AD-A254 263

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1992	XXXXXXXXDISSERTATION

4. TITLE AND SUBTITLE
Do First Language Writing and Second Language Reading Equal Second Language Reading Comprehension? An Assessment Dilemma

5. FUNDING NUMBERS

6. AUTHOR(S)
Judith E. Brisbois, Major

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Student Attending: Ohio State University

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA- 92-008D

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFIT/CI
Wright-Patterson AFB OH 45433-6583

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW 190-1
Distributed Unlimited
ERNEST A. HAYGOOD, Captain, USAF
Executive Officer

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

14. SUBJECT TERMS	15. NUMBER OF PAGES 233
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

DO FIRST LANGUAGE WRITING AND SECOND LANGUAGE READING EQUAL SECOND LANGUAGE READING COMPREHENSION? AN ASSESSMENT DILEMMA

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By

Judith E. Brisbois, B.A., M.A.

*****

The Ohio State University

1992

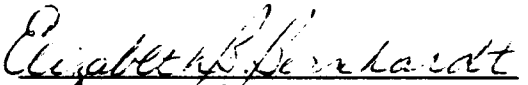
Dissertation Committee:

Elizabeth B. Bernhardt

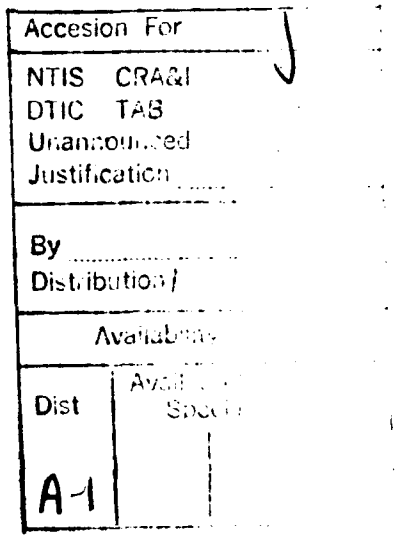
George E. Newell

Keiko K. Samimy

Approved by



Adviser
College of Education



To My Parents

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor, Dr. Elizabeth Bernhardt, whose research insights, guidance, and support proved invaluable in this endeavor. My gratitude also goes to Dr. George E. Newell, Dr. Keiko Samimy, and Dr. Michael Kamil for their perceptive comments regarding my research. Grateful acknowledgement is also extended to my colleagues, particularly Karen Angell, Mary Lavin-Crerand, Frank Gahren, Srivalli Sastry-Kuppa, Amy Burrows, and Mike Nolan for their help in scoring the many tests involved in this study, and to Craig Deville and Micheline Chalhoub for their assistance with the statistical design. Special thanks go to my parents, and to Susan Knight, Karen and John Angell, and Ellen Collie for their unending support, encouragement, and tolerance. My gratitude also goes to the organization that sponsored my doctoral program and in which I am honored to serve: the United States Air Force.

VITA

October 16, 1956.....	Born - Worcester, Massachusetts
1978.....	B.A., Framingham State College, Framingham, Massachusetts
1979-1981.....	Fighter Intercept Control Officer, USAF, Tyndall AFB, Florida
1981-1982.....	Fighter Intercept Control Officer, Osan AB, Korea
1982-1985.....	Fighter Intercept Control Instructor, Tyndall AFB, Florida
1985.....	M.A., University of Oklahoma, Norman, Oklahoma
1985-1989.....	Instructor of French, USAF Academy, Colorado Springs, Colorado

FIELDS OF STUDY

Major Field: Education

Professors Elizabeth B. Bernhardt, Gilbert A. Jarvis, and Keiko K. Samimy

Satellite Area: Instructional Design and Technology

Professors John C. Belland and Marjorie A. Cambre

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURES PAGE

CHAPTER I

THE PROBLEM

Introduction

During the past decade, the back-to-basics movement in education and its foreign language (L2) corollary, the proficiency movement, have stimulated an increased emphasis on testing. The National Assessment of Educational Progress (Alexander & James, 1987), the National Teacher Examination, and the Governors' Task Force on Teaching (National Governors' Association, 1986) all point to the use of tests to measure the achievement of the nation's educational system and to improve its performance. In the L2 realm, the American Council on the Teaching of Foreign Languages echoes the call for increased assessment through the use of the Proficiency Guidelines. According to Liskin-Gasparro (1984), the Guidelines promise to serve as the force behind "proficiency-based curricula, materials, evaluation instruments, and articulation plans" (p. 13).

Because of this focus on assessment, tests must be constructed with an eye toward increased validity. In the L2 arena, research is needed that will investigate L2 reading test validity and lead to its enhancement.

In recent years, increased discussion has ensued regarding the validity of L2 reading tests. One facet of the assessment debate centers on the decrease in construct validity that results when readers are required to write their responses on L2 reading assessment measures in their L2s. Rather than representing a true measure of reading comprehension, the resulting scores reflect a combination of L2 reading skill and L2 writing ability. Hock and Poh (1979), Lee (1986), and Shohamy (1984) document that native language (L1) responses yield a clearer picture of what has been comprehended.

The potential problem inherent in requiring L1 responses, however, constitutes a variation on the L2 reading-L2 writing assessment dilemma. Do the scores on L2 reading comprehension measures requiring L1 responses depict a composite of L2 reading ability and L1 writing skill? Research is needed that will determine the contribution of L1 writing ability to the scores on L2 reading comprehension tests.

A number of studies have demonstrated that reading assessment measures requiring target language responses reveal less evidence of comprehension than do those that call for responses in the native language (Hock & Poh, 1979; Lee, 1986; Shohamy, 1984). According to Weber (1991), when L2 readers are asked to respond "in their

weaker second language, they may be limited in their ability to demonstrate their level of understanding" (p. 109). The practice of requiring readers to write their responses in the L2 results in a confounding of the productive skill of L2 writing and the receptive skill of L2 reading. The resulting test scores render a combination of L2 reading comprehension and L2 writing ability.

Although L2 reading assessment measures requiring responses written in the native language yield scores that more accurately represent what has been comprehended from L2 text (Bernhardt, 1991; Hock & Poh, 1979; Lee, 1986; Shohamy, 1984; Weber, 1991), the reception-production dilemma may still influence test scores. The fusion of L2 reading comprehension and L1 writing ability could result in such scores representing a blend of both factors.

Statement of the Problem

Validity, which refers to the degree to which a test "measures the ability or knowledge that it is purported to measure" (Henning, 1987, p. 196), is "the most important feature of a test" (Bernhardt, 1991, p. 192). In order for a reading comprehension test to be valid, it must accurately and appropriately tap the construct, or unobservable trait, known as reading comprehension. According to Messick (1988), "although construct-related

evidence may not be the whole of validity, there can be no validity without it" (p. 35).

Although research documents that L2 reading assessment measures calling for written L2 responses lack construct validity (Hock & Poh, 1979; Lee, 1986; Shohamy, 1984), requiring readers to respond in their L1s may not completely alleviate the problem. Indeed, do L1 literacy abilities, such as writing capacity, contribute to the scores on L2 reading assessment measures? If so, what proportion of each score is attributable to L1 literacy variables, such as reading capability and writing skill? Furthermore, what proportion is attributable to L2 literacy components such as vocabulary knowledge and grammatical ability? In order to better understand what is actually represented by the scores on L2 reading assessment measures, research is needed that will investigate the unique contributions of selected L1 and L2 literacy variables to L2 reading comprehension scores.

A stronger grasp of what constitutes the scores on reading assessment measures will lead researchers to insights into the reading comprehension process, and will better equip instructors, administrators, and curriculum developers to generate sound pedagogical decisions. Expansion of the research knowledge base will point to enhanced instructional practices, to improvements in

curriculum development and, hence, to increased learning. Furthermore, a more sophisticated understanding of what L2 reading test scores reveal will lead to more accurate diagnoses of and solutions to learning difficulties, and to the more appropriate placement of students into course levels. As insight into these issues develops, researchers and instructors will be able to more accurately interpret the scores on L2 reading comprehension measures, and to formulate more expedient decisions based on their interpretations.

Purpose of the Study

The research questions addressed by the present study are as follows:

1. What is the unique contribution to L2 reading comprehension of each of the following:

a. L1 reading comprehension as measured by the score on the comprehension portion of the Nelson-Denny Reading Test, Form E (Brown, Bennett, & Hanna, 1981), and then by the mean score on three recall protocols.

b. L1 writing ability as measured by the combined holistic scores on three writing samples.

c. Passage-specific L2 vocabulary knowledge as measured by the score on a supply-definition test.

d. L2 grammatical skill as measured by the score on the grammar section of the United States Air Force

Academy's (USAFA's) French Placement Validation Test (PLAVAL).

2. Does the unique contribution of each independent variable differ by level of L2 study?

Definition of Terms

Calibration Session: A "training session during which raters read and rate sample essays that correspond to the levels of the scale to be used for holistic scoring" (Tedick, 1988, p. 27).

L2 Grammatical Knowledge: The subjects' ability to recognize and use French grammatical forms as determined by the scores on the grammar section of USAFA's French PLAVAL.

Holistic Scoring: A rapid method of writing assessment in which raters assign a score based upon their overall impression of the text.

Passage-Specific L2 Vocabulary Knowledge: The proportion of vocabulary words from the French recall passages, appearing on a supply-definition test, for which the subjects were able to write either a definition (in English) or an English equivalent.

Pausal Unit: The words surrounded by locations in a text that are "acceptable for pausing to catch a breath, give emphasis to a story, or to enhance meaning" (Johnson, 1970, p. 13).

Placement Validation Test (PLAVAL): A test of foreign language ability administered to most freshman cadets entering USAFA. The scores are used to determine student placement into course levels. Cadets scoring high enough on the PLAVAL are exempted (validated) from USAFA's foreign language requirement (Westfall, 1988).

Reading Comprehension: The score generated by the combined values of the weighted pausal units recalled after reading a passage, scored according to Johnson's propositional analysis system (Johnson, 1970), or the score attained on the comprehension portion of the NDRT, Form E (Brown et al., 1981).

Recall Protocol: A comprehension measure for which subjects write in their native language all that they can remember about a given text/passage.

Text/Passage: A reading of approximately 250 words.

Writing Ability: The quality of the recall protocols that the subjects wrote, in English, after reading three passages written in English. The overall quality was determined by combining the holistic scores of the three protocols. The holistic scores were arrived at using the Test of Written English Scoring Guide (Educational Testing Service, 1990).

Theoretical Considerations

Test Theory

Test validity is the focus of the present study because this study sheds light on the assessment dilemma centering on what L2 reading comprehension test scores actually represent. "Validity is the most important consideration in test evaluation," according to a joint commission of the American Psychological Association, the American Educational Research Association, and the National Council of Mathematics Educators (in Bernhardt, 1991, p. 192). Messick (1988) attests to the crucial nature of validity to the testing process in what he refers to as a "unified view" of validity (p. 33):

> Validity is an overall evaluative judgment, founded on empirical evidence and theoretical rationales, of the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores. As such, validity is an inductive summary of both the adequacy of empirical evidence for and the appropriateness of test interpretation and use (pp. 33-34).

According to Messick (1988), the unifying force of the "unified view of validity" (p. 35) is "empirically grounded construct interpretation" (p. 35). Thus, because construct validity serves as the cornerstone of test validity, it is impossible for test validity to exist without "construct-related evidence" (Messick, 1988, p. 35). In other words, in order to serve as a successful measurement

device, a test must accurately and appropriately measure the construct that it purports to measure. In order for a test to provide construct validity, it must "match as closely as possible what is known about a process at any point in time" (Bernhardt, 1991, p. 192).

An instrument that claims to measure the construct of L2 reading comprehension, therefore, must tap it in a manner that accurately reflects current theory- and research-based knowledge of the L2 reading comprehension process. Messick (1988) continues, "The more systematically and empirically grounded that the construct understanding of the criterion domain is, the more rational foundation is afforded for forecasting likely relationships" (p. 36). Thus, the accuracy of the interpretations and the appropriateness of the decisions that are based on the results of L2 reading comprehension tests hinge on the degree of construct validity that these assessment measures provide.

Although test validity is the principal focus of this study, the importance of reliability cannot be ignored. A test is reliable if it measures the construct(s) under investigation in an accurate, consistent, and dependable manner (Henning, 1987). According to Green (1981), "Validity is the ultimate index of utility, but validity is impossible without reliability" (p. 1004). Thus, ideally,

the more reliable the test, the more valid its results and the more fruitful the interpretations based on those results.

The present study has investigated whether or not the practice of requiring written L1 responses on L2 reading comprehension tests results in decreased construct validity. In other words, does L1 writing ability interfere with the successful measurement of L2 reading skill by contributing to the scores on L2 reading comprehension tests?

If so, L1 writing ability, along with other literacy variables such as L1 reading skill, L2 vocabulary knowledge, and L2 grammatical ability, could act as confounding variables prohibiting both valid interpretations of L2 reading scores and productive decisions based on them. Investigation of the unique contributions of such factors will aid in clarifying the extent to which they affect the scores on L2 reading comprehension measures.

Literacy Transfer

Cummins's (1981) Interdependence Hypothesis, which asserts that literacy skills are able to transfer between languages, buttresses the proposition that L1 writing ability constitutes a portion of the scores on L2 reading assessment measures. The results of numerous studies on

students in bilingual programs pointed Cummins (1981) to the idea that a common underlying language proficiency aids in the development of a bilingual's literacy skills in both the L1 and the L2. The Interdependence Hypothesis maintains that experience with either language promotes development of the capacities underlying both languages:

> To the extent that instruction in Lx [i.e., Language x] is effective in promoting proficiency in Lx, transfer of this proficiency to Ly will occur provided there is adequate exposure to Ly (either in school or environment) and adequate motivation to learn Ly. (p. 29)

In other words, the literacy abilities that learners develop in one language are able to transfer to other languages and vice versa.

Research on language transfer supports Cummins's (1981) Interdependence Hypothesis and upholds his assertion that a common underlying proficiency aids in language learning. Carson, Carrell, Silberstein, Kroll, and Keuhn (1990), for example, discovered a positive correlation between the L1 and L2 reading comprehension scores of adult Chinese and Japanese English as a second language (ESL) learners. According to Carson et al. (1990), these data point to a transfer of reading skills from the L1 to the L2.

In an investigation of the transfer of writing ability, Canale, Frenette, and Bélanger (1988) examined the writing skills of 9th and 10th graders enrolled in L1 French and L2 English courses in Ontario. Again, the data indicate a positive correlation between L1 and L2 writing scores. According to Canale et al. (1988), these results uncover a common underlying proficiency in writing skills across languages.

Literacy transfer has not been found to occur in all cases, however, supporting Cummins's (1981) assertion that transfer capability is possible only after learners attain a threshold level of L2 proficiency. Clarke (1980), for example, contends that exposure to difficult tasks can result in a "short circuit" in L2 reading comprehension skills, causing proficient readers to rely on the strategies that less proficient readers normally use. Similarly, Cziko (1978) argues for the existence of a developmental order in the ability of the L2 reader to use syntactic, semantic, and discourse constraints. Thus, some threshold level of L2 language proficiency appears to be necessary for skilled L1 readers to maintain their proficiency while reading in the L2.

Both theoretical knowledge and research evidence support the idea of literacy skills transfer. Thus, it is expected that L1 literacy abilities, such as writing

skills, contribute to the scores on L2 reading assessment measures requiring written L1 responses. Furthermore, because a developmental threshold appears to affect the degree to which literacy abilities are able to transfer, the contribution of L1 writing skill to L2 reading comprehension test scores should differ by course level. Research is needed that will examine the extent of this contribution and demonstrate whether or not it changes by level of L2 study.

Reading-Writing Relationship

The notion of a reading-writing relationship has long piqued the interest of scholars, theorists, and researchers. The Sophists of Ancient Greece were the first to consider the existence of a relationship between reading ability and writing skill. Theoretical knowledge regarding the reading-writing relationship is relevant to the present study because if literacy abilities transfer between languages, and if reading skill and writing capacity transfer to each other, then L1 writing ability is expected to contribute to the scores on L2 reading assessment measures.

Currently, three schools of thought exist regarding the nature of the reading-writing relationship. The first argues that through the printed page, readers will absorb the "style, usage, and background knowledge" (Bélanger,

1987, p. 10) necessary to become proficient writers. The second asserts that writing development is "a sufficient, though not a necessary, precursor to the successful acquisition of reading ability" (Shanahan & Lomax, 1988, p. 201). The third contends that reading and writing are complementary facets of the same general language ability and, thus, are affected by the same influences. The specific nature of the reading-writing relationship, however, remains largely undiscovered (Bèlanger, 1987).

Through the years, reading and writing have been perceived as opposite ends of a language continuum stretching from the receptive extreme (reading) to the productive extreme (writing). As a result, researchers and theorists have tended to treat the two abilities as separate, albeit related skills. Correlational techniques have frequently been used to determine the extent to which skilled readers are proficient writers, to learn whether or not training in one skill leads to improved performance in the other, and to explore the notion that increased experience with one skill leads to improved performance in the other (Bèlanger, 1987; Stotsky, 1983).

The correlational studies of the past three decades have attempted to gauge the relationship between product-based measures of reading and writing in order to explore the hypothesis that the same cognitive processes

account for proficiency in both areas. This vein of research, mainly involving children, has consistently revealed a positive correlation between reading and writing skills, pointing to the existence of common underlying processes (Carson et al., 1990; Juel, Griffith, & Gough, 1986; Langer, 1984; Schewe & Froese, 1986). The evidence that these studies have provided has supported and contributed to the development of a theoretical framework (Bélanger, 1987; Tierney & Shanahan, 1991).

Theorists such as Wittrock (1983) and Squire (1983) view reading skill and writing ability as employing the same cognitive processes. According to Wittrock (1983), both reading and writing involve "generative cognitive processes" (p. 600) that allow the learner to construct meaning by relating the text to personal experience. Readers generate meaning by relating the text to prior knowledge while writers induce meaning by relating prior knowledge to the text. The same cognitive processes appear to be used in both reading and writing (Wittrock, 1983). Squire (1983) evinces agreement with Wittrock by stating that composing and comprehending are "two sides of the same basic process" (p. 581) of constructing and reconstructing ideas.

Tierney and Pearson (1983) further develop Wittrock's (1983) and Squire's (1983) notion of shared cognitive processes in reading and writing through their composing model of reading. According to this model, both reading and writing are essentially composing processes that include five principal characteristics: planning, drafting, aligning, revising, and monitoring. In the planning stage, goals are set and knowledge is mobilized. The drafting phase refers to the "refinement of meaning which occurs when readers and writers deal directly with the print on the page" (Tierney & Pearson, 1983, p. 571). At the alignment level, readers and writers adopt a stance. When refining, they examine their developing views. Finally, during the monitoring phase, readers and writers step back "from the texts they have created to evaluate what they have developed" (Tierney & Pearson, 1983, p. 577). Thus, the model illustrates the idea that the same general cognitive processes are used in both reading and writing (Tierney & Pearson, 1983).

Shanahan and Lomax (1986, 1988) investigated the pedagogical value of three models of the reading-writing relationship: (a) reading affects writing, (b) writing affects reading, and (c) the Interactive Model, which asserts that reading and writing influence each other. The data indicate that the Interactive Model affords the best

description of the reading-writing relationship (Shanahan & Lomax, 1986, 1988). In other words, the results lend further credence to the notion that the same general cognitive processes are employed in both reading and writing.

In summary, both theoretical and empirical evidence point to the existence of a relationship between reading skill and writing ability. Evidence of a reading-writing relationship, coupled with that of the transfer of skills between languages, lead to the hypothesis that L1 writing ability contributes to the scores on L2 reading comprehension measures. Research is needed that will investigate the magnitude of this contribution, and examine whether or not it changes by L2 course level.

The L2 Reading Comprehension Process

Bernhardt's (1986) Constructivist Reading Model contributes a valuable theoretical foundation to this study because it takes into account the roles in the L2 reading comprehension process of two of the literacy variables that were explored: L2 vocabulary knowledge and L2 grammatical ability. This six-part model, generated from German, French, and Spanish recall protocol data, integrates both text-based and extratext-based components. The text-based elements include phonemic/graphemic features, syntactic feature recognition (grammatical ability), and word

recognition (vocabulary knowledge), while the extratext-based features comprise intratextual perceptions, prior knowledge, and metacognition. Because the model is non-linear, the six elements are free to interact and to meld together in varying combinations in order to forge the construct of L2 reading comprehension (Bernhardt, 1986).

This model is particularly relevant to the present study because although it attests to the importance of word recognition (vocabulary knowledge) and syntactic feature recognition (grammatical ability), the unique contributions of those elements to the L2 reading comprehension process remain unspecified. Such flexibility is crucial because the contributions may differ by language, by orthography, by course level, or by a host of other variables. The components may influence one another in diverse ways and to different degrees depending on the situation (Everson, 1986). The present study has aided in the testing of this model because it has explored the extent to which two of its elements--word recognition and syntactic feature recognition--contribute to the L2 reading comprehension scores of adult native English speakers. Additionally, this study has examined whether or not the magnitudes of the contributions differ by course level.

Assumptions

The present study was based on the following assumptions:

1. The subjects would perform all tasks to the best of their ability.

2. The subjects' placement into beginning, intermediate, or advanced courses accurately reflected their language ability levels.

3. The recall protocol scores would accurately reflect reading comprehension.

4. The NDRT comprehension scores would accurately depict reading comprehension.

5. The holistic writing scores would accurately represent writing ability.

6. The results of the passage-specific vocabulary test would accurately indicate the subjects' passage-specific vocabulary knowledge.

7. The scores on the grammar portion of the PLAVAL would accurately reflect the subjects' knowledge of French grammar.

Limitations

1. Because hierarchical regression was used, replication is necessary in which the data will be entered into the statistical model in different orders.

2. Replication using native speakers of languages other than English and learners of languages other than French is advisable.

3. Because the English and French reading passages are all historical in nature, replication using other text genres is recommended.

4. Because the topics of the English and French reading texts all pertain to French history, replication using alternate topics is desirable.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

In order to examine the validity of second language (L2) reading comprehension measures, this study sought to establish the unique contribution of each of the following literacy variables to L2 reading test scores: (a) native language (L1) reading comprehension, (b) L1 writing ability, (c) L2 vocabulary knowledge, and (d) L2 grammatical skill. Of particular relevance to the present study are those experiments relating to the following areas: (a) the effect of the language of response on L2 reading comprehension test scores, (b) the transfer of literacy skills between languages, (c) the reading-writing relationship, and (d) the effect of vocabulary knowledge and grammatical skill on reading comprehension. Although L2 research is emphasized in this review, selected L1 studies will be considered because they provide the theoretical and empirical framework upon which many of the L2 studies have been based.

Language of Response on L2 Reading Assessment Measures

The language of response on L2 reading assessment measures is of central importance to the present study because of its focus on the construct validity issues that ensue when reading test scores reflect a combination of writing ability and reading comprehension. Hock and Poh (1979), Shohamy (1984), and Lee (1986) argue that the use of the L1 as the language of response affords a clearer picture of what has been comprehended than does the use of the L2. L1 responses, though yielding more information about what has been understood than L2 responses, may not completely rectify the construct validity problem.

In a study of 39 Malaysian adults enrolled in an English as a foreign language (EFL) program, Hock and Poh (1979) administered two multiple choice reading comprehension tests written in English, and one in Malay. Subjects scored higher on the test written in their native language. According to Hock and Poh (1979), multiple choice test items written in the L1 "seem to measure more accurately the understanding of the students" (p. 87).

Shohamy (1984) administered Hebrew and English versions of multiple choice and open-ended reading comprehension tests to 2,000 adult native Hebrew speakers

enrolled in an EFL program. The subjects, particularly those at the beginning level, scored higher on the test versions that were written in the L1 and that required written L1 responses. Shohamy's (1984) explanation for the difference in scores is that the use of the L2 causes "an unnecessary source of difficulty associated with the unknown vocabulary in the questions and distractors" (p. 157).

Lee's (1986) replication of Carrell's (1983) study also illustrates that the language of response influences the results of L2 reading comprehension measures. Carrell (1983) required a group of English as a second language (ESL) learners to read an English text and then to write their recall protocols in English. In contrast, Lee (1986) allowed the subjects to respond in their L1s, which resulted in different patterns of recall. According to Lee (1986), "Assessing comprehension with the *native* language allows learners to more fully demonstrate their comprehension" (p. 353).

In summary, empirical evidence documents that readers are able to express more fully what they have comprehended from L2 texts when they are allowed to respond in their native languages. Research in this area demonstrates that tests requiring readers to write their responses in their weaker L2s lead to a distorted picture

of what has actually been comprehended from text. This practice decreases construct validity and diminishes the meaningfulness of the resulting test scores because the scores actually represent a fusion of L2 writing ability and L2 reading comprehension. The present study has taken this assessment dilemma one step further by clarifying the extent to which native language writing ability contributes to the scores on L2 reading assessment measures, and by ascertaining the degree to which this contribution differs by course level.

Transfer of Literacy Skills Between Languages

Previous research in the domain of literacy transfer serves as a key building block in the foundation of the present study. If literacy skills transfer between languages, and if reading skill and writing ability transfer to one another, L1 reading capacity and L1 writing ability should contribute to L2 reading comprehension test scores.

L1 Reading to L2 Reading

Research provides ample evidence that L1 reading skills transfer to the L2. Groebel (1980), Carson, Carrell, Silberstein, Kroll, and Keuhn (1990), Goldman, Reyes, and Varnhagen (1984), Reyes (1987), and Wagner, Spratt, and Ezzaki (1989) discovered positive correlations in their subjects' scores on L1 and L2

reading comprehension measures. Cziko (1978) and Clarke (1978, 1980) maintain that reading skills can transfer, but only after the reader has attained a certain threshold of L2 reading ability.

In an examination of the relationship between L1 and L2 reading skills, Groebel (1980) found a positive correlation between the scores on standardized L1 and L2 reading tests administered to 383 native Hebrew-speaking ESL learners. Both the Hebrew and English assessment measures consisted of three texts of general interest, each followed by five multiple choice questions. Groebel (1980) links her results to the notion of language transfer by suggesting that it might "be helpful to spend some time improving reading skills in the native language in order to determine whether this could improve reading comprehension in the target language" (p. 59).

Like Groebel (1980), Carson et al. (1990) discovered positive correlations between their subjects' L1 and L2 reading comprehension test scores. Forty-eight Chinese and 57 Japanese ESL learners completed cloze reading passages in English and in their L1s. The researchers assert that the results lend credence to the supposition that reading skills can transfer.

Goldman et al. (1984) compared the reading comprehension scores of Spanish-English bilingual fourth graders in both English and Spanish. Twelve subjects produced oral recall protocols for two stories in Spanish and two in English. The children were instructed to recall in their language of choice. Like Groebel (1980) and Carson et al. (1990), Goldman et al. (1984) discovered a positive correlation between the scores on the protocols based on L1 readings and those based on L2 readings. According to Goldman et al. (1980), the data support the notion of a transfer of reading skills from the L1 to the L2.

In a study of 48 Spanish-English bilingual fourth graders, Reyes (1987) required her subjects to read two texts in English and two in Spanish, and then to produce oral recall protocols for each. A strong positive correlation ensued between the proportions of information recalled from the English and Spanish reading passages. According to Reyes (1987), "The reading skills already available in Spanish are easily transferred to English" (p. 125).

Like Goldman et al. (1984) and Reyes (1987), Wagner et al. (1989) compared the L1 and L2 reading test scores of children. In a longitudinal study, Wagner et al. (1989) administered multiple choice Arabic (language of

first literacy) and French reading measures to a total of 166 Moroccan children when they were in the third grade and then when they were in the fifth grade. Positive correlations emerged between the Arabic and French reading scores, and increased along with grade level. According to Wagner et al. (1989), the results "support the hypothesis that first literacy provides an important underlying structure on which to build second literacy acquisition" (p. 45).

Although the results of the reading transfer studies support Cummins's (1981) Interdependence Hypothesis, and point to the existence of a common underlying proficiency, additional inquiry such as the present study is needed in order to examine the unique contributions of selected L1 literacy abilities to the scores on L2 reading comprehension measures. The reading assessment instruments used by Groebel (1980), Wagner et al. (1989), and Carson et al. (1990) may not have been completely accurate in their measurements of L2 reading comprehension. Groebel (1980) and Wagner et al. (1989) used multiple choice tests, which are often passage independent. According to Bernhardt (1991), readers are frequently able to answer multiple choice test questions without reading the associated passage(s). Additionally, because of the difficulty involved in developing multiple

choice questions that tap knowledge at the integrative level, discrete-point information is often tested. Furthermore, although it is unclear whether or not Groebel's L1 passages were authentic, she states that the L2 passages consisted of translations of L1 texts from the alternate form of the L1 reading test. In order to attain a more realistic depiction of text comprehension, authentic texts should be used (Bernhardt, 1991).

The use of cloze tests by Carson et al. (1990) may not have garnered scores that accurately reflect the construct of reading comprehension. Shanahan, Kamil, and Tobin (1982), Markham (1985), and Kamil, Smith-Burke, and Rodriguez-Brown (1986) document that cloze testing does not measure reading comprehension at the inter-sentential level. Additionally, it may not have been appropriate to cloze the Chinese or Japanese passages "because there is no clear definition of a word as an orthographic unit in each of these writing systems" (Carson et al., 1990, p. 252).

Goldman et al.'s (1984) and Reyes's (1987) use of the recall protocol procedure to assess reading comprehension alleviates the problems associated with passage independence, the testing of discrete-point knowledge, and the measurement of reading skill only at the intra-sentential level (Bernhardt, 1991). In order

to expand upon their results, however, the present study used a test genre other than narrative.

Additionally, unlike Goldman et al. (1984), Reyes (1987), and Wagner et al. (1989), the present study featured adult subjects. The results of studies examining children's literacy abilities are somewhat limited because children are still developing cognitively and are in the process of becoming literate. Data gathered from more mature subjects offer a clearer picture of the relationship between L1 and L2 reading ability because adults, whose cognitive development has stabilized, are already fully literate.

Although Groebel (1980), Carson et al. (1990), Goldman et al. (1984), Reyes (1987), and Wagner et al. (1989) discovered positive correlations between L1 and L2 reading abilities, Cziko (1978) and Clarke (1978, 1980) contend that this relationship is limited by language ability level. In a study of 96 native English-speaking elementary school learners of French as a second language, Cziko (1978) found that proficient L1 readers experience difficulties using L2 syntactic, semantic, and discourse constraints until they reach a sufficiently high level of L2 reading skill.

In a case study of one proficient and one poor native Spanish-speaking adult ESL learners of similar L2 ability, Clarke (1978) found that proficient L1 readers are also proficient L2 readers. According to Clarke (1978), however, the fact that the good L1 reader comprehended more in the L1 than in the L2 points to the probability that a threshold of L2 proficiency is needed before skilled L1 readers are able to maintain their proficiency in the L2.

Similarly, in a study of 21 native Spanish-speaking adult ESL learners, Clarke (1980) discovered that although proficient L1 readers can transfer some skills to the L2, exposure to difficult L2 tasks can result in excessive reliance on the syntactic clues normally used by less proficient readers. Clarke (1980) asserts that as task difficulty increases, L1 reading capacity may "short circuit," causing skilled readers to revert to the comprehension strategies that less proficient readers normally employ.

In summary, extensive research evidence reveals that L1 reading skills transfer to the L2. According to Carson et al. (1990), Goldman et al. (1984), Groebel (1980), Reyes (1987), and Wagner et al. (1989), the positive correlations between the scores on L1 and L2 reading comprehension tests indicate that reading skills

can and do transfer. Cziko (1978) and Clarke (1978, 1980) argue that although reading skills can transfer, the process is affected by the level of L2 reading ability. The present study has expanded upon these findings by accomplishing the following: (a) assessing reading comprehension using more sophisticated assessment measures than cloze or multiple choice tests, (b) involving adult subjects, (c) ascertaining the contribution of L1 literacy skills to the scores on L2 reading tests, and (d) examining whether or not those contributions differ by course level.

L1 Writing to L2 Writing

Empirical evidence reveals that writing ability transfers from the L1 to the L2. The writing studies reviewed can be divided into two principal areas: (a) product and (b) process. In the product-oriented domain, Canale, Frenette, and Bèlanger (1988) and Carson et al. (1990) correlated the scores on their subjects' L1 and L2 writing samples. In explorations of the writing process, Edelsky (1982), Jones and Tetroe (1987), Lay (1982), and Arndt (1987) investigated the similarities and differences in the strategies and procedures that their subjects used while writing in the L1 versus the L2.

In an examination of L1 and L2 writing products, Canale et al. (1988) collected writing samples from 32 French-speaking 9th and 10th graders enrolled in an ESL program. Each subject wrote two essays in English and two in French. After the 128 writing samples were scored both holistically and analytically, positive correlations emerged between the L1 and L2 essay scores. According to Canale et al. (1988), the data point to the transfer of writing ability between languages.

Also exploring the product domain, Carson et al. (1990) analyzed the writing samples of 48 Chinese and 57 Japanese adult ESL learners. Each subject was required to write one essay in English and one in his or her L1. Although a significant positive correlation emerged between the scores on the L1 and L2 essays for the Japanese subjects, the correlation for the Chinese participants was low and insignificant. Carson et al. (1990) conclude that although the transfer of writing ability appears to manifest itself, the degree to which it occurs varies by language background.

Paralleling the writing product studies (Canale et al., 1988; Carson et al., 1990), the writing process studies (Arndt, 1982; Edelsky, 1982; Jones & Tetroe, 1987; Lay, 1982) also point to the transfer of literacy skills between languages. In the process realm, Edelsky

(1982) analyzed 477 Spanish and 49 English writing samples of nine native Spanish-speaking first, second, and third graders enrolled in an ESL program. After analyzing the similarities and differences between the L1 and L2 writing samples, Edelsky asserts that the data "support the perspective that what a young writer knows about writing in the first language forms the basis of new hypotheses rather than interferes with writing in another language" (1982, p. 227). Edelsky (1982) concludes that certain L1 writing processes transfer to L2 writing.

Lay (1982) studied the composing processes of four university level Chinese ESL learners in order to determine whether or not the L1 is used in the L2 composing process. Each subject wrote one essay in Chinese and two in English while thinking aloud. According to Lay (1982), the subjects used many of the strategies that native English speakers use, such as rereading topics, reevaluating organization, and modifying vocabulary. They also translated key words and expressions into Chinese in order to obtain "a stronger impression and association of ideas" (Lay, 1982, p. 406). The more frequently the subjects used their native language, the better was the quality of their L2 essays.

Like Edelsky (1982), Lay asserts that elements of the L1 writing process transfer to L2 writing.

Jones and Tetroe (1987) used think-aloud protocols to examine the planning patterns of six adult Spanish-speaking ESL learners. Each subject produced four English and two Spanish writing samples, using the think-aloud procedure. The subjects consistently applied L1 knowledge in planning, even when writing in the L2. Jones and Tetroe (1987) interpret this finding as "strong, direct data for the transfer of first-language skill to the second language" (p. 55). Like Edelsky (1982) and Lay (1982), Jones and Tetroe (1987) conclude that L1 writing processes transfer to the L2.

Akin to Lay (1982) and Jones and Tetroe (1987), Arndt (1987) required her subjects to think aloud as they wrote. Six Chinese graduate students enrolled in an EFL program wrote one essay in Chinese and one in English on topics designed to elicit similar discourse patterns. The subjects wrote while thinking aloud and were later asked to recount their composing strategies. Arndt (1987) states, "The composing strategies of each individual writer were found to remain consistent across languages" (p. 257). According to Arndt (1987), this finding points to writing skills transfer.

In summary, numerous investigations of both writing products and processes uphold Cummins's (1981) Interdependence Hypothesis and support the notion of a common underlying language proficiency. In the product realm, Canale et al. (1988) argue for automatic transfer while Carson et al. (1990) assert that although transfer occurs, the extent to which it does depends on language background. In the writing process arena, Edelsky (1982), Lay (1982), Jones and Tetroe (1987), and Arndt (1987) contend that L1 writing practices carry over to the L2. The present study took this vein of inquiry one step further by measuring the unique contribution of L1 writing ability to L2 reading comprehension test scores, and by demonstrating whether or not the contribution changes by course level.

The Reading-Writing Relationship

Research into the reading-writing relationship is germane to this review because one of the main goals of the present study was to gauge the contribution of L1 writing skill to the scores on L2 reading comprehension measures. The studies reviewed overwhelmingly indicate that writing ability transfers to reading skill.

L1 Studies

For decades, theoretical and empirical evidence in the L1 domain has demonstrated that writing skill and reading ability transfer to one another because they share many of the same cognitive processes (Tierney & Shanahan, 1991). Both product- and process-oriented studies generally point to the existence of a relationship between reading ability and writing skill, though the nature of the relationship may be inconsistent.

In the product-based domain, Loban (1964) compared the scores of 220 subjects across 12 grade levels for reading skill and writing capacity. The Stanford Achievement Test was used to determine reading level, and holistic scores were used to indicate writing ability. According to Loban (1964), the correlations were "so striking to be beyond question" (p. 212). Tierney and Shanahan (1991) point out, however, that Loban's overall results are tenuous because only one reading test and one writing measure were used. Additionally, although the correlations were high for the subjects who scored very well or very poorly, variations emerged in the results of the average subjects. Furthermore, differences in subject performance could have been attributable to individual instructional histories, rather than to any

relationship between reading and writing abilities (Tierney & Shanahan, 1991).

According to Tierney and Shanahan (1991), recent research casts doubt on Loban's "straightforward interpretation" (p. 248). In Martin's (1977) case study of six Australian children, only two scored similarly on both the reading and writing measures. Martin (1977) concludes that "the evidence suggests that reading and writing are intertwined, but in ways that are not easily predictable" (p. 52) (in Tierney & Shanahan, 1991).

In the 1980s, researchers began to explore the particular kinds of knowledge that reading and writing share. In a longitudinal study across grades one and two, Juel, Griffith, and Gough (1986) compared their subjects' test scores on word recognition, spelling, reading comprehension, and writing quality. Strong correlations emerged between some of the variables, but not between others. Shanahan (1984) and Shanahan and Lomax (1986, 1988) examined the test scores of second and fifth graders on reading skills such as comprehension, word analysis, and vocabulary knowledge, and on writing skills such as spelling, vocabulary diversity, syntax, and story structure. Complementing the results of Juel et al. (1986), some of the reading abilities were shown to correlate with some of the writing skills, but not

with others. According to Tierney and Shanahan (1991), the results of these studies indicate that although it is likely that knowledge is shared in reading and writing, the sharing process is "neither as simple nor as complete as was once assumed" (p. 249).

Like the more recent writing product studies, the writing process studies have attempted to shed light on the precise nature of the reading-writing relationship. In these studies, information about cognitive processes is normally collected using techniques such as the think-aloud procedure, interviews, and observations. Wittrock (1984) and Squire (1984) propose that reading and writing are both manifestations of the same cognitive processes. Tierney and Pearson (1983), in their composing model of reading, build upon Wittrock's (1984) and Squire's (1984) theories by suggesting that both reading and writing are composing processes that share the same fundamental properties: planning, drafting, aligning, revising, and monitoring. (See Chapter I, Theoretical Considerations.)

Kucer (1985), in his theory of text world production, asserts that "cognitive efficiency would demand a sharing of strategies in the two processes" (p. 319). According to Kucer (1985), reading and writing processes, therefore, probably run in parallel and most

likely use many of the same basic mechanisms. Thus, "reading and writing would each become one instance of text world production, drawing from a common pool of cognitive and linguistic operations" (Kucer, 1985, p. 319).

According to Tierney and Shanahan (1991), recent research expands upon the notion that reading and writing share the same cognitive processes such as gathering ideas, questioning, and hypothesizing (p. 215). Reading and writing appear to differ, however, by the extent to which the various cognitive processes and strategies are used, and by the specific aspects of reading and writing that elicit their use. Ryan (1985), for example, asked eight above-average fifth graders to read and write both narrative and expository texts. Although she identified six thinking strategies common to reading and writing, she discovered that the patterns of use of these strategies differ by text genre (in Tierney & Shanahan, 1991).

Kirby (1986) and Martin (1987) discovered similar phenomena. Kirby (1986) observed five high-risk college freshmen as they read and wrote in a similar text genre, and later conducted retrospective interviews with them. She found that the subjects tended to use similar strategies in both reading and writing, and that

limitations in strategy use were generally parallel in both the reading and writing arenas. The subjects, for example, who did little planning in writing did little previewing before reading (in Tierney & Shanahan, 1991).

After observing the reading and writing behaviors of seven 12th graders and later examining their think-aloud protocols, Martin (1987) identified eight categories of meaning-making strategies. Although the same strategies were used during both reading and writing, the extent to which each was enlisted differed for reading versus writing. Questioning and rereading strategies, for example, were used more frequently in writing than in reading (in Tierney & Shanahan, 1991).

Langer (1986) examined the reading and writing behaviors of 13 third graders, 36 sixth graders, and 15 ninth graders as they read and wrote stories and reports. Results indicate that although the same cognitive processes appear to be used in reading and writing, differences are also evident. In both reading and writing, the subjects focused on meaning construction and used strategies of questioning, hypothesizing, and validating. With respect to differences, the subjects were more concerned with citing evidence in support of their interpretations (validating) while reading, and tended to do more hypothesizing while writing. Further

differences emerged according to text genre and subject age. Langer (1986) asserts that "the different purposes underlying reading and writing invoke different behaviors and approaches, even when the tasks and topics are parallel" (p. 96).

Tierney and Shanahan (1991) note that discrepancies in the writing process studies include variable selection and description, assessment measures of untested reliability, conclusions based solely on introspective, retrospective, or think-aloud data, small sample sizes, and limited information about the subjects' instructional histories. These studies, however, have generated a great deal of information about reading, writing, and the relationship between the two in less than a decade (Tierney & Shanahan, 1991).

In summary, L1 research into the reading-writing relationship reveals that reading and writing capacities are interrelated, although the precise nature of the relationship remains unspecified. The present study has expanded upon this research domain by revealing the unique contribution of L1 writing ability to the scores on L2 reading assessment measures, and by examining whether or not this contribution differs by course level.

L2 Studies

Although for decades research has pointed to the existence of a relationship between L1 reading skill and L1 writing ability, little has been done to explore this connection for L2 learners. Only three studies, Janopolous (1986), Carson et al. (1990), and Carrell and Connor (1991) have been published in which the primary goal was to examine the relationship between L2 writing ability and L2 reading skill. Additionally, Scholz, Hendricks, Spurling, Johnson, and Vandenburg (1983) produced a correlational matrix depicting the relationship between the scores on three standardized tests of English reading proficiency and three of writing ability.

Janopolous (1986) correlated the holistic scores on English essays written by 79 graduate ESL students of varying L1 backgrounds, with the amount of time they reported having spent reading for pleasure in their L1s and in English. The results indicate that the subjects who had spent more time reading L2 texts for pleasure were more likely to be proficient in L2 writing. According to Janopolous (1986), the data provide evidence of a connection between L2 reading skill and L2 writing proficiency.

Carson et al. (1990) correlated the English reading and writing test scores of 47 Chinese and 58 Japanese adult ESL learners. The data reveal a higher correlation for the Chinese subjects than they do for the Japanese participants. Carson et al. (1990) assert that this study, like the Janopolous (1986) study, points to a link between L2 reading and writing, though the relationship is stronger for some native language groups than it is for others.

Carrell and Connor (1991) correlated the English reading and writing scores of 23 undergraduate and 10 graduate ESL students of mixed language backgrounds. Subjects were required to read two passages--one persuasive and one descriptive--and then to write a recall protocol and answer eight multiple choice questions for each. Additionally, the subjects wrote one persuasive and one descriptive essay, which were scored both holistically and qualitatively. Of the eight possible correlations, three were significant. Those correlations ensued between the recall protocol scores and the holistic writing scores for both the persuasive and descriptive genres, and between the multiple choice test scores and the holistic writing scores for the persuasive genre. According to Carrell and Connor

(1991), the nature of the relationship between L2 reading skill and L2 writing ability varies by text genre.

Scholz et al. (1983) correlated the scores of 182 adult ESL learners of varying L1 backgrounds on three standardized English reading proficiency tests and three English writing tests. The three reading examinations comprised multiple choice and cloze tests while the three writing measures consisted of a multiple choice test and an essay, which was scored both holistically and analytically. The data reflecting all combinations of reading and writing tests indicate positive correlations.

Although these studies reveal a link between reading skill and writing ability, additional research such as the present study is needed. The self-reported data that Janopolous (1986) used to determine the amount of time spent reading for pleasure were based on the subjects' memories and may not have been completely accurate. Additionally, because some subjects might read faster than others, the time spent reading might not serve as an accurate indication of the amount read. Furthermore, the subjects were required to write only one essay, severely limiting the potential pool of student writing. The results, therefore, could be based on less than completely fruitful data.

The potential discrepancies apparent in the Carson et al. (1990) study relate to the rubric used to score the Chinese and Japanese writing samples, and to the method of assessing reading comprehension. Although the Test of Written English Scoring Guide was used to rate the English essays, the Chinese and Japanese raters developed their own scoring rubric, which may not have been equivalent. This inconsistency could render any comparison of the L1 and L2 writing scores less than entirely meaningful. Cloze tests were used to measure reading comprehension in all three languages. According to Bernhardt (1991), this testing method does not adequately measure reading comprehension because it "clearly has little if anything to do with a reader's understanding of connected discourse" (p. 198).

Several potential complications emanate from Carrell and Carson's (1991) study. First, the subjects were required to write their recall protocols in the L2, resulting in scores that represent a combination of writing ability and reading comprehension. Second, the use of multiple choice questions to measure reading comprehension opens the way to the validity issues associated with that testing technique. Third, the subjects' backgrounds included diverse languages and orthographies, thus limiting generalizability.

The results of Scholz et al.'s (1983) correlational matrix appear to serve as a straightforward indicator of the existence of a L2 reading-L2 writing relationship. The reading assessment measures, however, consisted of multiple choice and cloze, which as previously mentioned, could have resulted in less than fully accurate measures of reading comprehension. Additionally, the use of ESL subjects of mixed L1 backgrounds by Scholz et al. (1983) and by Janopolous (1986) may have resulted in reduced generalizability because of the combinations of native languages, orthographies, and cultures.

In summary, the results of the L2 studies reviewed clearly point to the existence of a relationship between reading skill and writing ability. The potential validity problems resulting from the testing techniques used in Scholz et al. (1983), however, coupled with the inconsistencies in the Janopolous (1986) study and in Carson et al. (1990) and Carrell and Connor (1991), point to the need for further exploration. The present study has provided insight into the reading-writing relationship by examining the contribution of writing skill to the scores on reading comprehension measures, and by investigating whether or not this contribution changes by course level.

The Relationship of L2 Vocabulary Knowledge and L2 Grammatical Skill to L2 Reading Comprehension

The findings in this vein of research underpin the present study, which has measured the unique contributions of L2 vocabulary knowledge and L2 grammatical skill to the scores on L2 reading comprehension tests. In the L2 realm, inquiry into the contribution of the various literacy skills to reading comprehension is still in its infancy.

Henning (1975) and Koda (1990) conducted the only other multiple regression studies to date that gauge the contributions of L2 vocabulary knowledge and L2 grammatical skill to L2 reading comprehension. In a study of 27 female Iranian college students enrolled in second-semester English, Henning (1975) administered cloze and multiple choice tests to asses vocabulary knowledge, and cloze, multiple choice, and short answer tests to measure grammatical skill. The sum of all of the test scores was used as the measure of reading comprehension. According to Henning (1975), the results indicate that vocabulary knowledge is the predominant predictor of intermediate-level text comprehension. For low difficulty material, however, it serves as an "invalid predictor of reading comprehension" (Henning, 1975, p. 113). Grammatical skill contributes negligibly

(only four percent of the variance) to L2 reading comprehension scores. Thus, Henning (1975) maintains that "student grammar performance should not be taken into account in the scoring of reading comprehension examination [sic]" (p. 113).

Koda (1990) administered a battery of language proficiency and reading comprehension tests to 39 college-level Japanese as a foreign language learners of varying L1 backgrounds. The language proficiency tests measured grammatical skills and vocabulary knowledge through a translation exercise, a multiple choice test, and a sentence completion task. Cloze and multiple choice tests were used to assess reading comprehension. The data indicate that while vocabulary skills contributed significantly to the scores on the multiple choice test, grammatical skills contributed highly to the scores on the cloze test.

Additional research using multiple regression, such as the present study, is necessary in order to achieve a clearer picture of the contributions of vocabulary knowledge and grammatical skill to the scores on L2 reading assessment measures. Because Henning's (1975) sample consisted solely of females, male subjects were tested in order to enhance the generalizability of results. Additionally, neither Henning (1975) nor Koda

(1990) used sample sizes considered large enough for the effective use of multiple regression analysis. Kamil, Langer, and Shanahan (1985) state that a minimum of 20 subjects is needed per independent variable in order for the results of such analyses to be meaningful. According to Kamil et al. (1985), "Results from studies with low subjects-to-measures ratios are suspect, as they have a tendency to be highly unreliable" (p. 149). The present study included 131 subjects, well exceeding the minimum number necessary. Furthermore, Henning (1975) and Koda (1990) used cloze and multiple choice tests to assess reading comprehension, opening the way to the validity problems related to those measures. In the present study, reading comprehension was assessed using the recall protocol, a more sensitive and sophisticated measure than multiple choice or cloze.

Pike (1979) and Barnett (1986) also explored the relationship between selected L2 literacy variables and L2 reading comprehension, though they did not measure the unique contributions. Pike's (1979) correlational matrix of the scores on the Test of English as a Foreign Language subtests evince high correlations between the vocabulary knowledge and reading comprehension portions, and between the English Structure and reading comprehension subtests.

Barnett (1986) compared her subjects' scores for grammatical skill and vocabulary ability to their scores for reading comprehension. One hundred twenty-four native English-speaking subjects enrolled in a college French course were administered two multiple choice, rational-deletion cloze tests intended to measure grammatical skill and vocabulary knowledge. Recall protocols on a single text were used to measure reading comprehension. According to Barnett (1986), the data indicate that reading comprehension increases along with vocabulary knowledge and grammatical skill. These results, like those of Pike (1979), point to the hypothesis that vocabulary knowledge and grammatical skill would contribute to reading comprehension scores.

Koda (1989) sought to determine the relationship between vocabulary knowledge and L2 reading comprehension in a study of 24 college students of mixed L1 backgrounds enrolled in a first-year Japanese course. Vocabulary knowledge was measured using a translation exercise, sentence completion tasks, and multiple choice word grouping exercises. Reading comprehension was assessed using a cloze test and short-answer questions on five short paragraphs. Like Pike (1979) and Barnett (1986), Koda (1989) discovered a relationship between vocabulary knowledge and reading comprehension.

In a study comparing the silent versus oral reading comprehension scores of 14 subjects enrolled in college-level German courses, Bernhardt (1983) grouped her subjects according to their scores on a standardized test of grammatical skill. The subjects at the higher level of grammatical skill were also the most proficient readers. Thus, the data lend credence to the expectation that grammatical skill would contribute to the scores on L2 reading comprehension measures.

The results of Pike (1979), Barnett (1986), Koda (1989), and Bernhardt (1983) point to a connection between grammatical skill and vocabulary knowledge, and reading comprehension. Further inquiry such as the present study, however, is still needed to explore this link. Koda's (1989) use of cloze and multiple choice tests to assess reading comprehension could have resulted in scores affected by the validity problems associated with those measurement techniques. Although Barnett (1986) avoided such pitfalls by using the recall protocol, she administered only one reading passage per subject without first testing for background knowledge. Thus, the subjects possessing more background knowledge may have garnered higher reading comprehension scores without necessarily understanding more of the passage content. In addition, neither validity nor reliability

data exist for the rational-deletion cloze test that she used to measure vocabulary knowledge and grammatical skill. Furthermore, because the subjects in Pike (1979) and Koda (1989) were of mixed language backgrounds, their linguistic, orthographic, and cultural differences could have led to less than completely valid results. Bernhardt's (1983) small sample size limits the generalizability of her data.

In summary, the results of the studies reviewed in this section indicate that vocabulary knowledge and grammatical skill are related to reading skill. The present study has expanded upon this line of inquiry by exploring the contribution of the L2 variables to the scores on L2 reading comprehension assessment measures, and by investigating whether or not the contributions differ by course level. Furthermore, the present study accomplished the following: (a) measured reading comprehension using a more sophisticated testing method than multiple choice or cloze, (b) employed multiple reading passages in order to curb the effects of background knowledge, and (c) featured a sufficient number of subjects to ensure meaningful multiple regression data.

Chapter Summary

The studies reviewed in this chapter comprise an important part of the foundation underlying the hypothesis that L1 reading comprehension, L1 writing ability, L2 vocabulary knowledge, and L2 grammatical skill would contribute to the scores on L2 reading assessment measures. Previous research, however, has not clearly determined the unique contribution of each variable to L2 reading test scores, or whether or not the contributions differ by course level.

The present study has explored this issue using more sophisticated testing techniques. The recall protocol, rather than multiple choice, cloze, or short-answer questions, was used to measure reading comprehension. Multiple reading and writing samples were collected in order to guard against the effects of background knowledge, and to allow for a wider sampling of subject performance. Authentic texts, rather than translated or simplified versions, were used in order to achieve a more realistic view of what had been comprehended.

The subjects shared the same L1 background, thus eliminating the problems of generalizability that a melange of cultures and orthographies can cause. Additionally, because the subjects were adults whose cognitive development had stabilized and who were already

fully literate, the results provide a clearer picture of the nature of the relationships between literacy abilities. Furthermore, the number of subjects in the present study was sufficient to render meaningful results from multiple regression analysis.

CHAPTER III

PROCEDURES

Population and Sample

The population consisted of all cadets enrolled in French courses at the United States Air Force Academy (USAFA), Colorado Springs, Colorado. Cadets completing the Academy's four-year program earn bachelor of science degrees and commissions as military officers. Applicants are selected based on their high school grade point averages, their Scholastic Aptitude Test or American College Test scores, and their leadership potential as demonstrated by participation in extracurricular activities (Westfall, 1988).

While at the Academy, nearly all cadets must complete at least one year of foreign language study. Cadets are placed into course levels according to their scores on the Placement Validation Test (PLAVAL) designed and administered by the Department of Foreign Languages, USAFA. Approximately 15 per cent of all cadets earn PLAVAL scores that are high enough for waiver of the foreign language requirement (Westfall, 1988).

The foreign language minor, instituted in 1986, allows cadets in any of the Academy's 25 academic majors to declare as a minor area one of the seven foreign languages offered. Approximately 12 cadets per year earn a minor in French.

Although USAFA's rigorous selection process somewhat decreases this population's heterogeneity, the student body represents all 50 states and several allied nations, and the cadets' backgrounds reflect a wide variety of socio-economic levels. Furthermore, the Academy's minority recruiting program ensures the representation of cadets from diverse cultures and environments (Westfall, 1988).

The sample consisted of all cadets enrolled in French 141, 222, 321, and 365. According to the USAFA Curriculum Handbook (1990-1991), French 141 consists of "Accelerated basic foreign language study with emphasis on comprehension, communicative skills, and structure" (p. 239). This language course is designed for cadets who have already had one or two years of language study, and comprised six class sections of approximately 15 subjects each for a total of 88.

French 222, the follow-on course to French 141, features "Continued refinement of grammatical and structural knowledge" and emphasizes "conversational

practice and aural comprehension" (USAFA Curriculum Handbook, 1990-1991, p. 240). This course consisted of one section of 21 subjects.

French 321 and 365 are advanced-level courses. The former, a "capstone skills development course" focusing on "Intensive practice in listening, speaking, and reading" (USAFA Curriculum Handbook, 1990-1991, p. 241), included a single section of 15 subjects. The latter, which emphasizes oral discussion of contemporary issues, comprised one section of seven subjects.

Research Design

In the present study, multiple regression was used to ascertain the amount of variance uniquely contributed by each independent variable to the variance in the independent variable (Kerlinger & Pedhazur, 1973). In other words, regression "explains or predicts the amount of variance in a dependent variable on the basis of its correlation with two or more independent variables" (Kamil, Langer, & Shanahan, 1985, p. 132). Because the sequence of data entry can "influence the relative importance of the variables" (Kamil et al., 1985, p. 140), a theory-based hierarchical regression model was selected.

Independent Variables

English Reading Comprehension

Two separate testing instruments were used to assess native language (L1) reading comprehension: (a) the comprehension portion of the Nelson-Denny Reading Test (NDRT), Fcrm E (Brown, Bennett, & Hanna, 1981), and (b) the mean percentage score on three immediate timed written recall protocols.

NDRT comprehension section.

The NDRT is a multiple choice reading assessment measure that generates four scores: (a) vocabulary knowledge, (b) comprehension, (c) reading rate, and (d) overall. The comprehension portion consists of eight reading passages, each followed by a series of multiple choice questions. The test was administered to the subjects when they entered the Air Force Academy. Test scores were obtained through the Department of English, USAFA.

Recall protocols.

When the recall protocols were administered, each subject received three passages in English, between 200 and 250 words in length. Subjects had a total of five minutes per passage to read each and then to write everything they could remember about each. The recall

protocols were scored by two trained raters, using Johnson's propositional analysis system (1970).

L1 Writing Ability

This variable was measured using the combined holistic scores on the same three recall protocols that were employed to assess L1 reading comprehension. (See Recall Protocols.) In order to arrive at the scores for writing ability, four trained, experienced raters used the Test of Written English Scoring Guide (Educational Testing Service, 1990).

French Grammatical Skill

Scores for French grammatical skill were obtained from the grammar portion of USAFA's French PLAVAL. This section of the placement examination features multiple choice and cloze test items.

Passage-Specific French Vocabulary Knowledge

A supply-definition test was used to ascertain the proportion of vocabulary words, appearing in the three French reading passages, that the subjects were able to define. For each French word, the subjects were required to write, in English, either the definition or the English equivalent.

Dependent Variable

Second Language (L2) Reading Comprehension

This variable was measured using the mean percentage score on three immediate free written recall protocols. Each subject was given three French passages between 200 and 250 words in length. The subjects were instructed to read each passage as many times as they needed to and then to write, in English, everything they could remember about each.

Instrumentation: Independent Variables

L1 Reading Comprehension

NDRT comprehension scores.

L1 reading comprehension was measured using scores from the comprehension portion of the NDRT, Form E (Brown, et al., 1981). The comprehension section comprises eight reading passages followed by a total of 36 multiple choice items. Scores, which were acquired through the Department of English, USAFA, were unavailable for subjects in their senior year.

Recall protocol scores.

Immediate timed written recall protocols were employed to measure L1 reading comprehension. The recall protocol was selected for use as a reading comprehension assessment measure because, unlike multiple choice or short answer tests, it offers the subjects no information

pertaining to passage content. Bernhardt (1983) recommends this instrument because it requires the reader to integrate not only the components of the reading process, but the information contained in the passage. Three reading passages were used in order to limit the effects of background knowledge. The passages, selected from the Encyclopedia Britannica (1984), were each between 200 and 250 words in length, and historical in nature. The following French historical topics were included: (a) Louis IX, (b) Napoleon, and (c) De Gaulle. (See Appendix A for the texts and Appendix B for the scoring templates.)

A five-minute time limit was imposed for each text in order to decrease the possibility of the subjects memorizing the passages. Five trained native English speakers divided the texts into pausal units, and then rated each pausal unit on a scale of one to four for its importance to the overall meaning of the text, according to Johnson's propositional analysis system (1970). Two trained native English-speaking raters scored the protocols. Interrater reliability, measured using the Pearson product-moment correlation procedure, was .94.

Holistic Scores for L1 Writing Ability

Four trained, experienced raters used the Test of Written English (TWE) Scoring Guide (Educational Testing Service, 1990) to assess writing ability on a scale of one (demonstrates incompetence in writing) to six (demonstrates clear competence in writing on both the rhetorical and syntactic levels). The TWE, which is the writing component of the Test of English as a Foreign Language, is "an indirect test of knowledge of important structural and grammatical points of standard written English" (Educational Testing Service, 1989, p. 1).

The writing samples were scored during one major scoring session for which each rater was paid 10 dollars per hour. The writing samples based on the Louis IX text were scored first, followed by those based on Napoleon and then De Gaulle. Calibration sessions, each lasting approximately one hour, were held before scoring the protocols for each text. These sessions ensured that all scorers understood and agreed upon the criteria stated in the TWE Scoring Guide. Overall interrater reliability, as measured by the Spearman-Brown Prophecy Formula (Henning, 1987), was .95. The overall percentage of agreement was 64%.

French Grammatical Ability

For beginning-level subjects, scores on the grammar portion of the French PLAVAL were acquired from the Department of Foreign Languages, USAFA. Because the Academy does not maintain PLAVAL scores for more than one year, the grammar portion was readministered to the subjects enrolled in French 222, 321, and 365 during the third week of the fall semester, 1991.

Because the grammar section also contained several items intended to assess vocabulary skill, two experienced instructors from USAFA's French Division identified and agreed upon the questions that were designed to test grammatical ability and those meant to assess vocabulary knowledge. Although all of the subjects were tested on all of the items, only the items that were intended to measure grammatical skill were counted toward the scores for grammatical ability.

Passage-Specific Supply-Definition Vocabulary Test

Two weeks before the recall tasks were administered, the subjects completed a vocabulary test in which they wrote either the definition (in English) or an English equivalent for each of the 50 French words presented. (See Appendix C for the vocabulary test and scoring guide.) Two French language experts scored each answer on a two-point scale. The two-point scale consisted of

the following: (a) zero points for an incorrect response, (b) one point for a partially correct response, and (c) two points for a correct response. Interrater reliability, as measured by the Pearson product-moment correlation coefficient, was .99.

Instrumentation: Dependent Variable

Recall Protocols Based on L2 Reading Passages and Written in the L1

Immediate free written recall protocols were employed to measure L2 reading comprehension. Three reading texts were used in order to constrain the effects of background knowledge. The three French passages, each between 200 and 250 words in length and historical in nature, were selected from the Encyclopedia Universalis (1968), the French counterpart to the Encyclopedia Britannica. The topics included the following elements of French history: (a) the Gauls, (b) Louis XIV, and (c) the colonial empire. (See Appendix D for the texts, Appendix E for the English translations, and Appendix F for the scoring templates.) Texts of the same genre and on the same general topic as those used in the L1 recall procedure were selected in order to maintain consistency in discourse type. Five trained French language experts divided the texts into pausal units and then rated them for their importance to the overall meaning of the text

on a scale of one to four, according to Johnson's propositional analysis system (1970). Two trained French language experts scored the protocols. Interrater reliability, measured using the Pearson product-moment correlation procedure, was .94.

Procedures and Data Collection

The experiment took place at USAFA during the fall semester of 1991.

Passage-Specific French Vocabulary Knowledge Test Before the Experiment

In July 1991, two experienced French language instructors from USAFA selected the words from the French passages that the beginning-level subjects would probably be unable to define. (The beginning level served as the base line.) The instructors then selected from the list of words, the 50 words that the beginning-level subjects would find the most difficult to define. These words appeared in a randomly ordered list on the Passage-Specific French Vocabulary Knowledge Test. Three weeks before L2 reading comprehension data collection took place, a sufficient number of copies of the vocabulary test were given to the Chief of the French Division at USAFA. The vocabulary test was administered by the course instructors two weeks before the L2 reading comprehension data were collected.

During the Experiment

The instructors of French 141, 222, 321, and 365 distributed one copy of the vocabulary test to each subject and assigned the test to be completed outside of class. The subjects were told that their test performance would not be reflected in their course grades and that they would be under the provisions of the Cadet Honor Code not to use dictionaries or to seek help in defining the words. They were allowed a maximum of 50 minutes (one minute per word) to complete the test. The course instructors collected the tests during the following class period.

French Grammatical Knowledge

PLAVAL scores were on file only for the beginning-level subjects, to whom the examination was administered in July 1991 as part of USAFA's admissions procedure. Because the Department of Foreign Languages does not retain the PLAVAL scores for cadets from previous years, the researcher re-administered the grammar portion to the intermediate- and advanced-level subjects during the first week of September 1991. Identical time limits were imposed for all subjects.

NDRT Comprehension Portion

USAFA personnel administer the NDRT, Form E (Brown et al., 1981), to all cadets entering the Academy as part of the admissions procedure. Scores, which were obtained through the Department of English, USAFA, were unavailable for subjects in their senior year.

Recall Protocols

The recall protocols were collected over a four-day period during the week of September 22, 1991. Because the USAFA-wide class schedule is composed of two-day cycles, some of the French classes meet the first day of the cycle while the others meet the second. The subjects completed the recall protocols based on L1 passages during the first two-day cycle and those based on L2 texts during the second cycle.

Recall Protocols Based on L1 Reading Passages and Written in the L1

Before the experiment.

During the second week of September 1991, the researcher used the class rosters provided by the French instructors to determine the number of cadets at each language level, and then assembled the appropriate number of research packets. The texts were arranged in differing orders (1, 2, 3; 2, 3, 1; 3, 1, 2; 1, 3, 2; 2, 1, 3; 3, 2, 1) in order to limit the effects of fatigue.

Subjects were selected through the use of a table of random numbers (Snedecor & Cochran, 1967) to receive the various arrangements. The packets were coded to reflect passage arrangement, course level, course section, and recipient, and consisted of seven pages: (a) a coded page, (b) the first passage, (c) a blank sheet for the recall protocol, (d) the second passage, (e) a blank sheet for the recall protocol, e) the third passage, and (f) a blank sheet for recall protocol.

During the experiment.

The researcher advised the subjects not to open the research packets until instructed to do so and distributed the packages. The subjects were told that the research packets contained three exercises and that they would have five minutes to complete each one. Additionally, the subjects were advised that upon opening the packet to the first passage, they would have a total of five minutes to read it and, without referring back to it, to write, in English, everything they could remember about it. A two-minute warning was announced after the subjects had spent three minutes on each text. The researcher used a stopwatch to time the exercises. After the first exercise was completed, the same procedure was used for the second and third.

Recall Protocols Based on L2 Reading Passages and Written in the L1

Before the experiment.

During the second week of September 1991, the researcher prepared the French reading recall packets in the same manner in which the English recall packets were prepared. Subjects were selected through the use of a table of random numbers (Snedecor & Cochran, 1967) to receive the various packet arrangements.

During the experiment.

The researcher followed the same procedures that were used in the English reading recall portion of the experiment with two exceptions. First, the subjects had the entire 50-minute class period to read and recall the French texts. Second, because the subjects were working at their own pace, they were reminded to complete the readings in the order given and that once they completed one recall exercise and began to work on another, they could not go back to previous protocols. After each subject had completed the three recall protocols, his or her packet was collected.

Pilot Studies

Pilot Study

The first pilot study was conducted at USAFA in September 1990. One hundred twenty-six cadets enrolled in basic, intermediate, and advanced French courses read three French texts and wrote protocols for each in English, which resulted in 119 usable research packets. This test of experimental procedures allowed for the clarification of directions, provided the researcher with information regarding the amount of time needed for test administration, and showed that, in order to ensure standardization, the same person should administer the tests to all classes. The 119 research packets included 238 reading recall protocols that were used for practice in holistic scoring, and in scoring for reading comprehension according to Johnson's propositional analysis system (1970).

Of the 119 research packets collected, 10 from each course level were selected for scoring, for a total of 30 research packets. Each subject's holistic score for writing ability represents the combined score of two protocols. Each subject's score for L2 reading comprehension (Johnson, 1970) represents the average reading comprehension score on the same two protocols. The data, analyzed using the Pearson product-moment

correlation coefficient, reveal an overall correlation of .75 between the subjects' holistic scores for writing ability and their L2 reading comprehension scores. By level, the correlations were .49 for beginners, .70 for intermediate subjects, and .68 for advanced subjects.

Pilot Series

A series of three pilot studies was carried out in July and August 1991 in order to test the passage-specific vocabulary test, the L1 recall protocol procedure, and the L2 reading assessment measure.

In July 1991, the passage-specific vocabulary measure was administered to 12 students enrolled in French 102 at the Ohio State University. The 102 level was selected because it comprised students at a level of French similar to that of the USAFA cadets at the beginning level. Because beginning learners are less experienced in the language, it was expected that administering the test to these students would uncover more problem areas. This pilot study pointed to the need for an improved test format and provided the researcher with information relating to the amount of time needed for test administration. The completed tests were used for practice scoring.

The recall protocols used to measure L1 reading comprehension were administered to 19 French 102 students in August 1991. This process reinforced the importance of careful adherence to the five-minute time limit and revealed an area in which the directions needed clarification.

The L2 reading comprehension measure, also pilot tested in August 1991, was administered to 16 students enrolled in French 102. The process provided the researcher with valuable information regarding the time needed for test administration. The protocols were used for practice in scoring for reading comprehension according to Johnson's (1970) propositional analysis system.

Data Analysis

When using multiple regression, sample sizes need to be sufficiently large in order to keep standard error small and thus to increase the dependability of the results (Kerlinger & Pedhazur, 1973). According to Kamil et al. (1985) a minimum of 20 subjects is needed per independent variable. A total of 131 subjects participated in the present study, approximately 32 per independent variable. The Statistical Analysis System was used to perform the multiple regression analyses.

In all of the hierarchical regression programs used in the present study, the scores for the independent variables were entered into the model in the following order: (a) L1 reading comprehension, (b) L1 writing ability, (c) L2 vocabulary knowledge, and (d) L2 grammatical skill. The L1 variables were inserted first because all of the subjects were native English speakers who had amassed more experience with the L1 than they had with the L2. L1 reading scores were introduced before the L1 writing data because, according to Shanahan and Lomax (1986), "more reading information is used in writing than vice versa" (p. 122). L2 vocabulary scores were entered before those for L2 grammar because L2 vocabulary knowledge accounted for more of the variance in L2 reading comprehension scores than did L2 grammatical skill in the stepwise multiple regression analyses accomplished by Henning (1975) and Koda (1990).

First Research Question

The scores for all of the subjects were subjected to programs of hierarchical multiple regression in order to examine the unique contribution of the variance of the scores for each independent variable (L1 reading comprehension, L1 writing ability, L2 vocabulary knowledge, and L2 grammatical ability) to the variance in the scores for the dependent variable (L2 reading

comprehension). Because two separate instruments were used to test L1 reading comprehension, two individual programs were run, first including the NDRT comprehension scores, and then featuring the recall protocol scores.

Second Research Question

The data were placed into levels (Beginners and Non-Beginners) and subjected to a series of multiple regression programs in order to investigate the unique contribution of the variance in the scores for each independent variable to the variance in the scores for the dependent variable. Two separate analyses were accomplished per course level, first including the NDRT comprehension scores and then containing the recall protocol scores. The results were examined for differences, by course level, in the unique contribution of the variance in the scores for each of the independent variables to the variance in the scores for the dependent variable.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

In recent years, enhanced test validity has emerged as a central concern among foreign language (L2) researchers and educators. Increased validity of L2 reading measures will lead researchers to an expanded understanding of the reading comprehension process and will aid instructors in diagnosing and solving reading problems, and in placing students into course levels. In order to improve test validity, the contributions of the factors constituting L2 reading test scores must be established. The purpose of the present study was to measure the unique contributions of selected literacy variables--native language (L1) reading comprehension, L1 writing ability, L2 vocabulary knowledge, and L2 grammatical skill--to the scores on L2 reading comprehension measures. Additionally, this study has ascertained whether or not the contributions of the independent variables to the dependent variable differ by level of L2 study.

The four independent variables and the dependent variable were measured using a number of assessment methods. Scores from two separate testing instruments were used to measure the first independent variable, L1 reading comprehension: (a) the comprehension portion of the Nelson-Denny Reading Test (NDRT), Form E (Brown, Bennett, & Hanna, 1981), and (b) the mean percentage score of three immediate timed written recall protocols. The second independent variable, L1 writing ability, was assessed using the combined holistic score generated from the same three recall protocols used to measure L1 reading comprehension. A supply-definition test consisting of vocabulary words specific to the L2 reading passages was employed to measure the third independent variable, L2 vocabulary knowledge. L2 grammatical ability, the fourth independent variable, was measured using the grammar section of the Air Force Academy's French placement test. The dependent variable, L2 reading comprehension, was measured using the mean percentage score of three immediate free written recall protocols. The range of potential scores for the comprehension portion of the NDRT extends from 0 to 36, for L1 writing ability from 3 to 18, and for L2 grammatical skill from 0 to 61. A scale of 0 to 100 was used to measure L1 reading comprehension as assessed by

recall protocols, L2 reading comprehension, and L2 vocabulary knowledge.

The purpose of the present chapter is to report the analysis of the data after they have been subjected to a program of hierarchical multiple regression. The results will be reported separately according to the genre of L1 reading comprehension assessment measure used (NDRT comprehension section or recall protocol procedure).

Sample size fluctuates because of absences and the unavailability of certain test scores. NDRT scores were unavailable for five subjects. Scores for grammatical ability were inaccessible for two subjects. Because two subjects wrote their L1 recall protocols in French, their scores could not be used as measures of L1 reading comprehension. One subject was absent and thus unable to complete the vocabulary measure. L1 writing scores were unavailable for three subjects.

Rationale for Subject Grouping

Although the subjects were enrolled at beginning, intermediate, and advanced course levels, the intermediate- and advanced-level subjects were consolidated into a single non-beginner group for the purposes of the present study. This merger was accomplished for two reasons. First, grouping the

intermediate and advanced subjects together afforded a larger sample size: 43 Non-Beginners rather than 21 intermediate and 22 advanced subjects. Second, a series of Analyses of Variance (ANOVAs) revealed fewer significant differences between the scores on all of the variables featured in this study for intermediate- versus advanced-level subjects than they did between the scores for Beginners versus the consolidated group of intermediate and advanced subjects, or Non-Beginners.

Table 1 presents the results of a series of ANOVAs performed on the scores of intermediate-level versus advanced-level subjects. (ANOVA tables appear in Appendices G through L.) The results are significant for the scores on three of the variables and insignificant for the scores on the remaining three. Significant differences emanated between the scores on the following variables: (a) L1 reading comprehension as measured by the scores on recall protocols, $F(1, 40) = 4.60$, $p=.0380$, (b) L2 vocabulary knowledge, $F(1, 40) = 20.98$, $p=.0019$, and (c) L2 grammatical skill, $F(1, 40) = 11.08$, $p=.0001$. No significant differences emerged between the scores on these variables: (a) L1 reading skill as assessed by the scores on the comprehension portion of the NDRT, $F(1, 34), = 2.83$, $p=.1020$, (b) L1 writing ability, $F(1, 36) =$

.02, p=.8863, and (c) L2 reading comprehension, F(1, 40) = 2.07, p=.1580.

Table 2 delineates the results of a series of ANOVAs performed on the scores of Beginners versus upper-level subjects (Non-Beginners). (ANOVA tables appear in Appendices M through R.) The results are significant for the scores on every variable except for L1 writing ability. Significant differences emanated as follows: (a) L1 reading skill as assessed by scores on the comprehension portion of the NDRT, F(1, 76) = 5.47, p=.0219, (b) L1 reading comprehension as measured by the scores on recall protocols, F(1, 82) = 11.48, p=.0011, (c) L2 vocabulary knowledge, F(1, 82) = 136.10, p=.0001, (d) L2 grammatical skill, F(1, 82) = 317.02, p=.0001, and (e) L2 reading comprehension, F(1, 82) = 114.49, p=.0001. The results for L1 writing ability were insignificant: F(1, 76) = 2.74, p=.1021.

Table 1

Results of ANOVAs for Intermediate versus Advanced Subjects

Variables	N	F	p
1. L1 Reading			
a. NDRT	36	2.83	.1020
b. Protocols	42	4.60	.0380
2. L1 Writing	38	.02	.8863
3. L2 Vocabulary	42	20.98	.0019
4. L2 Grammar	42	11.08	.0001
5. L2 Reading	42	2.07	.1580

Table 2

Results of ANOVAs for Beginners versus Non-Beginners

Variables	N	F	p
1. L1 Reading			
a. NDRT	78	5.47	.0219
b. Protocols	84	11.48	.0011
2. L1 Writing	78	2.74	.1021
3. L2 Vocabulary	84	136.10	.0001
4. L2 Grammar	84	317.02	.0001
5. L2 Reading	84	114.49	.0001

Descriptive Data

Means and Standard Deviations

Table 3 presents the means, standard deviations, and minimum and maximum scores achieved for All Subjects and for all variables. The mean score for L1 reading ability as measured by the comprehension portion of the NDRT (N = 126) is 27.04 and the standard deviation is 5.58. Although the range of potential scores is 0 to 36, the lowest score attained is 8 while the highest is 35. The mean for L1 reading comprehension based on recall protocol scores (N = 129) is 38.96 and the standard deviation is 9.74. Because a percentage score was used for the recall protocols, the range of possible scores extends from 0 to 100. The lowest score obtained is 18 and the highest is 63. The average L1 writing score (N = 128) is 10.85 and the standard deviation is 2.70. Of a range of potential scores from 3 to 18, the lowest score achieved is 5, while the highest is 18. For L2 vocabulary knowledge (N = 130), the average score is 19.37 and the standard deviation is 17.11. Although the range of possible scores extends from 0 to 100, the lowest score obtained is 0 and the highest score is 81. The mean score for L2 grammatical ability (N = 129) is 28.90 and the standard deviation is 11.38. Of a range of possible scores from 0 to 61, the lowest score achieved

is 1 and the highest is 54. The mean L2 reading comprehension score (N = 131) is 25.90 and the standard deviation is 11.87. A percentage score was employed, resulting in a range of potential scores from 0 to 100. The lowest score is 7 and the highest is 62.

Table 4 reports the means, standard deviations, and minimum and maximum scores for Beginners. The L1 reading scores as measured by the comprehension portion of the NDRT (n = 87) average 25.87 and the standard deviation is 5.97. The minimum score obtained is 8, while the maximum is 34. The average L1 reading comprehension score based on recall protocols (n = 87) is 36.44 and the standard deviation is 8.71. The lowest score attained is 18 and the highest is 63. The mean L1 writing score (n = 87) is 10.31 and the standard deviation is 2.24. The lowest score achieved is 5 and the highest is 15. For L2 vocabulary knowledge (n = 88), the average score is 9.76 and the standard deviation is 6.34. The minimum score obtained is 0 and the maximum is 32. The mean score for L2 grammatical ability (n = 86) is 22.07 and the standard deviation is 6.13. The lowest score is 1, while the highest is 39. The average L2 reading comprehension score (n = 88) is 19.73 and the standard deviation is 6.53. The lowest score attained is 7 and the highest is 39.

Table 5 contains the means, standard deviations, and minimum and maximum scores obtained for Non-Beginners. The mean score for L1 reading as measured by the comprehension portion of the NDRT (n = 39) is 29.64 and the standard deviation is 3.44. The lowest score achieved is 21, while the highest is 35. The average L1 reading comprehension score (n = 42) as assessed using recall protocols is 44.19 and the standard deviation is 9.75. The minimum score obtained is 25 and the maximum is 62. For L1 writing (n = 41), the mean score is 12.00 and the standard deviation is 3.22. The lowest score attained is 5 and the highest is 18. The average score for L2 vocabulary knowledge (n = 42) is 39.50 and the standard deviation is 14.91. The minimum score is 10 and the maximum is 81. The mean score for L2 grammatical ability (n = 43) is 42.56 and the standard deviation is 5.65. The lowest score obtained is 29 while the highest is 54. For L2 reading comprehension (n = 43), the average score is 38.53 and the standard deviation is 10.22. The lowest score achieved is 21 and the highest is 62.

Table 3

Means and Standard Deviations of Literacy Variables for All Subjects

	N	Mean	SD	Minimum	Maximum
1. L1 Reading					
a. NDRT	126	27.04	5.58	8	35
b. Protocols	129	38.96	9.74	18	63
2. L1 Writing	128	10.85	2.70	5	18
3. L2 Vocabulary	130	19.37	17.11	0	81
4. L2 Grammar	129	28.90	11.38	1	54
5. L2 Reading	131	25.90	11.87	7	62

Table 4

Means and Standard Deviations for Literacy Variables for Beginners

	N	Mean	SD	Minimum	Maximum
1. L1 Reading					
a. NDRT	87	25.87	5.97	8	34
b. Protocols	87	36.44	8.71	18	63
2. L1 Writing	87	10.31	2.24	5	15
3. L2 Vocabulary	88	9.76	6.34	0	32
4. L2 Grammar	86	22.07	6.13	1	39
5. L2 Reading	88	19.73	6.53	7	39

Table 5

Means and Standard Deviations for Literacy Variables for Non-Beginners

	N	Mean	SD	Minimum	Maximum
1. L1 Reading					
a. NDRT	39	29.64	3.44	21	35
b. Protocols	42	44.19	9.75	25	62
2. L1 Writing	41	12.00	3.22	5	18
3. L2 Vocabulary	42	39.50	14.91	10	81
4. L2 Grammar	43	42.56	5.65	29	54
5. L2 Reading	43	38.53	10.22	21	62

Correlations

Table 6 delineates the correlations between the scores for all variables for All Subjects. Of the 14 correlations, 9 are significant at the .0001 level. The remaining five correlations are significant to at least the .0060 level: (a) the correlation between the scores for L1 reading comprehension as assessed by the NDRT and those for L1 writing ability, (b) the correlation between the scores for L1 reading comprehension as assessed by the NDRT and those for L2 vocabulary knowledge, (c) L1 reading scores as measured by recall protocols and the scores for L2 grammatical skill, (d) the correlation between L1 writing and L2 vocabulary scores, and (e) the correlation between the scores for L1 writing and L2 grammatical skill.

The correlations between the scores for L2 reading comprehension and those for the other variables are as follows: (a) L1 reading skill as measured by the scores on the comprehension portion of the NDRT (N = 126), .38, (b) L1 reading ability as assessed by recall protocol scores (N = 129), .52, (c) L1 writing ability (N = 128), .37, (d) L2 vocabulary knowledge (N = 130), .72, and (e) L2 grammatical skill (N = 126), .73. The correlations between L1 reading scores as assessed by the comprehension portion of the NDRT and the scores for L1

writing ability (N = 123), L2 vocabulary knowledge (N = 126), and L2 grammatical skill (N = 124) are as follows: .25, .32, and .34. The correlations between the scores for L1 reading comprehension as measured by recall protocols and those for L1 writing ability (N = 127), L2 vocabulary skill (N = 127), and L2 grammatical ability (N = 126) are: .39, .34, and .33. The correlations between the scores for L1 writing ability and those for L2 vocabulary knowledge (N = 127) and L2 grammatical skill (N = 126) are .29 and .28. The correlation between the scores for L2 vocabulary skill and L2 grammatical knowledge (N = 128) is .84.

Table 7 reflects the correlations between the scores for all variables for Beginners. Of the 14 correlations, 8 are significant. The correlation between the scores for L2 vocabulary knowledge and L2 grammatical skill are significant at the .0001 level. The remaining seven significant correlations (significant to at least the .0140 level) include those between L1 writing ability and L1 reading comprehension as measured by both the NDRT and the recall protocol procedure, between L1 writing ability and L2 reading comprehension, between L1 reading comprehension as measured by both the NDRT and the recall protocol procedure and L2 reading comprehension, between L2 vocabulary knowledge and L2 reading comprehension, and

between L2 grammatical skill and L2 reading comprehension. Insignificant correlations emanated between the L1 reading scores as measured by both the NDRT and the recall protocol procedure and L2 vocabulary knowledge and L2 grammatical skill, and between L1 writing ability and L2 vocabulary knowledge and L2 grammatical skill.

The correlations between the scores for L2 reading comprehension and those for the other variables are: (a) L1 reading skill as measured by the scores on the comprehension portion of the NDRT (n = 87), .27, (b) L1 reading comprehension as assessed by the recall protocol procedure (n = 87), .33, (c) L1 writing ability (n = 87), .28, (d) L2 vocabulary knowledge (n = 88), .36, and (e) L2 grammatical skill (n = 86), .27. The correlations between the scores for L1 reading ability as measured by the comprehension section of the NDRT and those for L1 writing ability (n = 86), L2 vocabulary knowledge (n = 87), and L2 grammatical skill (n = 85) are: .29, .14, and .10. The correlations between the scores for L1 reading as assessed using recall protocols and those for L1 writing ability (n = 87), L2 vocabulary knowledge (n = 87), and L2 grammatical ability (n = 85) are: .36, .14, and -.02. The correlations between the scores for L1 writing ability and those for L2 vocabulary knowledge

(n = 87) and L2 grammatical ability (n = 85) are .20 and .18. The correlation between the scores for L2 vocabulary knowledge and L2 grammatical skill (n = 86) is .56.

Table 8 displays the correlations between the scores for all variables for Non-Beginners. Of 14 possible correlations, 2 are significant, both of them having achieved significance to at least the .0011 level: (a) the scores for L2 reading comprehension and L1 reading ability as measured by recall protocols, and (b) the scores for L2 vocabulary knowledge and L2 grammatical skill.

The correlations between the scores for L2 reading comprehension and those for the other variables are: (a) L1 reading skill as measured by the scores on the comprehension portion of the NDRT (n = 39), .21, (b) L1 reading comprehension as assessed by the recall protocol procedure (n = 42), .49, (c) L1 writing ability (n = 41), .20, (d) L2 vocabulary knowledge (n = 42), .27, and (e) L2 grammatical skill (n = 43), .27. The correlations between the scores for L1 reading ability as measured by the comprehension section of the NDRT, and those for L1 writing ability (n = 37), L2 vocabulary knowledge (n = 39), and L2 grammatical skill (n = 39) are: -.09, .18, and .27. The correlations between the scores for L1

reading comprehension as assessed using recall protocols, and those for L1 writing ability (n = 41), L2 vocabulary knowledge (n = 41), and L2 grammatical skill (n = 42) are: .26, .04, and .03. The correlations between the scores for L1 writing ability and those for L2 vocabulary knowledge (n = 40) and L2 grammatical skill (n = 41) are .04 and -.20. The correlation between the scores for L2 vocabulary knowledge and L2 grammatical skill (n = 42) is .49.

Table 6

Correlations among Literacy Variables for All Subjects

	2	3	4	5
1. L1 Reading				
a. NDRT	.25**	.32**	.34*	.38*
	(123)	(126)	(124)	(126)
b. Protocols	.39*	.34*	.33**	.52*
	(127)	(127)	(126)	(129)
2. L1 Writing		.29**	.28**	.37*
		(127)	(126)	(128)
3. L2 Vocabulary			.84*	.72*
			(128)	(130)
4. L2 Grammar				.73*
				(129)
5. L2 Reading				

Note. Sample size is indicated in parentheses.

*p=.0001. **p≤.0060.

Table 7

Correlations among Literacy Variables for Beginners

	2	3	4	5
1. L1 Reading				
a. NDRT	.29**	.14	.10	.27**
	(86)	(87)	(85)	(87)
b. Protocols	.36**	.14	-.02	.33**
	(87)	(87)	(85)	(87)
2. L1 Writing		.20	.18	.28**
		(87)	(85)	(87)
3. L2 Vocabulary			.56*	.36**
			(86)	(88)
4. L2 Grammar				.27**
				(86)
5. L2 Reading				

Note. Sample size is indicated in parentheses.

*p=.0001. **p≤.0140.

Table 8

Correlations among Literacy Variables for Non-Beginners

	2	3	4	5
1. L1 Reading				
a. NDRT	-.09	.18	.27	.21
	(37)	(39)	(39)	(39)
b. Protocols	.26	.04	.03	.49*
	(41)	(41)	(42)	(42)
2. L1 Writing		.04	-.20	.20
		(40)	(41)	(41)
3. L2 Vocabulary			.49*	.27
			(42)	(42)
4. L2 Grammar				.27
				(43)
5. L2 Reading				

Note. Sample size is indicated in parentheses.

*$p \leq .0011$.

Inferential Data

First Research Question

What is the unique contribution to L2 reading comprehension of each of the following: (a) L1 reading comprehension as measured by the score on the comprehension portion of the NDRT, and then by the mean score on three recall protocols, (b) L1 writing ability as measured by the combined holistic scores on three writing samples, (c) passage-specific L2 vocabulary knowledge as measured by score on a supply-definition test, and (d) L2 grammatical skill as measured by the score on the grammar section of the United States Air Force Academy's French Placement Validation Test.

In all of the hierarchical regression programs used in the present study, the scores for the independent variables were entered into the statistical model in the following order: (a) L1 reading, (b) L1 writing, (c) L2 vocabulary, and (d) L2 grammar. The L1 variables were introduced first because the subjects, all native English speakers, are considerably more experienced with the English language than they are with the L2. L1 reading data were inserted before the L1 writing scores because, according to Shanahan and Lomax (1986), "more reading information is used in writing than vice versa" (p. 122). L2 vocabulary scores were entered before those for L2

grammar because L2 vocabulary skill consistently contributed more to L2 reading comprehension scores than did L2 grammatical knowledge in the stepwise multiple regression analyses performed by Henning (1975) and Koda (1990).

Table 9 portrays the results of the hierarchical regression program for All Subjects, using scores from the comprehension portion of the NDRT as the L1 reading measure. The scores for all of the independent variables contributed significantly to the variance of the dependent variable, L2 reading comprehension. L2 vocabulary (N = 123) provided the major contribution: 35.68% of the variance, $F(1, 122) = 58.27$, $p=.0001$. L1 reading (N = 126) offered the second highest contribution: 14.12% of the variance, $F(1, 125) = 20.38$, $p=.0001$. The cumulative contribution of L2 vocabulary and L1 reading was 49.80%. L1 writing (N = 123), the third highest contributor, supplied 7.03% of the variance, $F(1, 122) = 10.70$, $p=.0500$. The cumulative contribution of the L1 variables and L2 vocabulary was 56.83%. L2 grammar (N = 121) accorded the smallest contribution to the variance: 3.04%, $F(1, 120)$, $p<.0500$. The combined contribution of all four independent variables was 59.87%. The L1 variables accounted for 21.15%, while the L2 variables contributed 38.72%.

Table 10 portrays the results of the hierarchical regression program for All Subjects, using recall protocol scores to assess L1 reading comprehension. All of the independent variables contributed significantly to the variance of the scores for L2 reading comprehension. L2 vocabulary (N = 126) provided the main contribution: 30.91% of the variance, $F(1, 125) = 99.84$, $p=.0001$. L1 reading (N = 128) offered the second highest contribution: 27.49% of the variance, $F(1, 127) = 47.77$, $p=.0001$. The combined contribution of L2 vocabulary and L1 reading was 58.40%. L1 writing (N = 127), the third highest contributor, accorded 3.84% of the variance, $F(1, 126) = 6.93$, $p=.0200$. The cumulative contribution of the L1 variables and L2 vocabulary was 62.24%. L2 grammar (N = 124) supplied the smallest contribution to the variance: 3.19%, $F(1, 123) = 10.97$, $p=.0050$. The combined contribution of all four independent variables was 65.43%. The L1 variables accounted for 31.33%, while the L2 variables contributed 34.10%.

Table 9

Summary of Hierarchical Regression Analysis for All Subjects using NDRT Comprehension Scores

Steps and Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.1412	.1412	20.38	.0001
	(126)			
2. L1 Writing	.0703	.2115	10.70	.0500
	(123)			
3. L2 Vocabulary	.3568	.5683	58.27	.0001
	(123)			
4. L2 Grammar	.0304	.5987	8.78	.0500
	(121)			

Note. Sample size is indicated in parentheses.

Table 10

Summary of Hierarchical Regression Analysis for All Subjects using L1 Protocol Scores

Steps and Variables	R^2 Increment	Cum. R^2	F	p
1. L1 Reading	.2749	.2749	47.77	.0001
	(128)			
2. L1 Writing	.0384	.3133	6.93	.0200
	(127)			
3. L2 Vocabulary	.3091	.6224	99.84	.0001
	(126)			
4. L2 Grammar	.0319	.6543	10.97	.0050
	(124)			

Note. Sample size is indicated in parentheses.

Second Research Question

Does the unique contribution of each independent variable differ by level of L2 study?

In order to answer this research question, the results of the hierarchical regression programs for Beginners and Non-Beginners will first be reported and then be compared to each other. First, the regression models for Beginners and Non-Beginners using the comprehension scores from the NDRT as the L1 reading measure will be treated, followed by the regression models for both groups using recall protocol scores to assess L1 reading comprehension.

Table 11 reveals the results of the hierarchical regression program for Beginners using the comprehension portion of the NDRT to measure L1 reading skill. L2 vocabulary knowledge and L1 reading comprehension contributed significantly to the variance of the L2 reading comprehension scores, while L1 writing ability and L2 grammatical skill did not. L2 vocabulary ($n = 86$) provided the principal contribution: 8.50% of the variance, $F(1, 85) = 8.61$, $p=.0043$. L1 reading ($n = 87$) offered the next highest contribution: 7.06% of the variance, $F(1, 86) = 6.46$, $p=.0129$. The combined contributions of L2 vocabulary and L1 reading was 15.56%. L1 writing ($n = 86$) was the third highest contributor,

supplying 3.47% of the variance, $F(1, 85) = 3.22$, $p=.0600$. The cumulative contribution of the L1 variables and L2 vocabulary was 19.03%. L2 grammar (n = 84) accorded the smallest contribution to the variance: 1.24%, $F(1, 83) = 1.23$, $p=.2000$. The combined contribution of all four independent variables was 20.27%. The L1 variables accounted for 10.53% of the variance, while the L2 variables contributed 9.74%.

Table 12 discloses the results of the hierarchical regression program for Non-Beginners using the comprehension portion of the NDRT to assess L1 reading skill. None of the variables contributed significantly. L2 vocabulary (n = 37) provided the chief contribution: 8.67% of the variance, $F(1, 36) = 3.53$, $p=.0691$. L1 writing (n = 37) supplied the second highest contribution: 5.85%, $F(1, 36) = 2.21$, $p=.1500$. The cumulative contribution of L1 writing and L2 vocabulary was 14.52%. L1 reading (n = 39), the third highest contributor, supplied 4.42% of the variance, $F(1, 38) = 1.71$, $p=.1992$. The cumulative contribution of the L1 variables and L2 vocabulary was 18.94%. L2 grammar (n = 37) accorded the smallest contribution: 1.49%, $F(1, 36) = .60$, $p=.4444$. The combined contribution of all four independent variables was 20.43%. The L1 variables

accounted for 10.27% of the variance, while the L2 variables contributed 10.16%

Differences have been revealed in the magnitudes of the relative contributions of the independent variables to the scores for L2 reading comprehension for Beginners versus Non-Beginners. The major dissimilarity presented itself in order of the contributions of the independent variables. Although L2 vocabulary was the primary contributor for both groups, L1 writing emerged as the secondary contributor for Non-Beginners, while L1 reading comprehension did so for Beginners. The contribution of L1 writing was greater by 2.38% for Non-Beginners, and L1 reading offered 2.64% more of the variance for Beginners. The contributions of the L2 variables were highly similar. L2 vocabulary contributed nearly equally for Beginners (8.50%) and for Non-Beginners (8.67%), or .17% more to the variance for Non-Beginners. L2 grammar offered the smallest contribution for both groups: 1.24% for Beginners and 1.49% for Non-Beginners. Thus, L2 grammar contributed only .25% more to the variance for Non-Beginners.

Table 11

Summary of Hierarchical Regression Analysis for Beginners Using NDRT Comprehension Scores

Steps/Variables	R^2 Increase	Cum. R^2	F	p
1. L1 Reading	.0706	.0706	6.46	.0129
	(87)			
2. L1 Writing	.0347	.1053	3.22	.0600
	(86)			
3. L2 Vocabulary	.0850	.1903	8.61	.0043
	(86)			
4. L2 Grammar	.0124	.2027	1.23	.2000
	(84)			

Note. Sample size is indicated in parentheses.

Table 12

Summary of Hierarchical Regression Analysis for Non-Beginners Using NDRT Comprehension Scores

Steps/Variables	R^2 Increase	Cum. R^2	F	p
1. L1 Reading	.0442	.0442	1.71	.1992
	(39)			
2. L1 Writing	.0585	.1027	2.21	.1500
	(37)			
3. L2 Vocabulary	.0867	.1894	3.53	.0691
	(37)			
4. L2 Grammar	.0149	.2043	.60	.4444
	(37)			

Note. Sample size is indicated in parentheses.

Table 13 discloses the results of the hierarchical regression program for Beginners using recall protocol scores to measure L1 reading comprehension. L1 reading and L2 vocabulary contributed significantly, while L1 writing and L2 grammar did not. L1 reading (n = 87) provided the foremost contribution: 10.58% of the variance, $F(1, 86) = 10.06$, $p=.0021$. L2 vocabulary (n = 87) supplied the second highest contribution to the variance: 8.38%, $F(1, 86) = 8.91$, $p=.0037$. The cumulative contribution of L1 reading and L2 vocabulary was 18.96%. L1 writing (n = 87). the third highest contributor, accorded 3.00%, $F(1, 86) = 2.91$, $p=.0916$. The cumulative contribution of the L1 variables and L2 vocabulary was 21.96%. L2 grammar (n = 85) supplied the smallest contribution to the variance: 1.85%, $F(1, 84) = 1.95$, $p=.2000$. The combined contribution of all four independent variables was 23.81%. The L1 variables accounted for 13.58% of the variance, while the L2 variables contributed 10.23%

Table 14 discloses the results of the hierarchical regression program for Non-Beginners using recall protocol scores as the L1 reading comprehension measure. L1 reading (n = 42), the only variable to contribute significantly, provided the predominant contribution: 23.96% of the variance, $F(1, 41) = 12.60$, $p=.0010$. L2

vocabulary (n = 40) supplied the second highest contribution to the variance: 5.40%, $F(1, 39) = 2.74$, $p=.1500$. The cumulative contribution of L1 reading and L2 vocabulary was 29.36%. L2 grammar (n = 40) offered the third highest contribution: 2.69%, $F(1, 39) = 1.38$, $p=.2485$. The cumulative contribution of L1 reading and the L2 variables was 32.05%. L1 writing (n = 41) accorded the smallest contribution to the variance: .43%, $F(1, 40) = .21$, $p<.5000$. The combined contribution of all four independent variables is 32.48%. The L1 variables accounted for 24.39% of the variance, while the L2 variables contributed 8.09%.

Differences in the magnitudes of the relative contributions of the independent variables to the scores for L2 reading comprehension have been revealed for Beginners versus Non-Beginners. The differences manifested themselves primarily in the contribution of L1 reading comprehension. L1 reading contributed more than twice the variance to L2 reading comprehension for Non-Beginners (23.96%) than it did for Beginners (10.58%), a difference of 13.38%. L2 vocabulary contributed 8.38% for Beginners and 5.40% for Non-Beginners, offering 2.98% more to the variance for Beginners. L1 writing contributed 3.00% of the variance for Beginners and .48% for Non-Beginners, a difference of

2.52%. L2 grammar contributed nearly equally for Beginners (1.85%) and Non-Beginners (2.69%), supplying only .84% more of the variance for Non-Beginners.

Table 13

Summary of Hierarchical Regression Analysis for Beginners Using L1 Protocol Scores

Steps/Variables	R^2 Increase	Cum. R^2	F	p
1. L1 Reading	.1058	.1058	10.06	.0021
	(87)			
2. L1 Writing	.0300	.1358	2.91	.0916
	(87)			
3. L2 Vocabulary	.0838	.2196	8.91	.0037
	(87)			
4. L2 Grammar	.0185	.2381	1.95	.2000
	(85)			

Note. Sample size is indicated in parentheses.

Table 14

Summary of Hierarchical Regression Analysis for Non-Beginners Using L1 Protocol Scores

Steps/Variables	R^2 Increase	Cum. R^2	F	p
1. L1 Reading	.2396	.2396	12.60	.0010
	(42)			
2. L1 Writing	.0043	.2439	.21	.5000
	(41)			
3. L2 Vocabulary	.0540	.2979	2.74	.1500
	(40)			
4. L2 Grammar	.0269	.3248	1.38	.2485
	(40)			

Note. Sample size is indicated in parentheses.

Discussion

Means and Standard Deviations

The means, standard deviations, and minimum and maximum scores obtained for the literacy variables for all subjects reveal some interesting differences between the data for L1 reading comprehension as measured by the NDRT versus recall protocols. (See Table 3.) More subjects attained high scores on the NDRT comprehension section than they did on the recall protocols. This difference is evidenced by the fact that the mean NDRT comprehension score (27.04) was only 7.96 points below the maximum score achieved (35), while the average recall protocol score (38.96) was 24.04 points below the maximum obtained score (63). Additionally, the standard deviation for the NDRT comprehension scores (5.58) was lower than that of the recall protocol scores (9.74), indicating less variability in the NDRT comprehension scores. The frequency chart presented in Figure 1 clearly depicts a negative skewness for the NDRT comprehension scores, exhibiting a large cluster of high scores and little variability between scores. In contrast, the frequency chart of the recall protocol scores, shown in Figure 2, presents a near-normal curve demonstrating greater variability in scores. The integrative recall protocol is a more sensitive measure

of reading comprehension than is the discrete-point NDRT because the former is based on a continuous scale, while the latter is based on a discrete scale.

The means, standard deviations, and minimum and maximum scores on the literacy variables for Beginners versus Non-Beginners reflect several differences between the two groups. (See Tables 4 and 5.) The means for all variables were higher for Non-Beginners. The Beginners amassed all of the minimum scores, though the minimum score for L1 writing ability was the same for Beginners and Non-Beginners. The Non-Beginners achieved all but one of the maximum scores (L1 reading comprehension as assessed by the scores on recall protocols).

Additionally, the difference between the NDRT comprehension scores and the recall protocol scores that was seen in the data for All Subjects appeared once again for the Beginner and Non-Beginner groups. For both groups, more subjects achieved high scores on the NDRT comprehension section than they did on the recall protocols. For Beginners, the mean NDRT comprehension score (25.87) was only 8.13 points below the maximum score achieved (34), while the average recall protocol score (36.44) was 26.56 points below the maximum obtained score (63). For Non-Beginners, the average NDRT comprehension score (29.64) was only 5.36 points below

the maximum score obtained (35), while the mean L1 recall protocol score (44.19) was 17.81 points below the maximum score (62). Also, for both groups, the standard deviation for the NDRT comprehension scores was lower than that of the recall protocol scores, indicating less variability in the NDRT comprehension scores. For Beginners, the standard deviation for the NDRT comprehension scores was 5.97, and 8.71 for the recall protocol scores. For Non-Beginners, the standard deviation for the NDRT comprehension scores was 3.44, and 9.75 for the recall protocol scores. The frequency chart displayed in Figure 3 reflects the negatively skewed nature of the NDRT comprehension scores for Beginners, indicating a large cluster of high scores. In contrast, the frequency chart presented in Figure 4 demonstrates that the Beginners' recall protocol scores form a near-normal curve. Similarly, the frequency chart of NDRT comprehension scores exhibited in Figure 5 shows a profoundly negative skew for Non-Beginners, while the frequency chart displayed in Figure 6 indicates a near-normal curve for Non-Beginner's recall protocol scores. Once again, the continuous-scale recall protocol scores provided more variability than did the discrete-scale NDRT comprehension scores.

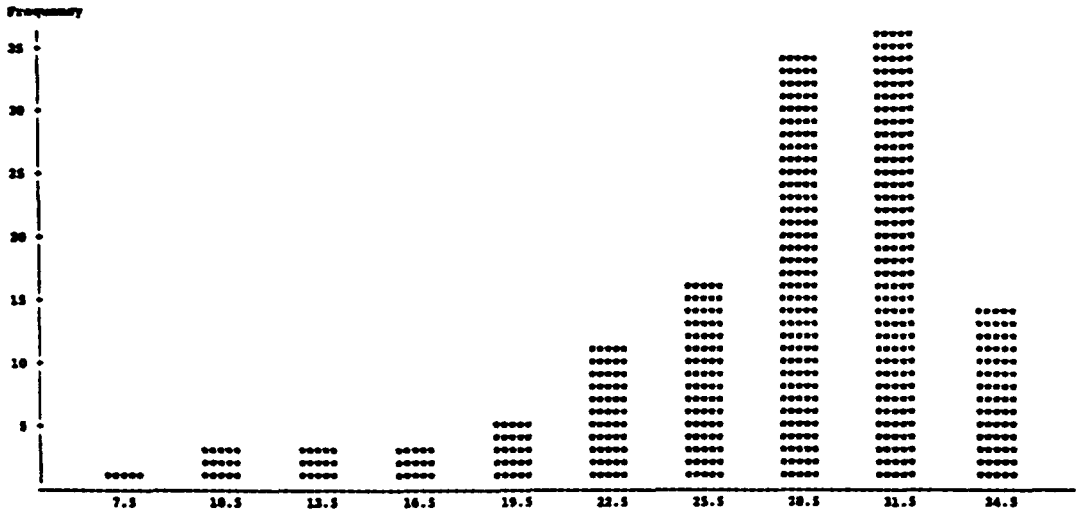


Figure 1. Frequencies of NDRT Comprehension Scores for All Subjects

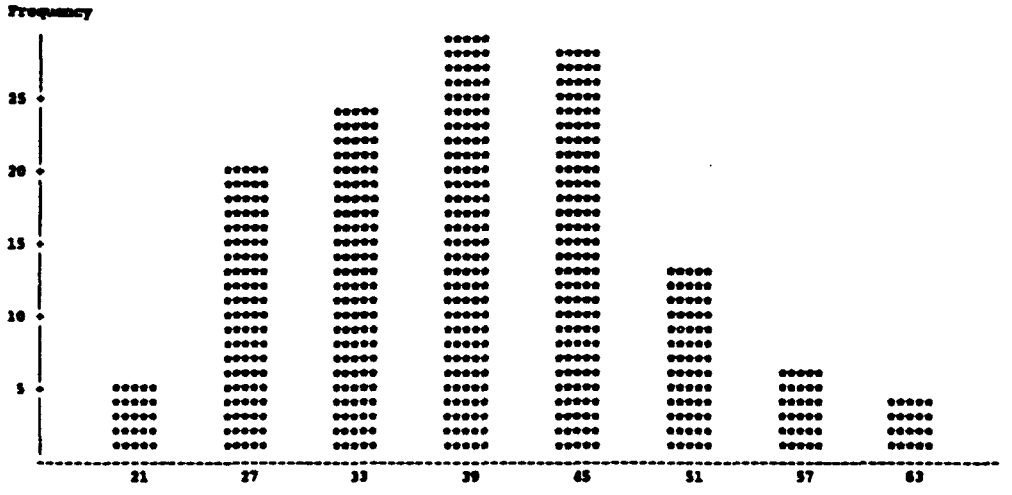


Figure 2. Frequencies of L1 Recall Protocol Scores for All Subjects

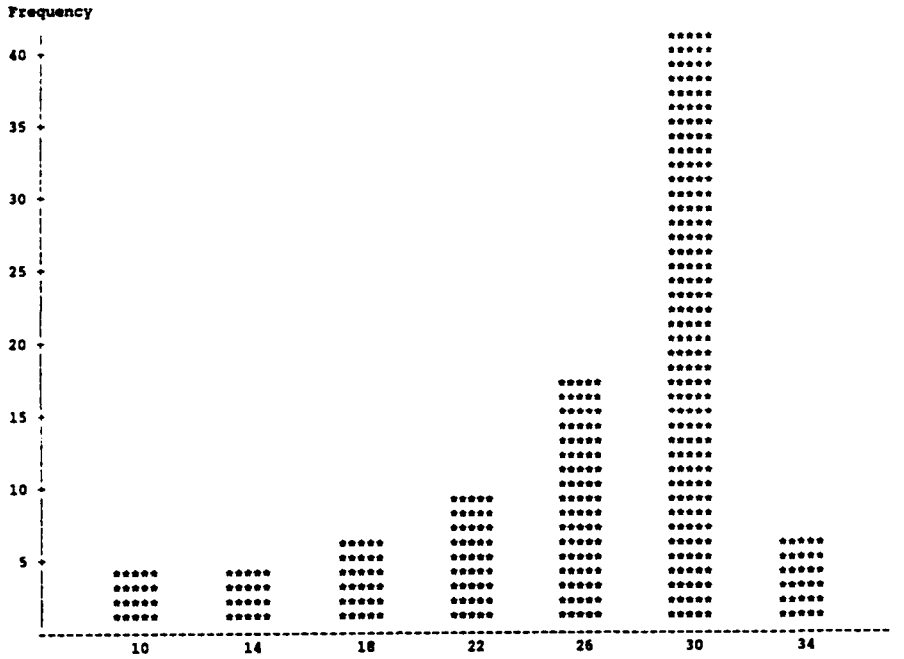


Figure 3. Frequencies of NDRT Comprehension Scores for Beginners

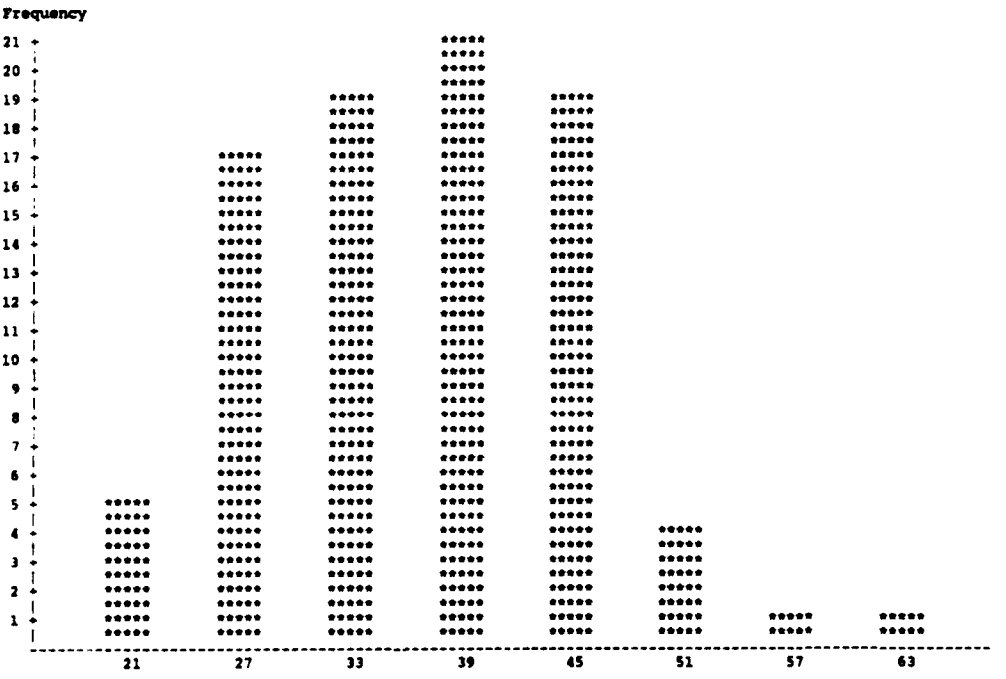


Figure 4. Frequencies of L1 Recall Protocol Scores for Beginners

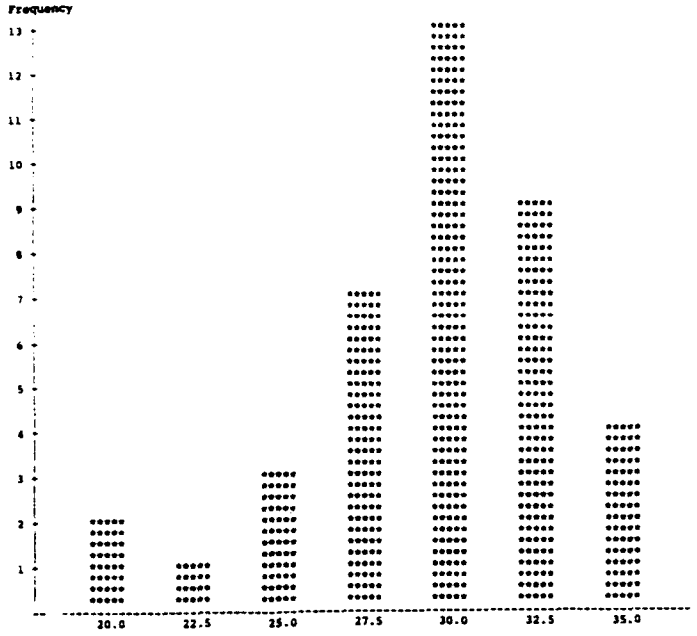


Figure 5 Frequencies of NDRT Comprehension Scores for Non-Beginners

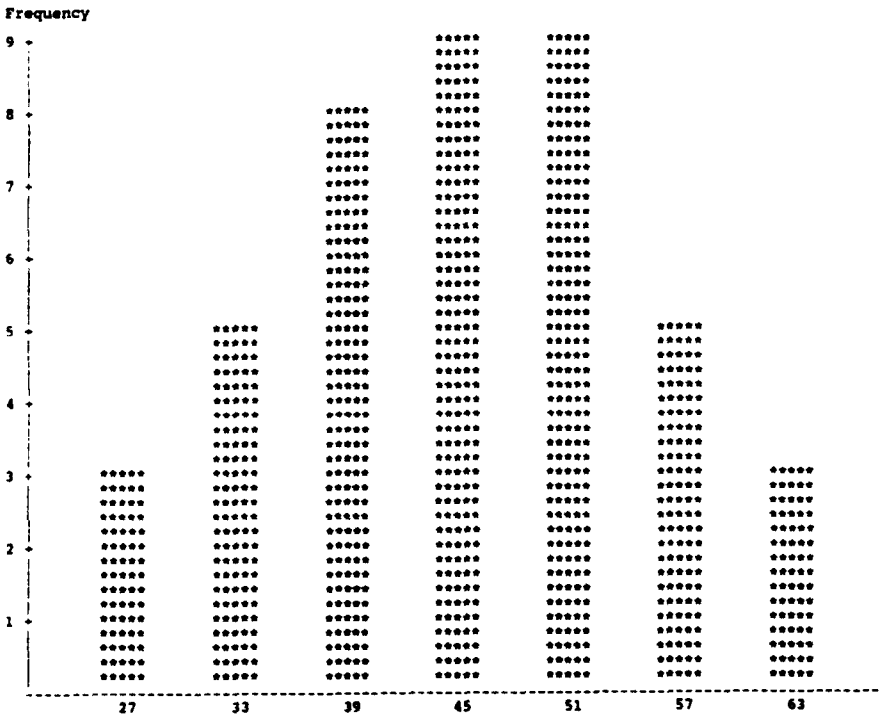


Figure 6. Frequencies of L1 Recall Protocol Scores for Non-Beginners

Correlations

Several interesting findings emerge from a comparison of the correlations among the variables for Beginners versus those for Non-Beginners. (See Tables 7 and 8.) A greater number of significant correlations arose between the variables for Beginners than for Non-Beginners. Of 14 possible correlations, 8 reached significance for Beginners and only 2 did so for Non-Beginners. The two significant correlations achieved by Non-Beginners were also significant for Beginners. Particularly interesting is the fact that for Non-Beginners, the only variable to correlate significantly with the scores for L2 reading comprehension was an L1 variable (L1 reading comprehension as measured by recall protocol scores) while all of the variables--both L1 and L2--did so for Beginners.

A difference between the groups was revealed regarding L1 reading comprehension scores as measured by the NDRT comprehension section versus recall protocols. For Beginners, the scores for both the NDRT comprehension section and the recall protocols correlated significantly with the scores for L2 reading comprehension, while only the recall protocol scores did so for Non-Beginners. Furthermore, for Beginners, the correlations between the

scores for L2 reading skill and those for L1 reading comprehension as measured by the NDRT comprehension section (.27) and by the recall protocols (.33) are similar in magnitude, differing by only .06. In contrast, for Non-Beginners, the magnitude of the correlation between the scores for L2 reading comprehension and those for L1 reading skill as measured by NDRT comprehension scores is .21, while the correlation between L2 reading comprehension scores and L1 recall protocols scores is .49, a difference of .29.

A dissimilarity was revealed between Beginners and Non-Beginners in the relationship between the scores for L1 reading comprehension and those for L1 writing ability. For Beginners, significant correlations emerged between the scores for L1 writing ability and those for L1 reading comprehension as assessed by both the NDRT comprehension section (.29) and the recall protocol procedure (.36). In contrast, for Non-Beginners, the correlations between the scores for L1 writing ability and those for L1 reading skill as measured by the NDRT comprehension section (-.09) and the recall protocol procedure (.26) are not significant.

Hierarchical Regression Models

Additional Findings: Beginners versus Non-Beginners using NDRT comprehension scores.

In addition to the contrasts between Beginners and Non-Beginners in the relative contribution of each independent variable to the dependent variable (See Second Research Question), several related differences were revealed in the results of the hierarchical regression programs. Two of the four independent variables achieved significance for Beginners (L2 vocabulary and L1 reading), while none did so for Non-Beginners.

A difference exists in the orders of importance of the contributions of the independent variables for Beginners versus Non-Beginners. For Beginners, the independent variables contributed in the following order: (a) L2 vocabulary (8.50%), (b) L1 reading (7.06%), (c) L1 writing (3.47%), and (d) L2 grammar (1.24%). For Non-Beginners, however, the order was: (a) L2 vocabulary (8.67%), (b) L1 writing (5.85%), (c) L1 reading (4.42%), and (d) L2 grammar (1.49%). Although L2 vocabulary contributed the most for both groups, and L2 grammar contributed the least, L1 reading contributed more for Beginners, while L1 writing did so for Non-Beginners.

For both groups, the L1 variables accounted for a slightly greater percentage of the variance than did the L2 variables. The contributions of the L1 variables were similar for both groups: 10.53% for Beginners and 10.27% for Non-Beginners, a difference of only .26%. The L2 variables also contributed nearly equally for Beginners (9.74%) and Non-Beginners (10.16%), .42% more for Non-Beginners.

The total variance accounted for was also nearly identical by group. For Beginners, 20.27% of the total variance was accounted for, versus 20.43% for Non-Beginners, .16% more for Non-Beginners. Furthermore, the amount of variance that remains to be accounted for is nearly equal for both groups. For Beginners, 79.73% of the variance is unaccounted for, versus 79.57% for Non-Beginners.

Additional findings: Beginners versus Non-Beginners using recall protocol scores.

In addition to the differences between Beginners and Non-Beginners in the relative contribution of each independent variable to the dependent variable (See Second Research Question), several other dissimilarities were revealed in the results of the hierarchical regression programs. Two of the four independent variables achieved significance for Beginners (L1 reading

and L2 vocabulary), while only L1 reading contributed significantly for Non-Beginners.

The orders of importance of the contributions differed for Beginners versus Non-Beginners. For Beginners, the order of importance was as follows: (a) L1 reading (10.58%), (b) L2 vocabulary (8.38%), (c) L1 writing (3.00%), and (d) L2 grammar (1.85%). In contrast, the order for Non-Beginners was: (a) L1 reading (23.96%), (b) L2 vocabulary (5.40%), c) L2 grammar (2.69%), and (d) L1 writing (.43%). For both groups, L1 reading was the chief contributor and L2 vocabulary accorded the second highest amount of variance. While L1 writing served as the third principal contributor for the Beginners, and grammar the lowest contributor, the opposite occurred for Non-Beginners.

For both groups, the L1 variables accounted for more of the variance in L2 reading comprehension scores than did the L2 variables. A difference exists, however, in the magnitudes of the contributions. For Non-Beginners, the contribution of the L1 variables was greater by 10.81%: 13.58% for Beginners and 24.39% for Non-Beginners. The L2 variables contributed similarly for both groups: 10.23% for Beginners and 8.09% for Non-Beginners, a difference of 2.14%.

The total variance accounted for also differed by group. For Beginners, 23.81% of the total variance was accounted for, versus 32.48% for Non-Beginners, a difference of 8.67%. Thus, the amount of variance that remains to be accounted for is greater for Beginners. For Non-Beginners, 67.52% of the variance is unaccounted for, versus 76.19% for Beginners.

A comparison of the hierarchical regression models for All Subjects: NDRT comprehension scores versus recall protocol scores.

Several similarities and differences emanated between the results of the hierarchical regression programs for all subjects using NDRT comprehension scores versus L1 protocol scores. (See Tables 9 and 10.) In terms of similarities, all of the independent variables achieved significance in both models. Additionally, the orders of importance of the contributions of the independent variables were identical for both models: (a) L2 vocabulary, (b) L1 reading, (c) L1 writing, and (d) L2 grammar.

The relative contributions of the independent variables, however, differed between models. The contribution of L2 vocabulary was greater by 4.77% in the model using NDRT comprehension scores (where it contributed 35.68%) than it was in the program using

recall protocol scores (where it supplied 30.91%). L1 reading in the model using recall protocols accounted for nearly twice the variance (27.49%) as it did in the model using the NDRT comprehension section (14.12%), a difference of 13.37%. The contribution of L1 writing was greater in the model using NDRT comprehension scores. In the model using NDRT comprehension scores, the contribution of L1 writing was 7.03%, and 3.84% in the model using recall protocols scores, a difference of 3.19%. The contribution of L2 grammar was also nearly identical regardless of the L1 reading assessment method used. L2 grammar offered 3.04% of the variance when NDRT comprehension scores were used and 3.19% when recall protocol scores were employed, a difference of only .15%.

The contribution of the combined L1 variables was greater by 10.18% in the model using recall protocol scores (31.33%) versus the model using NDRT comprehension scores (21.15%). The combined L2 variables contributed similarly, regardless of the L1 reading measure used: 38.72% using NDRT comprehension scores and 34.10% using recall protocol scores, a difference of 4.62%. The model using NDRT comprehension scores resulted in a slightly smaller overall contribution to the variance (59.87%) than the model featuring recall protocol scores (65.43%), a difference of 5.56%. Thus, the total variance

unaccounted for is greater in the model using NDRT comprehension scores (40.13%) than it is in the model using recall protocol scores (34.57%).

A comparison of the hierarchical regression models for Beginners: NDRT comprehension scores versus recall protocol scores.

The results of the hierarchical regression programs for Beginners reveal several similarities and differences. (See Tables 11 and 13.) In both models, L2 vocabulary and L1 reading were significant, while L1 writing and L2 grammar were not. The orders of importance of the contributions of the independent variables, however, were different. In the model featuring NDRT comprehension scores, the contributions emanated as follows: (a) L2 vocabulary, (b) L1 reading, (c) L1 writing, and (d) L2 grammar. The order of importance of the variables in the model using recall protocol scores was: (a) L1 reading, (b) L2 vocabulary, (c) L1 writing, and (d) L2 grammar.

The contributions of L2 vocabulary were nearly identical in the model using NDRT comprehension scores (8.50%) and in the model using recall protocol scores (8.38%), differing by only .12%. L1 reading scores also contributed similarly to the variance, regardless of the L1 assessment method employed: 10.58% using recall

protocols and 7.06% using NDRT comprehension scores, a difference of 3.52%. In the model using NDRT comprehension scores, L1 writing contributed 3.47%, while it supplied 3.00% of the variance in the model using recall protocol scores, a difference of only .47%. L2 grammar offered 1.24% of the variance in the model using NDRT comprehension scores and 1.85% in the model using recall protocol scores, a difference of only .61%.

The combined L1 variables contributed comparably to the variance in both models: 10.53% in the model using NDRT comprehension scores and 13.58% in the model using recall protocol scores, a difference of only 3.05%. The contribution of the L2 variables was also almost equal for both models: 9.74% using NDRT comprehension scores and 10.23% using recall protocol scores, a difference of only .49%. The cumulative contribution of all four independent variables was similar in the model using NDRT comprehension scores (20.27%) and in the model featuring recall protocol scores (23.81%), a difference of only 3.54%. Thus, the total variance unaccounted for is nearly the same: 79.73% in the model using NDRT comprehension scores and 76.19% in the model using recall protocol scores.

A comparison of the hierarchical regression models for Non-Beginners: NDRT comprehension scores versus recall protocol scores.

Several important differences have been revealed in the results of the hierarchical regression programs for Non-Beginners using the NDRT comprehension section versus recall protocols to assess L1 reading comprehension. (See Tables 12 and 14.) In the model using NDRT comprehension scores, none of the variables are significant. In contrast, using recall protocol scores, the contribution of L1 reading is significant, while the contributions of the other variables remain insignificant. The orders of importance of the contributions of the independent variables differ according to the L1 reading measure used. Using NDRT comprehension scores, the order of importance is: (a) L2 vocabulary, (b) L1 writing, (c) L1 reading, and (d) L2 grammar. In contrast, the order of importance using recall protocol scores is: (a) L1 reading (b) L2 vocabulary, (c) L2 grammar, and (d) L1 writing.

The major difference between models in the relative contributions of the independent variables focuses on L1 reading comprehension. L1 reading comprehension when assessed using recall protocol scores accounted for more than five times the variance (23.96%) than it did when

measured using the NDRT comprehension section (4.42%), a difference of 19.54%. The contribution of L1 writing was greater by 5.42% in the model using NDRT comprehension scores (in which it accounted for 5.85% of the variance) versus the model using recall protocol scores (in which it contributed .43). The contribution of L2 vocabulary was 8.67% using NDRT comprehension scores and 5.40% using recall protocol scores, a difference of 3.27%. The contribution of L2 grammar was nearly the same regardless of L1 reading assessment method: 1.49% using NDRT comprehension scores and 2.69% using recall protocol scores, a difference of only 1.20%.

The contribution of the combined L1 variables was greater by 14.12% in the model using recall protocol scores (24.39%) versus the model using NDRT comprehension scores (10.27%). In contrast, the contribution of the L2 variables was almost equal for both models: 10.16% using NDRT comprehension scores and 8.09% using recall protocol scores, a difference of only 2.07%. The cumulative contribution of all four independent variables was greater by 12.05% in the model using recall protocol scores were used (32.48%) versus the model using NDRT comprehension scores (20.43%). Thus, the variance remaining unaccounted for is greater in the model using

NDRT comprehension scores (79.57%) than it is in the model using recall protocol scores (67.52%).

Summary

In summary, the total variance accounted for in the hierarchical regression models differs by course level and according to the instrument used to measure L1 reading comprehension. For All Subjects, 65.43% of the variance was accounted for when recall protocol scores were used, versus 59.87% when NDRT comprehension scores were employed. For Beginners, 23.81% of the variance was accounted for using recall protocol scores, as opposed to 20.27% using NDRT comprehension scores. For Non-Beginners, 32.48% of the variance was accounted for when recall protocol scores were used, compared to 20.43% when NDRT comprehension scores were employed.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Overview of the Study

The present study has focused on enhancing the validity of second language (L2) reading comprehension assessment measures by establishing the contributions of selected native language (L1) and L2 literacy variables to the scores on L2 reading comprehension tests. This study has investigated the unique contributions of the scores on measures of L1 writing ability, L1 reading comprehension, L2 vocabulary knowledge, and L2 grammatical skill to L2 reading comprehension test scores, and has examined whether or not these contributions differ by language course level.

The subjects included 131 cadets (88 Beginners and 43 Non-Beginners) enrolled in French courses at the United States Air Force Academy in Colorado Springs, CO. Their L1 reading ability was assessed using two different testing instruments: (a) the comprehension portion of the Nelson-Denny Reading Test (NDRT), Form E (Brown, Bennett, & Hanna, 1981), and (b) the recall protocol

procedure. L1 writing performance was measured for overall quality using the Test of Written English holistic scoring guide (Educational Testing Service, 1990). A passage-specific supply-definition test was employed to assess L2 vocabulary knowledge. L2 grammatical ability was measured using the grammar portion of the Air Force Academy's French placement test. The dependent variable, L2 reading comprehension, was tested using recall protocols.

The data were subjected to several different programs of hierarchical multiple regression. Separate programs were run for All Subjects, for Beginners, and for Non-Beginners. Additionally, because L1 reading ability was measured using two different testing instruments, separate multiple regression programs were performed for each subject group, first using the scores from the comprehension portion of the NDRT, and then using the recall protocol scores.

Summary of Findings and Conclusions

Overview

This section is divided into two main parts: (a) Descriptive Data, containing information about the means, standard deviations, and correlations calculated in the present study and (b) Inferential Data, including information emanating from the hierarchical regression

programs. Individual summaries of findings and conclusions will be provided for the means and standard deviations, for the correlations, and for the hierarchical regression data.

Descriptive Data

Means and standard deviations.

In summary, three principal findings have emerged from an analysis of the means, standard deviations, and minimum and maximum scores. (See Chapter IV, Tables 3, 4, and 5.) First, the means are higher for Non-Beginners than for Beginners on all variables. Second, for both Beginners and Non-Beginners, the mean NDRT comprehension scores tend to be higher (concentrated closer to the maximum possible score) than the mean recall protocol scores. Third, the standard deviation of the NDRT comprehension scores is smaller for Non-Beginners than it is for Beginners, while the opposite is true for the standard deviation of the recall protocol scores. In other words, the Non-Beginners' NDRT comprehension scores are less variable than those of the Beginners, while the Non-Beginners' recall protocol scores are more variable.

As documented by the differences between the means for Beginners versus Non-Beginners, the Non-Beginners' performance was superior for all of the L1 and L2 variables. According to the series of analyses of

variance that was performed in order to group the subjects, the differences are in fact significant for all variables except L1 writing ability. (See Chapter IV, Table 2.) The dissimilarity between groups for the L2 variables (reading comprehension, vocabulary knowledge, and grammatical ability) is consistent with the course curriculum in which the subjects are increasingly exposed to the L2, and particularly to elements of L2 literacy. As expected, the Non-Beginners, possessing more L2 knowledge, scored higher on the L2 variables. Unexpected, however, was the significant difference between the groups for L1 reading comprehension. These data suggest that language students opting to continue their L2 studies comprise a self-selected group of exceptionally literate individuals. Furthermore, high proficiency in L1 reading appears to be consistent with high proficiency in L2 reading, pointing to the transfer of literacy abilities between languages.

The finding that the subjects' scores were generally higher on the comprehension portion of the NDRT versus the recall protocols is indeed thought-provoking. Because of the cluster of high NDRT comprehension scores, the distributions for All Subjects, Beginners, and Non-Beginners were profoundly negatively skewed. (See Chapter IV, Figures 1, 3, and 5.) In contrast, the

distributions of the recall protocol scores for the three subject groups were nearly normal. (See Chapter IV, Figures 2, 4, and 6.) The negatively skewed distributions of NDRT comprehension scores exhibit minimal variability, pointing to limitations in the test's discriminatory power. In contrast, the near-normal distributions of recall protocol scores display a much greater degree of variability, denoting a greater capability to discriminate. In other words, integrative tests clearly provide a more sensitive measure of reading comprehension than do discrete-point tests.

Furthermore, not only has the integrative measure proved to be a more sensitive testing instrument than the discrete-point test, but its sensitivity increases along with reading proficiency. In the present study, the Non-Beginners' scores on the comprehension portion of the NDRT were less variable than those of the Beginners, while the Non-Beginners' recall protocol scores were more variable. In other words, the less proficient readers' NDRT comprehension scores were more variable than were those of the more proficient readers. In contrast, the recall protocol scores of the more proficient readers were more variable than were those of the less proficient readers. Thus, as reading skill increases, the greater

sensitivity and enhanced discriminatory power of integrative tests come to the fore. Integrative testing methods, such as the recall protocol procedure, need to gain wider acceptance, particularly for high-proficiency readers.

Correlations.

Correlations between all of the literacy variables were calculated for All Subjects, for Beginners, and for Non-Beginners. (See Chapter IV, Tables 6, 7, and 8.) In the following summary, correlations will be referred to as high (above approximately .60), medium (from approximately .40 to .60), or low (below .40) according to the guidelines set forth by Kamil, Langer, and Shanahan (1985). For All Subjects, significant correlations emanated between all of the variables. Evidently, literacy abilities are interrelated regardless of language, pointing not only to the existence of a common underlying proficiency (Cummins, 1981), but to the presence of common underlying proficiencies.

High correlations appeared between the scores for L2 reading comprehension and those for L2 vocabulary knowledge and L2 grammatical ability, and between the scores for L2 vocabulary knowledge and L2 grammatical ability. A medium correlation emerged between the scores for L1 reading comprehension as measured by recall

protocol scores and those for L2 reading comprehension (See Figure 7). Low correlations ensued between the scores for L1 reading comprehension as assessed by the comprehension portion of the NDRT and those for L2 reading comprehension (See Figure 8), L1 writing ability (See Figure 9), L2 vocabulary knowledge, and L2 grammatical ability, and between the scores for L1 reading comprehension as assessed by recall protocols and those for L1 writing ability (See Figure 10), L2 vocabulary knowledge, and L2 grammatical ability, and between the scores for L1 writing ability and L2 reading comprehension (See Figure 11).

For Beginners, 8 of the 14 correlations were significant. A significant medium correlation emanated between L2 vocabulary knowledge and L2 grammatical ability. Significant low correlations evolved between the scores for L2 reading comprehension and those for L1 reading skill as assessed by both the comprehension portion of the NDRT and recall protocols (See Figures 12 and 13), L1 writing ability (See Figure 14), L2 vocabulary knowledge, and L2 grammatical ability, and between the scores for L1 writing ability and those for L1 reading comprehension as assessed by both the comprehension portion of the NDRT and recall protocols (See Figures 15 and 16). Insignificant correlations

appeared between the scores for L1 reading comprehension as assessed by both the comprehension portion of the NDRT and recall protocols, and those for L2 vocabulary knowledge and L2 grammatical skill, and between the scores for L1 writing ability and L2 vocabulary knowledge and L2 grammatical skill.

For Non-Beginners, only 2 of the 14 correlations achieved significance. Two medium significant correlations arose between the scores for L2 reading comprehension and those for L1 reading skill as assessed by recall protocols (See Figure 17), and between the scores for L2 vocabulary knowledge and L2 grammatical ability. The remaining correlations were insignificant.

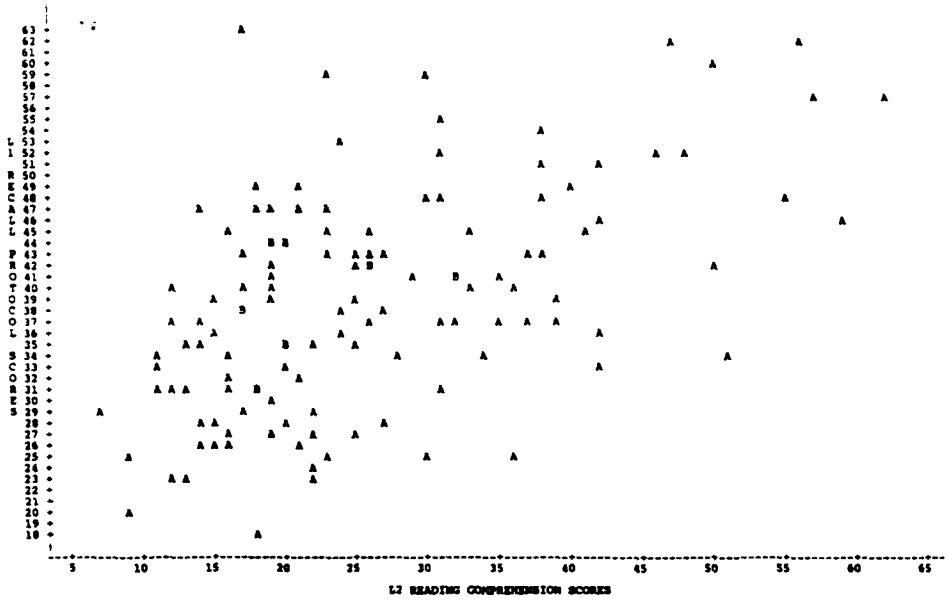


Figure 7. Scatterplot of the Correlation Between the Scores for L1 Recall Protocols and L2 Reading Comprehension for All Subjects

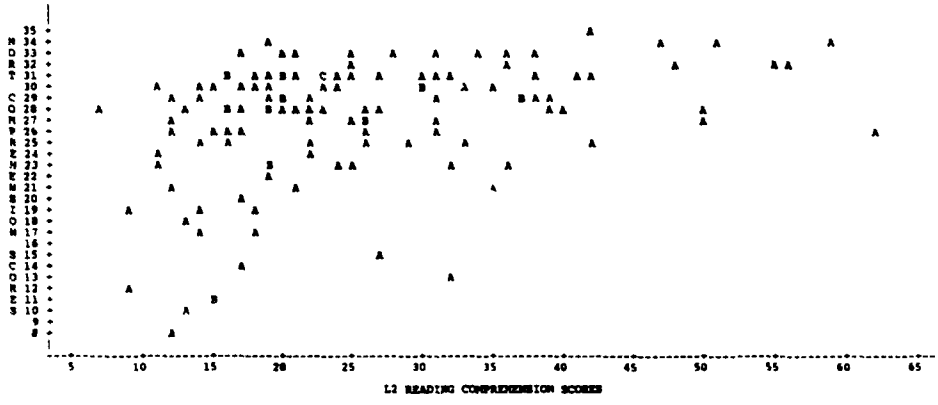


Figure 8. Scatterplot of the Correlation Between NDRT Comprehension Scores and L2 Reading Comprehension Scores for All Subjects

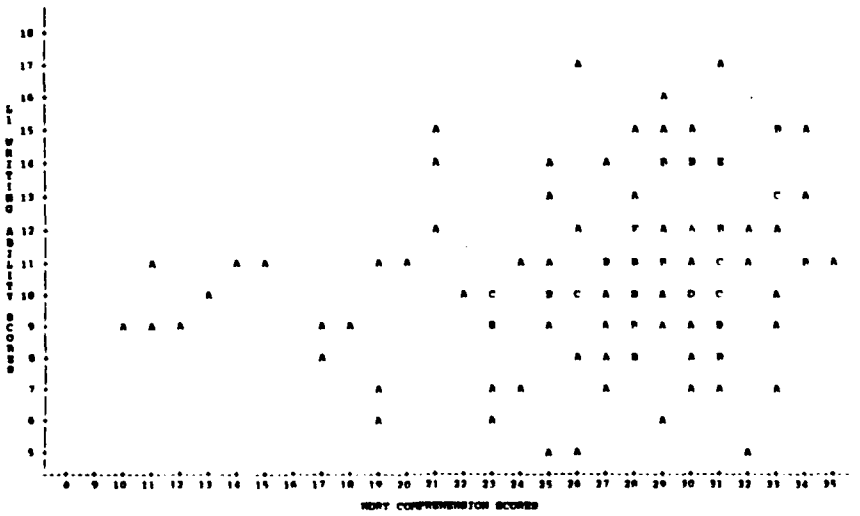


Figure 9. Scatterplot of the Correlation Between L1 Writing Ability Scores and NDRT Comprehension Scores for All Subjects

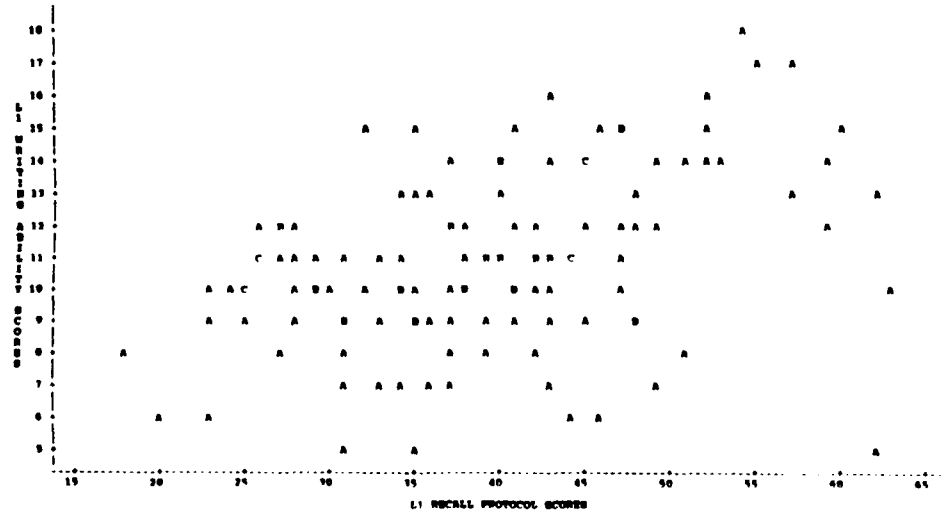


Figure 10. Scatterplot of the Correlation Between L1 Writing Ability Scores and L1 Recall Protocol Scores for All Subjects

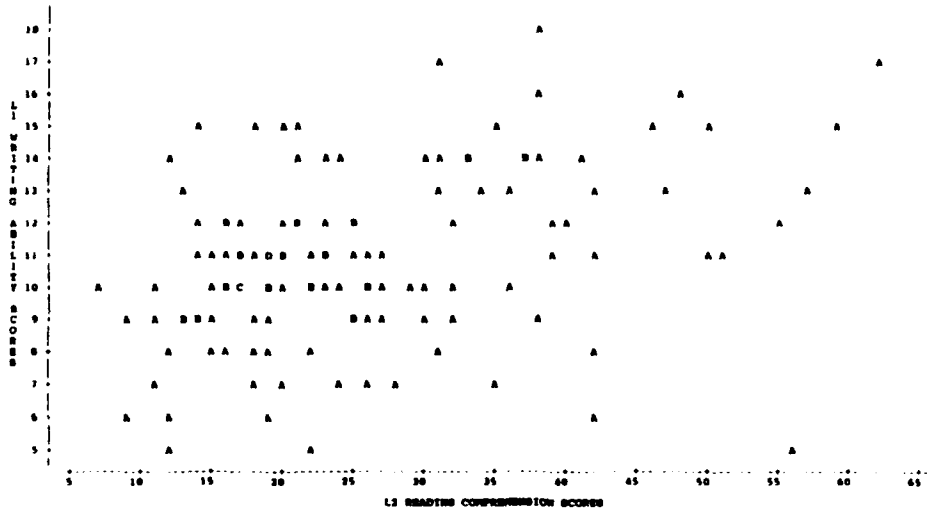


Figure 11. Scatterplot of the Correlation Between the Scores for L1 Writing Ability and L2 Reading Comprehension for All Subjects

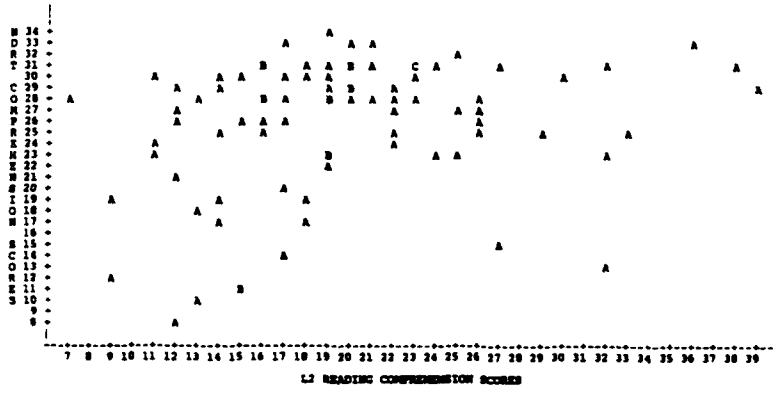


Figure 12. Scatterplot of the Correlation Between NDRT Comprehension Scores and L2 Reading Comprehension Scores for Beginners

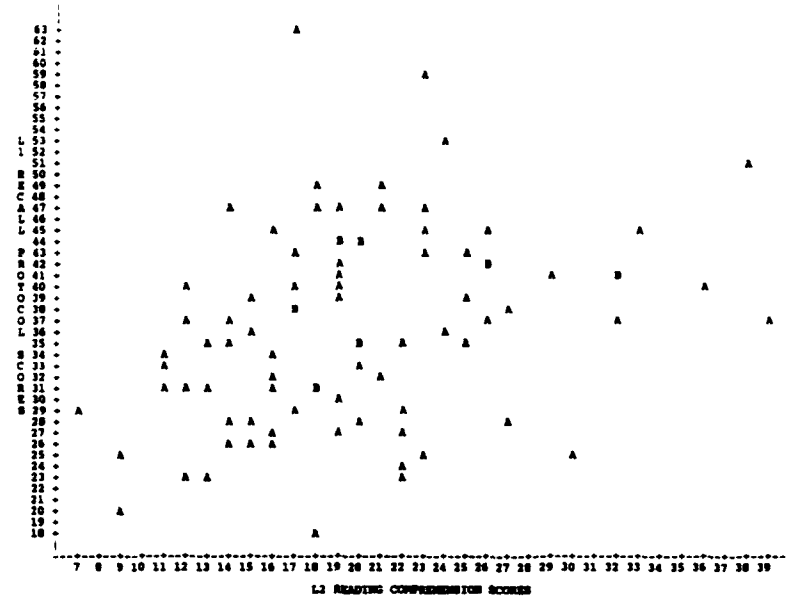


Figure 13. Scatterplot of the Correlation Between L1 Recall Protocol Scores and L2 Reading Comprehension Scores for Beginners

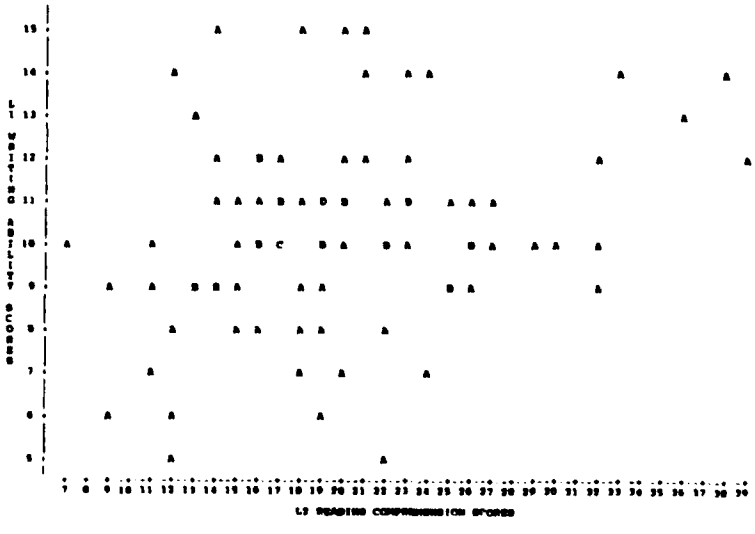


Figure 14. Scatterplot of the Correlation Between the Scores for L1 Writing Ability and L2 Reading Comprehension for Beginners

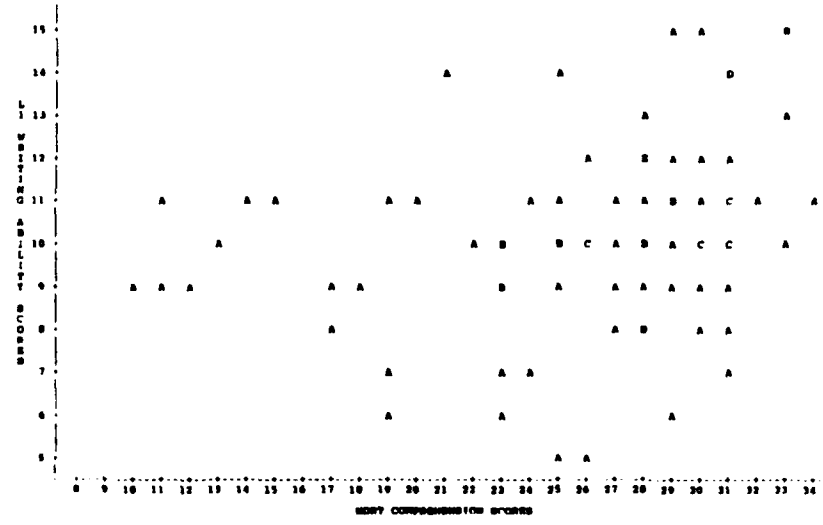


Figure 15. Scatterplot of the Correlation Between L1 Writing Ability Scores and NDRT Comprehension Scores for Beginners

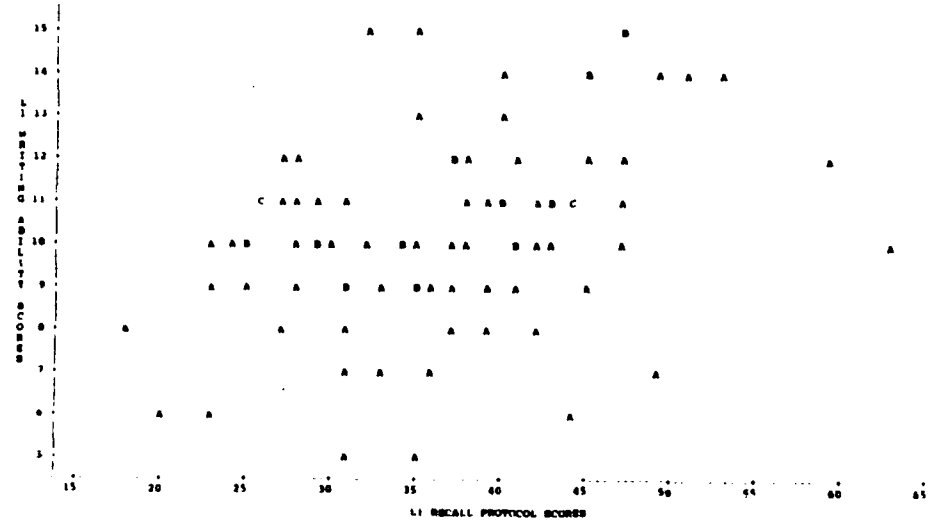


Figure 16. Scatterplot of the Correlation Between L1 Writing Ability Scores and L1 Recall Protocol Scores for Beginners

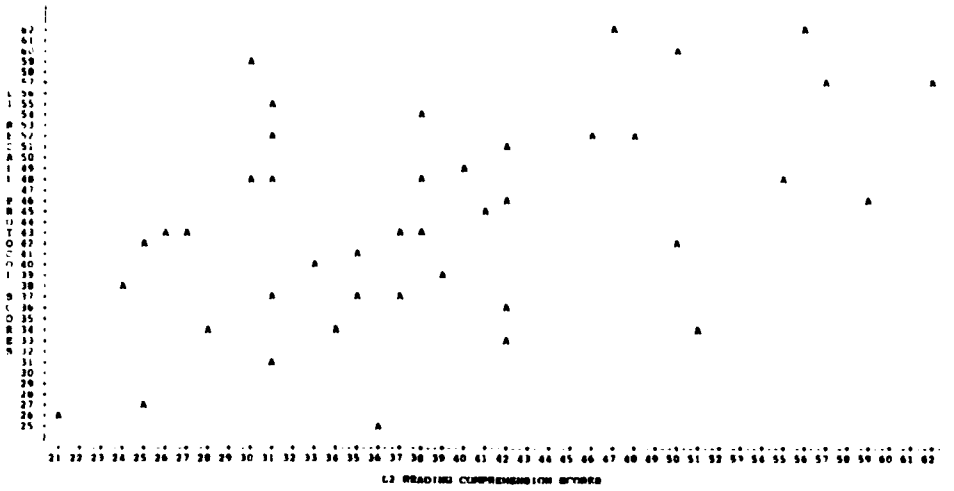


Figure 17. Scatterplot of the Correlation Between the Scores for L1 Recall Protocols and L2 Reading Comprehension for Non-Beginners

Because the theoretical bases underpinning the present study focus on Cummins's Interdependence Hypothesis and on the reading-writing relationship, the following correlations are of particular significance: (a) L1 and L2 reading comprehension, (b) L1 reading skill and L1 writing ability, and (c) L1 writing ability and L2 reading comprehension. Interestingly, the correlations between the scores for L1 and L2 reading comprehension differed depending on the testing instrument used to measure L1 reading ability.

For All Subjects, a medium correlation ensued when recall protocol scores were used (.52), but a low correlation resulted when the NDRT comprehension scores were employed (.38). The explanation for this difference

centers on the lack of variability in the NDRT comprehension scores. As evidenced by the acute negative skewness in the distribution of NDRT comprehension scores, a "floor effect" occurred. Unlike a ceiling effect, in which scores reach but rarely exceed a certain plateau, a floor effect takes place when most of the scores are overwhelmingly high and very few are low. Either situation results in reduced variability. The near-normal distribution of recall protocol scores, in contrast to that of the NDRT scores, demonstrates much more variability. Thus, the recall scores were able to correlate to an increased extent with the scores for L2 reading comprehension. Because the recall protocol scores were not hampered by a floor effect, their use resulted in a more accurate view of the relationship between L1 and L2 reading ability. The dissimilarity in the results points to the necessity of multiple measures in both research and pedagogy.

Additionally, the use of different L1 reading tests resulted in some dissimilarities in the correlations between the scores for L1 and L2 reading comprehension for Beginners versus Non-Beginners. For Beginners, the correlations between the scores for L2 reading comprehension and L1 reading skill as measured NDRT comprehension scores and by recall protocol scores were

both significant and similar in magnitude (.27 and .33, respectively). In contrast, the correlation for Non-Beginners was low and insignificant when NDRT comprehension scores were used (.21), but medium and significant when recall protocol scores were employed (.49). The explanation for the appearance of lower correlations for both Beginners and Non-Beginners when NDRT comprehension scores were used again focuses on the floor effect resulting from the negatively skewed distribution of scores. The restricted variability limited the extent to which the L1 reading scores could correlate with the L2 reading comprehension scores. Thus, particularly for the more proficient readers, the scores obtained through integrative testing offered a clearer picture of the relationship between L1 and L2 reading comprehension. As reading skill increases, so does the need for more discriminating testing instruments.

The fact that the Beginners achieved a lower correlation between the scores for L1 and L2 reading comprehension than did the Non-Beginners when recall protocol scores were used points to a ceiling effect on the Beginners' L2 proficiency. Because the Beginners were less proficient in the L2 than were the

Non-Beginners, their L2 comprehension scores remained relatively low. In contrast, the Non-Beginners, who had a larger store of L2 knowledge at their disposal, attained L2 reading scores that were able to correlate to a greater degree with their L1 recall protocol scores. Evidently, the type of reading assessment measure used made much less difference in the correlations for the less proficient readers than it did in those for the more capable readers. As reading skill develops, more discriminating testing instruments, such as the recall protocol, become increasingly valuable.

The correlations discovered in the present study between the scores for L1 and L2 reading comprehension are consistent with the results of numerous correlational studies examining reading ability transfer. In a study of adult Chinese and Japanese English as a second language (ESL) learners of mixed L2 ability levels, Carson et al. (1990) found a .37 correlation between the scores for L1 and L2 reading comprehension for the Chinese subjects and a .51 correlation for the Japanese subjects. The correlation for the Chinese subjects is almost identical to that of the present study for the mixed L2 ability group (All Subjects) using the NDRT (.38), and the correlation for the Japanese subjects is

almost identical to that of the present study for All Subjects using recall protocols (.52).

Using the recall protocol procedure, Reyes (1987) discovered a correlation of .62 between the L1 and L2 reading comprehension scores of fourth-grade native Spanish speakers at varying English proficiency levels enrolled in an ESL program. Reyes's data are consistent with the correlation in the present study for All Subjects using recall protocols (.52). The difference is greater, however, between Reyes's correlation and the one in the present study for All Subjects using the scores for the comprehension portion of the NDRT (.38). This greater difference probably came about because of the floor effect reflected in the distribution of NDRT comprehension scores.

Using standardized multiple choice tests of L1 Hebrew and L2 English reading comprehension, Groebel (1980), discovered a correlation of .32 between the scores of native Hebrew speakers whose "language proficiency varied considerably" (p. 56). This correlation is similar to the one found in the present study for All Subjects when reading was assessed using the NDRT (.33), but much lower than the one resulting from the use of the recall protocol scores (.52). Perhaps the native-language reading scores in Groebel's

Like the correlations between the scores for L1 and L2 reading comprehension, those between the scores for L1 reading ability and L1 writing skill differed depending upon the instrument used to test L1 reading ability. For All Subjects, for example, the correlation was .25 when using NDRT comprehension scores and .39 when using recall protocol scores. Additionally, the correlations exhibited more consistency with the results of previous studies when recall protocol scores were used than they did when NDRT comprehension scores were employed.

Juel, Griffith, and Gough (1986) discovered a .39 correlation between second graders' scores for reading comprehension and writing ability. Reading skill was assessed using the reading comprehension subtest of the IOWA Test, and writing ability was measured using holistic scores on a writing sample. The correlation found by Juel et al. (1986) is identical to the one in the present study for All Subjects using recall protocol scores (.39), but higher than the one that ensued when NDRT comprehension scores were employed (.25).

In a study of fourth graders, Schewe and Froese (1986) found a .45 correlation between the scores for reading comprehension as assessed by recall protocols and those for writing ability as measured using an analytic scoring system. Like the .39 correlation garnered by

Juel et al. (1986), Schewe and Froese's correlation is consistent with the one found in the present study obtained when using recall protocol scores (.39), but greater than the one attained using NDRT comprehension scores (.25).

Langer (1984) discovered a .43 correlation between 10th graders' scores for reading achievement and writing ability. Reading achievement was assessed using the California Test of Basic Skills, and writing ability was measured using holistic scores on a writing sample. The correlation discovered by Langer, like those found by Juel et al. (1986) and by Schewe and Froese (1986), is consistent with the findings in the present study using recall protocol scores (.39), but greater than the results obtained employing NDRT comprehension scores (.25).

In one of the few explorations of the reading-writing relationship featuring adult subjects, Carson et al. (1990) correlated the scores for native-language reading skill and writing ability for adult Chinese and Japanese ESL learners at mixed native-language ability levels. Reading comprehension was assessed using the cloze procedure, and writing ability was measured using holistic scores on a single writing sample. The correlations were .27 for the

Chinese subjects and .49 for the Japanese subjects. The correlation for the Chinese subjects is somewhat lower than those discovered in the studies of children, but consistent with the one found in the present study when NDRT comprehension scores were used (.25). Although the correlation for the Japanese subjects was slightly higher, the data are still reasonably consistent with those that emanated in the studies featuring children and in the present study when recall protocol scores were used (.39). Perhaps the consistency would have been greater, however, had Carson et al. (1990) used a testing instrument other than cloze, or the combined score on multiple writing samples.

The present study is the first to explore the relationship between the scores for L1 writing ability and L2 reading comprehension. If, as previous research has indicated, literacy abilities transfer between languages, and overlapping literacy abilities are involved in the processes of reading and writing, it is reasonable to assume that L1 writing ability transfers to L2 reading skill. The data in the present study point to a relationship between the two variables. Significant correlations emanated between the scores for L1 writing ability and L2 reading comprehension for All Subjects (.37) and for Beginners (.33), though the correlation for

Non-Beginners (.20) was positive but insignificant. Evidence of a relationship between L1 writing ability and L2 reading comprehension takes Cummins's Interdependence Hypothesis (1981) a step further by melding literacy transfer theory with reading-writing relationship theory. Perhaps what actually transfer between languages are the literacy abilities that underlie both reading skill and writing capacity.

Inferential Data: Hierarchical Regression Analyses

Overview.

In this section, the data generated from the hierarchical regression models will be summarized and discussed as follows: (a) for All Subjects, (b) for Beginners and Non-Beginners using the scores from the comprehension portion of the NDRT, and (c) for Beginners and Non-Beginners using recall protocol scores.

All subjects.

Some provocative findings emerged from the results of the hierarchical regression programs performed on the data for All Subjects. In summary, regardless of the instrument used to assess L1 reading comprehension, all of the independent variables contributed significantly to the scores for L2 reading comprehension, and in the same order of importance. Using NDRT comprehension scores, the independent variables contributed as follows: (a) L2

vocabulary knowledge (35.68%), (b) L1 reading comprehension (14.12%), (c) L1 writing ability (7.03%), and (d) L2 grammatical skill (3.04%). (See Figure 18.) When recall protocol scores were used, the contributions were: (a) L2 vocabulary knowledge (30.91%), (b) L1 reading comprehension (27.37%), (c) L1 writing ability (3.84%), and (d) L2 grammatical skill (3.19%). (See Figure 19.)

The finding that the scores for L2 vocabulary knowledge contributed more to the dependent variable than did the scores for any of the other independent variables, particularly grammatical ability, is consistent with the results of the stepwise regression studies performed by Henning (1976) and Koda (1990). In Henning's (1976) study of female Iranian subjects enrolled in English as a foreign language courses, grammaticality of response ranked far below the vocabulary variables, contributing less than four percent of the variance to that of the reading comprehension scores.

Koda (1990) administered L2 vocabulary, L2 grammar, and multiple choice and cloze L2 reading comprehension tests to college-level subjects enrolled in Japanese as a foreign language courses. While vocabulary skills contributed predominantly to the multiple choice

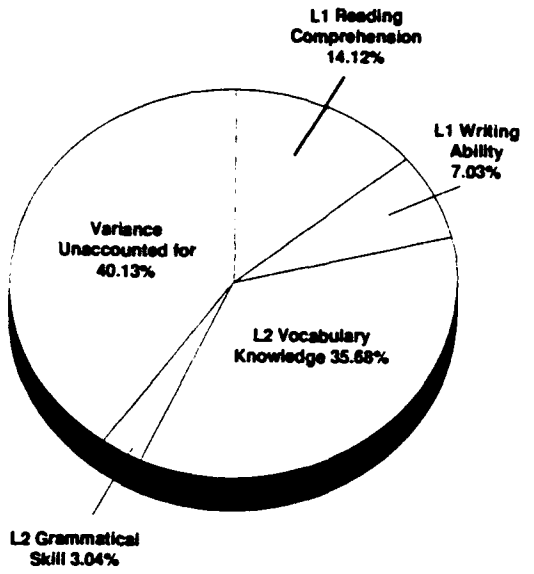


Figure 18. Chart of the Results of the Hierarchical Regression Program for All Subjects using NDRT Comprehension Scores

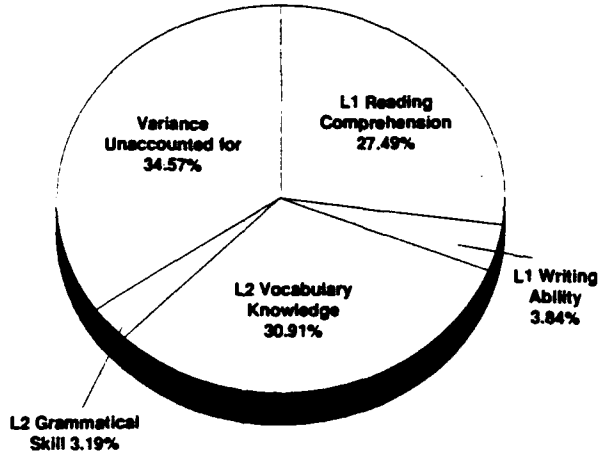


Figure 19. Chart of the Results of the Hierarchical Regression Program for All Subjects using L1 Recall Protocol Scores

reading test scores, grammatical ability did so for the cloze test scores. Koda's (1990) results using a multiple choice reading measure are consistent with the results of the present study, while her data using the cloze test are not. It is not surprising, however, that grammatical ability contributed more to cloze test scores because this test genre examines the ability to insert a syntactically appropriate word in a clause, rather the capacity to glean meaning from text (Bernhardt, 1991).

Although L2 vocabulary knowledge was the primary contributor, the L1 variables offered a substantial cumulative contribution in both models (21.15% using NDRT comprehension scores and 31.33% using L1 recall protocol scores). Clearly, L1 literacy skills, particularly L1 reading ability, contribute significantly to L2 reading comprehension. This finding not only bolsters Cummins's theory of literacy transfer, but lends support to the results of the correlational studies pointing to the transfer phenomenon between L1 and L2 reading skills (Carson et al., 1990; Groebel, 1980; Reyes, 1987; Wagner et al., 1989) and to those indicating that reading ability and writing skill transfer to each other (Carson et al., 1990; Juel et al., 1986; Langer, 1984; Schewe & Froese, 1986).

The results for All Subjects also indicate that some of the elements featured in Bernhardt's (1986) Constructivist Model contribute more to the reading comprehension process than do others. According to the data, L2 vocabulary knowledge offers a greater contribution to L2 reading comprehension than does L2 grammatical skill. Thus, in L2 programs focusing on reading, vocabulary knowledge should be emphasized and grammatical skill de-emphasized. Additionally, prior knowledge in the form of L1 literacy abilities contributes a great deal to L2 reading comprehension. Evidently, language students depend upon their L1 skills in order to make sense of the L2.

The most intriguing difference between the results of the two hierarchical regression programs performed for All Subjects is the dissimilarity in the amount of variance contributed by the L1 reading scores when assessed by the comprehension portion of the NDRT versus recall protocols. L1 reading comprehension as assessed by recall protocols contributed nearly twice the variance (27.49%) as it did when measured by NDRT comprehension scores (14.12%). This dissimilarity is once again a function of the lack of variability in the NDRT comprehension scores. The more normally distributed recall protocol scores had more variance to contribute to

the dependent variable. The use of the recall protocol resulted in a clearer view of the contribution of L1 reading skill to L2 reading comprehension. The fact that the overall picture of reading comprehension changes depending on the testing instrument used to measure it underscores the need for multiple measures in research and in the classroom.

Additionally, a greater percentage of the overall variance was accounted for in the regression model using recall protocol scores (65.43%) versus the one featuring NDRT comprehension scores (59.87%). This difference again relates to the greater variance in the recall protocol scores versus that of the NDRT comprehension scores. The recall protocol scores had more variance to offer, thus contributing more to the overall variance. The more sensitive the assessment measure is, the more accurate the information will be about the components of the L2 reading comprehension process.

Beginners versus Non-Beginners using NDRT comprehension scores.

The results of the hierarchical regression models featuring the scores for L1 reading skill as measured by NDRT comprehension scores point to some interesting similarities and differences between Beginners and

Non-Beginners. For Beginners, L2 vocabulary knowledge and L1 reading comprehension provided significant contributions, while none of the variables contributed significantly for Non-Beginners. The order of importance of the independent variables for Beginners was: (a) L2 vocabulary knowledge (8.50%), (b) L1 reading comprehension (7.06%), (c) L1 writing ability (3.47%), and L2 grammatical skill (1.24%). (See Figure 20.) For Non-Beginners, the independent variables contributed as follows: (a) L2 vocabulary knowledge (8.67%), (b) L1 writing ability (5.85%), (c) L1 reading comprehension (4.42%), and (d) L2 grammatical skill (1.49%). (See Figure 21.)

Consistent with the results for All Subjects, L2 vocabulary contributed more to the L2 reading scores than did any of the other independent variables, especially L2 grammatical skill. This finding coincides with the results of the regression studies performed by Henning (1976) and Koda (1990). Clearly, vocabulary knowledge is more important to the reading comprehension process than is grammatical skill. In language classes centering on reading, then, grammar study should be de-emphasized and vocabulary learning stressed. Unlike the results for All Subjects, however, the cumulative contribution of the L1 variables was greater than that of the L2 variables.

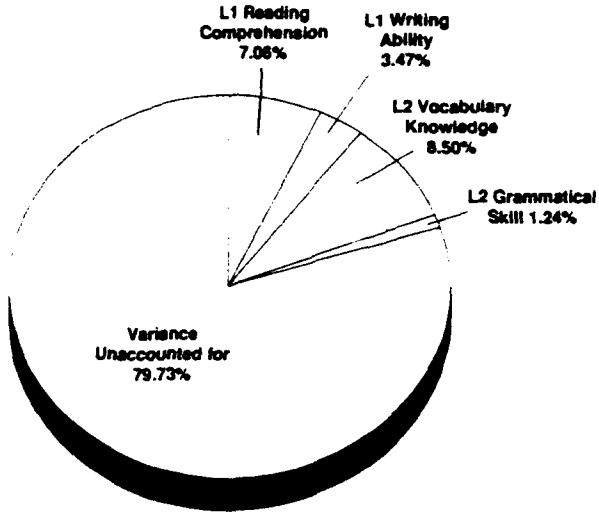


Figure 20. Chart of the Results of the Hierarchical Regression Program for Beginners using NDRT Comprehension Scores

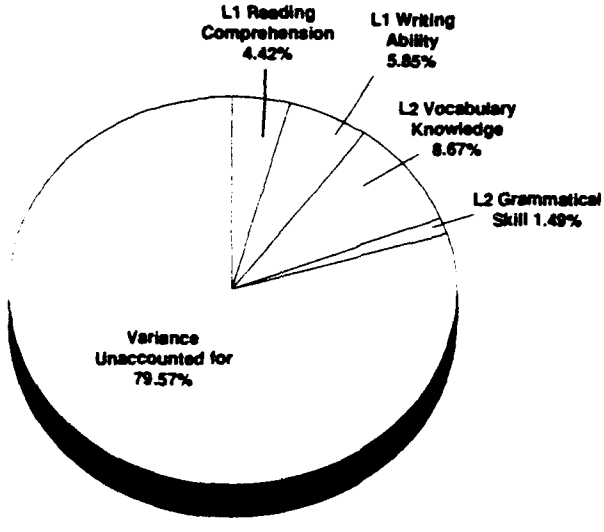


Figure 21. Chart of the Results of the Hierarchical Regression Program for Non-Beginners using NDRT Comprehension Scores

This finding reinforces Cummins's (1981) Interdependence Hypothesis.

The data for Beginners versus Non-Beginners aid in expanding Bernhardt's (1986) Constructivist Reading Model by exploring the contributions of several of the components of the model by course level. According to the results of the present study, vocabulary knowledge is a crucial element in the reading comprehension process, slightly more so for Non-Beginners (8.67%) than for Beginners (8.50%). In fact, the cumulative contribution of the L2 variables was greater for Non-Beginners than for Beginners, while the opposite was true for that of the L1 variables. Evidently, language students rely on prior knowledge in the form of L1 skills until their L2 abilities develop and thus become increasingly useful in the comprehension process.

The cumulative contribution of the L1 variables was greater than that of the L2 variables for both groups. In light of this finding, L2 instructors and curriculum and textbook developers should capitalize on language learners' already developed L1 abilities. Language transfer needs to be encouraged, especially at the beginning levels of L2 study. As L2 proficiency increases, however, its importance in the L2 reading comprehension process manifests itself. Instructors and

curriculum and textbook developers should consider the fact that upper-level learners become increasingly able to use the L2 as a source of prior knowledge aiding them in L2 study.

The finding that the L1 variables contribute less as L2 proficiency increases is inconsistent with the results of Wagner et al. (1989), whose hierarchical regression data indicate that L1 literacy abilities co ~ibute more to L2 literacy skills as L2 skills increase. The subjects in the Wagner et al. (1989) study, however, were children whose L1 and L2 abilities were still developing. Melding the results of Wagner et al. (1989) and of the present study, it becomes evident that after L1 abilities have stabilized, L2 abilities are able to come to the fore.

Additionally, both hierarchical regression models resulted in nearly equal portions of the total variance unaccounted for: 79.73% for Beginners and 79.57% for Non-Beginners. Evidently, L2 readers draw on components of the reading process that were not included in the present set of independent variables. These additional variables may include affective factors, as well as elements featured in Bernhardt's (1986) Constructivist Reading Model such as phonemic/graphemic features, metacognition, intratextual perception, prior knowledge

beyond the L1, or the recognition of words or syntactic features that were not used in the present study.

Beginners versus Non-Beginners using recall protocol scores.

The results of the hierarchical regression models using recall protocol scores to measure L1 reading comprehension also bring out several similarities and differences between Beginners and Non-Beginners. L1 reading comprehension and L2 vocabulary knowledge were significant contributors for Beginners, while for Non-Beginners, only L1 reading comprehension contributed significantly. For Beginners, the independent variables contributed as follows: (a) L1 reading comprehension (10.58%), (b) L2 vocabulary knowledge (8.38%), (c) L1 writing ability (3.00%), and (d) L2 grammatical ability (1.85%). (See Figure 22.) The order of importance of the variables for Non-Beginners was: (a) L1 reading comprehension (23.96%), (b) L2 vocabulary knowledge (5.40%), (c) L2 grammatical skill (2.69%), and (d) L1 writing ability (.43%). (See Figure 23.)

Unlike the regression models for Beginners and Non-Beginners using NDRT comprehension scores, L1 reading comprehension rather than L2 vocabulary knowledge was the major contributor for both groups. This difference results from the increased variance in the recall

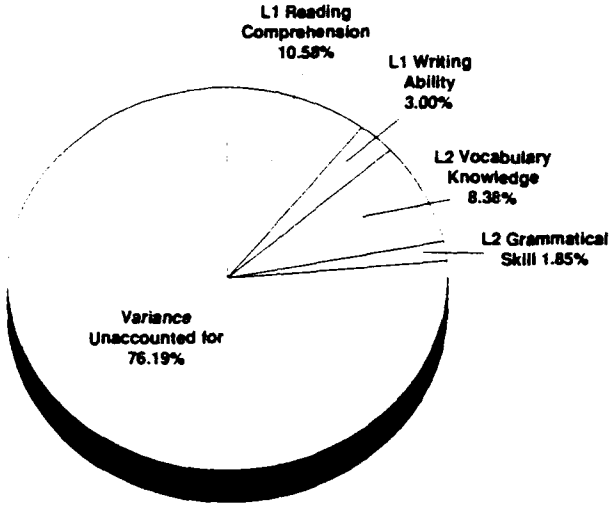


Figure 22. Chart of the Results of the Hierarchical Regression Program for Beginners using L1 Recall Protocol Scores

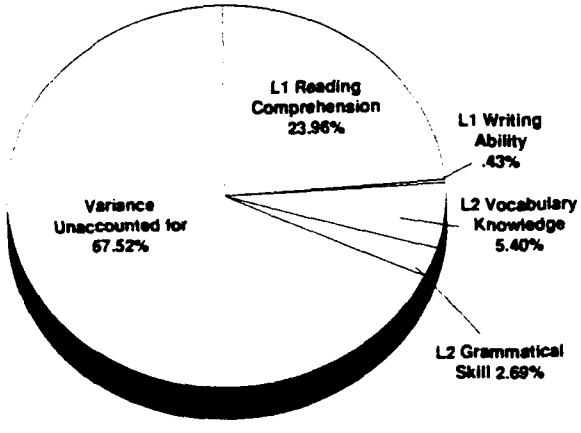


Figure 23. Chart of the Results of the Hierarchical Regression Program for Non-Beginners using L1 Recall Protocol Scores

protocol scores, particularly for Non-Beginners. The NDRT comprehension scores had less variance to contribute to the scores for L2 reading comprehension.

Relatedly, the most prominent difference between the results of the regression models using recall protocol scores for Beginners versus Non-Beginners lies in the magnitudes of the contributions of L1 reading comprehension. L1 reading ability contributed more than twice the variance for Non-Beginners (23.96%) as it did for Beginners (10.58%). This dissimilarity ensued because the Non-Beginners' recall protocol scores were even more variable than were those of the Beginners.

Consistent with the results of the programs for Beginners and Non-Beginners using NDRT comprehension scores, the regression models using recall protocol scores show a higher cumulative contribution for the L1 variables than for the L2 variables. Once again, this finding serves to support Cummins's (1981) Interdependence Hypothesis. Furthermore, and coinciding with the results of the other regression models used in this study, L2 vocabulary contributed more to the variance of the L2 reading scores than did L2 grammatical ability, as it did in the stepwise regression analyses performed by Henning (1976) and Koda (1990).

The regression models using recall protocol scores show 76.19% of the overall variance unaccounted for in the data for Beginners and 67.52% unaccounted for for Non-Beginners. Although for both groups less of the overall variance remains unaccounted for than is unaccounted for in the programs using NDRT comprehension scores, it is once again evident that L2 readers move beyond the present set of independent variables by drawing on components of the reading process that were not included.

Implications for Assessment and Pedagogy

The results of the present study provide significant insights into the realms of assessment and pedagogy. In the area of assessment, this study has opened a window on the dilemma regarding whether or not reading comprehension tests measure what they are supposed to measure by determining the specific contributions of selected L1 and L2 literacy variables to L2 reading comprehension test scores. Although the data resulting from the regression models for Beginners and Non-Beginners indicate that the effect of L1 writing ability on L2 reading test scores is insignificant, the results of the programs for All Subjects show it to be small but significant. Thus, although the contribution of L1 writing ability to the scores on L2 reading tests

appears to be slight, it should be taken into consideration when interpreting such test scores. More information is needed regarding the precise nature of the relationship between L1 writing ability and L2 reading comprehension.

The data in the present study will lead to enhanced test development and to improvements in diagnostic and placement testing. In terms of test development, the knowledge that L1 writing ability is a confounding factor, albeit a minor one, points to the need for the exploration of alternative testing methods that will decrease its effects. Additionally, because different results emanate from different testing instruments, the use of multiple measures is necessary in order to arrive at an accurate picture of what has been comprehended. For diagnostic testing, perhaps L1 literacy abilities need to be tested in order to gain insight into learner difficulties in L2 studies. Regarding placement testing, L1 literacy abilities, especially reading comprehension, have been identified as important predictors of L2 literacy ability. L1 literacy skills, therefore, should be considered when placing students into language course levels. Students who are highly literate in the L1 should perhaps be placed in accelerated L2 courses.

In the pedagogical domain, this study attests to the criticalness of vocabulary knowledge to L2 reading comprehension. Instructors and curriculum developers should continue to stress vocabulary skill as a crucial element in L2 study.

Additionally, the data underline the importance of capitalizing on students' already developed L1 knowledge. Instructors should draw attention to parallels and contrasts between the students' L1 and the L2, and encourage learners to transfer their skills between languages. Furthermore, educators must familiarize themselves with learners' native languages and develop an awareness of what the students know about their native tongues so that this knowledge can be used effectively in L2 study.

In programs in which the learners are already literate in their native languages, the importance of L1 literacy abilities to the development of L2 literacy should be increasingly reflected in L2 curriculum development. In such programs, beginning-level courses should aid students in going from the known to the unknown by emphasizing the use of L1 knowledge in L2 study. Follow-on courses should guide learners into increased use of their developing L2 knowledge. Furthermore, students experiencing difficulties in the L2

may benefit from further L1 instruction. If this were accomplished, perhaps more L2 learners would continue in L2 study.

Recommendations for Research

Further research is needed that will validate and expand upon the results of the present study. Replications of this research should include subjects of diverse ages (such as children or high school students), of language backgrounds other than L1 English and L2 French, or who are enrolled at a wider variety of L2 course levels.

Additionally, alternative testing methods should be used to measure the variables. Because gross measures of reading and writing were used in the present study, replications are needed in which finer assessment measures will be employed. The constructs of reading and writing, for example, could be divided into subparts and then examined in order to achieve a clearer view of exactly which facets of each construct contribute to the other. Furthermore, reading comprehension should be tested using passages featuring different rhetorical organizations, and using other modes of discourse.

Different data analysis techniques could also be used. The independent variables in the present study, for example, should be entered into programs of

hierarchical multiple regression in varying orders. Stepwise multiple regression programs could be employed. Qualitative methods could be used in order to further explore the construct of reading comprehension, subjecting the recall protocols to qualitative analyses.

Furthermore, because the variance contributing to that of the L2 reading scores is still not completely accounted for, additional independent variables should be considered. These variables should include affective factors and the elements of the reading comprehension process featured in Bernhardt's (1986) Constructivist Model such as phonemic/graphemic features, metacognition, intratextual perception, other sorts of prior knowledge besides previous L1 experience, and the recognition of words and syntactic features other than those included in the present study.

Looking beyond the findings specific to the present study, research needs to be done to profile the individuals who are most likely to self-select to continue their L2 study. To date, little has been done either to identify learners who are likely to continue, or to pinpoint their reasons for doing so.

In addition, integrative testing methods, such as the recall protocol procedure, need to gain wider acceptance as a measure of reading comprehension. Not

acceptance as a measure of reading comprehension. Not only is the recall protocol more sensitive than discrete-point tests, but its sensitivity becomes more pronounced as reading proficiency increases. In order for this procedure to attain wider use, however, the scoring process needs to be rendered less time consuming. Research into the automatization of this process would open the way to increased use of this testing method.

Furthermore, the results of this study underscore the need for multiple measures in research. The use of more than one testing instrument will lead to a clearer view of the constructs being measured, and hence to more astute research insights and to more perceptive theory-building.

Limitations

Several limitations ensued in the present study. First, the subject group used was fairly homogeneous in nature, as demonstrated by the negative skewness of the NDRT comprehension scores. The results would be more generalizable if subjects with more variable literacy skills had participated. Second, the number of subjects in the Non-Beginner group (n = 43) was not entirely adequate for the use of multiple regression. According to the guidelines set forth by Kamil, Langer, and Shanahan (1985), a minimum of 20 subjects is needed per

ensuing from the Non-Beginners must be considered carefully before it is used as the basis for generalizations. Third, the subjects were aware that the data resulting from their L1 and L2 protocols and from their L2 vocabulary knowledge tests were for research purposes. This knowledge may have had some effect on the quality of their responses. Fourth, the subjects' lack of familiarity with the recall protocol procedure might have affected the results.

APPENDIX A

ENGLISH READING PASSAGES

Louis was born April 25, 1214 at Poissy, in the modern *département* of Forêt des Yvelines. He was the fourth child of King Louis VIII and Queen Blanche of Castile, but since the first three died at an early age, Louis, who was to have seven more brothers and sisters, became heir to the throne. He was raised with particular care by his parents, especially his mother. Experienced horsemen taught him riding and the fine points of hunting. Tutors taught him biblical history, geography, and ancient literature. His mother instructed him in religion herself and educated him as a sincere, unbigoted Christian. Louis was a boisterous adolescent occasionally seized by fits of temper, which he made efforts to control. When his father, who succeeded Philip II Augustus, ascended the throne in 1223, the long struggle between the Capetian dynasty and the Plantagenets of England (who still had vast holdings in France) was still not settled, but there was a temporary lull since the English King, Henry III, was in no position to resume the war. In the south of France, the Albigensian heretics, who were in revolt against both church and state, had not been brought under control. Finally, there was ferment and the threat of revolt among the great nobles who had been kept in line by the firm hand of Philip Augustus.

Since the congress of Erfurt, the Russian emperor had shown himself less and less inclined to deal with Napoleon as a trusted partner. In the spring of 1812, therefore, Napoleon massed his forces in Poland to intimidate Alexander. After some last attempts at agreement, in late June his Grande Armée--about 453,000 men, including contingents extorted from Prussia and from Austria, began to cross the Nieman River. The Russians retreated, adopting a "scorched earth" policy. Napoleon's army did not reach the approaches to Moscow till the beginning of September. The Russian commander in chief, Mikhail I. Kutuzov, engaged it at Borodino on September 7. The fight was savage, bloody, and indecisive, but a week later Napoleon entered Moscow, which the Russians had abandoned. On that same day, a huge fire broke out, destroying the greater part of the town. Moreover, Alexander unexpectedly refused to treat with Napoleon. Withdrawal was necessary, and the premature onset of winter made it disastrous. After the difficult crossing of the Berezina River in November, fewer than 10,000 men fit for combat remained with Napoleon's main force.

This catastrophe heartened all the peoples of Europe to defy Napoleon. In Germany the news unleashed an outbreak of anti-French demonstrations. The Prussian

contingents deserted the Grande Armée in December and turned against the French.

De Gaulle's original call for resistance had attracted only a handful of Frenchmen who happened to be in Britain at the time. But as the British continued to fight, a trickle of volunteers from France began to find its way to his headquarters in London. De Gaulle promptly established an enterprise called Free France and in 1941 capped it with a body called the Comité National Français (French National Committee) for which he boldly claimed the status of legal government-in-exile. During the next three years, first in London and then (after 1943) in Algiers, he insisted on his right to speak for France and on France's right to be heard as a great power in the council of the Allies. His demands and his manner irked Churchill and Roosevelt and caused persistent tension. The American government unsuccessfully attempted in 1942 to sidetrack him in favor of Gen. Henri Giraud, who immediately after the Allied landings in North Africa was brought out of France to command the French armies in liberated North Africa and to assume a political role as well. De Gaulle arrived in Algiers in May 1943 and joined Giraud as co-president of a new French Committee of National Liberation. By the end of the year, he had outmanoeuvred Giraud and emerged as the unchallenged spokesman for French resisters everywhere. Even the Communists in 1943 grudgingly accepted his leadership.

APPENDIX B

SCORING TEMPLATES FOR ENGLISH READING PASSAGES

Scoring Template: Louis IX

Proposition	Points
Louis was born	4
April 25, 1214	3
at Poissy	3
in the modern *département*	1
of Forêt des Yvelines	1
He was the fourth child	4
of King Louis VIII	3
and Queen Blanche of Castile	3
but	1
since the first three	2
died at an early age	2
Louis	4
who was to have	1
seven more brothers and sisters	1
became heir	4
to the throne	3
He was raised	4
with particular care	4
by his parents	1
especially his mother	1
Experienced horsemen	3
taught him riding	3
and the fine points	1

of hunting	2
Tutors	3
taught him biblical history	3
geography	2
and ancient literature	2
His mother	2
instructed him	4
in religion herself	2
and educated him	2
as a sincere	1
unbigoted Christian	4
Louis was a boisterous adolescent	3
occasionally seized	1
by fits of temper	2
which he made efforts to control	1
When his father	4
who succeeded Philip II Augustus	1
ascended the throne	4
in 1223	1
the long struggle	4
between the Capetian dynasty	4
and the Plantagenets of England	4
(who still had vast holdings in France)	1
was still not settled	4
but there was a temporary lull	3

since the English king	2
Henry III	1
was in no position	3
to resume the war	3
In the south of France	3
the Albigensian heretics	4
who were in revolt	2
against both church and state	3
had not been brought	2
under control	2
Finally	1
there was ferment	3
and the threat of revolt	4
among the great nobles	3
who had been kept in line	2
by the firm hand	1
of Philip Augustus	2

Scoring Template: Napoleon

Proposition	Points
Since the congress of Erfurt	1
the Russian emperor	4
had shown himself	2
less and less inclined	4
to deal	3
with Napoleon	4
as a trusted partner	2
In the spring of 1812	3
therefore	1
Napoleon massed his forces	4
in Poland	2
to intimidate Alexander	4
After some last attempts at agreement	2
in late June	1
his Grande Armée	3
--about 453,000 men	1
including contingents extorted from Prussia	1
and from Austria	1
began to cross the Niemen River	3
The Russians retreated	4
adopting a "scorched earth" policy	1
Napoleon's army	3
did not reach	3

the approaches to Moscow	4
till the beginning of September	2
The Russian commander in chief	3
Mikhail I. Kutuzov	1
engaged it at Borodino	3
on September 7	1
The fight was savage	2
bloody	1
and indecisive	2
but a week later	1
Napoleon entered Moscow	4
which the Russians had abandoned	2
On that same day	1
a huge fire broke out	4
destroying the greater part of the town	3
Moreover	1
Alexander	3
unexpectedly refused to treat with Napoleon	3
Withdrawal was necessary	4
and the premature onset of winter	2
made it disastrous	4
After the difficult crossing	3
of the Berezina River	2
in November	1
fewer than 10,000 men	3

fit for combat	2
remained with Napoleon's main force	3
This catastrophe	2
heartened all the peoples of Europe	4
to defy Napoleon	4
In Germany	2
the news unleashed an outbreak	2
of anti-French demonstrations	3
The Prussian contingents	2
deserted the Grande Armée	4
in December	1
and turned against the French	4

Scoring Template: De Gaulle

Proposition	Points
De Gaulle's original call	4
for resistance	2
had attracted	2
only a handful of Frenchmen	3
who happened to be in Britain	2
at the time	1
But	1
as the British	2
continued to fight	3
a trickle of volunteers	4
from France	2
began to find its way	1
to his headquarters in London	4
De Gaulle	2
promptly established	4
an enterprise	1
called Free France	4
and in 1941	1
capped it	2
with a body	2
called the Comitè National Français	3
(French National Committee)	1
for which he boldly claimed	2

the status	1
of legal government-in-exile	4
During the next three years	1
first in London	2
and then	1
(after 1943)	1
in Algiers	2
he insisted	4
on his right to speak	4
for France	4
and on France's right	4
to be heard	3
as a great power	3
in the council of the Allies	1
His demands	3
and his manner	1
irked Churchill and Roosevelt	4
and caused persistent tension	3
The American government	3
unsuccessfully attempted	4
in 1942	1
to sidetrack him	3
in favor of Gen. Henri Giraud	3
who	1
immediately after the Allied landings	2

in North Africa	2
was brought out of France	2
to command the French armies	3
in liberated North Africa	2
and to assume a political role as well	2
De Gaulle arrived in Algiers	3
in May 1943	1
and joined Giraud	3
as co-president	1
of a new French Committee of National Liberation	3
By the end of the year	1
he had outmanoeuvred Giraud	4
and emerged	4
as the unchallenged spokesman	4
for French resisters everywhere	4
Even the Communists	3
in 1943	1
grudgingly accepted his leadership	3

APPENDIX C

FRENCH READING PASSAGES

Les gaulois et plus gènèralement les Celtes n'ont jamais été fermés sur eux-mêmes: dés le VIIe siécle av. J.-C., ils entretenaient des rapports commerciaux avec les peuples les plus civilisés du bassin méditerranéen. De tous les barbares, ils étaient sans doute les moins barbares. L'entrée de la Gaule dans le monde méditerranéen et hellénique a précédé de loin la conquête romaine et a préparé la voie á celle-ci. Il est en particulier abondamment établi que l'expansion commerciale et l'influence technique et culturelle de la colonie phocéenne de Marseille débordaient largement le cadre de ce qui allait devenir la province romaine de Narbonnaise. C'est du reste á l'appel des Marseillais, victimes de la piraterie ligure et des incursions de leurs voisins, que les Romains intervinrent en Gaule méridionale. Cette région servit ensuite de base de départ á Jules César pour conquérir l'ensemble du territoire; il ne faisait en cela que lutter de vitesse avec les Germains: si la Gaule n'était devenue romaine, elle fût á coup sûr devenue germanique.

La langue latine supplanta peu á peu les parlers celtiques; sur tout le territoire s'élevérent des villes, avec leurs réseaux d'aqueducs et de voies. Cependant, les Gaulois, en devenant gallo-romains, ne reniérent pas leurs traditions propres. Cette fidélité se manifeste

aussi bien dans le domaine religieux que dans le domaine artistique.

Louis XIV est un des personnages historiques sur lesquels l'attention demeure portée, sans que nul historien puisse prétendre donner de lui une image certaine et définitive. Qu'il ait influencé directement les destinées françaises et qu'á ce titre on ne puisse imaginer l'histoire de la France sans lui, nul doute. Mais, parce que son régne a curieusement associé une incontestable gloire á de trés lourds malheurs pour la nation, il a été extrêmement loué ou critiqué et ses historiens se sont souvent partagés entre apologistes et détracteurs. On doit observer qu'il est beaucoup plus malaisé á comprendre pour des hommes de la fin du XXe siécle que pour ceux du début, parce que les profondes mutations de la société française au cours de cette période ont fait disparaître des aspects de la mentalité collective qui demeuraient, il y a soixante-dix ans encore, relativement proches du XVIIe siécle. En revanche, les renouvellements de la méthode historique, surtout des études érudites sur les conditions de vie en France au temps de Louis XIV, autorisent une meilleure intelligence du pays sur lequel son action s'est exercée.

La période du régne personnel s'étend de 1661 á 1715, soit pendant cinquante-quatre ans, période du gouvernement effectif du souverain. C'est par le travail que l'on régne, disait Louis XIV; il a mis ce principe en

pratique, jour après jour, par son assiduité aux affaires.

Un élément de fierté pour ce courant patriotique: l'empire colonial. Commencée sous la Restauration, continuée sous le second Empire, l'expansion de la France outre-mer fut poursuivie par Jules Ferry sous la IIIe République. Pour le parti alors au pouvoir, il fallait donner á la France les éléments d'une puissance nouvelle. Les conquêtes continentales devenues impossibles, c'est par-delá les mers qu'il faut tenter l'aventure, aventure au demeurant conforme á l'action civilisatrice de la France. <<Rayonner sans agir, déclarait Jules Ferry en 1885, sans se mêler aux affaires du monde, vivre de cette sorte pour une grande nation, croyez-le bien, c'est abdiquer et dans un temps plus court que vous ne pouvez le croire, c'est descendre du premier rang au troisiéme ou au quatriéme.>> Pour les adversaires de Ferry, c'est au contraire détourner la France de la <<ligne bleue des Vosges>>. <<Mon patriotisme est en France!>> s'exclamait Clemenceau. Mais bientôt, seule l'extrême gauche devait continuer á protester contre l'expansion coloniale de la France.

En 1914, l'influence française s'étend á l'Afrique du Nord (départements algériens, protectorat sur le Maroc et la Tunisie), á l'Afrique occidentale et á l'Afrique équatoriale placées sous l'autorité de gouverneurs, á l'Indochine (Cochinchine, Annam, Tonkin, Cambodge et

Laos) dont l'organisation est fixée par un décret du 20 octobre 1911, á Madagascar, aux Antilles, á l'Océanie et aux comptoirs de l'Inde.

APPENDIX D

ENGLISH TRANSLATIONS OF FRENCH READING PASSAGES

Translation: The Gauls

The Gauls and more generally the Celts were never isolated: from the seventh century B.C., they maintained commercial rapport with the most civilized nations of the Mediterranean basin. Of all the barbarians, they were without doubt the least barbarous. The entry of the Gauls into the Mediterranean and Hellenic worlds far preceded the Roman conquest and prepared the way for it. In particular, it is abundantly established that the commercial expansion and technical and cultural influence of the Phoenician colony at Marseilles extended widely beyond the group of those who were going to become part of the Roman province of Narbonnaise. Heeding the call of the people of Marseilles, victims of organized piracy and of the raids of their neighbors, the Romans intervened in southern Gaul. This region served as the base of departure of Julius Caesar to conquer the main body of the territory; this only lead to brief struggles with the Germanic people. If Gaul hadn't become roman, it would certainly have become German.

Latin gradually replaced the Celtic languages; the entire territory elevated itself from the cities, with their networks of aqueducts and roads. Nevertheless, the Gauls, in becoming gallo-roman, did not deny their own

territories. This loyalty strongly manifests itself in the artistic domain.

Translation: Louis XIV

Louis XIV is one of the historical characters on whom attention remains focused, and to whom not one historian can pretend to give an absolute and definitive image. He has so directly influenced French destiny that one cannot imagine the history of France without him. But, because his reign witnessed a curious combination of incontestable glory and some very heavy misfortunes for the nation, he was either highly praised or criticized and historians were either his apologists or detractors. One must understand that his reign is much more difficult for the people of the end of the 20th century to understand than it is for those of the beginning because the deep changes of French society during this century erased the collective mentality that, 70 years ago, was relatively close to that of the 17th century. On the other hand, improvements in historians' methods, especially regarding the conditions of life in France during the time of Louis XIV, facilitate a better knowledge of the country on which his authority was exercised.

The period of personal reign, which extended from 1661 to 1715--54 years of effective government by the sovereign. It is through work that one reigns, said

Louis XIV; he put this principle into practice day after day by his devotion to his duties.

Translation: Colonial Empire

An element of pride for this patriotic current: the colonial empire. Begun under the Restoration and continued under the Second Empire, the expansion of France overseas was pursued by Jules Ferry under the Third Republic. For the party then in power, it was necessary to imbue France with a new strength. Because conquests of Continental Europe had become impossible, it was necessary to tempt adventure beyond the seas and spread French civilization. "For a great nation to expand without acting, declared Jules Ferry in 1885, without becoming involved in world affairs, is to abdicate, and in a shorter time than you can believe, it is to descend from the first rank to the third or fourth." The adversaries of Ferry wanted, on the contrary, to turn France from "the blue line of the Vosges." "My patriotism is in France!" exclaimed Clemenceau. But soon, only the extreme left continued to protest against the colonial expansion of France.

In 1914, French influence extended to North Africa (Algerian departments, protectorate of Morocco and Tunisia), to East Africa and to Equatorial Africa placed under the authority of governors, to Indochina (Cochinchine, Annam, Cambodia, and Laos) of which the government is fixed by a decree of October 20, 1911, to

Madagascar, to the Antilles, to Oceania, and to the riches of India.

APPENDIX E

SCORING TEMPLATES FOR FRENCH READING PASSAGES

Scoring Template: Les Gaulois

Proposition	Points
Les gaulois	4
et plus gènèralement les Celtes	2
n'ont jamais été fermés sur eux-mêmes:	4
dés le VIIe siécle av. J.-C.,	2
ils entretenaient des rapports commerciaux	4
avec les peuples les plus civilisés	3
du bassin méditerranéen.	3
De tous les barbares,	1
ils étaient sans doute	1
les moins barbares.	3
L'entrée de la Gaule	4
dans le monde méditerranéen	4
et hellénique	2
a précédé de loin	4
la conquête romaine	4
et a préparé la voie á celle-ci.	3
Il est en particulier	1
abondamment établi	1
que l'expansion commerciale	3
et l'influence technique et culturelle	3
de la colonie phocéenne de Marseille	4
débordaient largement le cadre	2
de ce qui allait devenir	1

la province romaine de Narbonnaise.	2
C'est du reste	1
á l'appel des Marseillais,	3
victimes de la piraterie ligure	2
et des incursions de leurs voisins,	2
que les Romains intervinrent	4
en Gaule mèridionale.	1
Cette règion servit ensuite	4
de base de dèpart á Jules Cèsar	4
pour conquèrir l'ensemble du territoire;	3
il ne faisait en cela	1
que lutter de vitesse	1
avec les Germains:	2
si la Gaule n'était devenue romaine,	3
elle fût á coup sûr	1
devenue germanique.	3
La langue latine	4
supplanta peu á peu	2
les parlers celtiques;	2
sur tout le territoire	1
s'èlevèrent des villes,	1
avec leurs rèseaux d'aqueducs	2
et de voies.	2
Cependant,	1
les Gaulois,	1

en devenant gallo-romains,	3
ne reniérent pas leurs traditions propres.	4
Cette fidélité	1
se manifeste aussi bien	2
dans le domaine religieux	3
que dans le domaine artistique.	3

Scoring Template: Louis XIV

Proposition	Points
Louis XIV	4
est un des personnages historiques	2
sur lesquels l'attention demeure portée,	3
sans que nul historien	3
puisse prétendre donner de lui	1
une image certaine et définitive.	3
Qu'il ait influencé directement	4
les destinées françaises	4
et qu'á ce titre	1
on ne puisse imaginer	3
l'histoire de la France sans lui,	4
nul doute.	1
Mais,	1
parce que son régne a curieusement associé	2
une incontestable gloire	4
á de trés lourds malheurs	4
pour la nation,	1
il a été extrêmement loué ou critiqué	4
et ses historiens se sont souvent partagés	2
entre apologistes et détracteurs.	2
On doit observer	1
qu'il est beaucoup plus malaisé á comprendre	4
pour des hommes de la fin du XXe siécle	3

que pour ceux du début,	1
parce que les profondes mutations	3
de la société française	2
au cours de cette période	1
ont fait disparaître	2
des aspects de la mentalité collective	3
qui demeuraient,	2
il y a soixante-dix ans encore,	2
relativement proches du XVIIe siécle.	1
En revanche,	1
les renouvellements de la méthode historique,	4
surtout des études érudites	1
sur les conditions de vie en France	2
au temps de Louis XIV,	2
autorisent une meilleure intelligence du pays	4
sur lequel son action s'est exercée.	3
La période du régne personnel	4
s'étend de 1661 á 1715,	3
soit pendant cinquante-quatre ans,	1
période du gouvernement	2
effectif du souverain.	1
C'est par le travail que l'on régne,	4
disait Louis XIV;	3
il a mis ce principe en pratique,	3
jour aprés jour,	2

par son assiduité aux affaires. 3

Scoring Template: l'Empire Colonial

Proposition	Points
Un élément de fierté	3
pour ce courant patriotique:	3
l'empire colonial.	4
Commencée sous la Restauration,	4
continuée sous le second Empire,	2
l'expansion de la France outre-mer	4
fut poursuivie	4
par Jules Ferry	4
sous la IIIe République.	3
Pour le parti alors au pouvoir,	2
il fallait donner á la France	3
les éléments d'une puissance nouvelle.	4
Les conquêtes continentales	3
devenues impossibles,	3
c'est par-delá les mers	4
qu'il faut tenter l'aventure,	2
aventure au demeurant conforme	2
á l'action civilisatrice de la France.	4
<<Rayonner sans agir,	4
déclarait Jules Ferry	2
en 1885,	1
sans se mêler aux affaires du monde,	2
vivre de cette sorte pour une grande nation,	2

croyez-le bien,	1
c'est abdiquer	4
et dans un temps plus court	2
que vous ne pouvez le croire,	1
c'est descendre du premier rang	3
au troisiéme ou au quatriéme.>>	2
Pour les adversaires de Ferry,	3
c'est au contraire	2
détourner la France	3
de la <<ligne bleue des Vosges>>.	2
<<Mon patriotisme est en France!>>	4
s'exclamait Clemenceau.	4
Mais bientôt,	1
seule l'extrême gauche	3
devait continuer á protester	4
contre l'expansion coloniale	2
de la France.	1
En 1914,	3
l'influence française	3
s'étend á l'Afrique du Nord	4
(départements algériens,	1
protectorat sur le Maroc	1
et la Tunisie),	1
á l'Afrique occidentale	3
et á l'Afrique équatoriale	3

placées sous l'autorité de gouverneurs,	1
á l'Indochine	4
(Cochinchine,	1
Annam,	1
Tonkin,	1
Cambodge	1
et Laos)	1
dont l'organisation est fixée	1
par un décret du 20 octobre 1911,	1
á Madagascar,	2
aux Antilles,	2
á l'Océanie	2
et aux comptoirs de l'Inde.	3

APPENDIX F

SCORING GUIDE: PASSAGE-SPECIFIC FRENCH VOCABULARY TEST

Scoring Guide: Passage-Specific French Vocabulary Test

1. fierté _____ pride
2. siécle _____ century
3. était _____ was
4. outre-mer _____ overseas
5. entretenaient _____ maintained
6. lesquels _____ the ones
7. poursuivie _____ pursued/sought
8. rapports _____ interactions
9. monde _____ world
10. la pouvoir _____ (the) power
11. voie _____ way
12. portée _____ carried
13. fallait _____ was necessary/used to be necessary
14. débordaient _____ extended beyond
15. renouvellements _____ improvements/reforms
16. puissance _____ power/strength
17. cadre _____ group
18. nul _____ no/not one/none
19. devenues _____ became
20. appel _____ call
21. puisse _____ can/is able to
22. tenter _____ to tempt
23. disait _____ said/was saying
24. voisins _____ neighbors

25. lourds _____ heavy _____

26. demeurant _____ remaining/living _____

27. mis _____ put _____

28. faisait _____ made/was making/did/was doing _

29. èrudites _____ scholarly/smart _____

30. loin _____ far _____

31. rayonner _____ to radiate/to spread out _____

32. meilleure _____ better/best _____

33. lutter _____ to fight _____

34. partagès _____ shared _____

35. agir _____ to act _____

36. exercée _____ exerted/exercised _____

37. s'èlevèrent _____ to lift (themselves) up/elevate

38. doit _____ must/should/has to _____

39. rang _____ rank/rung _____

40. soit _____ was _____

41. rèseaux _____ web/network _____

42. malaisè _____ difficult _____

43. devait _____ had to _____

44. renièrent _____ deny _____

45. dèbut _____ beginning _____

46. s'ètend _____ extends (itself) _____

47. croyez _____ believe _____

48. mêler _____ to mix/to mingle/to blend _____

49. proches _____close/near_____

50. comptoirs _____markets/counters/banks_____

APPENDIX G

ANOVA TABLE OF NDRT COMPREHENSION SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	30.2500	30.2500	2.83
S/A	34	364.0556	10.7075	
Total	35	394.3056		

$p < .1020$

APPENDIX H

ANOVA TABLE OF L1 RECALL PROTOCOL SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	402.3810	402.3810	4.60*
S/A	40	3496.0952	87.4024	
Total	41	3898.4762		

*p<.0380

APPENDIX I

ANOVA TABLE OF L1 WRITING ABILITY SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	0.2250	0.2250	.02
S/A	38	412.7500	10.8618	
Total	39	412.9750		

$p<.8863$

APPENDIX J

ANOVA TABLE OF FRENCH VOCABULARY KNOWLEDGE SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	3137.3571	3137.3571	20.98*
S/A	40	5981.1429	149.5286	
Total	41	9118.5000		

*p<.0001

APPENDIX K

ANOVA TABLE OF FRENCH GRAMMATICAL SKILL SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	288.0952	288.0952	11.08*
S/A	40	1039.8095	29.9952	
Total	41	1327.9047		

*p<.0019

APPENDIX L

ANOVA TABLE OF L2 READING COMPREHENSION SCORES FOR INTERMEDIATE VERSUS ADVANCED SUBJECTS

Source	df	SS	MS	F
A	1	214.8810	214.8810	2.07
S/A	40	4152.7619	103.8190	
Total	41	4367.6429		

$p<.1580$

APPENDIX M

ANOVA TABLE OF NDRT COMPREHENSION SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	99.2821	99.2821	5.47*
S/A	76	1378.2051	18.1343	
Total	77	1477.4872		

*p<.0219

APPENDIX N

ANOVA TABLE OF L1 RECALL PROTOCOL SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	973.7619	973.7619	11.48*
S/A	82	6958.3810	84.8583	
Total	83	7932.1429		

*p<.0011

APPENDIX O

ANOVA TABLE OF L1 WRITING ABILITY SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	23.1125	23.1125	2.74
S/A	78	658.8750	8.4471	
Total	79	681.9875		

$p < .1021$

APPENDIX P

ANOVA TABLE OF FRENCH VOCABULARY KNOWLEDGE SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	18870.0119	18870.0119	136.10*
S/A	82	11368.9762	138.6461	
Total	83	30238.9881		

*p<.0001

APPENDIX Q

ANOVA TABLE OF FRENCH GRAMMATICAL SKILL SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	9366.2976	9366.2976	317.02*
S/A	82	2422.6905	29.5450	
Total	83	11788.9881		

*p<.0001

APPENDIX R

ANOVA TABLE OF L2 READING COMPREHENSION SCORES FOR BEGINNERS VERSUS NON-BEGINNERS

Source	df	SS	MS	F
A	1	7895.0698	7895.0698	114.49*
S/A	82	5792.7442	68.9612	
Total	83	13687.8130		

*p<.0001

LIST OF REFERENCES

Alexander, L., & James, T. (1987). The nation's report card: Improving the assessment of student achievement: Report of the Study Group. Cambridge, MA: National Academy of Education.

Arndt, V. (1987). Six writers in search of text: A protocol-based study of L1 and L2 writing. English Language Teaching Journal, 41(4), 257-267.

Barnett, M. (1986). Syntactic and lexical/semantic skill in foreign language reading: Importance and interaction. Modern Language Journal, 70, 343-349.

Bèlanger, J. (1987). Theory and research into reading and writing connections: A critical review. Reading-Canada-Lecture, 5(1), 10-18.

Bernhardt, E. (1983). Three approaches to reading comprehension in intermediate German. Modern Language Journal, 67, 111-115.

Bernhardt, E. (1986). Reading in a foreign language. In B. Wing (Ed.), Listening, reading, writing: Analysis and application (pp. 93-115). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Bernhardt, E. (1991). Reading development in a second language. Norwood, NJ: Ablex Publishing Corporation.

Brown, J., Bennett, J., & Hanna, G. (1981). The Nelson-Denny Reading Test. Chicago: Riverside Publishing Co.

Canale, M., Frenette, N., & Bèlanger, M. (1988). Evaluation of minority student writing in first and second language. In J. Fine (Ed.), Second language discourse: A textbook of current research. Norwood, NJ: Ablex Publishing Corporation.

Carrell, P. (1983). Three components of background knowledge. Language Learning, 33, 183-207.

Carrell, P., & Connor, U. (1991). Reading and writing descriptive and persuasive texts. Modern Language Journal, 75(3), 314-324.

Carson, J., Carrell, P., Silberstein, S., Kroll, B., & Kuehn, P. (1990). Reading-writing relationships in first and second language. TESOL Quarterly, 24(2), 245-266.

Clarke, M. (1978). Reading in Spanish and English: Evidence from adult ESL students. Language Learning, 29(1), 121-150.

Clarke, M. (1980). The short circuit hypothesis of ESL reading: Or when language competence interferes with reading performance. Modern Language Journal, 64, 203-209.

Cummins, J. (1981). Bilingualism and minority-language children: Language and literacy series. Toronto: Ontario Institute for Studies in Education. (ERIC Document Reproduction Service No. ED 215 557)

Cziko, G. (1978). Differences in first- and second-language reading: The use of syntactic, semantic, and discourse constraints. Canadian Modern Language Review/La revue canadienne des languages vivantes, 34, 473-489.

Edelsky, C. (1982). Writing in a bilingual program: The relation of L1 and L2 texts. TESOL Quarterly, 16(2), 211-228.

Educational Testing Service (1989). TOEFL Test of Written English Guide. Princeton, NJ: Educational Testing Service.

Educational Testing Service (1990). Test of Written English (TWE) Scoring Guide. Princeton, NJ: Educational Testing Service.

Encyclopedia Britannica (1984). Chicago, IL: Encyclopedia Britannica, Inc.

Encyclopedia Universalis (1968). Clichy, France: Paul Dupont.

Everson, M. (1986). The effect of word-unit spacing upon the reading strategies of native and non-native readers of Chinese: An eye-tracking study. Unpublished doctoral dissertation. The Ohio State University, Columbus, OH.

Goldman, S., Reyes, M., & Varnhagen, C. (1984). Understanding fables in first and second languages. NABE Journal, 8(2), 35-66.

Green, B. (1981). A primer of testing. American Psychologist, 36(10), 1001-1011.

Groebel, L. (1980). A comparison of students' reading comprehension in their native language with their reading comprehension in the target language. English Language Teaching Journal, 35, 54-59.

Henning, G. (1975). Measuring foreign language reading comprehension. Language Learning, 25(1), 109-114.

Henning, G. (1987). A guide to language testing. Cambridge, MA: Newbury House.

Hock, T., & Poh, L. (1979). The performance of a group of Malay-medium students in an English reading comprehension test. Regional Language Centre Journal, 10(1), p. 81-87.

Janopolous, M. (1986). The relationship of pleasure reading and second language writing proficiency. TESOL Quarterly, 20(4), 763-768.

Johnson, R. (1970). Recall of prose as a function of the structural importance of linguistic units. Journal of Verbal Learning and Verbal Behavior, 9, 12-20.

Jones, S., & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), Writing in real time: Modeling production process. Norwood, NJ: Ablex Publishing Corporation.

Juel, C., Griffith, P., & Gough, P. (1986). Acquisition of literacy: A longitudinal study of the children in first and second grade. Journal of Educational Psychology, 78(4), 243-255.

Kerlinger, N., & Pedhazur, E. (1973). Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston.

Kamil, M., Langer, J., & Shanahan, T. (1985). Understanding research in reading and writing. Boston: Allyn and Bacon, Inc.

Kamil, M., Smith-Burke, M, & Rodriguez-Brown, F. (1986). The sensitivity of cloze to intersentential integration of information in Spanish bilingual populations. In J. Niles and R. Lalik (Eds.), Thirty-fifth yearbook of the National Reading Conference (pp. 334-338). Rochester, NY: National Reading Conference.

Kirby, K. (1986). Reading and writing processes of selected high-risk college freshmen. Unpublished doctoral dissertation. University of Georgia.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. Foreign Language Annals, 22(6), p. 529-540.

Koda, K. (1990). Factors affecting second language text comprehension. J. Zutell and S. McCormick (Eds.) Thirty-ninth Yearbook of the National Reading Conference (pp. 419-427). Chicago, IL: The National Reading Conference, Inc.

Kucer, S. (1985). The making of meaning: reading and writing as parallel processes. Written Communication, 20(3), 317-336.

Langer, J. (1984). The effects of available information on responses to school writing tasks. Research in the Teaching of English, 18(1), 27-44.

Langer, J. (1986). Reading, writing and understanding: An analysis of the construction of meaning. Written Communication, 3(2), 195-218.

Lay, N. (1982). Composing processes of adult ESL learners: A case study. TESOL Quarterly, 16(3), 406.

Lee, J. (1986). On the use of the recall task to measure L2 reading comprehension. Studies in second language acquisition, 8, 201-212.

Liskin-Gasparro, J. (1984). The ACTFL Proficiency Guidelines: A historical perspective. In T. Higgs (Ed.) Teaching for proficiency, the organizing principle. Lincolnwood, IL: National Textbook Co.

Loban, W. (1964). Language ability grades seven, eight, and nine (Cooperative Research Project No. 1131). Berkeley, CA: University of California.

Markham, P. (1985). The rational deletion cloze and global comprehension in German. Language Learning, 35, 423-430.

Martin, S. (1987). The meaning-making strategies reported by proficient readers and writers. Paper presented at the meeting of the National Reading Conference, St. Petersburg, FL.

Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. Braun (Eds.), Test validity (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum.

Mohan, B., & Lo, W. (1985). Academic writing and Chinese students: Transfer and developmental factors. TESOL Quarterly, 19(3), 515-534.

National Governors' Association (1986). Time for results: The governors' 1991 report on education. Washington DC: National Governors' Association. (ERIC Document Reproduction Service No. ED 279 608)

Pike, L. (1979). An evaluation of alternative item formats for testing English as a foreign language. Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED 206 627)

Reyes, M. (1987). Comprehension of content area passages: A study of Spanish/English readers in third and fourth grade. In S. Goldman and H. Trueba (Eds.), Becoming literate in English as a second language. Norwood, NJ: Ablex Publishing Corporation.

Ryan, S. (1985). An examination of reading and writing strategies of selected fifth-grade children. In J. Niles and R. Lalik (Eds.), Thirty-fourth yearbook of the National Reading conference. Rochester, NY: National Reading Conference, Inc.

Schewe, A., & Froese, V. (1986, December). Relating reading and writing via comprehension, quality, and structure. Paper presented at the annual meeting of the National Research Conference, Austin, TX. (ERIC Document Reproduction Service No. ED 281 141)

Scholtz, G., Hendricks, D., Spurling, R., Johnson, M., & Vandenburg, L. (1983). In J. Oller and K. Perkins (Eds.), Research in Language Testing (pp. 24-33), Rowley, MA: Newbury House.

Shanahan, T. (1984). Nature of the reading-writing relationship: An exploratory multivariate analysis. Journal of Educational Psychology, 76(3), 466-477.

Shanahan, T., Kamil M., & Tobin, A. (1982). Cloze as a measure of intersentential comprehension. Reading Research Quarterly, 17, 229-255.

Shanahan, T., & Lomax, R. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. Journal of Educational Psychology, 78(2), 116-123.

Shanahan, T., & Lomax, R. (1988). A developmental comparison of three theoretical models of the reading-writing relationship. Research in the Teaching of English, 22(2), 196-212.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. Language Testing, 1, 147-170.

Snedecor, G, & Chochran, W. (1967). Statistical Methods. Ames, IA: Iowa State University Press.

Squire, J. (1983). Composing and comprehending: Two sides of the same basic process. Language Arts, 60(5), 581-589.

Squire, J. (1984). Composing and comprehending: Two sides of the same basic processes. In J. Jensen (Ed.), Composing and comprehending (pp. 23-31). Urbana, IL: National Council of Teachers of English.

Stotsky, S. (1983). Research on reading/writing relationships: A synthesis and suggested new directions. Language Arts, 60(5), 627-642.

Tedick, D. (1988). The effects of topic familiarity on the writing performance of non-native writers of English at the graduate level. Unpublished doctoral dissertation. The Ohio State University, Columbus, OH.

Tierney, R., & Pearson, D. (1983). Toward a composing model of reading. Language Arts, 60(5), 568-586.

Tierney, R., & Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. Kamil, P. Mosenthal, and P. Pearson (Eds.), Handbook of reading research. New York: Longman.

USAFA Curriculum Handbook (1990-1991). Colorado Springs, CO: Curriculum Division, Office of the Registrar, USAFA.

Wagner, D, Spratt, J., & Ezzaki, A. (1989). Does learning to read in a second language always put the child at a disadvantage? Some counterevidence from Morocco. Applied Psycholinguistics, 10, 31-48.

Weber, R. (1991). Linguistic diversity and reading in American society. In R. Barr, M. Kamil, P. Mosenthal, and P. Pearson (Eds.), Handbook of reading research (vol 2), 97-119. New York: Longman.

Westfall, P. (1988). Re-estimation of student ability in foreign language using the Rausch Model. Unpublished doctoral dissertation. The Ohio State University.

Wittrock, M. (1983). Writing and the teaching of reading. Language Arts, 60(5), 600-606.

Wittrock, M. (1984). Writing and the teaching of reading. In J. Jensen (Ed.), Composing and comprehending (pp. 77-83). Urbana, IL: National Council of Teachers of English.