201.

———. 1989. The role of grammar instruction in a communicative approach. NE Conference Newsletter. 22-5.

———; Magdalena Andrade; Jeanne Egasse; and Elías Miguel Muñoz. 1986c. Dos mundos. New York: Random House.

Tulving, Endel. 1976. Ecphoric processes in recall and recognition. Recall and recognition, ed. by John Brown, 37-74. New York: John Wiley.

Tversky, Barbara. 1973. Encoding processes in recognition and recall. Cognitive Psychology 5.275-87.

Twaddel, F. 1973. Vocabulary expansion in the ESOL classroom. TESOL Quarterly 7.61-78.

VanPatten, Bill. 1984. Communicative values and information processing in L2 acquisition. On TESOL '84. A brave new world for TESOL. (Selected papers from the annual TESOL convention.)

-----. 1987. On babies and bathwater: Input in foreign language learning. Modern Language Journal 71.156-64.

-----. 1989. Can learners attend to form and content while processing input? Hispania 72.409-17.

Weltens, Bert. 1987. The attrition of foreign language skills: A literature review. Applied Linguistics 8.22-38.

-----; Kees De Bot; and Theo Van Els (eds.) 1986. Language attrition in progress. Dordrecht: Foris.

-----; Theo Van Els, and Erik Schils. 1989. The long-term retention of French by Dutch students. Studies in Second Language Acquisition 11.205-15.

Wingfield, Arthur, and Dennis L. Byrnes. 1981. The psychology of human memory. New York: Academic Press.

Winitz, Harris (ed.) 1981. The comprehension approach to foreign language instruction. Rowley, MA: Newbury House.