THE EFFECT OF A COMPUTERIZED HELP RESOURCE ON THE LEXICAL INFERENCING ABILITY OF ENGLISH AS A SECOND LANGUAGE (ESL) LEARNERS DURING CONTEXTUALIZED READING

By

DEBRA KIRCHGASSNER ABBOTT

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2000

This thesis is dedicated to my late mother, Marjorie L. Kirchgassner, for her incredible love and support and to my wonderful husband, Colin Abbott, who enabled me to complete this project.

ACKNOWLEDGMENTS

I am most grateful to my committee chair, Dr. Clemens Hallman, who provided dependable assistance over the years it took to complete this project. A special note of thanks is owing to Dr. Sebastian Foti, a member of my supervisory committee, for his creative and unique perspective on issues concerning virtually every aspect of my research. I am thankful to Dr. Lee Mullally, another member of my supervisory committee, for his wonderfully detailed editorial advice and encouragement. I also appreciate the help of my other committee members, Dr. Douglas Dankel and Dr. Roger Thompson, who were willing to give of their time and offer sound guidance.

My sincere appreciation goes to Dean Linda Crocker for her wise preparatory advice and excellent critique of my proposal. Dr. Candace Harper and Dr. John Bengston also provided invaluable advice during the formative phase of this project, and I owe them many thanks.

I am indebted to Ruth Henderson and the staff of the English House in Gainesville for the use of their facilities and for providing me with an opportunity to interact with the target population of this study. If it were not for Ruth's support of English as a Second Language (ESL) research, this study would not have been possible.

Finally, I wish to thank Brad Moulder for his patient and skillful assistance with quantitative data analysis and SPSS.

TABLE OF CONTENTS

LIST OF TABLES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE EFFECT OF A COMPUTERIZED HELP RESOURCE ON THE LEXICAL INFERENCING ABILITY OF ENGLISH AS A SECOND LANGUAGE (ESL) LEARNERS DURING CONTEXTUALIZED READING

By

Debra Kirchgassner Abbott

December, 2000

Chairman: Dr. Clemens Hallman
Major Department: School of Teaching and Learning

Although most vocabulary is picked up incidentally through the environment and not through formal instruction, the details of this process are poorly understood. Lexical guesses or inferences about the meaning of unknown words are likely the primary means by which new vocabulary is acquired, and this process has been shown to be especially difficult for foreign language learners. In this repeated measures study, thirteen advanced students of English as a second language (ESL) were exposed to two contextualized reading conditions on the computer. The first condition, passage only, involved reading passages alone, and the second, passage plus Word Hints, offered passages with the addition of a hypertextual reading aid. This help resource provided both semantic and graphophonic clues, but glosses and definitions were withheld. Repeated measures ANOVA results showed that a significantly greater number of guesses and more accurate

guesses were made in the presence of Word Hints. Repetition of the study showed that this effect increased over time. Behavioral data were analyzed to determine patterns of use with regard to Word Hints. A comparison of computer usage tracking data to lexical guessing data indicated that proficient readers who made the most accurate lexical guesses were also more likely to ignore graphophonic clues and request semantic help more selectively than others. Individuals who divided their attention between both levels of help in Word Hints tended to have lower scores for guessing accuracy. Overall, the data suggest that readers were competent judges of the kind of help they needed, and they also knew when they needed to obtain further clues about the meaning of an unknown word.

The study found that it is possible to enhance second language readers' own inferencing processes through the availability of specially modified input during contextualized reading. Participants seemed especially capable of exploiting additional semantic clues during reading, and it appears that helpful contexts are indeed a key factor in second language vocabulary acquisition. Research suggests, however, that several conditions must be met before lexical inferencing should be considered for use as an independent study strategy.

CHAPTER 1
INTRODUCTION

Vocabulary Acquisition and the Problem of Inferring Meaning from Context

Foreign students who wish to achieve native-like proficiency in academic English must learn many thousands of word families (words similar in form and meaning) and an even greater number of specific word forms (also called lexical units). Studies by Laufer (1992), Nation (1990; 1993), and Hirsh and Nation (1992) indicate that a learner would need to know over 3,000 word families or about 5,000 -7,000 word forms in order to achieve 95% coverage of words in a typical academic text. Other estimates suggest that the threshold required for proficient academic English may be even higher (Anglin, 1993). In any case, the sizable nature of second language (L2) lexical demands is also apparent to the learners. Students view their own ability to handle vocabulary study as a measure of achievement in a foreign language: Jones (1995) finds that "positive versus negative experiences of coping with vocabulary are strongly related to overall perceptions of success versus failure" among learners who use self-study strategies (p. 197).

The present study examines the effect of a tool designed to extend and support the learner's own lexical inferencing process during reading. Readers' inferences about the meanings of unknown words in a text are an important process in vocabulary acquisition. What is this relationship between vocabulary acquisition and lexical inferencing?

Moreover, how do proficient second language students learn thousands upon thousands of words?

One theory holds that constant, incremental vocabulary learning takes place from exposure to everyday contexts. Sternberg (1987) views contextualized vocabulary learning as a basic process of native language acquisition:

> First, although estimates of adult vocabulary sizes differ, there is no disagreement that the typical adult has in his or her vocabulary tens of thousands of words . . . Second, one's level of vocabulary is highly predictive, if not determinative, of one's level of reading comprehension . . .Third, vocabulary is probably the best single indicator of a person's overall level of intelligence . . There is a simple explanation to account for all three of these facts. The explanation is that most vocabulary is learned from context. (p. 89-90)

Regarding second language vocabulary acquisition, too, Coady (1993) asserts "current thinking is that the vast majority of these words have been learned through context rather than direct instruction." However, he further acknowledges that the literature on contextualized vocabulary learning is "confusing and contradictory," and in particular, it shows that contextualized learning through reading can be a problematic way to learn words (p. 16-17).

Unintentional word learning which occurs within some context and without the benefit of direct instruction is said to be incidentally learned, and inferences about the meanings of unknown words, or lexical guesses, appear to be the primary means by which incidental word learning is accomplished within reading contexts. Many authors have demonstrated that the "guessability factor" of certain kinds of words and texts is low. That is, many written texts typically lack context clues for a majority of words, or even when clues are provided, the contexts may be unreliable predictors of some words such as those which have multiple meanings (Haynes, 1993; Kelly, 1990; Laufer, 1990; Schatz & Baldwin, 1986). While some researchers have claimed that incidental

vocabulary learning rates are incremental and reliable (Konopak, Sheard, Longman, Lyman, Slaton, & Atkinson, 1987; Nagy, Herman, & Anderson, 1985), it appears more accurate to say that such rates among second language learners are typically haphazard -- and reliable only under optimal reading conditions (Coady, 1993, p. 17). In particular, these optimal reading conditions are probably most dependent upon the adequate provision of context clues and inferencing practice (Scherfer, 1993, p. 1149).

A study by Mondria and Wit-de Boer (1991) identifies the specific kinds of semantic context clues which aid lexical guessing in a second language, and other studies have begun to examine whether or not the presence of reading aids such as dictionaries or glossaries significantly boosts incidental vocabulary learning for foreign language students (Chun & Plass, 1996; Gildea, Miller, & Wurtenberg, 1990; Hulstijn, Hollander, & Greidanus, 1996; Jacobs, Dufon, & Hong, 1994; Knight, 1994; Krantz, 1991). In general, these studies show that rates of incidental vocabulary learning can be influenced by the presence of additional context clues, textual annotations, reading aids, and media resources. This is a highly significant discovery. Any gains exacted from a reliable, incremental acquisition process could potentially achieve remarkable cumulative effects over time. Stated another way, the long-term vocabulary gains of students might be significantly higher if incidental learning rates could be regularly boosted through the use of an appropriate reading aid.

However, there is a paradox involved in providing the reader with a given meaning (for example, a dictionary definition or gloss) during contextualized reading, an inferencing process. Research suggests that inferred meanings are better retained than given meanings (for a summary of this evidence, see Hulstijn, Hollander and Greidanus,

1996, p. 326), and yet all of the authors of the above-mentioned studies have provided readers with aids which effectively stop the inferencing process by giving them a word meaning. Why not provide a reading resource designed to stimulate and extend the inferencing (lexical guessing) process? If one could adequately control the risk that readers are going to make the wrong inferences about words, then would learners be attracted to such an "inference-guiding" tool and actually use it? Would the quality or quantity of their inferences improve? If so, then such a reading aid used over an extended period might boost word retention, and therefore incidental vocabulary learning, in a way that conventional resources such as dictionaries or glossaries do not.

The Aim of the Present Study

The present study investigated whether a computerized help resource could influence the rate and accuracy (or quantity and quality) of lexical guesses during an online contextualized reading task. The lexical aid considered shall be referred to as the Word Hints reading resource. This tool differed from a dictionary or a glossary in that it provided three kinds of hints or clues about an unknown word, semantic, graphophonic, and syntactic, without giving students explicit "answers" as to the meaning of a word. Each of these three kinds of lexical clues was available for a given word when students selected the resource during computerized reading. The tool utilized a specific combination of clues which were designed to help focus readers' attention on lexical form in addition to meaning if students chose to select secondary-layer (sound, spelling, and grammar) clues in addition to primary-layer (semantic) ones.

The characteristics of the semantic clues in the resource were derived from a model established by Mondria and Wit-de Boer (1991). Word Hints provided a

consistent quantity and quality of semantic information for every vocabulary word in the resource. The lexical guessing behavior of subjects was examined during normal contextualized reading on the computer, and this was compared to their lexical guessing behavior when both online reading passages and Word Hints was available.

The major issue of this investigation consisted of two parts: (1) whether the contextual and intralingual mediating variables that create difficulty during second language lexical inferencing have been correctly identified, and (2) whether these variables can be significantly influenced through the provision of a computerized lexical resource. Subjects' lexical guessing errors are typically an excellent subject of analysis since they reflect the linguistic and contextual elements that cause confusion or difficulty in reading. By examining the frequency (overall quantity) of subjects' lexical guesses, the study was designed to show whether the "passage plus Word Hints" condition had a significant and helpful influence over some key contextual and linguistic variables that contribute to reading difficulty. During lexical guessing, it was anticipated that the alleviation of certain adverse mediating factors (barriers to good inferencing) would result in less reader frustration and improved mental engagement thereby increasing inferencing ability (frequency or overall quantity). The measurement of readers' accuracy of lexical guessing reflected the level of quality and degree of appropriateness of the specific contextual and linguistic clues provided.

Need for the Study

The present study is necessary for several reasons. First, more research needs to focus on the question of whether or not reading interventions exist that can extend, support, or otherwise enhance the reader's own inferencing process. Second, while it has

been evident for some time that numerous contextual, interlexical, and intralexical factors make vocabulary inference-making difficult during second language reading (e.g. Haastrup, 1989, 1991; Laufer, 1990; Laufer and Sim, 1985), attention should now shift to the question of whether it is possible to significantly affect these adverse factors or mediating variables. Third, this study seeks to understand the influence of lexical inferencing behaviors on basic vocabulary acquisition and how this process varies among individuals. The act of providing readers with an aid to vocabulary acquisition is analogous to providing students with enhanced strategies or tools to achieve self-study in learning. Fourth, because this study is concerned with the specific elements necessary to construct an L2 reading aid, the outcome should be relevant to professionals in the area of lexicography, computer-assisted language learning, and curriculum design. Thus, this type of investigation is very pragmatic. Fifth, many experts believe that vocabulary development may be the single most important component of reading. For example, Laufer (1992) demonstrated that adequate vocabulary size is a more important factor in determining the success of L2 reading than general academic ability. Furthermore, Sternberg (1987) is so convinced of the centrality of vocabulary development that he calls his theory of vocabulary acquisition *knowledge acquisition*, and he points out that a good definition of intelligence is the ability to figure out the meanings of things (e.g. words) from context. From this perspective, lexical inferencing should be considered a fundamental topic in cognitive and L2 research.

Definition of Terms

In this investigation, Word Hints refers to the condition whereby a monolingual English lexical resource is provided as a supplement to a computerized reading passage.

This supplement operates within a hypertextual reading program. Help of this type is convenient as students only have to click a mouse on a word underlined in the computerized reading passage in order to receive additional clues for that word.

Clues are categorized as one of three types. *Semantic* refers to clues intended to help with the meaning of the word in a sentence. *Graphophonic* refers to clues designed to give some help with the sound-spelling patterns of a target word. *Syntactic* refers to clues which describe some grammatical characteristic of a word. Goodman (1974) calls these three types of clues "cue systems," and asserts that they constitute "psycholinguistic universals." These three distinct types of clues are impossible to completely separate in the presentation of the Word Hints resource. For instance, a "usage example," an additional sentence context which is designed to provide further semantic clues for a target word, may also function to provide syntactic clues. However, with the exception of the syntactic help, which was not examined in isolation in this study, it is possible to separate semantic from graphophonic clues.

Context is variously defined in experimental research, and two different kinds of "contexts" are used in this investigation. First, context can mean the enriched or suggestive sentence context (usage example) that surrounds target words in the passage plus Word Hints condition. Second, context refers to the reading passages that surround the target words in this study. These passages averaged 150 words in length, and the reading task involved asking participants to make lexical guesses on (approximately) one in every twenty words in the passage, or seven words per passage. This word density level is somewhat less dense than that used either by Bensoussan and Laufer (1984) in their L2 lexical guessing experiment or by Ittzes' (1991) L2 lexical guessing study.

Limitations and Assumptions

One limitation of this investigation is that outcomes may only be generalized to English as a second language readers. Any extrapolation of the results to the inferencing behaviors of other types of readers would be ill-advised.

Some of the theoretical assumptions of this investigation are as follows:

1. An important part of second language vocabulary acquisition involves learning the kinds of lexical nuances, complexities, restrictions, interrelationships, and multiple word meanings that native speakers know.
2. The tendency of second language learners to ignore difficult words, use generalized forms, or to engage in lexical simplification can be mediated.
3. The provision of additional lexical/semantic clues during reading helps more than it interferes with reading.
4. What it means to "know a word" is a matter of degree, function, and dimensionality--not an either/or situation. Thus, measurement of word knowledge should be carefully considered.

Given the fourth assumption, it is useful to note here that a coding scheme for subject responses was adapted from a study by Bensoussan and Laufer (1984) which categorizes subjects' guessing into several levels: contextually inappropriate guess, wild guess, wrong guess, approximate guess, and contextually correct guess.

Variables and Hypotheses

During all four lexical inferencing segments of this ABAB study, the dependent variable consisted of participants' lexical guesses that were coded for relative accuracy.

The independent variable was the presence of a stimuli called the Word Hints reading resource, a type of monolingual help designed to extend and support lexical inferencing by providing contextual and linguistic clues to readers but not explicit (given) meanings. The independent variable was measured on two distinct levels, the semantic (a

"top-down" reading process) and the graphophonic (a "bottom-up" reading process), and these will be discussed in Chapters 2 and 3.

There are some hypothesized contextual, interlingual, and intralingual mediating variables for which design and procedural controls were established in this investigation:

Proximity of context clues to the target word were controlled by ensuring that supplementary clues in Word Hints were always provided in close proximity to the target word. A study by Carnine, Kameenui, and Coyle (1984) shows that a close proximity of the context clue to the unknown word is, from the reader's point of view, the most helpful and effective location.

The *amount of context* was controlled by maintaining equal passage lengths in the two reading conditions, passage only and passage plus Word Hints.

The *quality of context and semantic clues* was regulated through the use of a method devised by Mondria and Wit-de Boer (1991) that is described in Chapter 2.

Part of speech affects lexical inferencing difficulty (Na & Nation, 1985). Equal numbers of nouns, verbs, adverbs, and adjectives were used in both conditions of the study.

Sound-spelling sources of lexical confusion were addressed by the provision of aural help and "synform" (similar lexical form) clues in the Word Hints condition.

The needs of students experiencing *syntactic confusion* were addressed by the provision of grammatical clues in the Word Hints condition.

Help was provided that showed readers *inflections and/or derivations* of each word in the resource. Only inflections and derivations sharing the same context-specific meaning as the target word were provided, however.

Passage topic was controlled by selecting ten topics and assigning them to both conditions of the study. Even though the same set of topics occurred in both the passage only and passage plus Word Hints conditions, each passage from the total number of twenty passages was constructed independently as a unique piece of writing. (Note the sample matched pair in Appendix B.)

This study relied on *expository texts*. Research by Bernhardt (1990) and Hulstijn, Hollander, and Greidanus (1996) suggests that the use of narrative texts may discourage the look-up behavior of a lexical resource during contextualized reading. That is, readers may be less willing to interrupt the flow of a narrative text when they need to look up the meaning of an unknown word. Therefore, expository style texts are preferred for vocabulary research.

Words that were extremely culturally specific (e.g. references to holidays, rituals, consumer items, etc.) were avoided as target words. This helps to control for the problem of *referential gaps* in understanding.

Polysemous (multimeaning) *words exhibiting contrastive ambiguity* were excluded as target words in the study (Pustejovsky & Boguraev, 1996).

Spanish-English cognates were excluded as target words.

In addition to contextual and linguistic variables, this study addresses several learner variables described below.

Differences in the background experiences of the readers were partially controlled by the fact that there were 20 different reading passages selected for the study on 10 widely differing expository topics. This is a large enough sample of topics to view the background effect as an "average" response for each student. In addition, the study

presented both the experimental reading condition and the control reading condition to each subject.

The *reading level* of the participants was early-advanced to advanced. Candidates for the study were given a pre-experimental screening to determine whether they could read English materials at a secondary level. The Informal Reading Inventory (IRI) of the English as a Second Language Assessment Battery (ESLAB), an instrument designed to measure the reading level of secondary school bilingual students, was administered (Rivera & Lombardo, 1979a, 1979b). The ESLAB IRI includes a variety of reading comprehension questions for each passage, and the validity and reliability of this instrument will be described in Chapter 3. Students who read the seventh and eighth grade level passages at a minimum of the instructional level (75% of comprehension questions correct) were admitted to the study. Students at the early-advanced and advanced level were targeted since studies in dictionary use suggest that learners gain only modest benefits from monolingual dictionaries until they exceed an intermediate proficiency level (Hartmann, 1992; Piotrowski, 1989). Advanced-level students and professional translators typically report the greatest benefits from and reliance upon monolingual dictionaries (Baxter, 1980; Bejoint, 1981).

Depth of processing or mental engagement with the task was examined by measuring the frequency of lexical guessing. Contextualized reading tasks that are too difficult or do not provide sufficient help tend to result in learners who give up on the reading passage. The findings of Bensoussan and Laufer (1984) as well as those of other researchers suggest that large numbers of inferences made during reading indicates that

readers are on task and able to process material in a meaningful way. Thus, the frequency of lexical guessing was considered in this study.

Some mediating variables in contextualized reading that were not controlled for in this study include such things as age, first language ability, and the general intelligence of the reader.

The first two research questions of this investigation compare the quantity and quality of subjects' inferences in the passage only condition to the same subjects' quantity and quality of inferences in the passage plus Word Hints condition:

> Is there any difference between readers' overall frequency of lexical guessing in the passage plus Word Hints condition and the overall frequency of guessing in the passage only condition?
>
> Is there any difference between readers' accuracy of lexical guessing in the passage plus Word Hints condition and the accuracy of guessing in the passage only condition?

The third through fifth research questions explore the different kinds of strategies that individual learners use when reading and requesting help in the passage plus Word Hints condition. This is an important area to investigate because dictionary look-up behavior has long been described as idiosyncratic and tremendously varied according to individual needs and preferences (Baxter, 1980; Bejoint, 1981; Davis & Lyman-Hager, 1997; Tomaszczyk, 1979). The pilot study of this investigation, described in Chapter 3, found that individuals differed greatly in their look-up behaviors in regard to Word Hints. While sources such as Hulstijn (1993) suggest that readers do know when they need to look up a word, it is less clear from other sources (such as Haastrup, 1989, 1991) whether most readers can reliably judge what level of lexical help they need (for example, top-down versus bottom-up lexical help which is discussed in more detail in Chapter 2).

These research questions attempt to characterize the outcomes of particular reading strategies that are anticipated in regard to the use of Word Hints:

> If some readers have an "intensive use" strategy (greater total number of look-ups) of Word Hints, do they make better quality (accuracy) or quantity (overall number) of inferences compared to other readers in the group who have a "minimal use" strategy (few or no look-ups)?
>
> If some readers choose to access predominantly semantic (top-down) information during the passage plus Word Hints condition, do they make many bottom-up lexical errors that might have been prevented had they consulted the graphoponic level of Word Hints?
>
> If some readers consult both levels of Word Hints (semantic and graphophonic), do they make better quality (accuracy) inferences about those words compared to other readers in the group who show a distinct preference for accessing the semantic layer only?

These five research questions can be operationalized as follows:

Ho: There will be no significant difference in the mean number of lexical guessing attempts in the passage only conditions compared to the passage plus Word Hints conditions.

Ho: There will be no significant difference in the mean number of accurate lexical guesses in the passage only conditions compared to the passage plus Word Hints conditions.

Ho: There will be no correlation between either the mean number of readers' accurate lexical guesses or the mean number of lexical guessing attempts and their rank according to frequency of lookups in Word Hints.

Ho: There will be no correlation between reader's incidence (rate) of subsemantic or bottom-level errors and their rank according to a ratio of semantic to graphophonic help requests.

Ho: There will be no correlation between the mean number of readers' accurate lexical guesses and their rank according to a ratio of semantic to graphophonic help requests.

Summary

In conclusion, it has been shown that lexical guessing refers to the way in which readers infer the meaning of unknown words in a passage. Also, lexical guessing is

hypothesized as the primary means by which native and non-native learners acquire vocabulary through contextualized reading. During the lexical inferencing process, ESL students must deal with additional sources of confusion, interference, and general difficulty that hamper "word guessability." This study investigates a computerized lexical resource designed to (1) extend the inferencing process of second language learners during reading and (2) encourage them to make more frequent attempts as well as more accurate attempts at guessing unknown word meanings. Also, the effectiveness of this lexical resource is to be judged according to how well it mitigates some of the adverse factors (mediating variables) that make word learning difficult for second language learners. Specific research studies that describe these issues in more detail will be examined in Chapter 2.

CHAPTER 2
REVIEW OF RELATED LITERATURE

Measurement of Vocabulary Knowledge: Receptive Versus Productive Understanding

Any attempt to measure a person's vocabulary knowledge is ideally an effort to characterize his or her mental lexicon or a subset of that lexicon. Miller (1991) describes the task in this way:

> The word lexicon has two senses. One is synonymous with dictionary . . . The other is more abstract: the words of a language, whether or not they have been written down. . . .
>
> The term mental lexicon introduces still a third, more personal, sense. What you know, your personal word knowledge, is but a subset of the abstract lexicon, the lexical component of the language. The abstract lexicon can be thought of as the sum total of all the different words in all the mental lexicons of people who know and use the language. Nobody knows every word, but somebody knows each one.
>
> What does it mean to say that someone knows a word? Does it mean that they use it in speaking? In writing? Does it mean that they can define it? Or does it mean merely that they have seen it before? There are many words that a person can recognize in reading and might even use in writing, but would never utter or expect to hear in ordinary conversation. In a printed dictionary, a word is either on the list or it is not; in a mental lexicon, the edges are fuzzy. (p. 128)

As this quote by Miller suggests, the structure of our mental lexicon is affected primarily by two things. The first is the context in which we encounter a word and the associations that we attach to it. The second is the functional level at which we are able to use a word. For example, do we recognize the word in listening or reading, or can we produce the word in speech and writing? It is generally acknowledged that individuals often require repeated exposures to a word before it passes from receptive to productive

use, and many words in our mental lexicons are recognizable but never pass into productive use (Aitcheson, 1994; Channel, 1990; Miller, 1991; Sternberg, 1987). The argument for differential store of words in the mind is well-expressed by Channel (1990):

> We must recognize that for both the L1 and the L2 user of a language, the two distinct processes of production (whether speaking or writing) and comprehension (whether listening or reading), make differential use of the store of words in the mind. Part of the production process must consist of selection of appropriate words according to the meaning to be conveyed. The word form is then converted into a phonological shape for onward processing into speech. Thus the direction of mapping is meaning --> sound. In comprehension, it is the other way and the direction of mapping is sound --> meaning. . .These differences mean that, for the mental word store, the optimal arrangement for production will be according to meaning, while the optimal arrangement for comprehension will be according to sound. (p. 22)

Aitcheson (1994) regards as oversimplified any model that depicts the interplay of perception and production as a "simple tug-of-war." Although she acknowledges that the mental lexicon consists of two major components, semantic-syntactic and phonological, the boundaries between these and other language-related "storage areas" in the brain overlap:

> The mental lexicon, therefore, is a cobbled-together compromise in which the needs of production, perception and memory are all partially satisfied. But on closer inspection, it proves even messier, because of the way in which it overlaps with other aspects of cognition and language. (p.225)

By "other aspects" Aitcheson is referring mainly to the interaction of vocabulary knowledge with general cognitive abilities. Although these various areas within the mental model interact and have fuzzy boundaries, the general distinction between receptive and productive lexical function is an extremely useful one to apply to vocabulary measurement. Another important distinction can be made regarding the linguistic components measured by a vocabulary test.

Word knowledge involves a certain dimensionality in terms of linguistic components. Laufer (1990a) lists six components:

1. Form: recognizing the spoken and the written form; being able to pronounce and spell the word correctly
2. Word structure: recognizing the basic free morpheme and the bound morphemes; being able to produce some derivations of the word
3. Syntactic pattern in a phrase/sentence
4. Meaning: referential, affective (the connotation of the word), pragmatic (the suitability of the word in a particular situation)
5. Lexical relations of the word with other words, such as synonymy, antonymy, hyponymy
6. Common collocations (p. 148)

The present study relies on lexical inferencing data that are primarily concerned with knowledge of word meaning. However, a contextualized lexical inference also tests a reader's understanding of word form, structure, syntax, collocations, and lexical relations indirectly. Also, this study is concerned with lexical inferencing during contextualized reading as a receptive form of lexical function.

Measures Used in Studies of Incidental Vocabulary Learning and Lexical Guessing

Nesi (1987) and many other language teachers such as Wesche and Paribakht (1994) justifiably regard a production task such as "use the word correctly in a sentence or composition" as the most discriminating vocabulary measure. Productive vocabulary use is beyond the scope of the present study, however.

Researchers of incidental vocabulary learning, a closely related topic, have been generally concerned with learners' progress through the levels of vocabulary function. That is, their studies have controlled for such things as the quality of lexical encoding, contextual suggestiveness, and the number of exposures required for a new word to pass from recognition memory to recall to accurate productive use (Hulstijn, 1989, 1992;

Hulstijn, Hollander, & Greidanus, 1996; Jenkins, Stein & Wysocki, 1984; Knight, 1994; Konopak, Sheard, Longman, Lyman, Slaton & Atkinson, 1987; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985). This body of research often attempts to characterize vocabulary acquisition over a period of time.

However, researchers of lexical inferencing are typically concerned with a briefer time period within vocabulary acquisition and a more narrowly defined issue. That is, they are concerned with the problems and pitfalls of the initial phase of comprehension (prerequisite to word knowledge) where a reader or listener makes conjectures and tests interpretations of the meaning of an unknown word from some context (Bensoussan & Laufer, 1984; Haastrup, 1991; Hulstijn, 1993; Mondria & Wit-De Boer, 1991; Na & Nation, 1985; Schatz & Baldwin, 1986; Schouten-Van Parreren, 1992). In L2 research, this question is particularly salient since the danger that second language learners will make an incorrect inference regarding the meaning of an unknown word is much greater than for native language learners (Ittzes, 1991; Kelly, 1990; Laufer & Sim, 1985; Nesi & Meara, 1994; Scherfer, 1993). The scope of a lexical guessing study is rather narrow, and therefore, the measures used should be of a type that Wesche and Paribakht (1994) classify as interpretation:

> Interpretation involves more precise semantic analysis including the relationship of target words with other words in given contexts (collocations, synonyms, and antonyms). This further semantic analysis again contributes to aspects of input comprehension and provides information for intake and integration of new items into lexical networks.
>
> Examples include the following:
>
> **finding the odd word** in a series of collocationally related words.
>
> **understanding the meanings and grammatical functions** of the target word in the text (i.e., in a given context) and recognizing words or phrases which could be substituted in the text.

classifying words according to their discourse functions (e.g., discourse connectives classified by type--cause and effect, contrast, addition).

multiple choice cloze exercises.

guessing the meaning of target words in context. (p. 10)

These authors envision their classification scheme of vocabulary measurement as a hierarchy with four progressively difficult receptive tasks: selective attention, recognition, manipulation and interpretation. The final and most difficult vocabulary task in this classification is, of course, production. Wesche and Paribakht's (1994) characterization of interpretation as the highest-level receptive task is consistent with research indicating that a reader's own lexical inferences are better retained than given meanings (Hulstijn, 1992; Hulstijn, Hollander & Greidanus, 1996). This phenomenon is congruent with the hierarchical model since given meanings would demand recognition memory, a lower-order receptive task.

The present study incorporates an interpretation measure, lexical inferences made during reading in context. Lexical inferences reveal not only the semantic word knowledge of readers, but incorrect inferences can be analyzed to determine the source of a reader's confusion regarding word form, structure, syntax, or lexical relations.

Summary

Vocabulary knowledge is matter of degree, dimensionality, and function. There is a general hierarchy of word knowledge (described by Wesche and Paribakht, 1994) with productive knowledge being more difficult to master than receptive knowledge. For example, word recognition tasks are easier than tasks requiring the appropriate use of a word in a specific context. Lexical inferencing involves measurement of vocabulary at the highest receptive level which Wesche and Paribakht (1994) call interpretation. The

present study makes use of lexical guessing data that fit the category of vocabulary interpretation, a type of receptive measure. Such measures test primarily for word meaning with indirect emphasis on word form, structure, syntax, and lexical relations.

Research Related to the Independent Variable: Overview of Lexical Inferencing

The independent variable in this study is the Word Hints computerized lexical resource, a tool constructed of several elements designed to affect the lexical inferencing process. What steps or components are involved in the inferencing process? The following discussion attempts to describe the inferencing process in more detail. The contextual, linguistic, and learner variables which help or hinder inferencing will be discussed in the next section.

As Haastrup (1991) points out, inferencing is a term used differently by various academic disciplines such as philosophy, psycholinguistics, and pragmatics. This author believes that "pragmatic inferencing is most closely related to inferencing in a second language" and therefore favors the following definition:

> . . . inferences are connections people make when attempting to reach an interpretation of what they read or hear. We have also suggested that the more interpretive 'work' the reader (hearer) has to undertake in arriving at a reasonable interpretation of what the writer (speaker) intended to convey, the more likely it is that there are inferences being made. (Brown/Yule, 1983, as cited in Haastrup, 1991, p. 21)

Lexical inferencing, then, refers specifically to inferences at the level of words and morphemes and not larger units of language. Lexical inferencing may occur during listening or reading, and when it occurs, a kind of synthesis happens which suggests that the process can be broken down into component parts:

> It (inferencing) is intended to refer to a process of identifying unfamiliar stimuli. In foreign language learning inferencing is concerned with the acquisition of new morphemes and vocables in 'natural contexts.' In inferencing,

> attributes and contexts that are familiar are utilized in recognizing what is not familiar. (A. S. Carton, 1971, as cited in Haastrup, 1991, p. 22)

Sternberg (1987) does an excellent job of establishing universal subcomponents within the inferencing process. He describes three interconnected processes of vocabulary acquisition that work together to help one make an inference from context. The first process is called selective encoding, or the way in which one separates relevant from irrelevant information for the purposes of formulating a definition. In reading, this would reflect the tendency of a learner who is trying to guess the meaning of a word to pinpoint the exact location of relevant semantic clues. A second acquisition process is called selective combination, and it refers to the way in which the learner combines relevant cues into a workable definition. A third process Sternberg identifies is called selective comparison, and it refers to the way in which the learner relates new information about a word to background knowledge, old information already stored in memory.

The three processes of knowledge acquisition "do not operate in a vacuum or at random," says Sternberg (p. 91). They operate on a relatively stable set of cues provided by the contexts that contain the words:

1. Temporal cues regarding the duration or frequency of X, the unknown word, or when X can occur.
2. Spatial cues regarding the location of X or the possible locations where X can sometimes be found.
3. Value cues regarding the worth or desirability of X or the kinds of affects X arouses.
4. Stative descriptive cues regarding properties of X such as size, shape, color, odor, feel, etc.
5. Functional descriptive cues regarding the possible purposes of X, the actions it can perform or potential uses.
6. Causal/enablement cues regarding possible causes or enabling conditions for X.

7. Class membership cues regarding one or more classes to which X belongs or other members of one or more classes of which X is a member.
8. Equivalence cues regarding the meaning of X or contrasts (such as antonymy) to the meaning of X. (R. J. Sternberg, 1987, p. 92)

While Sternberg is trying to identify universal cue types, Haastrup (1991) is concerned specifically with second language inferencing cues. She describes three sources of second language inferencing cues established by A. S. Carton. Language learners combine these three types of cue sources when making an inference:

intralingual (the target language)

interlingual (L1 and other foreign languages apart from the target language)

contextual (the linguistic context + knowledge of the world) (p. 22)

The Word Hints lexical resource presented here does offer both intralingual and contextual cues, but since it is a form of monolingual help, interlingual cues are not addressed. In fact, L1 cognates were avoided as target words in the study. The specific nature of the intralingual and contextual cues will be described in a later section.

Mediating Variables Affecting Lexical Inferencing

The factors that make it sometimes difficult to apply the three knowledge acquisition processes to the eight context cues are mediating variables. Sternberg (1987) identifies six:

1. Number of occurrences of the unknown word.
2. Variability of contexts in which multiple occurrences of the unknown word appear.
3. Importance of the unknown word to understanding the context in which it is embedded.
4. Helpfulness of surrounding context in understanding the meaning of the unknown word.
5. Density of unknown words.
6. Usefulness of previously known information in cue utilization. (p. 92-93)

Numbers one through five are a fairly exhaustive list of the mediating variables characterizing the relationship of an individual word to its context in speech or text. Item six is one of the most important individual difference variables, the previous experience of the reader or listener. To this we could add the effect of language proficiency on lexical inferencing that is described by Haastrup (1991) and Schouten-Van Parreren (1992).

An entirely different class of mediating variables is described by Laufer (1990c, 1990b, 1991). These are intra-lexical, "within word," sources of difficulty which can potentially cause lexical confusion and affect inferencing:

1. The pronounceability of a word: this factor largely depends on the learner's L1 phonemic system.
2. The part of speech of a word: adverbs are most difficult to learn, verbs and adjectives are of moderate difficulty, and nouns are the easiest.
3. The inflexional complexity of a word: features such as irregularity of plural, gender of inanimate nouns, and noun cases, make an item more difficult to learn than an item with no such complexity.
4. The derivational complexity of a word: lack of regularity with which morphemes can or cannot combine create meanings or the multiplicity of meanings which can confuse the learner.
5. The degree of abstractness of a word: concrete words are generally easier to learn.
6. The degree to which a word is general versus specific: learners of a second language prefer tend to overuse general terms and avoid highly specific ones.
7. The idiomaticity of a word: idioms, or non-literal meanings, can be a big obstacle to fluent comprehension.
8. The presence of register restrictions: neutral words which can be used in all registers will be easier to learn.
9. The presence of multiple meanings: word polysemy or homonymy can create serious difficulties.
10. Whether or not a word is a false cognate: this type of words is very likely to cause lexical confusion.
11. Whether or not a word is a synophone or synograph: words which look or sound similar to other words are a primary cause of confusion among learners of a second language.

Out of these many potential sources of intralexical confusion, some emerge as more influential than others during the inferencing process. Student responses in the following study showed that five of the eleven intralexical sources of confusion were most influential during lexical guessing.

Bensoussan and Laufer (1984) conducted a lexical guessing study where 60 advanced ESL learners in Israel were given 70 words to translate into their native language. One week later, students were asked to translate the same words. However, this time the words were incorporated as part of a 574 word reading passage. Students translated the target words in the passage and answered comprehension questions. A control group was not given the 70-word list but only the words in context. The researchers found no difference between the experimental and control groups on performance. Apparently, seeing the words in a list before seeing them in context does not increase the ability of students to guess the meanings.

A major finding of this study is that context is not very helpful for lexical guessing. Bensoussan and Laufer found that of the 70 words in the passage, clear contextual clues were available for only 13 words, indirect context clues were available for 28 words, and no clues at all were present for 29 words. Given these numbers, the researchers determined that some form of context clue was available for 41 of the 70 words (clear clues plus indirect clues). These 41 words were designated the "guessable" ones. Of the guessable words, students only translated 17 of them successfully. Context, therefore, was said to account for only 17 out of 70 words or 24% of the total. When the raw scores of the in context condition were compared to the scores in context minus the isolated scores, a Pearson correlation of $r = .84$ was found. However, the authors still

dismiss the differences in the context condition scores as a marginal gain. The provision of context alone was not found to be any great boon to lexical guessing.

The other major finding of this study related to the kinds of errors students made when translating words. Virtually all of the wrong answers in this study fit into one of five categories: wrong choice of meaning of a polyseme, mistranslation of a morphological troublemaker, mistranslation of an idiom, confusion of a synophone or synograph, or confusion of a false cognate. These five types of wrong answers accounted for a total of 1,209 responses. There were 1,092 items left blank or not attempted. Also, there were 1,176 words categorized as previously known by the students. (answered correctly in isolation on the list and also in context). Finally, the study data include 1,134 other types of responses such as correct guesses, approximate guesses, and wild guesses.

From the results of this study it can be seen that several of the many mediating variables affecting lexical guessing emerge as particularly influential in a second language learning context. Bensoussan and Laufer (1984) as well as other authors (Haynes, 1993; Kelly, 1990; Laufer, 1990a; Schatz & Baldwin, 1986) have found written contexts to be distinctly unhelpful to language learners at the lexical level, and this can be a major barrier to successful inferencing. Thus, Sternberg's fourth mediating variable, degree of helpfulness of context (1987, p. 92-93), is shown to be very important in regard to second language learning. Furthermore, of all the intralexical sources of confusion described in Bensoussan and Laufer's (1984) study, morphological troublemakers, polysemes, and synphones accounted for the greatest number of specific intralexical source errors. Morphological errors were particularly frequent (714 out of 4200 possible errors), and in this type of error, a learner defines a word (such as a compound word

"outline") as the sum of its parts (literally to mean "out of line"). Learners' problems with idioms and false cognates accounted for a small proportion of errors.

Summary

In this section, eighteen different mediating variables affecting lexical inferencing were identified from a combination of sources (Haastrup, 1991; Laufer, 1990c, 1991; Schouten-Van Parreren, 1992; Sternberg, 1987). These mediating variables operate on a continuum. That is, the same variable may either help or hinder lexical inferencing. Described in the positive scenario, the following is a list of the optimal or ideal conditions for each mediating variable. In other words, the following is a list of factors that make lexical inferencing easier:

1. There are repeated occurrences of the unknown word.
2. The word appears in a variety of contexts and examples (but not so many as to overwhelm the learner).
3. The unknown word is important to the context.
4. The context is helpful.
5. There is a low density of unknown words in the passage.
6. The learner's past experience is useful in figuring out the unknown word.
7. The learner has high general language proficiency.
8. The word can be pronounced without difficulty.
9. The word is not an adjective.
10. The word has simple inflections.
11. The word lacks derivational complexity, or it is not a "morphological troublemaker."
12. The word is concrete rather than abstract.
13. The word is not used in an extremely specific way.
14. The word is not an idiom or part of an idiomatic expression.
15. The word can be used across all registers.
16. The word does not have multiple meanings.
17. The word is not a false cognate.
18. There are few or no similar sounding/looking words in the language.

One ESL study in lexical inferencing, Bensoussan and Laufer (1984), found that numbers four, ten, thirteen, fifteen, sixteen, and seventeen above were the major factors accounting for lexical confusion on a word-guessing task. This suggests that it may be possible to influence lexical inferencing by controlling for some of the key mediating variables. Controlling for all of these variables at once may not be necessary to produce a positive affect on a lexical inferencing task. However, this study has sought to affect as many of these mediating variables as possible during online contextualized reading in order to optimize the lexical inferencing process for second language learners. This was accomplished mainly by presenting a tool, a lexical inferencing resource, designed with several key mediating variables in mind. That is, the independent variable of this study, the Word Hints resource, addresses several key mediating variables in its design. Many other contextual, intralingual, and learner variables were controlled for by the methodology described in Chapters 1 and 3. Two things will be described in the next section: (a) the specific elements of the computer interface of Word Hints as well as the mediating variables which these elements address; and (b) the particular research on language and technology which informs this design.

Design Elements of the Lexical Inferencing Resource Known as Word Hints

One goal in regard to the design of Word Hints was to get the learner to pay attention to the form of an unknown target word as well as its meaning. Haastrup (1989) suggests that learners with successful inferencing strategies tend to apply both "bottom-up" and "top-down" strategies when analyzing unknown words:

> The second useful parameter for describing word processing is bottom-ruled and top-ruled. These are well-known concepts from the area of comprehension at sentence and text level . . . So it is only natural that they should apply to word-level processing as well. A hierarchy of cue levels was established

> with "the top" constituted by context and semantics, and "the bottom" ranging from the smaller units (orthography/phonology) to larger (collocations), reflecting an increasing meaning focus as you move upwards:
>
> | | |
> |---|---|
> | Top level | context |
> | | semantics |
> | --- | --- |
> | Bottom level | collocations |
> | | syntax |
> | | word class |
> | | word origin |
> | | lexis |
> | | morphology |
> | | orthography/phonology |
>
> The hierarchy of cue levels grew out of the data . . .(p. 39)

In terms of Haastrup's hierarchy, Word Hints provides lexical inferencing cues on several different levels.

The primary layer of Word Hints is that information which is directed at the top-down, semantic processes of the reader. When clicking on a target word in the reading passage, all participants were presented with the primary layer of information and had to click on another button to select the secondary (bottom-up) layer of cues if they so desired. Contextual and semantic clues were the top-down form of help presented by Word Hints in the first layer of the tool. As Scherfer (1993) says:

> The construction of typical contexts is particularly ambitious, because it demands an exact prior analysis of the learning material on the level of semantics, valency, collocation, and the psycholinguistics of verbal perception. In this field a lot of basic research still has to be done. (p. 1149)

Context has been variously defined in studies of incidental vocabulary learning and lexical inferencing as "syntactic units . . . declarative sentences . . . as a short story . . .as a rather longer authentic text . . . and . . . as, among other things, a multiple choice task" (Scherfer, 1993, p. 1146).

To offset some of the difficulties and variability of the "typical contexts" problem, the present study employed a scheme for controlling the quantity and quality of cues in a reading context that shall be described in detail below.

In this study, readers viewed target words in the contexts of 20 expository prose passages from 150-200 words in length. (These passages covered ten different topics, and an experimentally-released version of the Degrees of Reading Power test published by Touchstone Applied Science Associates provided a general model for the creation of these brief expository passages.) Additional contexts were provided to readers who chose to look-up the word in Word Hints. In this resource, the same semantic form of the target word was used in the construction of two usage examples. The two usage examples incorporated low-frequency words and everyday contexts to the greatest extent possible for each word.

A scheme for controlling the number of semantic cues was adapted from a study by Mondria and Wit-de Boer (1991). These researchers wanted to test a theory that controlled for the degree and kind of contextual suggestiveness, or as they call it, "contextual pregnancy." Their subjects were 139 Dutch pupils learning French. Students were assigned to one of eight conditions that varied context type. For example, while being asked to determine the meaning of the word "arrosoir," students might receive one of the following types of sentences:

1. *Le jardinier remplit un arrosoir pour donner de l'eau aux plantes*. The gardener filled a watering can to give the plants water. (Includes suggestive or "pregnant" subject, verb, and function.)
2. *Le jardinier remplit un arrosoir pour finir son travail*. The gardener filled a watering can to finish his work (Includes suggestive subject and verb.)
3. *Le jardinier cherche un arrosoir pour donner de l'eau aux plants*. The gardener looked for a watering can to give the plants water. (Includes suggestive subject and function.)
4. *Le jardinier cherche un arrosoir pour finir son travail*. The gardener looked for a watering can to finish his work. (Includes suggestive subject only.)
5. *Je remplis un arrosoir pour donner de l'eau aux plantes*. I filled a watering can to give the plants water. (Includes suggestive verb and function.)
6. *Je remplis un arrosoir pour finir mon travail*. I filled a watering can to finish my work. (Includes suggestive verb only.)
7. *Je cherche un arros foir pour donner de l'eau aux plantes*. I looked for a watering can to give the plants water. (Includes suggestive function only.)
8. *Je cherche un arrosoir pour finir mon travail*. I looked for a watering can to finish my work. (Includes no suggestive factors.) (p. 256)

Then they completed three stages. First, students translated eight target words that appeared in sentences. In this first group of sentences, all eight sentences demonstrated one of the eight kinds of contextualization. They also completed a control test of eight target words where each of the targets appeared once in all of the eight context types. Second, students completed a learning stage where they were given ten minutes to study the correct meanings of the words they had previously guessed. Third, they completed a test stage where they translated the eight target words in new, decontextualized sentences. The results were intriguing. Mondria and Wit-de Boer showed that the suggestiveness of subject, verb, and function influences correct or incorrect interpretation of a word. All three of the context types were statistically significant. Suggestiveness of function was particularly likely to lead to correct guessing. Two and three-way interactions were not significant except in the case of subject and function. The combination of a highly

suggestive subject with a highly suggestive function has a lesser effect on the guessing score than might be expected on the basis of the separate effects of those factors.

Given this precise method for analyzing semantic cues, the Word Hints lexical resource incorporated either a subject, verb, or function cue in each of two usage examples provided for every target word. This criteria ensured that at least two helpful semantic clues were present for every target word. Participants in a pilot study indicated that they preferred two usage examples as opposed to just one. For some low-frequency, highly specific target words, it was difficult to provide a usage example with an everyday context. Thus, two usage examples gave the learner a better chance at reading the word in a familiar context.

In Haastrup's (1989) hierarchy, learners need bottom level cues as much as top-level ones. As cited earlier, Channel (1990) reminds us that the optimal storage in the mental lexicon for comprehension tasks is according to sound, so graphophonic cues are not unimportant in lexical inferencing as Ittzes (1991) data show:

> The errors that students made in the list version were especially interesting in what they show of L2 students' difficulties. Goodman's miscue categories of graphophonic, syntactic, and semantic errors for L1 reading (Freeman, Freeman, & Goodman, 1987) were helpful in analyzing the Hungarian music students' errors and indicate the universal qualities of the reading process. The students' translations revealed that at times they had miscued the English word (base translated as if it were bake; sailor as if it were tailor; taken on board as if it were to stand guard). Other wrong guesses showed graphophonic interferences across the languages (scene vs. szen 'coal'; taming vs. temazni 'to talk about a theme, to chat'). (p. 364)

The secondary layer of Word Hints presented bottom-level cues, and it offered a computerized sound recording of each target word. In addition, learners could compare the sound and spelling of the target to the sound and spelling of synphones or synographs for each word. These are the graphophonic features of Word Hints that were presented in

the secondary layer of help cues. Since this secondary layer of cues was only viewed by readers who selected it, a computer usage tracker, also known as a query log, recorded details of readers' interaction with this information. The graphophonic layer of Word Hints emphasized information at the level of morphology and lexis since it presented all inflections and/or derivations sharing the same meaning as the target word in context. In addition, Word Hints listed, on the level of word class, the part of speech of the word as it was given in the usage examples.

Many linguistic features of words are present only on occasion. Some special features that can affect lexical inferencing were not addressed by Word Hints since help for these items need not be provided in the case of each and every target word: cognates or false cognates, morphological troublemakers, polysemes such as "bank" which show contrastive ambiguity, and idioms. These types of words were eliminated as potential target words.

Summary

The Word Hints lexical resource consisted of cues on two levels: both "bottom-ruled" and "top-ruled" types of linguistic features. Bottom-ruled processes addressed by Word Hints included target word pronunciation, inflectional and/or derivational cues, help with synographs and synphones, word class cues, and the modeling of proper syntax through usage examples. In regard to top-ruled processes, particular attention was paid to controlling the number of semantic cues available in the usage examples, and a scheme for analyzing semantic cues was adapted from a study by Mondria and Wit-de Boer (1991). The cueing features supplied by Word Hints addressed many of the mediating variables related to context that were described in the previous section. Many intralexical

mediating variables were controlled for by eliminating words with special difficulties (e.g. morphological troublemakers, cognates, idioms, and the most difficult level of polysemes).

Relevant Data Regarding Use of Computerized Lexical Resources

Research shows that because of convenience, the use of online glossaries and dictionaries encourages lookup behavior among readers and increases the efficiency of lookup behavior (Aust, Kelley, & Roby, 1993).

Several studies shed light on the lookup behavior of ESL learners, in particular. A study by Krantz (1991) with advanced ESL learners showed that users of monolingual online dictionaries learned about as many words incidentally in a contextualized reading task as did users of bilingual dictionaries.

Hulstijn (1993) examined the specific nature of lookup behavior in ESL learners and found that readers not only seemed to know when they needed to look up a word, but tended to look up words that were deemed relevant to the reading goal. In addition, it was found that words whose meanings could easily be inferred from contextual information were looked up somewhat less frequently than words whose meaning could not easily be inferred. Students with greater vocabulary knowledge tended to look up fewer words than students with lesser vocabulary knowledge. Hulstijn noticed further individual differences from the fact that some subjects tended to have a “maximal” lookup strategy, looking up nearly all the words while others had a “minimal” strategy in looking up few words. However, no significant differences were found between maximal and minimal subjects on a test of English vocabulary knowledge or on a test of inferring ability (a lexical guessing task).

Summary

Altogether, these findings suggest that the convenience of an online, hypertextual lexical resource can encourage word lookups. In the lexical guessing task of this study, readers were asked to focus on guessing the meaning of unknown words. If the target words used in the study were truly unknown to the readers, then lookup behavior would likely be high based on the findings of Hulstijn (1993). The findings of Krantz (1991) show that online monolingual lexical help is about equally as useful as bilingual lexical help to advanced-level learners.

Implications Leading to the Present Study

Hulstijn (1992) conducted a series of five experiments on incidental vocabulary learning during contextualized reading. The most significant finding is described in the following way:

> If L2 learners/readers read an L2 text for comprehension of its content, but not with the intention to expand their L2 vocabulary, then they are more likely to remember the form and meaning of an unknown word in the text when they have inferred its meaning by themselves (high mental effort), than when the meaning has been given to them (low mental effort). Evidence for this conclusion stems from Experiments III, IV, and V, using a within-subject design. (p. 122)

Given this finding, it is surprising that in Hulstijn, Hollander and Greidanus' 1996 study, the issue of inferred meanings and the types of cues most beneficial to inferencing is abandoned. Instead the authors compare two reading contexts offering given meanings (marginal glosses and a dictionary) to a control group. Although controlling contextualized semantic cues is ambitious as Scherfer (1993) points out, it appears worthwhile so that we may further examine what effect the availability of enriched semantic cues have on the inferencing process when combined with other "bottom-ruled" cues in the hierarchy described by Haastrup (1989). This leads to the rationale behind the

design of the Word Hints lexical resource: (a) it must withhold giving meanings explicitly in order to encourage inferencing, and (b) it must present a variety of cue levels with careful attention to contextualized semantic cues. This combination is expected to help readers attend to both the form and the meaning of an unknown word--which is necessary in order to help reduce the chance of faulty guessing. Second language learners tend to stubbornly hang on to preconceived notions of the meanings of words as Laufer and Sim (1985), Ittzes (1991) and Bensoussan and Laufer (1984) point out, and the only remedies to this problem seem to be either (a) increasing readers' attention to form, (b) providing direct instruction, or (c) training readers to use bottom-ruled strategies as well as top-ruled ones. Generally speaking, researchers in this area are just beginning to acknowledge the importance of an attention to both form and meaning during the inferencing process.

Research by Laufer (1990c, 1991) and Sternberg (1987) has helped to define the many mediating variables that affect the inferencing process. Therefore, it is logical to conduct an investigation such as the present study that aims to present a stimulus designed to positively affect the inferencing process by influencing key mediating variables.

CHAPTER 3
METHODOLOGY

Design and Procedures of the Pilot Study

For the pilot study, two women and five men were selected, from English as a second language students at the English House, a private teaching facility in Gainesville, Florida. The native languages of the subjects were as follows:

3 subjects	Spanish
2 subjects	Portuguese
1 subject	Korean
1 subject	Turkish

The subjects were screened using an Informal Reading Inventory (IRI) designed for this study. The IRI contained five reading passages ranging from readability level grades three to eleven. Directions for the reading inventory asked students to circle all the words they didn't know, rate the overall difficulty of each reading passage on a Likert scale, and answer a comprehension question for each passage. Subjects' teachers were consulted regarding the overall English proficiency of the subjects. Three subjects were judged to be at an intermediate level, and four subjects were advanced learners of English.

Two sets of computerized reading materials were prepared for the pilot study. The four advanced subjects were given two reading passages at a higher readability level

than the ones provided to the three intermediate readers. Within each set of reading materials, both passages were matched according to readability level and length.

Approximately one week after the initial screening, subjects met individually with the researcher and were introduced to a sample reading passage on the computer with Word Hints available. After an introduction to the computer program, subjects read the first 300-350 word passage for which no help was available. Following the reading, the researcher asked each subject to make guesses (orally) regarding the meanings of all unknown words in the passage. Subjects' responses were recorded. Then subjects read a second 300-350 word passage with Word Hints available. Again subjects were asked to make oral guesses regarding the meanings of all unknown words in the passage. Then subjects were interviewed regarding their opinions and feelings about Word Hints. After the interview, the researcher explained the meanings of words for which subjects had made unsuccessful guesses, and subjects were debriefed.

Computer data from the reading sessions of the pilot study indicate that the average number of words looked up (including repeats of the same word) was 7.57 for the passage with Word Hints. In addition, the Word Hints help screen contained buttons that played sound and provided graphophonic word clues. Counting both the semantic and graphophonic levels, the average number of interactions was 8.85 for the passage with Word Hints available. There was considerable individual variation in lookups with Word Hints for the single passage. The smallest rate was for one student who looked up only three words. The highest was for two students who looked up fifteen words each.

Computer data indicate that several subjects spent a considerably longer time reading with Word Hints. The results below are indicated in seconds:

Table 1
Pilot Study Participant's Reading Times With and Without Word Hints

Subject	Passage #1 Only	Passage #2 + Hints	Percent Increase
1	281	1329	473%
2	224	307	37%
3	334	1295	388%
4	523	1324	253%
5	312	323	4%
6	193	329	170%
7	212	308	45%

Note: Reading times are displayed in seconds.

Again, there was considerable variation in reading times. Subject 1, a high-intermediate level student, spent almost five times longer reading the passage that had Word Hints available. The reader with the lowest level of English proficiency (low intermediate), Subject 5, looked up only four words and spent only 4% more time reading when Word Hints was available.

All subjects in the pilot study expressed positive comments regarding the design and utility of Word Hints. All subjects indicated that Word Hints was easy to use and understand. All seven subjects indicated that they preferred reading with Word Hints available and said it was generally helpful. When asked how the tool could be improved, most subjects said that they would like to see more sound included in the tool. Also, one subject indicated that some usage examples in Word Hints could have been simplified further. Another subject said that he liked Word Hints "because it helped him work a little harder to figure out the meanings of unknown words." Subjects were asked to speculate regarding how they would feel should Word Hints be "taken away" after

having been available. Subjects indicated that they preferred Word Hints but wouldn't mind too much if they had to guess the meanings of unknown words in the passage with Word Hints "taken away" after having been formerly available.

Some subjects reading with Word Hints available seemed to show intense concentration. In general, it was observed that Word Hints appeared to be successful in getting readers to pay attention to both the form and meaning of words. From the pilot study it was impossible to determine whether the presence of Word Hints might increase the accuracy of subjects' lexical guesses. The main reason for this was the presence of too many English-Spanish cognates (known words) in the pilot materials. This limited the number of lexical guesses obtained for each individual.

Summary

Due to the problems with English-Spanish cognates in the pilot study, a new group of potential target words were collected for use in the main study that excluded as many Latinate words as possible. Overall, it was observed that the presence of Word Hints in the pilot study stimulated interaction with the reading materials and increased reading times.

Design and Procedures of the Main Study

The main study utilized a single-group ABAB design. Borg and Gall (1989) describe the advantages of this experimental design:

> This ABAB experiment involved a single group of subjects, but it has much better internal validity than the single-group experiments described in the previous chapter: the one-shot case study; the one-group pretest-posttest design; and the time series design. One reason for its superiority is that the treatment was instituted twice during the experiment. The other reason is that the dependent variable is measured many times during the experiment . . . By simply introducing the treatment twice during an experiment, the researcher can convert a time-series

> design into an ABAB design, and thereby make a more internally valid test of a treatment's effectiveness. (p. 719-720)

In the context of computerized studies, the ABAB also helps detect novelty effects. If the scores of the second B treatment decline compared to the first B treatment, this may be due to the novelty effects. In addition, in the ABAB design, the study has a positive outcome in that it ends with the provision of a second treatment rather than a withdrawal of treatment as in an ABA design.

Compared to designs that utilize separate groups for the control and experimental condition, the single-group ABAB enables a researcher to make many measurements and gather a lot of data when sample sizes of a population are limited. Also, the ABAB is useful in situations where the experimental population is diverse (as is often the case with bilinguals and international students) and when it is impossible to match a control group and experimental group according to background characteristics.

In this experiment, during the baseline condition, A, subjects were presented with reading passages only on the computer and required to guess the meanings of selected target words they did not know. Then subjects recorded their responses in their native language. The experimental condition, B, included the same computerized reading as condition A <u>with</u> <u>the</u> <u>addition</u> of a form of word help called Word Hints. Students could access Word Hints by clicking the mouse on any of the target words in this condition. After examining the word help, students recorded their responses in their native language. The ABAB design provided an opportunity for a thorough characterization of what happens to an individual's lexical guessing behavior when the Word Hints condition was introduced.

Altogether, the experimental procedures of the main study occurred over a five-week period. The prescreening and preparatory phase occurred over the first three weeks while the data collection phase occurred during the last two weeks. The exact order of the procedures will be described in the remainder of this section.

The Informal Reading Inventory (IRI) of the English as A Second Language Assessment Battery (ESLAB) was administered to participants as a pre-screening device (Rivera & Lombardo, 1979a, 1979b). The validity and reliability of this instrument is described under the section, Instrumentation and Data Collection, in this chapter. From the group of eligible participants who were able to read materials at a minimum of the seventh and eighth-grade equivalent level, 17 men and women were selected. At the first meeting of the study, students were given a brief questionnaire asking them information such as their country of origin, how long they have studied English, how long they lived in the United States, and what sort of materials and topics they routinely read. Then subjects were shown the list of possible target words and asked to put a check beside any words that they knew. If a student checked a word as being "known," they were asked to write a definition beside that word in either English or their native language. Known words were excluded from the set of target words. Then students were trained on how to use the computer program in the study. The training showed them how to use a mouse in order to navigate through the reading passages. Then students were instructed in the use of the Word Hints lexical resource. Easy, known words were demonstrated in the sample Word Hints. Then subjects were tested on their understanding of the content of the sample Word Hints as well as their navigation ability through word hints. Afterward, the

participants were asked to orally explain the different features of Word Hints as well as their purpose.

At the first session of the lexical guessing segment of the study, learners were asked to read each passage in its entirety, guess the meaning of underlined words, and record their responses. At this baseline session, A, students completed five passages with seven underlined words in each passage for a total of 35 responses. Target word density was 1 in 15 words. Passages from an experimentally-released version of the Degrees of Reading Power (DRP), a standardized reading test, were used as expository models for the construction of passages for this study. Twenty expository passages averaging 150 words each were written for this study, and within these passages, ten topics were utilized twice. That is, each of ten topics was used once in the Word Hints condition and once in the passage only condition.

At the second lexical guessing session, B, subjects were presented with another five reading passages that each had seven underlined target words. For this session, Word Hints <u>was</u> available.

The third lexical guessing session was another baseline condition, A, that tested the students' guessing ability with another 35 words, and then the fourth lexical guessing session was another Word Hints or B condition that presented the final 35 of 140 total target words.

After the fourth lexical guessing session of the study, each subject was debriefed, informed about the computerized record of his or her reading, and provided with a list of definitions for all of the target words in the study. Thirteen of the seventeen participants completed the study.

Sample

The subject pool was selected from English as a second language (ESL) students who attended the English House in Gainesville, Florida. The English House is a non-profit resource center that offers conversation classes and media lab access to foreign students for a low fee. The English House has an extremely diverse group of native languages represented, and the total student enrollment varies seasonally.

Instrumentation and Data Collection

The prescreening measure consisted of a validated reading test for bilinguals, the Informal Reading Inventory (IRI) of the English as a Second Language Assessment Battery (ESLAB) (Rivera & Lombardo, 1979a, 1979b). The ESLAB was "developed to meet the need for a valid and reliable criterion-referenced instrument to assess the English language proficiency skills of secondary bilingual students (p.1)." The ESLAB Informal Reading Inventory (IRI) consists of a set of eight original passages ranging in length from 30 to 200 words. The authors used the Fry (1968) and Dale-Chall (1948) readability formulas to designate passage levels ranging from primer to eighth grade. In the ESLAB IRI, students read the graded stories silently and then write their answers to the questions at the end of each story. Students are advised to stop reading when they no longer comprehend.

The overall reading level is scored according to the following criteria:

Independent = 95 - 100%

Instructional = 75 - 94%

Frustration = 0 - 74%

The ESLAB IRI was validated on a sample of 59 inner-city Hispanic students in the seventh and eighth grade. These students ranged from ages 12 to 15. While the ESLAB IRI showed only a modest correlation with ESL grades (.24, significant at the $p < .05$ level; and .36, significant at the $p < .01$ level), its concurrent validity was good. The ESLAB IRI results correlated very well with the ESLAB Cloze Test (.74 at the $p < .01$ level) and also with the Stanford Diagnostic Reading Test (.62 at the $p < .01$ level). In addition, Rivera and Lombardo found the ESLAB IRI to be a very reliable test. In particular, the authors wanted to make sure the test was "measuring linguistic skills and not, for example, intellectual functioning, or attitudes (p.14)." The index of reliability for the ESLAB IRI was .83 for Hoyt's ANOVA and .74 for Cronbach's alpha.

During the lexical guessing segment of this study, students wrote down their guesses on paper (clipboard) as they participated in the reading activity at the computer. When Word Hints was available, students could click on any of the specially marked target words to receive additional help. Two "layers," or help windows were provided for each target word, and samples of these are included in Appendix C. The primary window layer offered two usage examples with semantic clues controlled by the method described in Chapter 2. At the bottom of the primary window layer were two buttons. One button enabled the user to return to reading. The other button enabled the user to select "more help." The more help button called up the secondary window layer that included graphophonic and grammatical clues described in Chapter 2. This information was arranged in two columns, with the part of speech of the target (grammatical clue) at the bottom. The left column presented inflections and/or derivatives of the target word with a sound button at the top that played a recording of these closely related words. The

right column presented synforms, words which are spelled alike or sound similar to the target but which do not share the same meaning. At the top of the right column was a sound button that played a recording of the synforms. At the bottom of the secondary window layer was a button that enabled users to return to the reading passage.

As subjects used Word Hints, the computer unobtrusively tracked record of everything the student clicked on. In addition, the computer created reports of total button clicks, the kind and number of particular kinds of buttons clicked, total reading time, time spent on each page, and time spent on word help.

Data Analysis

Participants' lexical guesses were recorded in the individual's native language and later coded on a scale adapted from the coding method used by Bensoussan and Laufer (1984). Two qualified independent raters were found for each language group represented in the study, and the raters scored the accuracy of target word guesses on a scale from zero to four. A score of zero reflected no attempt on the part of the reader to guess the meaning of the word, and a score of one was assigned to wild guesses having no apparent connection to the topic. Wrong guesses were assigned a score of two, and if a guess was judged close to the ideal meaning and contextually appropriate, a score of three was assigned. In order to achieve a score of four, a guess had to reflect the precise meaning suggested by the story context, and it had to be denotatively and connotatively correct. In addition to judging the accuracy of guesses, the paid raters recorded translations of wrong guesses for later analysis by the researcher. Each pair of independent raters was proficient in English as well as the native language of the participants, and the raters were provided with a key and scoring scheme in English. Each rater translated this key, scored the

participant's data, and then he or she met with the other rater to finalize any scores that did not concur. Although some raters complained that a few English words did not have an exact foreign equivalent, all pairs of raters were able to reach 100% concurrence through discussion of the intended meanings of the passages.

Participants' passage only, condition A, guesses were compared to passage plus Word Hints, condition B, guesses with a Repeated-Measures Analysis of Variance (ANOVA). The Repeated-Measures ANOVA was also used to compare frequency of lexical guessing in the A and B conditions. A paired t-test was used for *Post Hoc* analysis. Bivariate correlations were used to examine the secondary set of hypotheses concerning individual difference variables, and subject rankings were compiled that enabled comparison of the behavioral use of Word Hints to data concerning frequency and accuracy of guessing.

Summary

To summarize, the proposed study used a single group ABAB design. Subjects recorded their lexical guesses in their native language, and these were coded and analyzed for each subject. The computer recorded the look-up behavior of subjects in the Word Hints condition. Data regarding "number of button clicks" was analyzed. Each individual's total number of clicks within Word Hints, his or her total number of requests for help on different words, tally of the kind of help requested, the time spent on each lexical guessing session, and the total time spent on Word Hints were compiled. Strict procedures were followed to ensure that properties of the texts and target words were equivalent in both conditions of the study and across all sessions of the study. Some of the controls regarding text were described in Chapters 1 and 2. Subjects were selected

from advanced-level readers, those students who are most likely to benefit from lexical monolingual English help.

The purpose and design of this study are believed to be original. This investigation is not an exact replication of any other studies mentioned herein.

CHAPTER 4
RESULTS AND ANALYSIS

Primary Research Questions

The two primary research questions of this study sought to discover whether subjects made a greater number of lexical inferences as well as better quality inferences when help, passage plus Word Hints, was available. Participants' lexical guesses were recorded and scored according to the procedures described in Chapter 3. Lexical guessing scores were collected into tables for each individual participant, and the number of guessing attempts, story means, and moving averages was calculated. A preliminary analysis of the data tables of individual participants revealed no trends in moving averages that would indicate improvement due to story effects or time. Since story effects were not found, data analysis for the group proceeded as planned.

This study utilized a within-subjects ABAB design, and the first baseline and treatment, AB, occurred in Week 1. The second baseline and treatment, AB, occurred in Week 2. This structure meant that the time-series data of the study were nested into four sessions or blocks of interest: No Help (passage only) Week 1; Help (passage plus Word Hints) Week 1; No Help (passage only) Week 2; and Help (passage plus Word Hints) Week 2. In order to better manage the repeated measures data for the group, block means were calculated for both the number of guessing attempts and scores (accuracy).

For the first hypothesis concerning frequency of guessing, results of the data analysis found the mean number of participants' guessing attempts to be higher in both

Help conditions. A maximum of seven guesses was possible for each of twenty stories. Story means were calculated, and means for each session or block were averaged. Table 2 shows the means and standard deviations for each session in chronological order:

Table 2
Means and Standard Deviations of Lexical Guessing Attempts By Session

Session	Mean	Standard Deviation	N
A^1 No Help Week 1	5.74	1.03	13
B^1 Help Week 1	6.14	0.88	13
A^2 No Help Week 2	5.60	1.34	13
B^2 Help Week 2	6.54	0.55	13

A Repeated Measures Analysis of Variance found that a significantly greater number of guessing attempts was made when help in the form of Word Hints was available. Table 3 shows that Help was significant, $F(1,12) = 17.79$ at the $p < .01$ level:

Table 3
Repeated Measures ANOVA Results for Guessing Attempts by Help and Time (Week) Variables

Source	Df	Type III SS	MS	F
Help	1	5.82	5.82	17.79*
Error (Help)	12	3.93	0.33	
Week	1	0.22	0.22	0.61
Error (Week)	12	4.41	0.37	
Help * Week	1	0.94	0.94	2.68
Error (Help * Week)	12	4.21	0.35	

*$p < .01$

The mean number of attempts at guessing was not significantly different between Weeks 1 and 2. Also, there was no significant interaction of the help condition with scores for Week 1 and 2.

Results of the data analysis for the second hypothesis found that participants' mean scores for accuracy in guessing were higher for both Help conditions of Week 1 and 2. A maximum of 4 points was possible for each of the 140 unknown words in the story. A score of 4 represented a highly accurate guess, 3 was awarded for a near guess, 2 represented a wrong guess, 1 designated a wild guess and 0 meant that the student made no attempt to guess. Story means were calculated, and then story means for each block were averaged. Table 4 shows the means and standard deviations for each session or block in chronological order:

Table 4
Means and Standard Deviations of Lexical Guessing Accuracy By Session

Session	Mean	Standard Deviation	N
A^1 No Help Week 1	1.86	0.36	13
B^1 Help Week 1	2.19	0.41	13
A^2 No Help Week 2	1.78	0.48	13
B^2 Help Week 2	2.56	0.37	13

As with guessing attempts, the means for accuracy are higher in both Help conditions, and when Word Hints was "taken away" for Session 3, No Help during Week 2, participants' mean scores drop to 1.78, the lowest average of the four sessions.

A Repeated Measures Analysis of Variance found that participants made significantly more accurate guesses when Word Hints was available. Table 5 shows that Help was significant, $F(1,12) = 75.28$ at the $p < .01$ level:

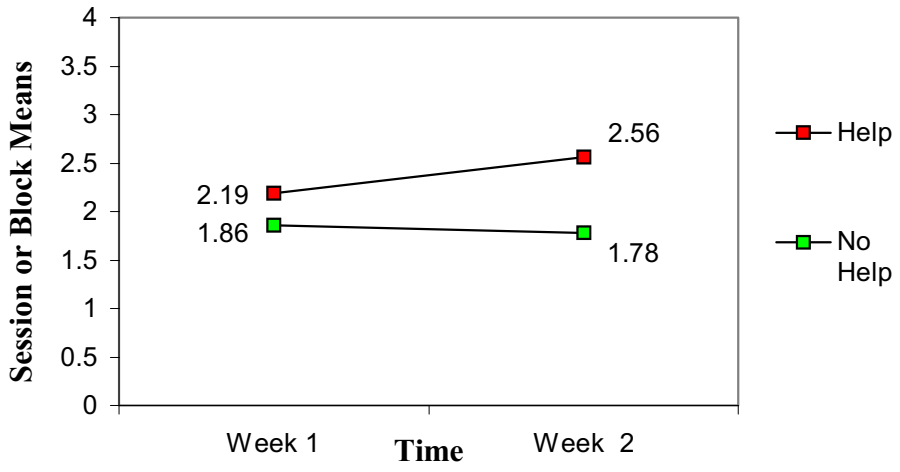
Table 5
Repeated Measures ANOVA Results for Guessing Accuracy by Help and Time (Week) Variables

Source	Df	Type III SS	MS	F
Help	1	4.00	4.00	75.28*
Error (Help)	12	0.64	5.32	
Week	1	0.27	0.27	4.43
Error (Week)	12	0.73	6.07	
Help * Week	1	0.65	0.65	18.02*
Error (Help * Week)	12	0.43	3.59	

*p < .01

Subjects' accuracy of lexical guessing was not significantly higher overall during Week 2 compared to Week 1. However, accuracy scores for help improved during Week 2 as the No Help scores decreased over the same period. Figure 1 shows the interaction of lexical guessing accuracy scores for the Help and No Help conditions by time (Week):

Figure 1



Accuracy of Lexical Guesses over Time: Help by Week Interaction Effect

To further examine this interaction, several paired samples T-tests were conducted for *post hoc* examination. Table 6 shows that 3 out of 4 pairs of session or block means were significantly different at the p < .01 level:

Table 6
Post Hoc Comparison of Paired Samples T- Tests Describing Help by Week Interaction

Paired Differences	Df	T	Cohen's D	Power
Help Week 1 & Help Week 2	12	5.62*	0.92	0.64
No Help Week 1 & No Help Week 2	12	-0.77	0.19	0.08
No Help Week 2 & Help Week 2	12	9.29*	1.78	0.99
No Help Week 1 & Help Week 1	12	4.06*	0.86	0.56

*p < .01

The results of the paired samples t-tests indicate that the greatest difference in means, t (12) = 9.29 at the p < .01 level, occurred between the accuracy scores of Sessions 3 and 4, the No Help and Help sessions presented in Week 2. Effect size differences between these two means was very large with Cohen's D = 1.78, (Cohen, 1988), and power = .99. Overall performance on accuracy dropped during Session 3, but when help was restored during Session 4, the group means were significantly higher than they were for the first Help session in Week 1. Therefore, the second largest difference in means, t (12) = 5.62 at the p <.01 level, occurred between Help in Week 1 and Help in Week 2. Again, since the mean accuracy scores of Help in Week 2 were very high, Cohen's D was equal to .92 and power was .64 in comparison with Help during Week 1. Another large difference was observed, t (12) = 4.06 at the p <.01 level, when the mean accuracy scores were compared for both sessions of the first week, Help Session 1 and No Help Session 2. Cohen's D for this comparison was equal to a respectably large .86 and power was equal

to .56. The difference between the two No Help conditions, Sessions 1 and 3, was not significant, and this lends strong support to the interpretation that improvement in participants' accuracy scores was not due to the effect of time (Week) or practice but a result of the presence of the Help condition.

Summary

The first null hypothesis of the study stated that there would be no significant difference in the mean number of lexical guessing attempts in the passage only (No Help) conditions compared to the passage plus Word Hints (Help) conditions. This hypothesis was rejected due to the results of a repeated measures ANOVA that showed guessing attempts were significantly higher when Help was offered in the passage plus Word Hints condition, $F(1,12) = 17.79$, $p < .01$.

The second null hypothesis of the study stated that there would be no significant difference in the mean number of accurate lexical guesses in the passage only (No Help) conditions compared to the passage plus Word Hints (Help) conditions. This hypothesis was also rejected due to the highly significant result of a repeated measures ANOVA, $F(1,12) = 75.28$, $p < .01$, which indicated participants' lexical guessing scores were much higher in the Help conditions and therefore their guesses were more accurate. A *post hoc* analysis of means was conducted utilizing several paired samples t-tests. Results of the *post hoc* revealed that lexical guessing accuracy scores were higher during the Help session of Week 2 compared to the Help session of Week 1, $t(12) = 5.62$, $p < .01$. Therefore, participants were not only able to make more accurate guesses when Word Hints was available, but they were able to obtain increased benefit from exposure to

Word Hints over time. Conversely, exposure to the passage only conditions did not result in more accurate lexical guessing and provided no benefit over time.

Secondary Research Questions

The remaining research questions of this study concerned the effect of individuals' reading strategies upon guessing accuracy and frequency.

The third hypothesis of the study asked whether there would be a significant correlation between either the mean number of participants' accurate lexical guesses or the mean number of lexical guessing attempts and participants' rank according to frequency of lookups in Word Hints. In other words, would frequent use of Word Hints (as defined by many clicks on the tool) correlate with either a high or low number of guesses or accuracy of guesses? From the data that were collected about reading behaviors, it was possible to calculate participants' total Word Hints lookups according to two different methods. The first method, total clicks, calculated frequency of lookups by counting the total number of an individual's computer clicks on both levels, the semantic and graphophonic, of Word Hints. The second method, unique information clicks, calculated frequency of lookups according to the number of different words an individual click on from within the semantic and graphophonic levels of Word Hints. This unique information method of compilation excluded multiple clicks or "practices" on the same word, a strategy that several participants seemed to prefer during lexical guessing. Although, the second method is a more conservative measure of lookup frequencies, relatively minor differences in ranking resulted from the two methods. Table 7 shows the frequency of lookup and rank of each participant according to the two methods of compilation:

Table 7
Subjects' Frequency of Word Hints Lookup and Rank by Compilation Method

Subject	Total Clicks	Rank According to Total Clicks	Unique Info Clicks	Rank by Unique Info Clicks
1	116	6	84	6
2	169	3	104	2
3	100	9	59	12
4	80	12	64	11
5	107	7	74	7
6	161	4	92	5
7	94	11	73	8
8	95	10	71	9
9	251	1	129	1
10	100	8	69	10
11	160	5	95	4
12	61	13	44	13
13	184	2	102	3

Regardless of the method used for ranking, frequent lookup behavior within Word Hints correlated significantly with increased mean guessing attempts during Session 2, the first time that Word Hints was deployed, but correlation was not significant during Session 4, the second time Word Hints was available in Week 2. Highly accurate guessing scores were negatively correlated with frequent lookup behavior in Help Sessions 2 and 4. Table 8 shows the correlations for both accuracy and frequency with ranking by two methods:

Table 8

Subjects' Rank According to Frequency of Word Hints Use Correlated with Lexical Guessing Attempts and Accuracy

Mean Scores w/ Word Hints	Rank by Total Clicks	Rank by Unique Info Clicks
Lexical Guessing Attempts		
Session 2, Week 1	0.60*	0.64*
Session 4, Week 2	0.42	0.43
Lexical Guessing Accuracy		
Session 2, Week 1	-0.26	-0.26
Session 4, Week 2	-0.63*	-0.63*

*p < .05

It is interesting to note that while there was a significant correlation between mean guessing attempts in Session 2 and rank according to frequency of lookup, there was not a significant correlation in Session 4. A plausible interpretation is that the initial introduction of Word Hints in Session 2 was associated with both a high level of computer mouse-clicking behavior and a high number of guessing attempts for some participants due to novelty effects. As Table 2 indicates, the overall mean number of guessing attempts for the group was actually highest in Session 4. However, since frequency of lookup was only moderately correlated with the number of guessing attempts made in Session 4, some subjects in this study apparently developed a more strategic or streamlined use of Word Hints as time went on.

In regard to accuracy of lexical guessing, subjects who clicked less within Word Hints (lower rank on clicking behavior) had a higher mean score on accuracy of guessing. These subjects still utilized Word Hints effectively and achieved more accurate guesses

when it was present. However, computer tracking data indicate that these subjects were less likely to click on the graphophonic level—presumably because they were better able to process contextual information at the semantic level with reduced need for lower-level clues. If the more proficient readers were aware that they did not need graphophonic cues from Word Hints, then it would follow that those who struggled with syntax and various lower-level issues were aware that they did need help at the graphophonic level. In fact, the computer usage tracking data confirm that subjects who experienced difficulty with reading passages as evidenced by lower guessing accuracy scores accessed graphophonic cues more often than the others. These individuals ranked high on frequency of Word Hints lookup but also had less accurate guesses about the meanings of unknown words.

The fourth hypothesis of the study questioned whether or not there would be a correlation between a reader's rate of bottom-level or subsemantic errors and his or her rank according to a ratio of the rate of semantic level accesses of Word Hints over the number of subsemantic accesses. The subsemantic level was represented by the sound and spelling cues of the graphophonic layer of Word Hints. As with the third hypothesis above, two methods were utilized for compiling computer behavior clicks. A ranking was calculated whereby a subject's preference for semantic level access was estimated by dividing his or her total number of computer clicks on the semantic layer of help by the total number of clicks on the graphophonic layer. Likewise, a similar ranking was calculated by dividing the total number of unique information (different word) requests on the semantic layer with the total number of unique information requests on the graphophonic layer. A bottom-level error score was calculated for each individual by

counting the total number of guessing attempts made during Help Sessions 2 and 4 and dividing this into the number of subsemantic errors. Subsemantic errors constituted a subset of all error types compiled in the study, and these included mistakes in grammar, syntax, or word class; confusion of synophones or synographs; morphological troublemakers, e.g. a word like skyscraper; lack of recognition of collocational phrase restrictions; and translation of a collocate or word other than the target word. Utilizing the unique information method of counting computer clicks, a Pearson correlation score of $r = .11$ ($p < .05$) was found to exist between participants' rank on semantic divided by subsemantic accesses and their rate of bottom-level error. Likewise, when the total click method of counting was used, an equally low Pearson correlation score of $r = .13$ ($p < .05$) occurred between participants' rank on semantic divided by subsemantic accesses and their rate of bottom-level error. One prediction concerning the fourth hypothesis was that readers who exhibit a strong preference for the semantic level (high semantic layer clicks over low graphophonic layer clicks) might make an increased number of bottom-level errors that could have been prevented had they accessed this level more often. This expectation proved to be false, and no relationship was shown between rate of subsermantic level error and behavioral preference for accessing the semantic level. This result supports the conclusion found in the third hypothesis, above, that the more proficient readers processed information well at the semantic level, made accurate lexical guesses, and often ignored the graphophonic layer of Word Hints. It is not surprising, therefore, that they failed to make many subsemantic errors. Another reason this correlation may have proven to be low is that the group as a whole achieved a comparatively moderate rate of subsemantic errors compared to errors as a whole. Table

9 shows the incidence of error types for all subjects in both the Word Hints or Help condition and also the No Help condition:

Table 9
Total Number of Lexical Guessing Errors by Error Type and Condition

Score Code and Error Type	With Help	No Help	Total Errors
2a: incorrect grammar, syntax, or word class	52	93	145
2b: confusion of synophone or synograph	20	30	50
2c: morphological troublemaker	2	4	6
2d: wrong choice of polyseme	0	0	0
2e: doesn't recognize pos./neg. connotation	15	15	30
2f: superordinate – too general, catch-all word	49	58	107
2g: cohyponyms – too specific word	2	3	5
2h: translating the opposite meaning	13	10	23
2i: uses functional substitute in the story	233	217	450
2j: ignores collocational phrase restriction	27	47	74
2k: translates keyword other than target word	22	35	57
2l: semantic relations – doesn't fit structure	10	21	31
Total Errors	445	533	978

All bottom-level errors, type 2a, 2b, 2c, 2j, and 2k, comprised 18.2% of all possible responses and 34% of participants' total errors. The most frequently made lexical guessing error was the use of a functional substitute, a word different from the target word but which nonetheless fits the reading context. Functional substitute guesses comprised 46% of all errors, and this error type occurred slightly more frequently during the Word Hints or help condition. Most other errors were slightly less frequent during the presence of Word Hints. In fact, it appears that Word Hints was useful in helping

some subjects avoid errors relating to grammar, word class, confusion of synophones and synographs, superordinates, collocations, and semantic relationships. Nevertheless, as could be expected with a lexical guessing study, error rates in guessing were high overall with the 978 lexical guessing errors comprising 53.7% of all possible responses by participants.

The fifth hypothesis was the final question pertaining to individual strategies and lexical guessing, and it investigated whether there would be any correlation between the mean number of a reader's accurate lexical guesses and his or her rank according to a ratio of semantic divided by graphophonic help requests. The subjects of interest in regard to this question were those who ranked low in a ratio of semantic to graphophonic help requests. If an individual ranked comparatively low on this ratio, then he or she was using both the semantic and graphophonic levels of Word Hints more evenly in comparison to those with a high rank. Did these low-ranking individuals make more accurate inferences?

Subjects' rankings were compiled according to the two methods described in the third and fourth hypothesis, the unique word method of compilation and the total click method. With regard to both compilation methods, the ratio of words accessed at the semantic level was divided by words accessed at the graphophonic level. The rankings were compared to subjects' lexical guessing accuracy scores for Help Sessions 2 and 4. A lower ranking number, e.g. one, meant that a participant used the semantic level of Word Hints much more frequently than the graphophonic level. Table 10 shows the Pearson correlations between ranking method and lexical guessing accuracy scores in the

two Help sessions. All correlations in the table are negative because lower rankings equated with higher scores for lexical guessing accuracy:

Table 10

Subjects' Rank According to the Ratio of Semantic to Graphophonic Lookups Correlated with Lexical Guessing Accuracy

Mean Word Hints Scores	Semantic Total Clicks Divided by Graphophonic Total Clicks	Semantic Unique Info Clicks Divided by Unique Info Graphophonic Clicks
Lexical Guessing Accuracy		
Session 2, Week 1	-0.64*	-0.66*
Session 4, Week 2	-0.68**	-0.62*

*p < .05; **p <.01

Regardless of the method of compilation, total clicks or unique information clicks, a rather large effect is observed with the various product moment correlation coefficients exceeding .60 for both Help conditions, Sessions 2 and 4, at the $p < .05$ level. Subjects who clicked almost exclusively on the semantic level were more likely to have higher scores for lexical guessing accuracy. Conversely, subjects who clicked a lot on both levels of Word Hints were likely to have lower scores on lexical guessing accuracy.

In the course of examining this relationship between levels of help accessed and scores, it was observed that the same five individuals ranked high on semantic over graphophonic access regardless of whether the total click method or the unique information method of compilation was used. Likewise, another five individuals ranked low on semantic over graphophonic access regardless of the method of compilation. These two groups had noticeably different scores on accuracy of lexical guessing. A paired sample t-test of lexical guessing mean scores revealed that $t(9) = 8.04$ at the $p < .01$ level. Although the sample size of this test was small, the effect was unusually large.

For this t observed score Cohen's D = 2.34, and this corresponds to a Pearson product-moment correlation coefficient of .75. Therefore, both t-test and correlation data affirm the connection between proficient lexical guessing and requests for predominantly semantic level information.

Summary

The secondary hypotheses of this study were concerned with the relationship between participants' accuracy and frequency of lexical guessing and the manner in which they used Word Hints. It was shown that readers who consulted Word Hints less frequently than average were more likely to consult the semantic level word help exclusively and also have higher scores for lexical guessing accuracy. Individuals who used both graphophonic and semantic level help from Word Hints tended to have lower scores. No relationship was found between the level of word help requested, semantic versus graphophonic, and the rate of graphophonic errors. However, several kinds of graphophonic and semantic errors were less frequent when Word Hints was available. In particular, fewer errors with grammar, word class, synophones and synographs, superordinates, collocations, and semantic relations occurred in the presence of Word Hints. Overall, the data suggest that readers were competent judges of the kind of help they needed, semantic or graphophonic, and they also knew when they needed to get clues from Word Hints.

CHAPTER 5
CONCLUSION

Future studies in lexical inferencing could explore many different avenues that would develop our understanding of basic vocabulary acquisition processes. For example, it remains to be seen whether prolonged periods of reading with an inferencing support tool like the one described in this study could improve students' retention of incidentally learned vocabulary. Also, further research could clarify the differences between contextualized reading and contextualized listening activities and the nature of incidental vocabulary learning. In addition, little is understood about individual difference variables among ESL and EFL students and this process. Further study needs to clarify the relationship between various primary languages and inferencing in a second language, and although we could speculate that language proficiency and lexical inferencing ability is positively correlated, not enough is known about learners' proficiency level in L1 and L2 and the lexical inferencing process. Readers and listeners have very distinct individualized learning strategies and preferences, and further investigation could explore the ways in which individuals learn to use contextualized lexical inferencing as an independent study strategy.

In this chapter, it will be argued that lexical inferencing activities could potentially be used in a meaningful way to extend classroom learning. The readers in this study, it should be remembered, made intelligent inferences about the meanings of many unknown words in a passage with no assistance other than the reading passage and Word Hints to guide them. Without the advantage of a teacher or even a dictionary, they made

many successful guesses about unknown words in a second language. Even under improved guessing conditions though, some of the inferences the readers made were still incorrect or they were unable to guess at all. So, how do we put into perspective the results of this study, and under what conditions might we best utilize lexical inferencing as an independent study strategy?

One clear implication of this study is that the results strongly support Mondria and Wit de Boer's (1991) finding that the suggestiveness of subject, verb, and function in a sentence influences one's correct or incorrect interpretation of a word. It was the carefully controlled nature of Word Hints usage examples that enabled participants of this study to view each target word in a total of three different contexts. That is, they read the target word in the original passage plus had two usage examples that illustrated the same sense of the word. Careful construction of the usage examples was essential to providing second language learners with contextualized meaning that was truly helpful and suggestive. The fact that participants in this study achieved a greater number of lexical guesses and more accurate guesses with Word Hints suggests that the vocabulary acquisition processes of second language learners can be influenced through the use of controlled or modified reading materials. And this is a finding that may serve to rekindle interest in promoting vocabulary acquisition through contextualized reading.

Contextualized approaches, especially non-instructional ones that advocate extended reading as a means to learn new vocabulary, have been subject to a good deal of criticism, and since the obstacles L2 learners can face during reading are numerous, much of this criticism is well-deserved. Huckin & Haynes (1993), frustrated with the difficulties that authentic texts pose for second language learners, effectively dismissed

inferencing activities as a viable strategy for L2 vocabulary learning:

> Thus, it appears that heavy reliance on contextual guessing may not be the optimal learning strategy for nonnative speakers. There is some controversy about whether contextual guessing is the optimal approach for *first*-language learners (see Pressley, Leven & McDaniel, 1987 versus Sternberg, 1987). And even if first-language learners do rely mainly on contextual guesswork (which is an unproven assumption), there is good reason to think that second language learners should not. A heavily top-down approach to reading and vocabulary building would invite a host of bottom-up problems. (p. 291)

Several contemporary researchers have chosen to bypass the difficulties that inferencing activities present in favor of investigating what constitutes the most effective means of using glosses or definitions during L2 contextualized reading to improve vocabulary learning (Davis & Lyman-Hager, 1997; Hulstijn, 1992; Jacobs, Dufon, & Hong, 1994; Watanabe, 1997). The risk with this approach is that providing readers with definitions for unknown words may lead to shallow processing of that material, and therefore, the prospect of achieving long-term gains as a result of exposure to marginal glosses during reading is unpredictable. In Davis and Lyman-Hager's (1977) study as well as Jacobs, Dufon, and Hong's (1994) study, no effect on comprehension was observed as a result of providing glosses. However, Watanabe (1997) found that use of regular and multiple-choice glosses could achieve significant long-term vocabulary gains with 231 Japanese college students learning English as a second language. Hulstijn's (1992) series of five incidental vocabulary learning studies achieved mixed results, but what is noteworthy here is that this researcher implemented a number of experimental controls designed to help resolve the "given meaning" versus "inferred meaning" debate:

> a) If learners/readers read an L2 text for comprehension of its content, and not with the intention to expand their L2 vocabulary, then they are more likely to remember the form and meaning of an unknown word in the text when they have inferred its meaning by themselves (high mental effort), than when the meaning has been given to them (low mental effort). Evidence for this conclusion stems from Experiments III, IV, and V, using a within-subject design. In experiments I

> and II, using a (weaker) independent-groups design, no evidence for this proposition was found.
>
> b) L2 learners/readers are more likely to infer an incorrect meaning of an unknown L2 word in an L2 text when no cue as to its meaning has been given than when a cue has been given (Experiments I and II).
>
> c) On the basis of (a) and (b), we argue that the discussion in foreign-language pedagogy should not focus on the question of whether it is better to give the meaning of an unknown word than let learners infer word meanings themselves without a cue ('giving versus guessing'). The discussion should rather focus on the question of which cue procedures are most effective (e.g. translation into L1, synonym in L2, concise sample sentence, multiple choice, separately or combined). (p. 122-123)

In Hulstijn's research, it mattered a great deal in terms of recall and performance whether or not learners thought they were going to be tested on target vocabulary in a reading exercise. For growth of vocabulary during casual reading or "non-intentional" study as described above, lexical inferencing appears to be a desirable option as a study strategy if several conditions are met.

The first criteria that should be applied before lexical inferencing can be considered a viable independent study technique is that listening or reading materials and usage examples of the type that were created for Word Hints need to be carefully constructed, screened, or modified to ensure that contextual clues are helpful and also that difficulty levels are appropriately matched to learners' abilities. A recent study by Horst, Cobb, and Meara (1998), supports this approach by demonstrating that a carefully controlled book-length reading treatment resulted in improved rates of incidental word learning. In a somewhat neglected area of research, vocabulary acquisition through listening, a study by Ellis (1995) found that more word meanings were acquired from an involved listening task if input was modified to allow for interactive questioning on the part of participants. Another form of specially constructed directions called premodified

input was found to result in a faster rate of acquisition in words per minute. Both of these findings led support to the notion that constructed, screened, or modified texts and aural inputs are highly desirable tools for increasing vocabulary acquisition.

The second criteria that should be applied to lexical inferencing as an independent reading activity is that the type of help and the reading level of students should be considered in advance. The results of the present study support the conclusions of Baxter (1980), Bejoint (1981), Hartmann (1992), and Piotrowski (1989) that indicate advanced-level readers will benefit the most from monolingual word help. The participants of this study were all early advanced to advanced-level readers, and overall, they were able to benefit from the addition of monolingual help sentences that functioned as usage examples. In addition, the results of the present study suggest that lexical guessing improved when readers where able to process information exclusively at the semantic level without the need to consult subsemantic levels of word help. Intermediate-level readers will be less likely to process monolingual word help successfully on the semantic level, and although they may experience some kind of benefits from the provision of graphophonic or other subsemantic lexical cues, such cues may have little or no effect on the lexical inferencing process.

A study by Davis and Lyman-Hager (1997) with third-semester students of college French (presumably intermediate-level readers) utilized bilingual help and offered both definitions and other kinds of information to students. They found that college students chose to click on word definitions almost exclusively and underutilized the sound cues, grammatical help, explanations of word antecedents or referents, explanations of character relations, and cultural references that were built into the

computer help. As a result, they suggested that students need to be trained in the use and benefits of non-definitional word help, and they also concluded that withholding access to word definitions is potentially useful. Future studies will need to determine precisely what kind of benefits intermediate-level readers obtain from non-definitional information.

Overall, regarding the type of word help and reading level, the evidence suggests three things: first, that advanced-level readers benefit from contextualized semantic help while reading and do not appear to need help cues at subsemantic levels; second, both intermediate and advanced-level readers stand to benefit from having word definitions withheld; third, intermediate-level readers will probably perform better with bilingual rather than monolingual help.

Third, based on the fact that readers achieved greater gains in accuracy of lexical inferencing during Week 2 with Word Hints, it is reasonable to predict that this kind of activity could boost students' incidental word learning if used over an extended period of time. Horst, Cobb, and Meara (1998) agree that there is a case to be made for incidental vocabulary learning through extended reading:

> As far as implications for vocabulary learning are concerned, the experiment makes a stronger case for incidental acquisition than was made in earlier *Clockwork Orange* replication studies. Subjects who read a full-length book recognized the meanings of new words at a higher rate than in previous studies with shorter texts, and build associations between new words as well. Unlike the same-day findings of earlier experiments, these vocabulary learning results represent knowledge that accumulated over a period of ten days. It seems likely that other vocabulary learning benefits accrued. For instance, a number of untested words were probably also learned (or partially learned) through exposure to *The Mayor of Casterbridge*, and the quality of vocabulary learning that occurred seems likely to have been high. Cobb (1997) found that encountering new words in multiple contexts resulted in a deeper, more transferable knowledge of words than the usual strategy of studying short definitions. (p.219)

A fourth criteria that should be applied in order for lexical inferencing to work as a study strategy is that it should be tried in combination with other study strategies or

instructional methods in order to promote vocabulary retention and ensure greater accuracy of learners' concepts about words. Even though the Word Hints resource improved the quality and quantity of readers' guesses about the meanings of unknown words, lexical guessing is still a probabilistic event, and it is natural for readers, especially L2 readers, to make incorrect inferences from time to time. A modular approach to study would be of interest where readers first exercise lexical inferencing during contextualized reading with a type of help such as Word Hints provided, second, they read material on one or more occasions with the same target words glossed, and third, they complete a productive exercise such as word mapping, story retelling, or relating words to previously known information. A recent study by Angela Joe (1998) found that story reading and retelling tasks promoted incidental vocabulary learning in English as a second language learners and that differences between two treatment groups, experimental and comparison, were insignificant:

> Many comparison group subjects employed generative strategies with or without recourse to the text even though they did not receive any explicit generative instruction. It appears that involvement in the read and retell task led the learners to generate unfamiliar words . . . These findings emphasize the importance of student factors in determining the extent of vocabulary learning. More importantly, these results suggest that prior instruction and presence or absence of text during recall did not produce measurable differences in vocabulary learning. It appeared to make no difference whether learners in the experimental groups received prior instruction on generative strategies or whether they recalled the passage with or without the aid of the text. The task itself seems to have stimulated the learners to engage in cognitive processing during recall and led the learners to perform similarly in the vocabulary tests. (p. 374)

Moreover, Joe also finds that in regard to individual differences and incidental vocabulary learning, acquisition is a "rich get richer" kind of scenario. She states that "the significant background knowledge effect indicates that individual student factors, such as prior vocabulary knowledge and generation, to a large extent determine the

vocabulary gains which will be made (p. 375).” Along a similar vein, Knight’s 1994 study profiles the incidental vocabulary learning and dictionary use of high and low ability foreign language learners, and she finds that while both groups learn words incidentally during reading, high verbal ability students learn more words than low verbal ability students. Likewise, Horst, Cobb, and Meara (1998) discuss the relationship between vocabulary size and vocabulary gains in their study of incidental vocabulary learning among English as a second language learners, and they characterize individual differences as follows:

> If a learner’s vocabulary is small, he or she may simply not know enough of the words in the text to be able to infer the meanings of unfamiliar words. However, if a learner’s vocabulary is large, learning gains may also be small because there are few new words available in the text to learn. Incidental uptake may also be low in learners with large vocabulary due to an effect Mondria and Wit-de Boer (1991) have observed: when surrounding contexts are easy to understand, new words are often not noticed (and hence not learned). In this study of low proficiency learners, growth seems more likely to have been limited by knowing too few words than by knowing too many. (p.218-219)

The findings of the present study are consistent with Horst, Cobb and Meara’s interpretation. That is, readers who were unable to process clues at the semantic level of Word Hints were more likely to click heavily on graphophonic clues and were less likely to make highly accurate lexical guesses. If further research confirms that ability level and vocabulary size are highly predictive of an individual’s rate of incidental vocabulary learning, then a mixed-method study approach --one that combines various cue procedures, language production activities, and generative strategies along with contextualized lexical guessing --would be highly promising.

A fifth criteria, reoccurrence of target vocabulary, should be applied before we attempt to embrace lexical guessing as a study strategy. Research shows that target vocabulary must be repeated many times within extended reading passages before

incidental learning rates are affected. Estimates vary as to how many occurrences of a word are necessary before that word can be incidentally learned through reading. In L1 research, Konopak et al. (1987) found that as few as two encounters with a word made a difference in rates of incidental word learning while Jenkins, Stein, and Wysocki (1984) found that as many as ten presentations of a target word to fifth grade students resulted in relatively low rates (16 to 27% over baseline) of word learning. Likewise, in second language learning research, Hulstijn, Hollander, and Greidanus (1996) found greater incidental word learning in groups that had read target words at least three times compared to words that had been seen only one time. The most controlled and thorough L2 data concerning the number of exposures needed to learn an unknown word come from Horst, Cobb, and Meara (1998):

> Generally, the text frequency data suggest that sizable learning gains can be expected to occur consistently for items that are repeated eight times or more. With fewer than eight repetitions, growth is much less predictable and the role of other factors becomes more apparent. There was a large amount of variation in gains on words that were repeated five times or less, including some instances of negative gains. (p. 215)

Once again, L2 research data suggest that a relatively simple type of input modification can yield measurable results for ESL students, and since these learners confront a greater number of obstacles to vocabulary acquisition than L1 learners (as described in Chapter 2), educators must experimentally test and pragmatically develop improved methods in order to help level the playing field.

Although further research is needed to adequately explain the relationship of lexical inferencing, L2 vocabulary acquisition, and individual differences, it may be possible to develop contextualized lexical guessing as a viable component of independent vocabulary study if the five basic controls or criteria mentioned above are met. In

addition, developing technologies need to be exploited in order to partially automate the development of specialized materials for second language reading and vocabulary acquisition.

Developments in Computer-Generated Corpora

Contemporary lexicographers have at their disposal a wide variety of digital resources for constructing both monolingual and bilingual dictionaries, and these include databases that include huge numbers of usage examples drawn from authentic texts (Berg, 1991; Cumming, Cropp, & Sussex, 1994; Fawcett, 1993). However, because of the large numbers of polysemous words in language, computerized matching of a target word in a reading passage with the same sense or meaning of the target in a sentence context has been problematic goal, one that has historically been better accomplished by human selection (Kerim-Zade & Pavlov, 1989; Louw, 1995). For the average ESL instructor, the process of hand-picking, screening, or writing the dozens or hundreds of usage examples needed to construct a contextualized resource for extended reading would be very time consuming and tedious, indeed. The task not only involves control of multiple word meanings, but it has already been shown that in order to provide helpful contexts to readers, the suggestiveness of subject, verb, and function needs to be taken into account, also. Fortunately, some developments in the use of computerized corpora and natural language processing may offer some hope in this area.

The work of Kita and Ogata (1997), shows some promise that the matching of appropriate translations, collocational phrases, and sentences to target words can be achieved by computers. Whereas polysemy or multiple meanings are difficult to compute, it is much easier to “infer” the correct context of a word by computing its

collocations, the company that a given word keeps. One of Kita and Ogata's tools for doing this is called a bilingual collocation concordancer, and this takes a collocational expression as input in one language and produces a pair of sentences (in L1 and L2) in which the given collocation occurs. This feat is achieved through a combination of methods. First, Kita and Ogata scan a corpus, and then they use a modified cost criteria formula which can "quantitatively estimate the extent to which processing is reduced by considering a word sequence as one unit (p.231)." In their modified formula, the cost of a word equals its frequency. This is combined with a tree-growing algorithm that produces collocation trees or hierarchical families of related phrases (p.233). The net result is that the bilingual collocation concordancer is capable of extracting complex predicate phrase patterns in two languages. This is an improvement over the old cost criteria method that tended to extract frozen phrase patterns such as "thank you very much." The upshot of Kita and Ogata's invention is that it not only works with words, phrases, and sentences because of the unique corpus, the ATR Dialogue Database, that it is built upon, but it could potentially be adapted for either monolingual or bilingual use. A system like this one that is capable of extracting multiple instances of complex collocations could theoretically be put to work extracting target words from families of collocate phrases and sentences, and this activity could supply all the components necessary to create the contents of a Word Hints resource in a more automated way.

Berleant et al. (1997) devised a somewhat less automated but fairly reliable context-matching tool with the creation of LEARN, a system that processes English unrestricted text by translating selected English words within it into foreign words. These are presented in the reading material to the learner. LEARN software "places

foreign words in meaningful contexts of interest to the reader," and it does so by the use of word experts:

> A word expert is a small expert system associated with a word or a few related words of a language, which can analyze occurrences of it in context. Word experts can be used for a number of tasks . . . . A LEARN word expert examines the context of its target word and then outputs its translation. (p. 110)

Once a word expert is created, it can be reused indefinitely in the system. However, real human beings are required to invest a lot of time at the outset in order to generate the disambiguation rules that a word expert relies upon. Once this work is done, word experts are capable of performing rather well. The Berleant group's word expert software can successfully translate the target word approximately 90% of the time (115).

One final study worth noting is that of McKinnon's (1993) multi-dimensional concordancer, a tool for mapping various kinds of three-dimensional views of key concepts in a book and the ways that the concepts interrelate. McKinnon used a correspondence analysis software package called SimCa to plot out spatial contexts, overall contexts, and vector-based role analysis of words in Kierkegaard's "Fear and Trembling." Word frequencies and occurrences in the text form the basis of the analysis:

> Correspondence analysis operates by calculating eigenvalues and eigenvectors or, more simply, by comparing points in terms of their respective "profiles." In this case the profile of a word is the percentage of all its occurrences in each of the eight chapters and that of a chapter the percentages of all occurrences of each of the 124 words in that chapter. Thus, each of these 124 words has an eight feature profile and each of these eight chapters a 124 feature one. (p.166)

The role and sense vectors constructed by McKinnon appear to be a particularly viable means of isolating clusters of interrelated keywords and collocates that one could easily match to target words in book-length material.

Three different computer programs have been described here, two concordancing programs and a system that uses disambiguation tasks to write rules of translation based on collocations. Developments in this field are showing increasing relevance to second language vocabulary learning, and in time, some combination of these techniques may enable computer programs to reliably match target words in contextualized reading passages with usage examples that consistently afford the same meanings as the selected targets.

In the field of computer-assisted language learning, it appears that little or no attention has been paid to the issue of creating a computer system that can sort sentences according to the level of suggestiveness of subject, verb, and function, the "word helpfulness" criteria established by Mondria and Wit-de Boer (1991). However, it should be possible to define syntax and structure rules that would enable a computer program to reliably isolate the more developed sentences such as "the gardener filled the watering can to give the plants water." This is a question that warrants further investigation.

Summary

This study has shown that readers' own inferencing processes can be enhanced during contextualized reading through the provision of a lexical resource that offers additional clues in the form of usage examples or sentence help, sound cues, word forms, and parts of speech. Participants seemed especially capable of exploiting additional semantic clues during reading, and it appears that helpful contexts are indeed a key factor in vocabulary acquisition. Lexical inferencing for second language learners could one day be used as a viable independent study strategy, but a preponderance of evidence suggests that this will not be the case unless a minimum of five criteria are observed:

reading or listening material and help contexts must be carefully controlled; for reading activities, the type of word help and the reading level of students must be taken into account; to achieve maximum benefit, inferencing activities must occur over an extended period of time; inferencing must be combined with other strategies; and target words for ESL students should reoccur in passages at least eight times.

In order to implement this study, materials were developed that consisted of controlled reading passages that (in the case of the experimental condition) contained target words linked directly to sentence contexts. These sentence contexts adhered to measurable criteria for suggestiveness of subject, verb, and function. In the real world, it is not an easy task for lexicographers much less educators to develop these kinds of materials on a large scale. However, recent developments in computer-assisted language learning suggest that specialized programming and new techniques in the use of computerized corpora could be exploited to facilitate rapid development of these kinds of materials.

APPENDIX A
TARGET WORD LIST

Words presented without help:

abeam
acroliths
ballyhoo
barnacled
behooves
bloated
bungle
calipers
cerulean
cinch
crevice
durmast
easel
ensconce
flabbergasted
floundering
foolscap
gamut
gerrymandering
gibberish
gorges
gossamer
haggle
harbingers
hubbub
humdrum
insouciant
inured
jagged
kiln
knavish
lackadaisical
loiter
maelstrom
mayhem
mewling
miscreants
musty
oozed
ousted
platen
pouting
prowling
quagmires
quay
quoins
rollicking
rumpled
scrims
scruffy
shenanigans
shoals
smudge
snipe
sojourning
soothsayers
stalemates
stopgap
surly
tenacity
tepid
trenchant
veer
vellum
vouches
vying
wallow
warmongering
wince
writhed

Words presented with Word Hints:

acumen
abutment
akimbo
ampersand
amuck
barges
bauble
bedlam
bevels
bleary
blithe
bulges
cantankerous
chisels
conundrums
coxswains
dawdling
delve
despondent
doggerel
enthralled
fettle
fidget
fiends
fledgling
flustered
goad
grapple
hackneyed
hallmarks
harangue
hobnobbed
huddle
imbue
jostling
kedge
laggards
nubs
poignancy
quavering
quibbling
pander
proffered
pulley
rakish
ramshackle
rant
recoilless
rehash
ribald
savvy
scaffold
scowling
serendipitous
shambles
shoddy
steadfast
swivel
tether
throngs
trowel
turret
twangy
unabashedly
unkempt
wafting
weltering
withering
yelp
zealots

APPENDIX B
SAMPLE PASSAGES

From the passage only condition:

Passage #12: Fashion Photography

For professionals, learning to photograph fashion models can be a ***cinch**** as long as a few crucial points are kept in mind. First, attempt to ***ensconce**** the subject in light, and at the same time, optimize the interplay between the fabric and light source. Second, keep multi-colored ***scrims**** on hand so that backdrops can be changed in an instant. Third, if shooting for advertising layouts, remember to use a consistent method of color checks. ***Cerulean**** blue works well for background - foreground matching tests. Fourth, run the ***gamut**** of styles in your work form funky to playful to serious. Fifth, remember to photograph ***pouting**** models as well as smiling ones. It is often an unexpected contrast of expression which makes a great photo. Finally, it ***behooves**** the photographer to develop a good working relationship with models. Don't forget that looking good is hard work! Remember these points, and you're on the way to becoming a fashion photographer.

From the passage plus Word Hints condition:

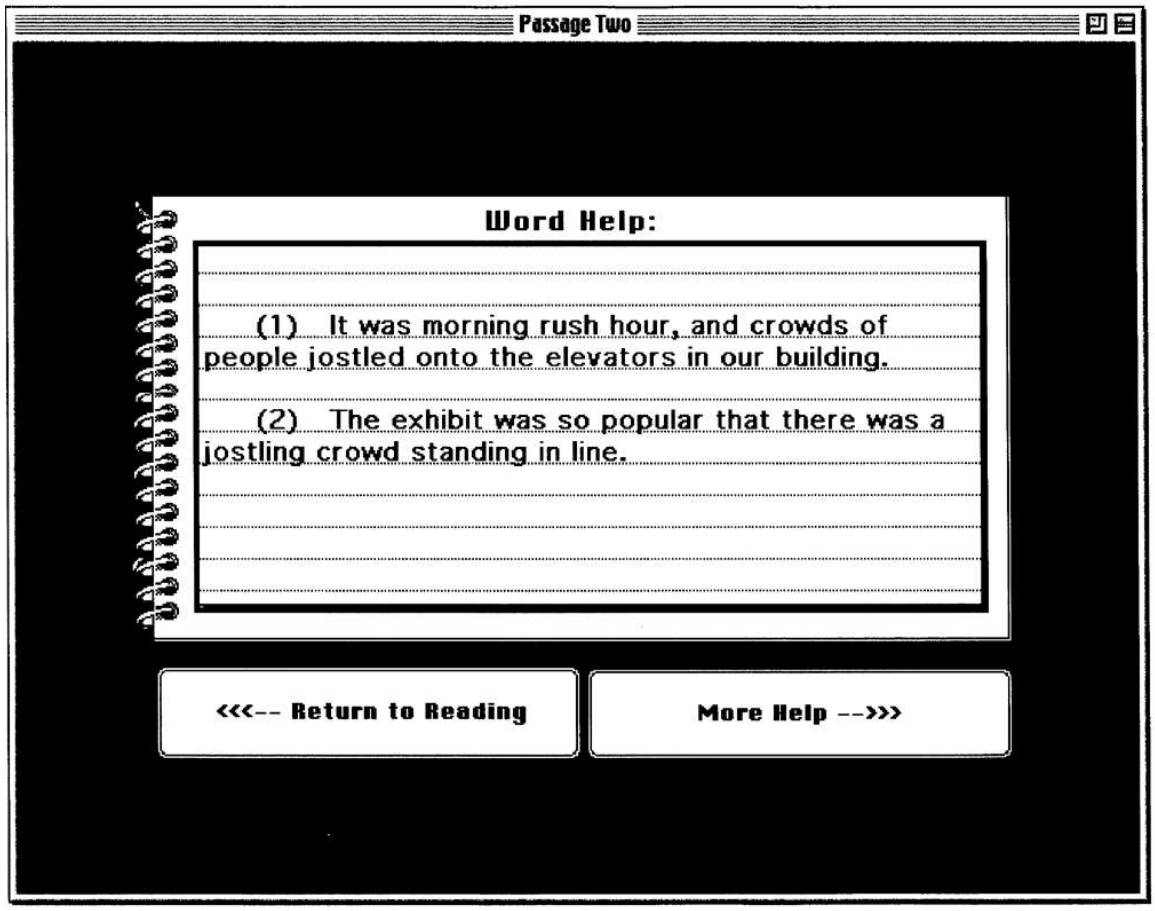
Passage #7: Portrait Studios

Many photographers try starting up portrait studios to get extra income, but many lack the ***savvy**** to succeed. This is because they underestimate the challenges involved in satisfying customers.

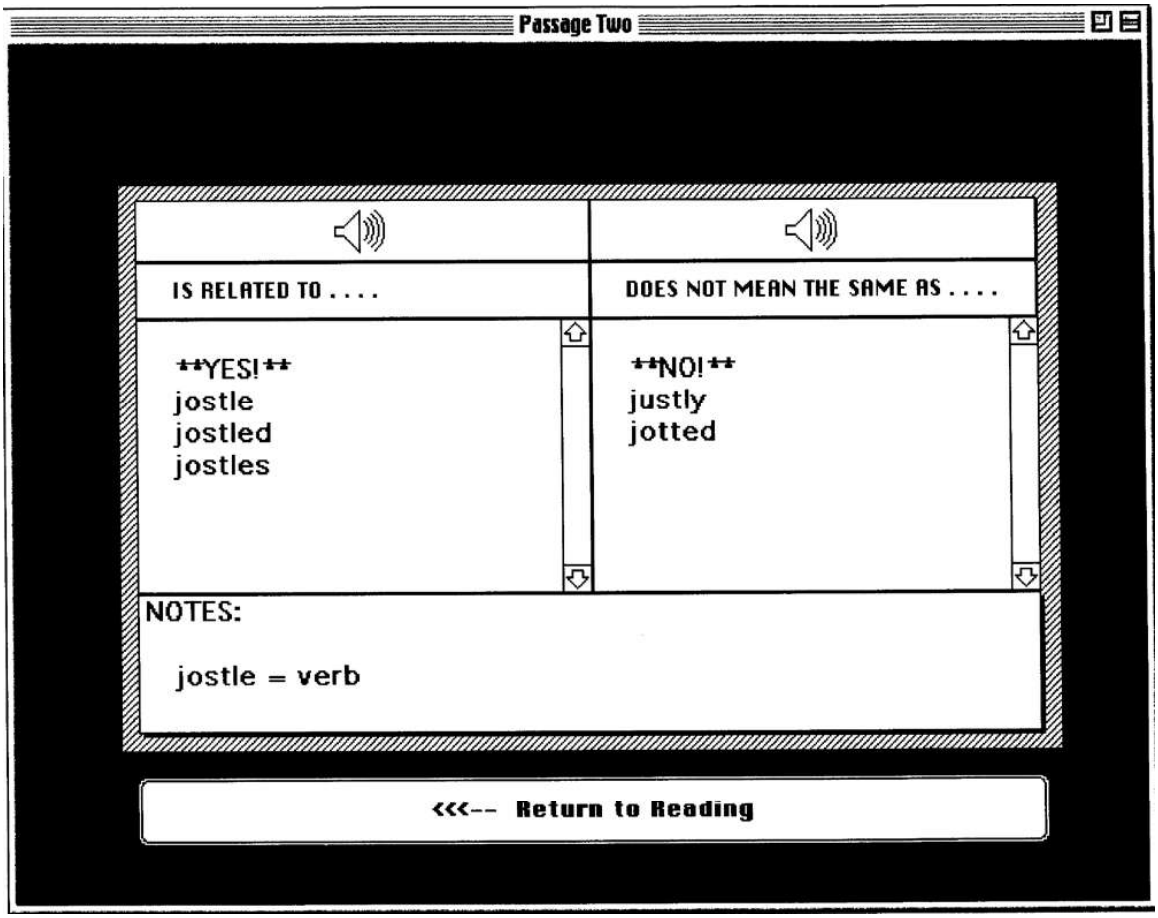
Some subjects are easily ***flustered**** by the presence of a camera and tend to freeze up. Others are reluctant to smile and seem to have been born ***scowling***.* A few customers will complain when they look unexpectedly old, disheveled, or ***bleary****-eyed in their photos. The worst subjects, children, tend to complain little yet exasperate photographers a lot. Parents are extremely delighted when a photo captures the ***blithe**** expressions of innocence. However, children ***fidget**** the most, and every good portrait photographer must be prepared to dangle one interesting ***bauble**** after another to try to get the attention of infants and to make older children laugh. Photographers who have great social skills make the best portrait photographers.

APPENDIX C
SAMPLE HELP FROM WORD HINTS

Semantic help is the first level of Word Hints that appears when the reader clicks upon a target word:



Graphophonic help is the second level of Word Hints that appears when the reader clicks the “More Help” button from the semantic or first level:



LIST OF REFERENCES

Aitchison, J. (1994a). Words in the mind: An introduction to the mental lexicon. (2nd ed.). Oxford: Blackwell.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. Monographs of the Society for Research in Child Development, 58, (10, Serial No. 238).

Aust, R., Kelley, M. J. & Roby, W. (1993). The use of hyper-reference and conventional dictionaries. Educational Technology Research and Development, 41 (4), 63-73.

Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? TESOL Quarterly, 14 (3), 325-336.

Bejoint, H. (1981). The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills. Applied Linguistics, 2 (3), 207-221.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. Journal of Research in Reading, 7 (1), 15-32.

Berg, D. L. (1991). The research potential of the electronic OED database at the University of Waterloo: A case study. Library Hi Tech, 36 (9), 37-52.

Berleant, D., Shi, L., Wei, X., Viswanathan, K., Chai, C., Majid, N., Qu, Y., & Sunkara, P. (1997). LEARN: Software for foreign language vocabulary acquisition from English unrestricted text. Computer-Assisted Language Learning, 10 (2), 107-120.

Bernhardt, E. B. (1990). A model of L2 text reconstruction: The recall of literary text by learners of German. In A. Labarca & L. M. Bailey (Eds.), Issues in L2: Theory as practice/practice as theory (pp. 21-43). Norwood, NJ: Ablex.

Carnine, D., Kameenui, E. J., & Coyle, G. (1984). Utilization of contextual information in determining the meaning of unfamiliar words. Reading Research Quarterly, 19 (2), 188-204.

Channell, J. M. (1990). Vocabulary acquisition and the mental lexicon. In J. Tomaszczyk & B. Lewandowska-Tomaszczyk (Eds.), Meaning and lexicography (pp. 21-30). Amsterdam: John Benjamins.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. The Modern Language Journal, 80 (2), 183-198.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), Second language reading and vocabulary learning (pp. 217-228). Norwood, NJ: Ablex.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), Second language reading and vocabulary learning (pp. 217-228). Norwood, NJ: Ablex.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cumming, G., Cropp, S., & Sussex, R. (1994). On-line lexical resources for language learners: Assessment of some approaches to word definition. System, 22 (3), 369-377.

Davis, J. N., & Lyman-Hager, M. A. (1997). Computers and L2 reading: Student performance, student attitudes. Foreign Language Annals, 30 (1),58-71.

Fawcett, H. (1993). The new Oxford English dictionary project. Technical Communication, 40 (3), 379-382.

Gildea, P. M., Miller, G. A., & Wurtenberg, C. L. (1990). Contextual enrichment by videodisc. In D. Nix & R. Spiro (Eds.), Cognition, education, and multimedia: Exploring ideas in high technology (pp.1-29). Hillsdale, N.J.: Erlbaum.

Goodman, K. S. (1974). Psycholinguistic universals in the reading process. Journal of Typographic Research, 4 (2), 103-110.

Haastrup, K. (1989). The learner as word processor. AILA Review, 6, 34-46.

Haastrup, K. (1991). Lexical inferencing procedures or talking about words. Tubingen, Germany: Gunter Narr Verlag.

Hartmann, R. R. K. (1992). Lexicography, with particular reference to English learners' dictionaries. Language Teaching, 25 (3), 151-159.

Haynes, M. (1993). The patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes, & J. Coady (Eds.), Second language reading and vocabulary learning (pp. 46-63). Norwood, NJ: Ablex.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? Reading in a Foreign Language, 8 (2), 689-696.

Horst, M., Cobb, T. & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring a second language vocabulary through reading. Reading in a Foreign Language, 11 (2), 207-223.

Huckin, T. & Haynes, M. (1993). Summary and future directions. In T. Huckin, M. Haynes, & J. Coady (Eds.), Second language reading and vocabulary learning (pp. 289 – 297). Norwood, NJ: Ablex.

Hulstijn, J. H. (1989). Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H. W. Cechert & M. Raupach (Eds.), Interlingual processes (pp. 49-73). Tubingen, Germany: Gunter Narr Verlag.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Bejoint (Eds.), Vocabulary and applied linguistics (pp. 113-125). London: MacMillan.

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. Modern Language Journal, 77 (2), 139-147.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurance of unknown words. Modern Language Journal, 80 (3), 327-338.

Ittzes, K. (1991). Lexical guessing in isolation and context. Journal of Reading, 34 (5), 360-366.

Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. Journal of Research in Reading, 17 (1), 19-28.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. American Educational Research Journal, 21 (4), 767-787.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? Applied Linguistics, 19 (3), 357-377.

Jones, F. R. (1995). Learning an alien lexicon: A teach-yourself case study. Second Language Research, 11 (2), 95-111.

Kelly, P. (1990). Guessing: No substitute for systematic learning of lexis. System, 18 (2), 199-207.

Kerim-Zade, I. & Pavlov, V. (1989). The semantico-functional variability of words and the teaching of vocabulary to advanced EFL students. Applied Linguistics, 10 (4), 382-391.

Kita, K. & Ogata, H. (1997). Collocations in language learning: Corpus-based automatic compilation of collocations and bilingual collocation concordancer. Computer-Assisted Language Learning, 10 (3), 229-238.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. Modern Language Journal, 78 (3), 285-299.

Konopak, B., Sheard, C., Longman, D., Lyman, B., Slaton, E, & Atkinson, R. (1987). Incidental versus intentional word learning from context. Reading Psychology: An International Quarterly, 8 (7), 7-21.

Krantz, G. (1991). Learning vocabulary in a foreign language: A study of reading strategies. Goteborg, Sweden: Acta Universitatis Gothoburgensis.

Laufer, B. (1990a). Ease and difficulty in vocabulary learning: Some teaching implications. Foreign Language Annals, 23 (2), 147-155.

Laufer, B. (1990b). 'Sequence' and 'order' in the development of L2 lexis: Some evidence from lexical confusions. Applied Linguistics, 11 (3), 281-296.

Laufer, B. (1990c). Why are some words more difficult than others?--Some intralexical factors that affect the learning of words. International Review of Applied Linguistics in Language Teaching, 28 (4), 293-307.

Laufer, B. (1991). Some properties of the foreign language learner's lexicon as evidenced by lexical confusions. International Review of Applied Linguistics in Language Teaching, 29 (4), 317-330.

Laufer, B. (1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability. Journal of Research in Reading,15 (2), 95-103.

Laufer, B., & Sim, D. D. (1985). Taking the easy way out: Non-use and misuse of clues in EFL reading. English Teaching Forum, 23 (2), 7-10.

Louw, J. P. (1995). How many meanings to a word? In B. B. Kachru & H. Kahane (Eds.), Cultures, ideologies, and the dictionary (pp. 357-365). Tubingen: Max Niemeyer Verlag.

McKinnon, A. (1993). The multi-dimensional concordance: A new tool for literary research. Computers and the Humanities, 27, 165-183.

Miller, G. A. (1991). The science of words. New York: Scientific American Library.

Mondria, J., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. Applied Linguistics, 12 (3), 249-267.

Na, L., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. RELC Journal, 16, 33-42.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. American Educational Research Journal 24 (2), 237-270.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. Reading Research Quarterly, 20 (2), 233-253.

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

Nation, I. S. P. (1993). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.), The bilingual lexicon (pp. 115-134). Philadelphia: John Benjamins

Nesi, H. (1987). Do dictionaries help students write? In Bloor, T. & Norrish, J. (Eds.), Written language (pp. 85-97). London: Centre for Information on Language Teaching and Research.

Nesi, H., & Meara, P. (1994). Patterns of misinterpretation in the productive use of EFL dictionary definitions. System, 22 (1), 1-15.

Piotrowski, T. (1989). Monolingual and bilingual dictionaries: Fundamental differences. In M. L. Tickoo (Ed.), Learners' dictionaries: State of the art (pp. 72-83). Singapore: Southeast Asian Ministers of Education Organization. (ERIC Document Reproduction Service No. ED 342 250).

Pustejovsky, J., & Boguraev, B. (1996). Lexical semantics: The problem of polysemy. Oxford, UK: Clarendon Press.

Rivera, C., & Lombardo, M. (1979a). The validation of the English As A Second Language Assessment Battery. (ERIC Document Delivery Service No. ED 185073)

Rivera, C., & Lombardo, M. (1979b). The English As A Second Language Assessment Battery (ESLAB) [Microfiche]. (Available in Tests in Microfiche, No. 11103, from Educational Testing Service, Princeton, N.J.)

Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. Reading Research Quarterly, 21 (4), 439-453.

Scherfer, P. (1993). Indirect L2-vocabulary learning. Linguistics, 31, 1141-1153.

Schouten-Van Parreren, C. (1992). Individual differences in vocabulary acquisition: A qualitative experiment in the first phase of secondary education. In P. J. L. Arnaud & H. Bejoint (Eds.), Vocabulary and applied linguistics (pp. 94-101). London: MacMillan.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown, & M. E. Curtis (Eds.), The nature of vocabulary acquisition (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum.

Tomaszczyk, J. (1979). Dictionaries: Users and uses. Glottodidactica, 12, 103-119.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. Studies in Second Language Acquisition, 19, 287-307.

Wesche, M., & Paribakht, T. (1994). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. Baltimore, MD: Paper presented at the Annual Meeting of the American Association for Applied Linguistics. (ERIC Document Delivery Service No. ED 369 291).

BIOGRAPHICAL SKETCH

Debra Kirchgassner Abbott was born in Washington, D.C., and attended grade school in Montgomery County, Maryland. Debra graduated from Damascus High School in 1980, and in May of 1984 at the age of 20, she received her B.A. from Sarah Lawrence College in New York. Following undergraduate studies, Debra attended the Catholic University of America in Washington, D.C., where she received a scholarship and student assistantship. In 1995, Debra completed a teaching certification program from Hood College in Frederick, Maryland, and in November of 1986, she moved to Florida. By 1988, Debra was a secondary school teacher in the Florida public school system. In 1990, she returned to Washington, D.C., for her master's degree examination, and in October 1990, she graduated with an M.A. in English rhetoric and composition. In the fall of 1992, Debra left public school teaching to pursue doctoral studies at the University of Florida. In the summer of 1993, she was awarded a full fellowship in multilingual, multicultural education sponsored by the Federal Department of Education. Debra's fellowship program at U.F. was a multidisciplinary one that combined language learning studies, technology research, and special topics in education. Since beginning her studies at U.F., Debra has been employed as a medical librarian, an instructional software designer, a research coordinator, and presently, she is employed full-time as Assistant Director for the Center for Instruction and Research Technology at the University of North Florida. Debra married Colin Abbott in December, 1998, and they enjoy living with a menagerie that includes three cats and a dog.