# Component Processes of Second Language Reading

Catherine Poulsen

A Thesis

in

The TESL Centre

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Arts at

Concordia University

Montréal, Québec, Canada

March 1992

ISBN 0-315-73628-3

## ABSTRACT

Component Processes of Second Language Reading

Catherine Poulsen

Second language (L2) readers, even fluent bilinguals, often read more slowly in their L2 than in their first language (L1). Reading involves a wide range of component skills, from low-level processes, such as word recognition and propositional assembly, to higher level processes, such as propositional integration and text modelling. Using a multiple regression paradigm to examine a wide range of component skills simultaneously, this study explored the relative role of these components in L1 and L2 reading in an attempt to determine which components are responsible for more effortful reading in the L2. Fifteen L1 and 15 L2 readers read four experimental texts sentence by sentence while their sentence reading times were recorded. Twelve characteristics of these 118 sentences, hypothesized to place differential demands on underlying lower and higher level components, were measured and used to predict the sentence reading times. The mean L1 and mean L2 unstandardized regression coefficients associated with each of the predictor variables and used as estimators of the amount of cognitive resources allocated to each component, were compared. Results revealed that both L1 and L2 readers devoted considerable resources to predictors associated with extracting and assembling propositions from the surface structure of the texts and with integrating propositions into a mental representation in memory. However, L2 readers were significantly slower than L1 readers only at the lower level of propositional extraction and assembly. The potential implications of these results for L2 reading instruction are considered, as are the methodological issues raised by this experimental paradigm.

ACKNOWLEDGEMENTS

I'm deeply indebted to my supervisor, Dr. Norman Segalowitz, for his inspiration, guidance, and support throughout all stages of this thesis.

I'd also like to thank my readers, Dr. Patsy Lightbown and Dr. Jack Upshur, for always keeping their doors open for me and for their ready assistance whenever needed.

Finally, a big thank you goes to my family, friends, and colleagues for sticking with me through the highs and lows on the road to completion; for sharing my enthusiasm, on the one hand, and for extending patience and encouragement at times of frustration, on the other.

TABLE OF CONTENTS

## LIST OF FIGURES

LIST OF TABLES

# INTRODUCTION

Skilled reading in English is an important goal for many users of English as a second language. In addition to the needs of immigrants to English-speaking countries who must function on a daily basis in their second language, students and professionals around the world in such fields as the pure and applied sciences and technologies often wish to access some of the latest developments in their fields in English-language journals and reports. Helping these and other individuals become proficient readers in English is, therefore, an important goal for applied linguists.

Sound pedagogy in reading ultimately must be based on an understanding of both what skilful reading is, that is, the component skills and processes involved, and how skill in reading develops. A major source of contention has been the often less than successful application of findings from studies examining the reading process to strategies for reading pedagogy. Whether this has been the result of researchers overextending implications of their findings, or practitioners anxious for guidelines for designing learning tasks, a word of caution is advised. Explication of the reading process at various stages of reading skill can provide a very useful description of the skills to be attained by the reader and thus forms a necessary first step in developing effective training; it cannot, however, reveal in and of itself *how* these skills are acquired. The goal of this thesis was to shed additional light on the reading process as applied to second-language reading; namely, to examine in which aspects of the reading process skilled bilinguals who read more slowly in English, their second language, differ from skilled first-language readers of English. While the present study attempts to understand reading by examining the contributing roles of its component processes, it does not mean to suggest that the best way to develop these skills is necessarily in isolation. Other research should guide this decision.

*Models of Reading*

A vast amount of research on reading has been conducted during the past twenty years by researchers in the fields of cognitive psychology, education, and linguistics. Most of the research has examined first language (L1) reading, but many of the findings are also of interest to second language (L2) reading investigators, and indicate useful sources of investigation for problems specific to L2 reading.

Research into reading has had to take on the challenge of investigating a highly complex skill. While reading is often referred to as a single skill, along with listening, speaking, and writing, it is itself made up of many closely interacting component skills. These include low-level, local, processing skills involving word and sentence units, such as word recognition, lexical access, syntactic parsing, and propositional encoding, as well as higher level processing skills involving global text features, such as propositional integration, inferencing, and the construction of a mental representation of the text in memory, among others. Initially, much research focused on specific, usually isolated, cognitive processes involved in reading. Recently, however, attention has been directed toward developing and testing theories which integrate the findings of these diverse studies. During this development, much debate has revolved around the relative importance of these component processes in reading and their differential involvement in skilled versus less skilled reading.

For some years a bottom-up model of reading was dominant (Gough, 1972). In this model component processes of reading are viewed as sequential operations. The output of low-level perceptual processes of letter and word recognition act as input to higher level comprehension processes. Higher level contextual factors inherent in the text or constructed by the reader on the basis of background knowledge are held to have no influence on low-level processes such as word recognition. The reader is, thus, viewed as a somewhat passive decoder of linguistic structure.

This bottom-up view of reading was rejected by both Goodman (1967) and Smith (1973), who adopted a top-down perspective of reading. According to these authors, readers are actively involved in a "psycholinguistic guessing game" using their knowledge of syntax, the context, and prior knowledge of the topic to anticipate and facilitate the decoding of words. The reader builds hypotheses about upcoming words and simply confirms them as the words are encountered. This model emphasizes the role of high-level conceptual knowledge in reading and places little importance on low-level decoding skills.

More recent evidence on reading does not support the view that skilled readers operate primarily in a top-down fashion. Results from numerous eye movement studies suggest that low-level processing is of primary importance to reading. They reveal that skilled readers fixate at least 80 percent of content words and that the duration of these fixations are dependent on word characteristics such as length, reflective of perceptual processing demands, and frequency, reflective of lexical access demands (Just & Carpenter, 1980; Just & Carpenter, 1987). The influence of context on word identification processes in reading also appears to be extremely limited, particularly in skilled reading. Top-down processing models claim that as skilled reading develops the reader increasingly makes use of context to facilitate low-level processing. In contrast to this premise, studies have shown that poor readers show much stronger context effects than skilled readers (Perfetti, 1985), even though skilled readers are better at using context to anticipate upcoming words or to generate meanings to unknown words (Daneman & Green, 1986).

More consistent with this and other empirical evidence, subsequent models of reading took an interactive view of the reading process (e.g., Rumelhart, 1977; Stanovich, 1980). These models suggest that processes at various levels influence each other by providing information from ongoing simultaneous processing which constrains the processing options at any given level. For example, a low-level process such as word recognition may be facilitated by syntactic and semantic contextual cues, but is not dependent upon or directed by them. Stanovich (1980) proposed the interactive-compensatory model of reading to account for individual differences in reading skill. According to this model, deficient skills at one level of processing may be compensated for by additional processing at another level. Thus, a weak reader could compensate for poor low-level decoding skills by making greater use of contextual cues in the text and prior knowledge of the text topic. The skilled reader, on the other hand, who possesses fast and efficient decoding skills would not need to draw on more effortful top-down processing to aid these low-level processes.

Perfetti's verbal-efficiency theory (1985) places even more emphasis on the efficient operation of low-level, local processing. His theory suggests that the capacity of working memory limits the ability of a reader to overcome poor low-level processing by any compensatory mechanism without some reading impairment. The basic assumption of the verbal-efficiency model is that there are limited resources in working memory which must be shared across all ongoing processing. High-level processes, Perfetti argues, will always demand considerable resources since they cannot operate independently of strategic control. Low-level processes, which can vary greatly in resource demands and have the potential of attaining automatic or near-automatic functioning, are considered by Perfetti to be central to individual differences in reading ability. Skilled readers, who have developed efficient low-level processing, have more resources available to devote to the demanding high-level comprehension processes. Poor readers, on the other hand, may have to draw resources away from high-level components in order to compensate for weak low-level skills. In some cases, this may necessitate slowing down or rereading sections; in others, it may result in a serious disruption of comprehension. The verbal-efficiency theory focuses on four local processes: lexical access, semantic encoding, phonological recoding and propositional encoding. Lexical access, semantic encoding and phonological recoding can all become highly efficient and automatic. Propositional encoding, even in a highly-skilled reader will probably never attain complete automaticity, but this process can nevertheless vary from extremely high to fairly low in cognitive resource demands. Developing efficiency at this level of propositional text-work would free working memory to perform the more effortful processes of propositional integration and text modeling. Ultimately, therefore, individual differences in reading are attributed to individual differences in the efficient operation of lexical access, semantic encoding, phonological recoding, and local propositional encoding.

Most recently, on the basis of Fodor's modular model of cognitive processing (1983), Stanovich (1991) suggests that as reading skill develops, low-level word recognition skills become modularised by the cognitive system such that eventually they can no longer be influenced by higher level processes.

Intuitively, one might expect that low-level processing would be more accurate if context and prior knowledge contributed additional information. However, Stanovich argues that the independent, isolated functioning of modular systems, termed information encapsulation, would have two distinct advantages in word recognition. First, it ensures that the cognitive system receives accurate information, unadulterated by potentially erroneous expectations. Second, it allows for fast, automatic processing, and is not reliant upon slow, strategic decisions. "We have in reading precisely the situation where an enormous advantage accrues to encapsulation: the potential specificity of stimulus-analyzing mechanisms (word recognition) is great relative to the diagnosticity of the background information that might potentially be recruited to aid recognition" (p. 431).

Thus, according to recent formulations of the reading process, skilled reading is characterised in part by acquired modularity of low-level processes, such as word recognition, which enables the formation and availability of fast and accurate mental representations of the incoming visual stimulus (words) for subsequent higher level processing. This means that a reader who is deficient in lower level components could be disadvantaged in two ways. First, the less skilled reader may be slower and require a greater portion of limited resources from centralized working memory to perform these low-level processes. Second, poor word recognition skills may provide degraded or incomplete codes to working memory, placing still further demands on higher level processing resources in order to compensate. While studies do indicate that contextually predictable words are identified more quickly, the probability of correct predictions is only between .20 and .35, and lower for less skilled than skilled readers (Gough, 1983; Perfetti, Goldman, & Hogaboam, 1979). In the absence of accurate decoding processes, then, the potential for 'misreading' of the text due to inaccurate predictions could be great.

*L2 Reading Instruction*

Recent methodologies for developing reading skill in both the first and second language have been strongly influenced by a top-down view of the reading process. In second language teaching, this 'meaning-based' approach to reading, which stresses the constructive role of the reader during comprehension, fits in well with the general movement in language teaching toward a communicative, learner-focused approach. Consistent with this view, therefore, many current methods of second language reading pedagogy characterize second language readers as too focused on the decoding of words, syntax, and other local aspects at the word and sentence level. They encourage the reader to utilize higher level processing in a principled guessing approach to reading assumed to characterize skilled readers, and to avoid dependence on decoding the text word by word or letter by letter (Williams, 1984). The reading lesson often begins with a pre-reading discussion on the text topic and attempts to draw upon what knowledge of the topic the learner already possesses. Many lessons contain exercises to be completed during the reading task, such as completing diagrams, taking notes, or answering questions inserted at natural breaks between text sections. These 'while-reading' exercises focus on the overall structure of the text, often requiring the reader to summarize recently read material and to anticipate what will follow. Exercises which focus on vocabulary also encourage the reader to use both prior knowledge and the context surrounding the use of the new word as clues to its meaning.

While these exercises facilitate the learner's task by providing rich contextual background and may serve important motivational purposes, they are based on two rather shaky assumptions regarding reading. The first assumption is that this top-down model is an accurate depiction of skilled reading. As discussed above, recent evidence refutes this view. The second assumption is that the way to train less skilled readers is to encourage them to imitate this assumed top-down approach of skilled readers by devoting more attention to higher level aspects of the text. However, examination of current evidence on low-level processing skill in L2 reading suggests that the less skilled L2 reader might not have the requisite component skills to do this. (Segalowitz, 1986). Slower processing of local components may not be maladaptive, just characteristic of an earlier stage of reading in which low-level skills have not yet developed fluency.

*Evidence of Deficient Component Skills in L2 Reading*

The development of reading processes involved in translating surface structure into an internal representation of the text may be particularly problematic for L2 readers since these processes rely strongly upon knowledge of and experience with the language in which the text was written. Language-specific knowledge and skilful use of this knowledge is implicated in such low-level processes as word recognition, lexical access, syntactic parsing, and semantic activation. This includes knowledge of the orthographic regularities of the language, spelling to sound correspondences, syntactic relationships and constraints, and word meanings. There is considerable evidence indicating that even skilled bilinguals are deficient in many of these areas in their second language.

Both visual and phonological codes may be implicated in basic word recognition. Skill in the use of orthographic redundancies, such as the positional frequency of letters or other spelling patterns, is closely related to fast and accurate visual recognition of words. Favreau, Komoda, and Segalowitz (1980) found that even skilled bilinguals are less efficient in their use of orthographic redundancies in their L2 than in their L1. In their L1, readers were more accurate in the identification of target letters when they appeared in a word (a word superiority effect), indicating an ability to make use of the positional frequency of letters in their L1; this word superiority effect, however, was only found in their L2 if more time was permitted to allow the effect to appear. Thus, even though these skilled bilinguals were knowledgeable about the orthographic redundancies of the L2, they were less efficient in their ability to make use of this knowledge.

While the role of phonological recoding in word recognition and lexical access has recently been the subject of considerable debate, there is little disagreement that phonological codes are generated while reading and that skill in phonological recoding is a distinguishing characteristic of advanced readers. Whether or not phonological recoding is used in lexical access, phonological codes, once generated, provide a relatively stable, verbatim, representation of recently read material in working memory for processes such as propositional assembly and integration (Perfetti, 1985; Segalowitz & Hébert, 1990). Segalowitz and Hébert (1990) used a sentence verification task contrasting phonologically-ambiguous sentences involving homophones with unambiguous sentences (those containing no homophones) to compare bilingual subjects' skill in using L1 and L2 phonological codes in working memory. The subject's task was to decide whether the sentence made sense (was meaningful). Their results suggested that bilinguals who nevertheless read more slowly in L2 have more difficulty with the use of phonological codes in their L2. This deficit places a heavier load on working memory and, given the role of phonological codes in providing a stable memory trace, may further weaken information available for higher level processing.

Closely related to word recognition and lexical access is the reader's representation and use of semantic information. The theory of spreading activation of semantic processing (Collins & Loftus, 1975) proposes that retrieval of a word in long-term memory automatically spreads its activation to other related concepts priming them for subsequent retrieval. For example, the word 'red' would spread its activation to other colours and to associated objects such as fire engine, which in turn might activate truck, ambulance, etc.

The extent and nature of semantic spreading activation has typically been studied in single-word lexical decision tasks using a priming procedure (Neely, 1977) to assess whether this activation is automatic or controlled. Favreau and Segalowitz (1983) found that while semantic activation of words related to the primes was fast and automatic (not under control of strategic predictions) in highly skilled reading, fluent bilinguals who nevertheless read more slowly in their L2 exhibited slow, controlled processing. The quality of semantic activation was further assessed by Vasos (1983). She used a paradigm in which the name of a category (e.g., weapon) was followed by the presentation of a picture. The subject's task was to decide whether this picture was or was not a member of the category prime. She found that in moderate bilinguals spreading activation of category names was less extensive in L2 than L1, spreading primarily to highly prototypical category members. Furthermore, this priming advantage, or spreading activation, was found to decay almost immediately. This suggests that the L2 reader may be left with more restricted and weaker contextual priming to aid subsequent word recognition. These results also suggest that intervening words may perhaps weaken other connections the reader has made with prior knowledge, thus making text integration and inferencing more difficult.

A number of studies have examined the role of syntactic knowledge in reading. Models such as the Augmented Transition Network (ATN) parser (Kaplan, 1975; Stevens & Rumelhart, 1975) and the Minimal Attachment strategy (Ferreira & Clifton, 1986) suggest that readers' knowledge of syntax permits them to predict the syntactic class of upcoming words, thereby limiting the number and categories of words to be considered in word recognition, as well as facilitating the parsing of sentences for the purpose of propositional assembly. Morrow (1985) provided evidence that syntactic knowledge can influence not only low-level processes, but also high-level processing of global text structure and interpretation. He found that grammatical morphemes such as prepositions, verb aspect and demonstratives helped to organize the reader's interpretation and mental representation of the overall text structure of narrative texts.

Limited ability in syntactic parsing or prediction coupled with slow word recognition could compound L2 readers' difficulties. First of all, the poorer word recognition skills would already be placing increased reliance on syntactic contextual cues. Secondly, if readers are also weak in syntactic knowledge, they would be unable to make effective compensatory use of that information to aid these poor word recognition skills. Few studies have examined the ability of L2 readers to make use of syntactic knowledge while reading. Cziko (1978) examined the speed at which beginner, intermediate, and advanced L2 readers of French could read short texts aloud under three conditions: *random* positioning of words in the text; semantically *anomolous* but syntactically correct positioning of words; and *meaningful* composition, both semantically and syntactically correct. As compared to the random text condition, all groups had faster reading times for the anomolous texts, and the advanced L2 and native readers, but not the beginner or intermediate L2 readers, also showed additional facilitation with meaningful texts. Cziko, therefore, concluded that even beginner-level L2 readers are sensitive to syntactic cues and able to make use of them in decoding.

More recent research examining the compensatory use of syntactic and semantic context when the focus of the task is to read for meaning and readers encounter unknown words, found that syntactic constraints appear to be overlooked as clues for guessing at the meaning of these unknown words. Both Haynes (1984) and Laufer and Sim (1985) found that rather than attending to syntactic cues, L2 readers often restructured the sentence to conform to guesses, often erroneous, based on word analysis, graphemic cues, and local context. Indeed, the use of context as a compensatory strategy for guessing at the meaning of *unknown* words can be ineffective and misleading. Using passages from real-life texts, Schatz and Baldwin (1986) found no significant difference between L1 readers' ability to infer the meaning of unknown words in context or in isolation. Haynes (1984) found that L2 readers were able to make use of local context clues (those within the same sentence) to guess at meaning, but could profit little from global (passage level) clues. Together, these findings suggest that not only does top-down facilitation of word recognition from, for example, syntactic or semantic contextual cues, require time; for the L2 reader it may also be an inadequate source of information due to insufficient understanding of the surrounding local or global context.

In summary, prior research suggests that L2 readers are deficient in several low-level component skills of reading. They evidence less skill in the use of orthographic redundancies and phonological codes, less automaticity in the semantic activation of words, and less extensive and durable spreading activation of word meanings. The ability of the L2 reader to compensate for these weaknesses by top-down syntactic or semantic contextual processing also appears to be limited. The reason for this may be two-fold. First, the quality of the reader's understanding of contextual cues may be lacking precisely because of the low-level processing problems. Secondly, deficiencies at the lower level of processing may be drawing much needed resources away from higher level components such that the reader has too few cognitive resources left over for high-level processing of the text.

*Assessing the Relative Importance of Component Skills*

It is difficult to determine the relative influence of skill at different levels from studies which focus on a single component or processing level. For example, an experimental task focusing on lexical access may demonstrate that skilled and less skilled readers differ on this low-level processing skill, but it cannot determine whether this difference plays a significant role overall in their differential reading ability. Likewise, studies which examine only higher level comprehension processes may overestimate the role of these components since they typically do not partial out differences in low-level ability; as discussed earlier, deficiency in higher level processing may be due to deficiencies at low-level processing, either because this low-level processing diverts needed resources or because it provides inaccurate or insufficient information needed to support higher level integration.

Multiple regression techniques have recently been applied to the study of reading in an attempt to overcome the problems associated with studies which focus on an isolated process by assessing the relative contribution of processing at several levels within a single study (Graesser, Hoffman, & Clark, 1980; Graesser & Riha, 1984). In these studies, the subjects' task is simply to read through a number of passages one sentence at a time on a video monitor while the reading times for each sentence are recorded (for applications of this basic design to word reading times see Haberlandt, 1984; Haberlandt & Graesser, 1985; Haberlandt, Graesser, & Schneider, 1989). Various properties of the sentences are measured and used to predict the variance in reading times across sentences. Two basic assumptions underlie this design and direct the selection of text characteristics. First, the time required to read and comprehend a given sentence is assumed to reflect the combined resource requirements of all ongoing processing of that sentence. Second, since it is impossible to observe directly the resources allocated to each of the component processes, an approximation of the demands at various levels of processing is reached by measuring those text characteristics which are likely to correspond to specific underlying processes or levels of processing (e.g., word frequency used as an index of lexical access demands). Some of the text characteristics chosen, such as number of propositions or syntactic complexity, correspond to lower level processing components. Others, such as number of coreferents or sentence importance, correspond to higher level, or global, processing components of reading. Using the properties of the sentences as predictor variables, standard multiple regression analysis estimates the unique contribution of a process associated with a given predictor by calculating the proportion of total predicted reading time variance accounted for by that single predictor after the effects of all other predictors have been partialled out.

As with any analytic method, multiple regression and its application to this experimental design has strengths and limitations. Perhaps the most serious limitation is that the size and significance of the regression coefficients used to estimate the relative importance of various processes can be strongly influenced by the variability of and intercorrelations among the predictors selected. This means that the pattern of results for a given reader or group of readers will differ depending on the texts used in the experiment. While this makes direct assessment of distribution of resources problematic, it is still useful to make comparative assessments of resource distribution on the same texts, for example, between readers of different abilities or with different reading goals, since the influence of the text characteristics would be held constant. (For a fuller examination of the limitations of multiple regression analysis of reading times see the discussion section of this thesis; see also, Graesser & Riha, 1984; Knight, 1984).

This method also has a number of potential advantages over more traditional analysis of variance (ANOVA) designs employing experimental manipulation and control. First, rather than covarying out or experimentally controlling for extraneous variables to examine the effects of a selected few, many variables can be entered into a single regression equation and their relative effects assessed. This means that reading passages need not be artificially constructed for experimental purposes; rather, any reading passage from the real world can be used, thus improving the ecological validity of the study. Second, once the sentence reading time data are collected, the sentence reading times can be reanalysed as new predictors are discovered or better measures of old ones are developed. The text sentences need only be scaled for the additional variables of interest and these new preditors entered into the regression equation.

Graesser, Hoffman, and Clark (1980) used the above method to compare how cognitive resources are distributed among low and high-level components of reading in fast versus slow L1 readers of English. Allocation of resources to low-level processing was assessed through three local text attributes: number of words, number of propositions, and syntactic complexity. Three global text attributes represented high-level processing: number of new argument nouns, passage familiarity, and passage narrativity. They found that, although all readers allocated most resources to global or high-level processing, the crucial difference between fast and slow readers was on local or low-level components. Furthermore, in a second experiment they found that variation in reading goals (preparation for an essay versus a multiple choice exam) influenced only high-level processing. Subjects in the essay condition allocated more resources to high-level components than subjects in the multiple choice condition. The amount of time devoted to processing local components, however, did not differ between the two reading goal conditions. "The speed of executing microstructure analyses (low-level processing) was different for fast versus slow readers, but was relatively constant when readers read for different purposes" (Graesser, Hoffman, & Clark, 1980:145). These findings suggest that not only is low-level processing crucial to skilled reading, its resource demands are also unaffected by strategic control. Tasks, therefore, designed to assist the reader with global comprehension and text integration may actually have little effect on the improvement of deficient low-level processes.

To summarize, results of L1 reading research from various paradigms, including eye movement studies and studies using multiple regression analysis of reading times, suggest that while high-level processing is very effortful, skilled low-level processing is both essential to reading ability and a primary characteristic of efficient, skilled readers. While studies on isolated components have found that, like less skilled L1 readers, L2 readers are often deficient in low-level skills, no studies have been conducted to date that examine the relative contribution of such deficiencies to overall differences in reading ability between skilled L1 readers and less skilled L2 readers. The purpose of the present study is to examine this question.

*Assessing Component Skills in L2 Reading: The Present Study*

This study on L2 reading addresses the following question: *Which components of reading skill are primarily responsible for slower reading in the L2?* More specifically, *to what extent are differences in reading efficiency between L1 and L2 reading due to high-level or low-level processing, or both?* It is hypothesized that L2 readers will require more time to perform low-level processing than L1 readers since they may not have had sufficient reading experience in the target language to develop efficiency in such components as lexical access and local propositional assembly. In addition, it is possible that L2 readers will exhibit some difficulties in high-level processing as well, since inefficiencies at the lower level may draw needed resources away from high-level integrative processing, leading the reader to read more locally.

To examine this question, the present study used a multiple regression design similar to that employed by Graesser, Hoffman, and Clark (1980) described above. L1 and L2 readers of English, both groups skilled L1 readers, read through a set of English texts one sentence at a time while their sentence reading times were recorded. The sentences were measured on a number of text characteristics. These became the predictor variables used to predict reading times in subsequent regression analyses. The text characteristics were chosen in light of previous research and theoretical models of reading relating them to processing effort. The processing effort in the present study was assessed by the size of the regression slope associated with a given text characteristic predictor variable. The regression slopes of the L2 readers were then compared with those of the L1 readers in order to assess the source of differences in allocation of processing resources across these various text features and their underlying reading components. The selected text features were conceptually divided into three levels of processing: a) *within-sentence* -- processing information locally, within the boundaries of a sentence; b) *between-sentence* -- connecting information across sentences; and c) *text-level* -- understanding the text globally and integrating it with prior knowledge.

The predictors at the within-sentence level were expected to show the most marked differences between L1 and L2 readers. There were four within-sentence predictors:

*Number of words.* One of the basic units a reader must process is the word. Clearly, sentences comprised of many words will take longer to read than short sentences and, therefore, number of words is a likely predictor of sentence reading time. Number of words read per minute is a common measure of reading ability and is typically calculated by dividing the total reading time of a passage by the number of words in the passage. This measure alone, however, is not a good estimate of word decoding demands since averaged into this are all text processing demands on the reader. Several processes take place at the word level and may be implicated in this predictor, including visual decoding, phonological recoding, and lexical access. Measuring such word attributes as orthographic regularity and grapheme-phoneme regularity might help to partial out the relative processing demands of visual and phonological decoding, but would be more appropriate predictors in a paradigm with word reading times as the criterion variable, and were, therefore, not included in the present study. It was expected that sentence reading times would increase as a function of number of words and that L2 readers would have larger regression coefficients than L1 readers for this variable.

*Mean word log frequency.* Eye movement studies have found that words which occur less frequently in a language are processed more slowly than highly frequent words (Just & Carpenter, 1980; Mitchell & Green, 1978). This facilitation may be due to increased skill in handling a frequent subset of words, including skill at visual decoding, phonological recoding, and lexical access of these words. These and other studies relating word reading times to frequency are usually based on the norms found in Kucera and Francis (1967), a compilation of word frequencies from a 1,014,232 word corpus of English. These norms are typically converted into their natural logarithmic values since small differences in frequency among common words have a comparable effect on reading times to large differences among rarer words (Just & Carpenter, 1980). In the present experiment, log word frequencies were averaged across the words in each sentence. It was predicted that as sentences increase in mean log word frequency, sentence reading time would decrease, thus yielding a negative regression slope. It is difficult to predict the outcome of the comparison between L1 and L2 readers. A larger slope for L1 than L2 readers might suggest that they are better able to benefit from highly frequent words due to modularized automaticity. However, it is also possible that a larger slope will be obtained for L2 readers. In this case, the most consistent interpretation would be that L2 readers who operate in a slower, controlled processing mode are more sensitive to the relative frequency of a word.

*Number of propositions.* In comprehending a text, the reader builds up an internal representation of the text in memory. Kintsch and van Dijk (1978; Kintsch, 1974; van Dijk & Kintsch, 1983) developed a model in which the basic units of this representation are characterized as propositions. A proposition is the smallest meaningful idea unit which can be verified as true or false, and is made up of a predicate and one or more arguments. A predicate can be a verb, or a defining relation, such as modifier or conjunction, linking the arguments. According to Kintsch's model, a reader extracts the underlying meaning of sentences in the form of micropropositions, builds up a coherent text base by linking these propositions through argument repetition, and condenses the detailed micropropositional representation into macropropositions through a set of rules called macro-operators.

Numerous experiments have been conducted which help to support the psychological validity of propositional representations. For example, Kintsch and Vipond (1979) were able to account for readability of various texts previously only poorly captured by more traditional readability formulas by taking into account potential individual differences among readers in their working memory capacity for processing propositions. In a study designed to determine the relative influence of number of words and number of propositions on reading time by varying number of words while keeping number of propositions constant, Kintsch and Keenan (1973) found that reading time was positively correlated with number of propositions to be processed, but not number of words.

In the present study, a propositional analysis was performed on each text following guidelines provided by Turner and Greene (1977), and the total number of propositions per sentence was counted. Number of propositions per sentence then served as a predictor variable in the regression analyses. Sentence reading time was expected to increase as a function of the number of propositions per sentence, and L2 readers were expected to require more processing time per proposition, resulting larger regression coefficients than those of L1 readers.

*Syntactic predictability.* The complexity of the syntactic structure can seriously affect the ease with which a reader processes and comprehends a sentence. Highly unpredictable syntax, such as garden path sentences often result in confusion and may necessitate rereading a sentence or phrase. Numerous models exist for assessing the syntactic complexity of a sentence (for reviews, see Huggins & Adams, 1980; Mitchell, 1987) including its derivational complexity, defined as the number of transformations necessary to convert the surface structure of a sentence into its underlying conceptual level, or deep structure; the number of open, unresolved syntactic units the reader must store in memory (Yngve, 1961); and the extent to which a sentence corresponds to series of interpretive strategies applied to parsing (Bever, 1970).

For the purposes of a cognitive processing approach to understanding reading, a model which relates syntactic complexity to ease of ongoing processing is most relevant and useful. Mitchell (1987) reviews evidence which suggests that readers adopt an on-line guess and backtrack approach to parsing such that as each word is encountered the reader makes an initial guess at parsing it without first looking ahead to upcoming words to verify the correctness of this first guess. If it subsequently turns out to be incorrectly parsed, the reader then backtracks to rectify this interpretation, as happens with garden-path sentences. The model most consistent with this evidence is the Augmented Transition Networks (ATNs) parser (Kaplan, 1975; Stevens & Rumelhart, 1975). At any given point in a sentence (referred to as a state), this model lists all grammatically correct possibilities for the syntactic class of the following word and the order by which these predictions should be attempted. In an experiment requiring readers to predict upcoming words of a sentence it was found that, like the ATN parser, readers had highly fixed orders of expectancy regarding the syntactic class of the following word. Furthermore, this model was able to predict points of reader processing problems due to misattribution of syntactic class. Where the following word was of a class less likely to occur, readers asked to read texts aloud made significantly more errors, and their substitutions corresponded to the syntactic class most expected based on the state of the ATN at the preceding word (Stevens and Rumelhart, 1975).

In order to assess the ease of syntactic parsing of a sentence, therefore, the present experiment used a predictive approach based on the ATN parser. Since ATNs are not yet able to account for all syntactic parsing of the English language, the extent to which the syntax of each sentence of the experimental texts conformed to actual predictions made by readers was used in this study as an index of the ease of syntactic processing of each sentence. It was hypothesized that an increase in the syntactic predictability of a sentence would be associated with a decrease in sentence reading time. Two different outcomes are possible for the comparison of L1 to L2 regression slopes. A larger slope for L1 readers than L2 readers may result if L1 readers are better able to profit from simple syntactic structures, whereas a larger slope for L2 readers may reflect greater difficulty with complex syntax.

Three between-sentence predictor variables were selected to tap processes involved in integrating ideas across sentences. According to Perfetti (1985), these processes will be particularly resource-costly and should evidence, therefore, relatively large standardized regression coefficients for both L1 and L2 readers. No differences between L1 and L2 readers, however, are expected at this level since the basic propositions will have already been extracted from the surface structure at the within-sentence level. The between-sentence variables were:

*Number of new argument nouns.* New argument nouns represent the required concept nodes forming the basic structure of an internal representation of a text. Other relationships and qualifiers are then attached to these basic nodes. As readers encounter each new argument noun, a new node must be set up, or foregrounded, in memory. Because this requires processing time, texts with more new argument nouns take longer to read than those with fewer but more highly elaborated concepts, even when the number of propositions is held constant (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). In the present study, therefore, a positive correlation between number of new argument nouns and reading time per sentence was expected, but no differences between L1 and L2 readers were expected.

*Number of intersentential coreferents.* In order to integrate information across sentences, a reader must link together words and information that pertain to a given referent. Linguistic cues, such as pronouns, definite articles, and word repetition, are used to signal a coreferent and prompt the reader to search for its antecedent. Locating an antecedent within the same sentence, particularly the same clause, is done relatively quickly and demands minimal resources (Clark & Sengul, 1979). Anaphoric reference across sentences, however, is increasingly resource costly because the antecedent may have been displaced from working memory by more recent information and must be reinstated before a match can be made (Lesgold, Roth, & Curtis, 1979; Cirilo, 1981). The number of times a reader must perform this process is likely to increase the overall time it takes to read the sentence. Therefore, reading time is expected to increase with the number of intersentential coreferent searches required. Comparison between L1 and L2 regression coefficient was expected to be non-significant.

*Maximum distance to antecedent.* While a sentence may take longer to read because of the sheer number of antecedent reinstatement searches required, coreferents for which the antecedent is very distant may be particularly demanding on the reader's processing resources. The more extensive a memory search the reader must perform in order to locate and reinstate the antecedent information, the more time, or resources, are required. Ideally, distance is best measured by the number of intervening propositions between the coreferent and its antecedent once the surface structure of the text has been processed and assembled into a propositional textbase in which propositions are connected on the basis of underlying meaning (McKoon & Ratcliff, 1980). While a complete, logical propositional textbase of a passage can be constructed following Kintsch and van Dijk's method of propositional analysis (1978), what this textbase would look like at a given point in a reader's processing of the passage is much more difficult to infer since it may vary dramatically as a function of the reader's processing capacity (Kintsch & Vipond, 1979). For the present study, therefore, distance to the coreferent was broadly estimated by counting the number of sentences from the coreferent back to its most recent antecedent. If a sentence had more than one coreferent, that sentence was assigned the value of the coreferent with the largest number of intervening sentences to its antecedent. It was hypothesized that reading times of sentences with a distant antecedent search would be greater than those with antecedents found in recently read sentences. No differences were expected between the mean regression coefficient of L1 and L2 readers.

Four text-level predictor variables were selected as indicators of the resources allocated to global text processing. Sensitivity to any of the text-level features would be evidenced by larger slope coefficients to these predictors. It was hypothesized that for any differences observed between L1 and L2 readers at this level, the L2 readers would have significantly smaller regression coefficents than the L1 readers. In other words, if L2 readers were allocating most of their cognitive resources to local processing components, they might fail to take notice of the larger structural elements of the texts, and consequently would not adjust their distribution of resources as a function of these attributes. The text-level predictor variables were:

*Sentence importance.* Readers have been shown to devote more processing time to sentences central to the main idea of a text (Cirilo & Foss, 1980). These sentences may contain redundant information, particularly if they are summarizing earlier text, making them less demanding at a local processing level. However, it is hypothesized that readers attending to high-level text integration will recognize the importance of this information and process it more fully, resulting in longer sentence reading times. It was expected therefore, that L1 readers would show stronger positive regression slopes than L2 readers. L2 readers are expected to show no effect at all or a smaller, but positive, effect if they are somewhat sensitive to sentence importance.

*Section-initiating sentences.* Sentences which introduce a new section of a text are generally associated with increased reading time (Lorch, Lorch, & Mogan, 1987; Lorch, Lorch, & Matthews, 1985). Again, while this sentence may not be more difficult to process locally, a reader sensitive to the larger structure of the text will recognize this shift and may devote additional resources to such tasks as predicting upcoming information and preparing a schema structure to accomodate it. In the present study, it was predicted that sentences which introduced a shift in the topic of the text would be read more slowly than other sentences. It was also predicted that this effect would be more pronounced for L1 readers than L2, who may indeed fail to recognize shifts if processing primarily at a local level.

*Section-concluding sentences.* It has been suggested that sentences which conclude a section of a text may also associated with increased reading time (Haberlandt, Berian, & Sandson, 1980). Readers devoting processing attention to higher level comprehension are more likely to recognize the concluding nature of this sentence and may, therefore, choose to allocate additional resources to consolidating the information presented in this section. This would result in an increase in reading times associated with these sentences.

In a later study, Haberlandt (1984) found that the effect for his 'boundary' predictor variable which included both section-initial and section-concluding sentences was signficant only for section-initial sentences. Since the issue of section-concluding wrap-up still remains a possibility despite Haberlandt's later finding, it was decided to leave it in the analysis and let the outcome speak to this question. It was predicted that, if this factor did influence reading time, sentences which conclude a section would be read more slowly than sentences internal to a section. Again, L1 readers were expected to evidence larger regression slopes than L2 readers.

*Narrativity.* Empirical studies have found that narrative texts are read more quickly and are better retained than expository texts (Graesser, Hauft-Smith, Cohen, & Pyles, 1980; Graesser, Hoffman, & Clark, 1980; Haberlandt, Berian, & Sandson, 1980). Graesser and Riha (1984) attempted to relate narrativity judgements to the degree to which sentences expressed active versus static propositions, but this measure was unable to capture the reading time variance accounted for by the narrativity ratings. The reading time advantage for narrative texts may be due more to overall schema processing differences than any local attribute such as staticness of information conveyed. Considerable work has been done to differentiate expository, descriptive texts from narratives in terms of their schematic structures (for a review, see Meyer

& Rice, 1984). However, this text feature still lacks clear theoretical links to underlying cognitive processes. While one could hypothesize that readers may be more familiar with narrative episodes or that this structure is more consistent with the way we process day to day events, further research is required in order to determine the processes responsible for the relative ease in reading narrative texts. Hopefully, future developments in text processing will provide a better measure to capture the underlying processes implicated by this factor. In the meantime, however, since the narrativity variable based on subjective ratings proved to be such a powerful predictor in the Graesser, Hoffman, and Clark (1980) study, it was considered important to include it here. Unlike the Graesser et al. texts which were either highly expository or highly narrative in nature, the texts used in the present study interleaved both narrative and expository elements. Consequently, rather than rating the entire passage as did Graesser, Hoffman, and Clark (1980) and applying that rating to each of the sentences in the entire text, the present study rated each paragraph for narrativity and applied that rating to each of the sentences in that paragraph. Sentences from paragraphs rated high in narrativity were expected to take less time to read than sentences from paragraphs rated low in narrativity. Without a clearer understanding of the underlying processes, it is difficult to predict differences between L1 and L2 readers. If facilitation for narrative texts operates through the use of mental scripts or schemas, and L2 readers are less skilled in the activation of appropriate schemas, a possible outcome would be smaller regression slopes for L2 than L1 readers.

A final variable was included to control for the effects of the serial position of a sentence in the text. This control predictor variable was:

*Serial position*. Sentences early in a text tend to be read more slowly than sentences toward the end of a text (Graesser, Hoffman, & Clark, 1980; Haberlandt & Graesser, 1985). This may be due to increased familiarity with the style of the text or topic matter, with later information being more easily integrated once an initial textbase exists in the reader's working memory. It may also reflect an initial settling in to the text and reading task. Negative regression coefficients are therefore predicted for both L1 and L2 readers. However, difficulty with these factors are more likely to be experienced by L2 readers and, therefore, their mean regression coefficient is expected to be larger than that of the L1 group.

Several other predictor variables were considered for inclusion but were ultimately dropped. Of particular interest are those which relate to the reader's integration of the text with prior knowledge, such as the familiarity of the text or the number of inferences drawn while reading. The first of these variables, familiarity, used by Graesser, Hoffman, and Clark (1980) has two problems. First, this variable lacks any solid theoretical rationale and cannot be clearly linked to specific underlying processes, thus making interpretation of any effects found particularly problematic. Secondly, familiarity is likely to vary depending on the reader, and considerable measurement error is introduced if the same rating of familiarity is applied across all subjects as in Graesser, Hoffman, and Clark (1980). The second variable which was finally dropped, construction of inferences, is an important component of reading and may demand considerable resources (van Dijk & Kintsch, 1983; O'Brien, Shank, Myers, & Rayner, 1988; Seifert, Robertson, & Black, 1985). It was finally excluded as a variable in this study since the number and type of inferences drawn on-line is difficult to predict on the basis of text analysis alone, particularly with natural texts, and would likely vary from reader to reader as a function of their prior knowledge and interest. As mentioned earlier, one advantage of this design is that as future research reveals new predictors and better measures of predictors, new regression analyses can be conducted on the sentence reading times already collected. Hopefully, future research developments will enable a reanaysis of this and other data incorporating some of these additional predictors of interest.

In summary, the present study attempts to determine which components of reading are responsible for slower, more effortful reading in the L2. To estimate the amount of resources allocated across various processes and levels, selected text characteristics were used to predict sentence reading times of L1 and L2 readers via multiple regression analyses. The regression coefficients of each text characteristic were used as an index of the resources allocated to the component reading process(es) associated with that characteristic, and served as the basis of comparison between L2 and L1 readers. It was hypothesized that L2 readers would devote significantly more resources to local, within-sentence processing than L1 readers.

# METHOD

## *Participants*

Fifteen anglophone and 15 francophone university student volunteers from Montreal participated in the experiment. They were remunerated at a rate of $5.00 per hour.

Participants were selected and retained for analyses according to the following criteria. First, all potential participants took a screening test to assess their reading ability in both English and French. This screening test was developed at Concordia University for the purposes of this and other studies examining cognitive processes in reading and is an adapted version of the one described in Favreau & Segalowitz (1982). From this participant pool, only those who in their L1 scored high in comprehension (above 75%) and were among the top 50% of readers in terms of speed were selected for participation in the present experiment. Furthermore, only those francophones were chosen who also scored high in comprehension (above 75%) in English, their L2, but read English at least 20% more slowly than French. Thus, both anglophone and francophone participants were highly skilled readers in their L1, while the francophones read English, their L2, considerably more slowly while maintaining high comprehension.

Secondly, since the dependent measure in this experiment was the time it took participants to read and *understand* the texts, any participants who scored below 60% comprehension were eliminated.

As a final descriptive measure, the 15 francophone participants who completed the experiment were also administered the Reading Comprehension and Vocabulary section of the Test of English as a Foreign Language (TOEFL; Educational Testing Service, 1984). Their mean performance on the TOEFL was 52.1/60 with a standard deviation of 5.1. This score is slightly lower than the Educational Testing Service (1987) reported mean of 55/60 ($N$ = 14,914) for native French-speaking applicants to United States eductional institutions from July 1984 through June 1986. It is also just below that of nonnative English-speaking Canadians applying for entrance to United States educational institutions during the same period, whose mean was also 55/60 ($N$ = 1438).

*Design*

The basic design of the study was as follows. Participants read through a set of texts, sentence by sentence. Reading time per sentence for each participant was thus obtained. Selected characteristics of each sentence were measured (e.g., syntactic predictability, importance of the sentence to the overall gist of the text, etc). These characteristics were chosen on the basis of the underlying processing components of reading they are assumed to map onto. As both the characteristics and reading times varied freely across sentences, the values of the characteristics were then used to predict the reading times in multiple regression analyses. The relative processing demands underlying each characteristic were estimated in terms of the resultant regression coefficients of each predictor. Finally, the regression coefficients of L2 readers were compared with those of L1 readers in order to estimate L1/L2 differences in allocation of resources to the selected text characteristics and the component processes assumed to underlie them.

*Materials*

Texts

Five written texts, one warm-up and four experimental, were selected from articles which appeared in popular science magazines. Only minor adaptations were made to the articles. In some cases, the length of the article was condensed through deletion of some sections to insure that the reading passages were of similar length. In other cases, some vocabulary was simplified. This was done to ensure that reading times would reflect normal reading processes, rather than time spent struggling over an unfamiliar word. Grammatical structures were not altered or simplified. The resulting passages were natural, authentic samples of English writing, constructed neither for the purposes of ESL instruction nor for experimentation.

The four experimental texts dealt with the following topics: *Whales*, the rescue of whales from an ice jam ("Whale and Farewell," 1985); *Incubators*, the early development of incubators for premature infants ("When incubator babies," 1985); *Skingrafts,* the treatment of burn victims with cultured skin grafts (Langone, 1984); and *Tornadoes,* the use of a new type of radar in the detection of tornadoes (Olson, 1984). The warm-up passage, *Camels,* discussed raising camels on farms for their meat (Katz, 1982). Descriptive statistics of the texts can be found in Table 1 (see Appendix A for the complete experimental texts).

Text characteristics

Sentences were measured on a number of characteristics. These became the predictor variables used to predict reading times in subsequent regression analyses. A description of how each of the variables was measured is given below (see Appendix B for a complete list of the predictor values for each sentence).

Within-sentence predictor variables:

*Number of words (WORDS).* For this predictor, a count of the total number of words in each sentence was taken. Numbers, hyphenated words, and conjunctions were all treated as single words.

*Mean log word frequency (FREQ).* The frequency assigned to each word of the texts was based on Kucera and Francis (1967) norms. Raw frequencies were transformed to natural logarithmic frequencies. Words which did not

Table 1

*Descriptive Statistics of Reading Texts*

| Text | No. words | No. sent. | No. words/sentence | |
|---|---|---|---|---|
| | | | Mean | *SD* |
| Warm-up | | | | |
| Camels | 739 | 35 | 21.1 | 10.9 |
| Experimental | | | | |
| Whales | 545 | 26 | 21.0 | 12.7 |
| Incubators | 689 | 28 | 24.6 | 12.3 |
| Skingrafts | 594 | 28 | 21.2 | 11.3 |
| Tornadoes | 719 | 36 | 20.0 | 8.6 |
| Mean | 637 | 29.5 | 21.6 | 11.1 |

appear in Kucera and Francis, received a frequency of .9999 (yielding a log frequency of -.0001). The log transformed word frequencies of individual words were then averaged across each sentence to yield a mean log word frequency value per sentence.

*Number of propositions (PROPS).* A propositional analysis was performed on each experimental text using Kintsch and van Dijk's (1978; Kintsch, 1974) notational system and following the guidelines laid out by Turner and Greene (1977). The total number of propositions per sentence was then counted (see Appendix D for an illustrative example of the propositional analysis).

*Syntactic predictability (SYNTAX).* In a separate study, 11 native English readers progressively predicted each subsequent word of the texts. In this study, participants were seated before a personal computer. Upon presentation of the first word of the text on the video monitor, they were requested to predict the following word, type in their prediction, and press return. The actual word occurring in the text then appeared on the screen and the participant predicted the subsequent word. This procedure continued in a cyclical fashion until the final word of each passage. Participants were never required to predict the first word of a sentence. Instead, after a brief pause following the appearance of the final word of a sentence, the entire sentence was cleared from the screen and the first word of the next sentence automatically appeared. Participants' predictions were stored in a computer file for later analysis. Both the actual words of the text and the participants' predictions were assigned syntactic categories based on the system outlined in Quirk, Greenbaum, Leech, and Svartvik (1985). The participants' predictions were then scored for whether they corresponded to the *syntactic class* of the actual word in the text (see Appendix C, Table C-1 for an illustrative sample of categorization and scoring). Since the participants predicted only one word at a time, occasionally the syntactic class of the word predicted was ambiguous. In these cases, the participant was given the benefit of the doubt and the syntactic category was scored as correct. For each word of the texts, the proportion of correct predictions was then calculated. Each sentence was assigned a syntactic predictability score equal to the mean of the proportions of correct predictions of the individual words making up that sentence (see Appendix C, Table C-2 for a sample tally sheet for calculation of these values.)

Between-sentence predictor variables:

*Number of new argument nouns (NANS).* For each text, every noun which introduced a new concept into the text was identified as a new argument noun. This was not simply a count of every new noun. For each noun, the underlying referent concept was determined. If it was the first time that concept was mentioned, that noun was considered a new argument noun. If the same concept was referred to using various words (e.g., icebreaker, Moskva, vessel, boat), only the first mention was counted. Conversely, if identical nouns were used to refer, nevertheless, to different underlying referents (e.g., the word *whales* was used to refer on one occasion to whales in general, and on another, to the actual whales trapped in the strait), each mention was considered a new argument noun. After this analysis was completed, the number of new argument nouns per sentence was tallied.

*Number of intersentential coreferents (COREFS).* Linguistic cues, such as pronouns, definite articles, and word repetition, were used to identify coreferents. Any coreferent whose antecedent was found within the same sentence was not counted since these matches are made with relatively little effort (Clark & Sengul, 1979). The number of intersentential coreferents were tallied for each sentence.

*Maximum distance to antecedent (DIST).* Distance to the antecedent coreferent was determined by counting the number of sentences from the coreferent back to its most recent antecedent. If a sentence had more than one coreferent, that sentence was assigned the value of the coreferent with the largest number of intervening sentences between that coreferent and its antecedent.

Text-level predictor variables:

*Sentence importance (IMPORT).* The values for this predictor were determined in a separate study. Eleven participants were asked to reduce each text to its most essential ideas. First, they were required to write one sentence that summarized the main idea of the text. This served to focus their attention on what they considered the essential information. Then they were asked to successively delete sentences which did not contain the main ideas. This deletion was done in three stages, one quarter of the sentences at a time; the sentences which remained were those the participant considered the most essential to the gist of the passage. The participants were told not to worry if the deletion disrupted the flow of the passage as the remaining sentences would be rewritten to produce a coherent summary. For each participant, each sentence received an importance rank of 0, 1, 2, or 3 corresponding to its retention after the first, second, or third round of deletions (e.g., 3 was assigned to those sentences remaining after all three rounds of deletions). Finally, each sentence was assigned an importance ratio defined as the total score summed across participants divided by the total possible score of 33 (a rank of 3 from all 11 participants).

*Section beginning (BEGIN).* The values for this predictor variable were also determined within a separate study. Eleven native English readers were given each experimental text printed out as a list of sentences. They were asked to partition the text into its component sections by drawing a line immediately above each sentence which introduced a shift in topic within the text, thus creating six to ten divisions per text. The participants' responses were compiled, and divisions were assigned to those locations where a majority of the participants (minimum six participants) had drawn a line. Sentences which initiated a new section within the text were assigned a categorical value of 1, while sentences within a section or ending a section received a value of 0. This method of dummy coding categorical variables can be found in Pedhazur (1982).

*Section ending (END).* For this predictor, the same section divisions were used as those for BEGIN. In this case, however, the last sentence of a section was assigned a categorical value of 1, while sentences initiating a section or internal to a section received a value of 0.

*Narrativity (NARR).* The ratings of narrativity assigned to the sentences were determined in a separate study in which participants were asked to rate on a 7-point scale the degree to which each paragraph was in a narrative genre (tells a story) or an expository genre (conveys static information). The low end of the scale was labeled, 'passage conveys static information', and the high end was labeled, 'passage conveys active information with actions and events unfolding in time'. The mean rating across participants for a given paragraph was then assigned to each of the sentences comprising that paragraph.

*Serial position (POSIT).* Values for this predictor were obtained by numbering sentences sequentially from start to finish of each text.

Comprehension measures

For each passage, 23 true/false verification statements were constructed to assess comprehension (see Appendix E for a complete list of these statements). Ten of these statements were true, 10 were false, and 3 were indeterminate (there was insufficient information provided in the article to verify their truthfulness). In their development, these statements were pilot tested to ensure that no answers were obviously true or false without having read the text, and that readers with the text available for verification consistently agreed on the correct answer.

As a second comprehension measure, participants were asked to retell the texts in their own words. These verbal retells were audiotaped, but not analysed for the purpose of this thesis.

Other materials and equipment

A consent form was prepared for each participant, a copy of which can be found in Appendix F. Written instructions for each experimental task were constructed and can be found in Appendix G. Reading texts and true/false statements were presented on an Apple IIe monitor. Participants advanced through the texts and selected their answers on a three-button response panel. Reading times and responses were recorded by an Apple IIe microprocessor using a specially prepared software routine.

*Procedure*

All participants were screened in a separate session prior to their participation in the experiment. For both the screening and the experiment, participants were tested individually in a private room.

The basic experimental procedure was as follows. Texts were presented on an Apple IIe monitor one sentence at a time. Participants advanced to each subsequent sentence when ready by pressing a button. The computer recorded the time between button presses (the time spent reading each sentence). Upon completion of reading, participants rated their familiarity with the topic of the text and with the particular story highlighted. Two comprehension tasks followed. First, participants retold the text in their own words while being audiotaped. They then verified 23 true/false statements presented on the computer monitor one at a time.

All instructions were given to the participants in writing, and clarified if necessary with the experimenter. Participants were asked to read as quickly as possible without sacrificing comprehension. They were also asked not to take any pauses during the reading section of the experiment since their sentence reading time was being recorded. The familiarity ratings were included for the possible development of an additional predictor variable assessing the impact of prior knowledge on reading time, but were not analysed for the present thesis. Both comprehension tasks were assigned to ensure that participants were reading for comprehension. Performance on the text retell task was recorded, but not assessed; its primary function was to encourage participants to process the text more fully and at a more global level than would be necessary for the subsequent factual recall task (T/F statement verification). Criterion performance on the verification task was used to eliminate those participants who read too quickly to comprehend the text or who did not attend to the task. This manipulation check was essential given that the assumption underlying the dependent measure was that the time taken to read a sentence reflected the time needed to process and comprehend that sentence. To reduce a 50% correct guessing probability rate on the verification task, participants selected their response from among four choices: TRUE, FALSE, CAN'T REMEMBER, and NOT ENOUGH INFORMATION. These were presented as equally viable options; indeed, for three out of the 23 statements there was truly insufficient information in the text to allow the participant to respond correctly with either true or false. Participants were told to take their time verifying the statements as only their responses and not their decision-making time was being recorded.

All participants went through the complete experimental procedure on a warm-up text first to familiarize them with the equipment and tasks. Any clarifications necessary were made before proceeding. The participants then read and completed the comprehension tasks for each of the four experimental texts. Order of presentation of the experimental texts was counterbalanced across participants.

Experimental sessions consisted of either two appointments, the first approximately one hour and the second 45 minutes, or of a single appointment of approximately two hours with a 15 minute break in the middle. The break between sessions was always given after completion of three texts, the warm-up and two experimental.

## RESULTS

In order to interpret appropriately the results of multiple regression analyses, descriptive statistics on all variables, zero-order correlations, and multicollinearity among predictors must also be examined. These will be reported first, followed by results of the multiple regression analyses, themselves. In reporting the results, several issues in multiple correlation analysis will be touched on. A more detailed exploration of these issues and their implications for interpretation of results can be found in the Discussion section of this thesis.

*Descriptive Statistics*

Table 2 lists the means (*M*s) and standard deviations (*SD*s) of the predictor variables across the 118 sentences. The *SD*s of the predictors are of particular interest as they can have a profound impact on the size and significance of the regression coefficients. In general, the larger the *SD* of a predictor the stronger its apparent effect on the criterion variable. The predictors in the present study differed greatly in variability. Compare, for example, FREQ ($M$ = 6.22, $SD$ = .60) with NANS ($M$ = 3.09, $SD$ = 2.53). Consideration of the influence of this issue on the results of this and previous studies employing this design will be addressed in the Discussion.

Table 2

*Descriptive Statistics of Predictor Variables*

| Predictor | Mean | *SD* |
|---|---|---|
| Within-Sentence | | |
| WORDS | 21.59 | 11.12 |
| FREQ | 6.22 | 0.60 |
| PROPS | 10.59 | 6.19 |
| SYNTAX | 0.57 | 0.09 |
| Between-Sentence | | |
| NANS | 3.09 | 2.53 |
| COREFS | 1.75 | 1.14 |
| DIST | 3.49 | 4.98 |
| Text-Level | | |
| IMPORT | 0.49 | 0.22 |
| BEGIN | 0.25 | 0.44 |
| END | 0.24 | 0.43 |
| NARR | 4.61 | 1.38 |
| Control | | |
| POSIT | 15.50 | 9.00 |

Table 3 provides the *M*s and *SD*s of the L1 and L2 sentence reading times. Although there is some overlap between the fastest L2 and the slowest L1 readers, on average, L2 readers read the experimental texts approximately 30% more slowly than L1 readers (8750 vs. 6284 msec per sentence, $t(28) = -2.897$, $p<.01$). This corresponds to a mean reading rate of 160 words per minute (wpm) for L2 readers and 212 wpm for L1 readers. L2 readers did maintain comparable levels of comprehension as measured by their scores on the T/F sentence verification task (L2, 77.33%, versus L1, 81.08%, ns). These results are comparable to those obtained on the screening test used to select the two groups of readers, and thus serve to confirm that the selection criteria were met.

*Zero-order Correlations and Collinearity*

Table 4 reports Pearson zero-order correlations among the predictor variables and between the predictors and the mean L1 and L2 sentence reading times. Applying Bonferroni-adjusted probabilities for multiple tests, reading times of both groups remained significantly correlated with four predictors: WORDS (L1, r=.941, p<.001; L2, r=.965, p<.001); PROPS (L1, r=.855, p<.001; L2, r=.877, p<.001); NANS (L1, r=.785, p<.001; L2, r=.766, p<.001); and IMPORT (L1, r=.494, p<.001; L2, r=.472, p<.001).

When predictors are completely orthogonal, the squared zero-order correlations express unique proportions of variance accounted for by the

Table 3

*Descriptive Statistics of Participants' Reading Times and Comprehension Scores*

| | Reading Time | | | | Comprehension | |
|---|---|---|---|---|---|---|
| | per sentence(msec) | | words per minute | | (%) | |
| Group[a] | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| L1 | 6284 | 1125 | 212 | 35.9 | 81.08 | 8.28 |
| L2 | 8750 | 3098 | 160 | 37.5 | 77.33 | 9.93 |

[a]*N* = 15 for each group.

Table 4

*Correlation Matrix of the Mean L1 and L2 Reading Times and the 12 Predictor Variables*

| Predictor | Reading Time | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | L1 | L2 | W | F | P | S | NA | C | D | I | B | E | N |
| WORDS (W) | .94* | .97* | | | | | | | | | | | |
| FREQUENCY (F) | -.24 | -.22 | -.10 | | | | | | | | | | |
| PROPOSITIONS (P) | .86* | .88* | .91* | -.16 | | | | | | | | | |
| SYNTAX (S) | -.26 | -.25 | -.22 | .31 | -.29 | | | | | | | | |
| NEW ARG NOUN (NA) | .79* | .77* | .75* | -.26 | .67* | -.19 | | | | | | | |
| COREFERENTS (C) | .20 | .20 | .26 | .00 | .20 | .01 | -.15 | | | | | | |
| DIST TO ANT (D) | .04 | .52 | .11 | .03 | .08 | -.10 | -.10 | .41* | | | | | |
| IMPORTANCE (I) | .49* | .47* | .47* | -.25 | .49* | -.24 | .31 | .23 | .12 | | | | |
| SECTION BEG (B) | .04 | .01 | -.02 | -.16 | -.06 | -.04 | .27 | -.26 | -.02 | -.03 | | | |
| SECTION ENDING (E) | -.01 | -.01 | .07 | .17 | .06 | -.01 | -.05 | .18 | -.05 | .02 | -.33* | | |
| NARRATIVITY (N) | .15 | .12 | .08 | -.01 | .10 | .25 | .06 | .24 | .01 | -.01 | .03 | .00 | |
| SERIAL POSITION (SP) | -.23 | -.21 | -.13 | .16 | -.12 | -.05 | -.30 | .15 | .34* | .09 | -.15 | .05 | -.34* |

*$p<.001$, Bonferroni-adjusted.

predictor variable. With non-manipulated variables this is rarely the case, and an examination of the correlations among predictors indicates that there is considerable intercollinearity among the predictors in this study. The most serious of these is the correlation between WORDS and PROPS ($r=.914$, $p<.001$). This means that WORDS and PROPS are highly redundant as predictors of reading time. Both these predictors are also correlated quite strongly with NANS (WORDS, $r=.752$, $p<.001$; PROPS, $r=.674$, $p<.001$), and IMPORT (WORDS, $r=.471$, $p<.001$; PROPS, $r=.491$, $p<.001$).

In order to assess further the degree of redundancy among predictors, a series of multiple regression analyses were performed in which each predictor in turn served as criterion variable. The multiple $R^2$ listed with each predictor in Table 5 indicates the amount of variance in that predictor which can be accounted for by the other predictors together. The variances in WORDS and PROPS are highly predictable by the other variables ($R^2$ WORDS = .910, $R^2$ PROPS = .860), further suggesting a problem of multicollinearity.

Because multiple regression analyses partial out redundant variance accounted for by the predictors, extreme multicollinearity can create some serious problems. It can lead to inaccurate regression coefficients, enlarged standard errors of coefficients, and small unique proportions of variance associated with those predictors which are collinear. In extreme cases, it may even reverse the expected sign of a regression coefficient (Pedhazur, 1982). Indeed, in preliminary standard multiple regression analyses on the mean L1

Table 5

*Multicollinearity among Predictor Variables: With and Without WORDS*

| Predictor | With WORDS $R^2$ | Without WORDS $R^2$ |
|---|---|---|
| WORDS | .910 | |
| FREQUENCY | .279 | .226 |
| PROPOSITIONS | .860 | .650 |
| SYNTAX | .259 | .255 |
| NEW ARGUMENT NOUNS | .772 | .631 |
| COREFERENTS | .509 | .392 |
| DISTANCE TO ANTECEDENT | .283 | .279 |
| IMPORTANCE | .322 | .317 |
| SECTION BEGINNING | .280 | .279 |
| SECTION ENDING | .154 | .153 |
| NARRATIVITY | .288 | .266 |
| SERIAL POSITION | .329 | .326 |

and mean L2 sentence reading times in which the complete set of predictors was entered, the signs of the regression coefficients associated with PROPS and with COREFS were negative when, on the basis of theory and zero-order correlations, positive signs were expected.

A commonly applied solution to multicollinearity problems, the elimination of variables, was used here. In order to minimize potential specification errors due to the omission of relevant predictors from the equation, only the most seriously collinear predictors, WORDS and PROPS, were considered for deletion. Clearly, in natural texts, number of words and propositions are bound to covary, and to a large extent, they reflect a similar processing demand, namely, volume of single information units processed and stored in working memory. Retaining two similar measures of the same underlying variable confounds the analysis and weakens the estimation of the effect of this variable.

Of the two, WORDS was eliminated for the following reasons. First, number of propositions is conceptually more meaningful. In comprehending the text, many models of reading propose that it is some form of these units of information that are processed in working memory and stored for later retrieval (for an overview, see Meyer & Rice, 1984). Furthermore, because function words are included in the total number of words, number of propositions is a better measure of the processing demands of the meaning-laden content words. In other words, propositions, not single words, appear to be the building blocks of reading comprehension, and thus reflect this low-level text work more accurately than words. Finally, elimination of WORDS did not remove the consideration of processing at the level of lexical access from the analysis. While lexical access does occur at the word-level, and is clearly an important first step in extracting meaning, number of words alone is a poor measure of this underlying process. Another predictor, FREQ, was considered to reflect better the sensitivity of the reader to demands of lexical access.

Thus, as a measure of the demands on working memory of low-level information units, PROPS was considered most tenable and therefore retained for all subsequent analyses, while WORDS was eliminated. This removed the most extreme collinearity problems. The second column of Table 5 shows the results of multicollinearity analyses with WORDS removed. The highest multiple $R^2$, that of PROPS (multiple $R^2$ = .650), was considered acceptable.

*Multiple Regression Analyses*

*Analysis of mean sentence reading times.* First, once WORDS was removed, the remaining predictors were regressed on the group means of the criterion variable, sentence reading time. Since there was no theoretical model to direct the order of entry of variables into a hierarchical regression equation, standard multiple regression was used. This treats each variable as if it were entered last into the equation. In other words, each coefficient is assigned only the unique proportion of variance it accounts for, variance of all other predictors having been partialled out. Having removed number of words, the signs for PROPS and COREFS were found to be in the direction expected.

Results of the regressions analyses on mean L1 and L2 sentence reading times are presented in Tables 6 and 7, respectively. In both these global analyses, the predictors were able to account for a significant 84.9% of the variance in reading time (L1 readers, $R^2 = .849$, $F(11, 106) = 54.387$, $p<.001$; L2 readers, $R^2 = .849$, $F(11, 106) = 53.972$, $p<.001$). The semipartial regression coefficient ($sr^2$) represents the proportion of variance accounted for uniquely by each predictor. Only a small proportion of the total prediction was accounted for by each of the predictors uniquely (L1 readers, $\Sigma sr^2 = .180$; L2 readers, $\Sigma sr^2 = .183$); the remainder constituted predicted variance which was shared by various combinations of predictors due to intercorrelations (L1 readers, .820; L2 readers, .817). Since the size of the regression coefficients was of primary interest, rather than the smallest number of significant predictors, no variables were eliminated just because they were unable to account for a significant unique proportion of the variance.

Three variables were significant predictors of both L1 and L2 mean sentence reading time: PROPS (L1, $t = 7.086$, $p<.001$; L2, $t = 8.648$, $p<.001$), NANS (L1, $t = 7.392$, $p<.001$; L2, $t = 6.276$, $p<.001$), and COREFS (L1, $t = 3.429$, $p<.01$; L2, $t = 2.931$, $p<.01$). In addition, IMPORT was a significant predictor of L1, but not L2, reading time (L1, $t = 2.033$, $p<.05$; L2, $t = 1.049$, ns).

Table 6

*Multiple Regression Analysis on Mean L1 Sentence Reading Times*[a]

| Predictor[b] | Coef | Std Error | Std Coef | $sr^2$ | *t* | *p* |
|---|---|---|---|---|---|---|
| FREQ | -66.355 | 255.135 | -0.011 | 0.000 | -0.260 | 0.795 |
| PROPS | 260.650 | 36.783 | 0.451 | 0.072 | 7.086 | 0.000 |
| SYNTAX | -1083.793 | 1691.270 | -0.028 | 0.001 | -0.641 | 0.523 |
| NANS | 647.921 | 87.658 | 0.459 | 0.078 | 7.392 | 0.000 |
| COREF | 519.254 | 151.416 | 0.166 | 0.017 | 3.429 | 0.001 |
| DIST | -16.355 | 31.884 | -0.023 | 0.000 | -0.513 | 0.609 |
| IMPORT | 1489.556 | 732.816 | 0.093 | 0.006 | 2.033 | 0.045 |
| BEGIN | -361.414 | 361.701 | -0.044 | 0.001 | -0.999 | 0.320 |
| END | -485.370 | 342.663 | -0.058 | 0.003 | -1.416 | 0.160 |
| NARR | 84.565 | 114.259 | 0.033 | 0.001 | 0.740 | 0.461 |
| POSIT | -20.702 | 18.232 | -0.052 | 0.002 | -1.135 | 0.259 |

[a]Multiple $R^2$ = .849. [b]*N* = 118.

Table 7

*Multiple Regression Analysis on Mean L2 Sentence Reading Times*[a]

| Predictor[b] | Coef | Std Error | Std Coef | $sr^2$ | *t* | *p* |
|---|---|---|---|---|---|---|
| FREQ | -113.865 | 343.935 | -0.014 | 0.000 | -0.331 | 0.741 |
| PROPS | 428.824 | 49.586 | 0.553 | 0.107 | 8.648 | 0.000 |
| SYNTAX | -432.740 | 2279.921 | -0.008 | 0.000 | -0.190 | 0.850 |
| NANS | 741.585 | 118.167 | 0.391 | 0.056 | 6.276 | 0.000 |
| COREF | 598.316 | 204.116 | 0.142 | 0.012 | 2.931 | 0.004 |
| DIST | -6.172 | 42.981 | -0.006 | 0.000 | -0.144 | 0.886 |
| IMPORT | 1036.633 | 987.875 | 0.048 | 0.002 | 1.049 | 0.296 |
| BEGIN | -579.285 | 487.592 | -0.053 | 0.002 | -1.188 | 0.237 |
| END | -667.553 | 461.928 | -0.059 | 0.003 | -1.445 | 0.151 |
| NARR | -4.947 | 154.027 | -0.001 | 0.000 | -0.032 | 0.974 |
| POSIT | -25.851 | 24.578 | -0.048 | 0.002 | -1.052 | 0.295 |

[a]Multiple $R^2 = .849$. [b]$N = 118$.

*Analysis of individual participants' sentence reading times.* Although analyses on sentence reading times averaged across subjects have been performed frequently in similar studies (e.g., Graesser, Hoffman, & Clark, 1980; Graesser & Riha, 1984; Haberlandt, 1984), this method fails to take into account the variance in reading times across subjects and increases the risk of Type 1 errors (Lorch & Myers, 1990). In order to take between-subject variability into account, a two-step method of analysis suggested by Lorch and Myers (1990) was used. The following procedure was carried out on both the L1 and L2 groups. First, individual regression equations were computed for each participant's reading times. Second, the unstandardized regression coefficients (*b*s) from each of the individual participants were entered into a table in which the participants labeled the rows and each predictor variable labeled a column. The group mean unstandardized regression coefficient of each predictor was then tested to see if it differed reliably from zero using single-group *t* tests. Finally, to compare L1 and L2 processing, an additional set of *t* tests were performed to determine if the coefficents of L1 readers and L2 readers differed significantly from each other. The unstandardized (*b*, slope) rather than the standardized regression (β) coefficients were used to compare across subjects since *b*s are fairly stable across populations whereas βs are sensitive to changes in the *SD*s of the predictor and criterion variables (Pedhazur, 1982; pp. 247-251).

An initial assessment of the stability of the effect of each predictor across participants involved calculating the proportion of participants in each group who showed the same direction of effect on each of the predictor coefficients as well as the proportion for whom the coefficients reached significance. Significance for individual regression coefficients was set at $p<.10$, since individual reading times are subject to much greater error variability due to, for example, lapses of attention. Table 8 lists the proportion of participants with positive and significant regression coefficients. The same three variables which were found significant in the global analyses, were again the most stable predictors of individuals' sentence reading times. Specifically, for both L1 and L2 readers, all participants showed a significant increase in reading time as number of propositions and number of new argument nouns increased, and all had positive slopes for COREFS, though only 67% of these in each group were significant. IMPORT and END were also fairly stable in the direction of their slopes (IMPORT: L1, 93% positive; L2, 73% positive; END: L1, 87% negative; L2, 93% negative); however, for few of the subjects did either IMPORT or END predict a significant proportion of the total variance in reading time (IMPORT: L1, 20% significant; L2, 7% significant; END: L1 and L2, 13% significant).

The mean L1 and L2 unstandardized coefficients are presented in Figure 1. *T* tests on the unstandardized coefficients revealed a number of similarities and differences in the influence of the predictors on L1 and L2

Table 8

*Proportion of Participants with Positive and Significant Coefficients in Individual Analyses*

| | L1[a] | | L2[b] | |
|---|---|---|---|---|
| Predictor | +ve coef | sig ($p$<.10) | +ve coef | sig ($p$<.10) |
| FREQ | .33 | .00 | .40 | .07 |
| PROPS | 1.00 | 1.00 | 1.00 | 1.00 |
| SYNTAX | .40 | .07 | .27 | .00 |
| NANS | 1.00 | 1.00 | 1.00 | 1.00 |
| COREF | 1.00 | .67 | 1.00 | .67 |
| DIST | .13 | .00 | .40 | .00 |
| IMPORT | .93 | .20 | .73 | .07 |
| BEGIN | .27 | .20 | .20 | .07 |
| END | .13 | .13 | .07 | .13 |
| NARR | .60 | .07 | .40 | .13 |
| POSIT | .33 | .07 | .00 | .13 |

[a]$N$ = 15. [b]$N$ = 15.

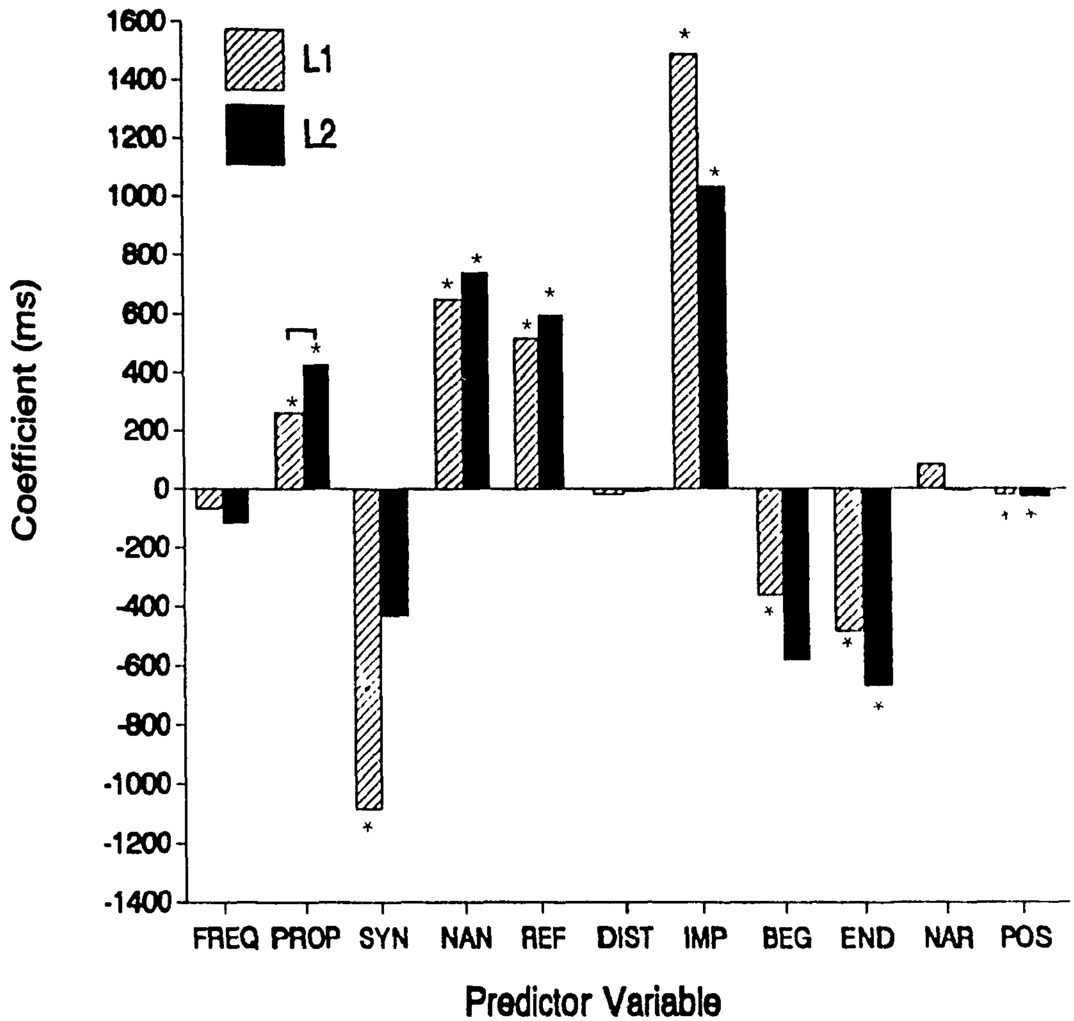


*Figure 1.* Mean unstandardized regression coefficients of individual L1 and L2 analyses.

reading times (see Table H-1 in Appendix H, for exact values). The mean regression coefficients of both L1 and L2 groups were significantly different from zero on PROPS (L1, $t(14) = 16.415$, $p<.01$; L2, $t(14) = 11.794$, $p<.01$), NANS (L1, $t(14) = 12.854$, $p<.01$; L2, $t(14) = 6.185$, $p<.01$), COREFS (L1, $t(14) = 10.811$, $p<.01$; L2, $t(14) = 6.011$, $p<.01$), IMPORT (L1, $t(14) = 4.961$, $p<.01$; L2, $t(14) = 2.441$, $p<.05$), END (L1, $t(14) = -4.084$, $p<.01$; L2, $t(14) = -2.751$, $p<.05$), and POSIT (L1, $t(14) = -2.147$, $p<.05$; L2, $t(14) = -6.919$, $p<.01$). That is, an increase in number of propositions per sentence, number of new argument nouns, number of coreferents, and sentence importance was associated with a significant increase in mean sentence reading times of both L1 and L2 readers; and there was a significant decrease in mean sentence reading time of both groups for sentences which concluded a section (END) and for later occurring sentences in the text (POSIT). The effect for END, however, was in the opposite direction from that anticipated.

In addition, L1, but not L2, sentence reading times were significantly predicted by SYNTAX (L1, $t(14) = -2.211$, $p<.05$; L2, $t(14) = -0.715$, ns) showing a significant decrease in reading time as syntactic category of words became more predictable, and by BEGIN (L1, $t(14) = -2.387$, $p<.05$; L2, $t(14) = -1.812$, ns). The direction of the effect for BEGIN, however, was also in the opposite direction from that predicted; sentences which introduced a new section within the text were read more quickly than those which did not.

There was one significant difference between the two groups in the comparison *t* tests. While both groups evidenced a significant increase in reading time as the number of propositions increased, L2 readers showed a significantly greater increase than L1 readers ($t(28) = -4.239$, $p<.001$).

The unstandardized regression coefficient associated with a given predictor represents the expected change in the criterion variable associated with one unit change in that predictor variable. For example, in the present case, for each proposition processed, L1 and L2 readers evidenced a mean increase in reading time of 261 milliseconds and 429 milliseconds, respectively. Clearly, since the units of measure differ for each of the predictors, it is inappropriate to use the size of the unstandardized coefficients to assess the relative importance of the predictors. Instead, the mean standardized regression coefficients, or beta weights, are often used for this purpose since they convert the values to a common, standard scale. However, betas must be interpreted with extreme caution since, although scale-free, they are affected by the variances and covariances of the variables in the analysis as well as those variables left out and subsumed by the error term. Because of these concerns, the mean standardized coefficients for each group were computed, but were not compared quantitatively. They are presented only in graphic form to provide a qualitative picture of the relative effect of the predictors on reading time. As illustrated in Figures 2 and 3, PROPS and NANS, followed by COREFS, seem to be most important in predicting the reading times of both L1 and L2 readers.

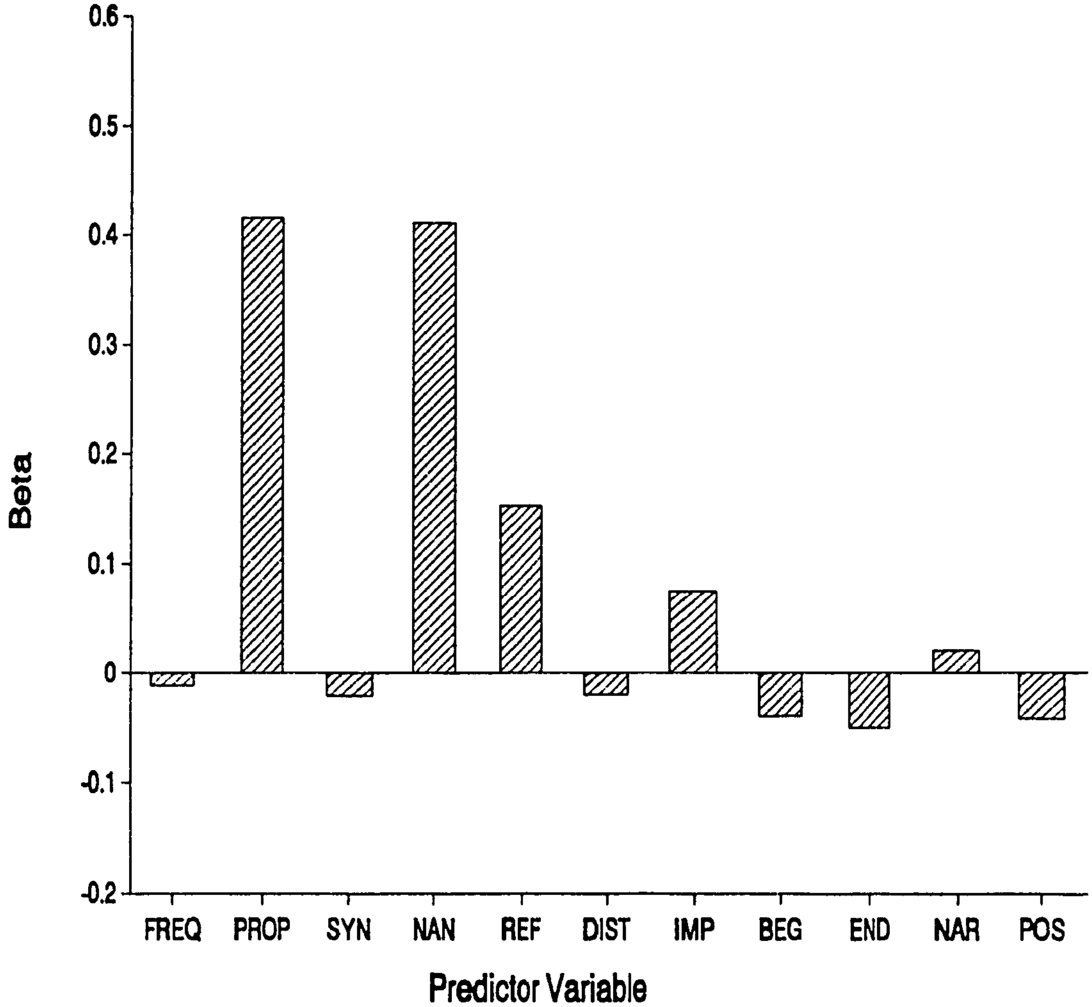


*Figure 2.* Mean L1 standardized regression coefficients.

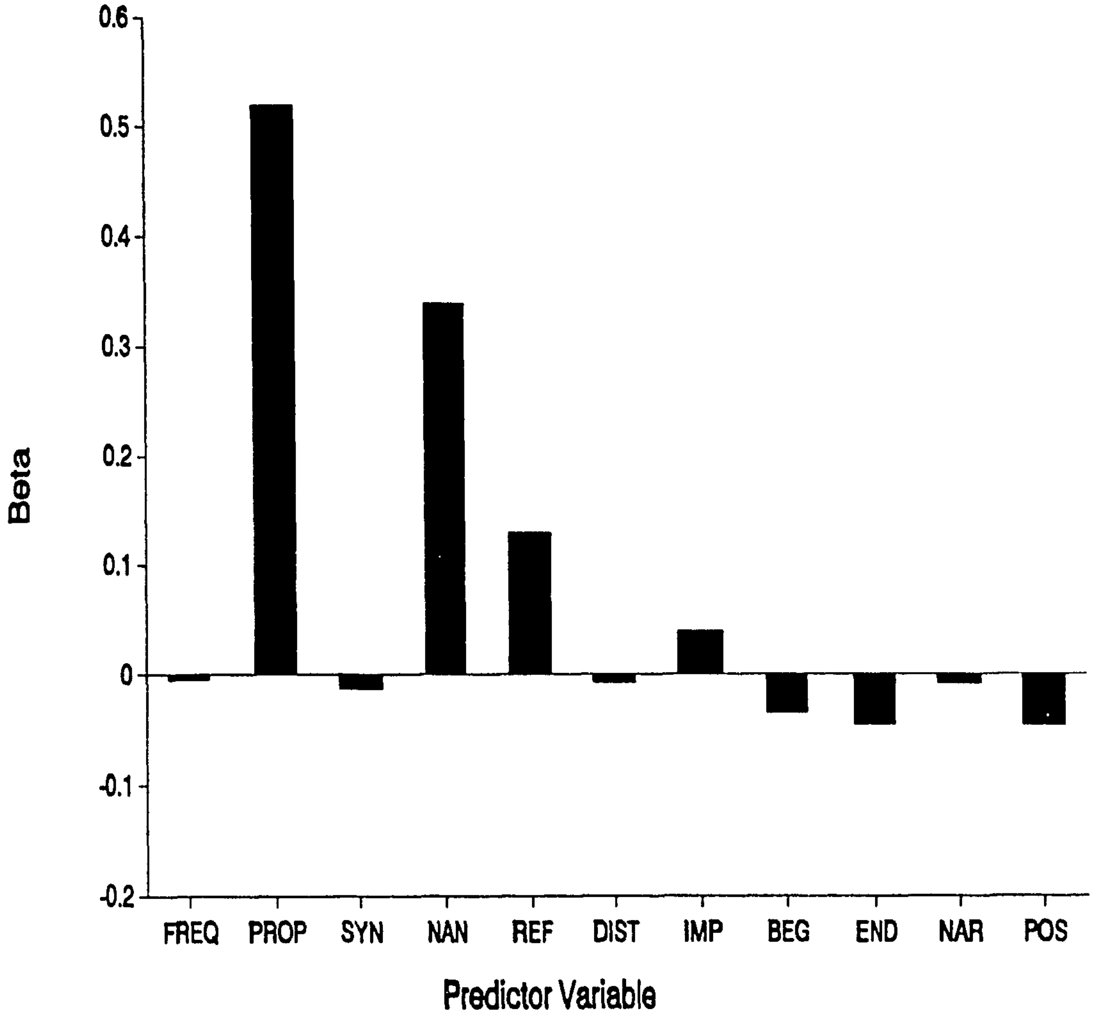


*Figure 3.* Mean L2 standardized regression coefficients.

# DISCUSSION

This study was designed to assess which aspects of the reading process are responsible for slower reading in the second language. More specifically, it compared skilled L1 readers to L2 readers in terms of their allocation of cognitive resources across selected component processes in reading. The discussion of the results will first address this question. A number of important methodological issues also arise from this study. How these issues influence the interpretation of the findings of this study and previous ones will be discussed second. Third, recommendations for future research in this area will be considered in light of both the current findings and the methodological advantages and limitations of this approach. Finally, the implications of the findings for L2 reading instruction will be discussed.

*Allocation of Cognitive Resources in L1 and L2*

The L2 readers in this study were very able readers of English. They scored high on the TOEFL test and attained similar comprehension levels to native readers on authentic English texts both on the screening and the experimental texts. To achieve this high comprehension on the experimental texts, however, these L2 readers had to compromise reading speed. The 30% difference in reading speed between L1 and L2 readers is of both statistical and practical significance since such a reduction in reading speed could put L2 readers at a disadvantage in educational or professional settings where extensive reading in English is required. The multiple regression analyses sought to determine the sources of this reading speed deficit.

L1 and L2 readers differed significantly on just one predictor, PROPS. While both groups showed a significant increase in reading time for each proposition processed, L2 readers clearly had more difficulty with processes involving propositions. As assessed by the slope of the regression line, which partials out the other predictors in the regression model, reading time per proposition was significantly longer for L2 readers than for L1 readers. They required, on average, 65% longer than L1 readers per proposition. This difference in propositional processing is large and a theoretically important one.

According to Perfetti's model of reading efficiency (Perfetti, 1985), propositional assembly and integration is critical to reading skill. Only the most elementary type of propositional encoding, the local assembly of two or three words into a proposition can become automatized, but the controlled processing in the encoding of complex propositions and the integration of propositions within and between sentences can vary considerably in efficiency and constitutes a large part of the work a reader must perform in constructing meaning from text. The present measure, total number of propositions per sentence, reflected both elementary propositional encoding and more complex within sentence assembly and integration of propositions. This would account for the smaller, yet significant, slope for L1 readers. It suggests that L1 readers have become highly skilled in propositional assembly and integration, perhaps automatic in elementary encoding, but that as an important, controlled process in reading, propositional integration continues to demand considerable cognitive resources.

In a study by Kintsch and Vipond (1979), it was demonstrated that the capacity of the reader for storing propositions in working memory (the buffer) and for handling incoming propositions (input size) was able to account for the relative readability of two different texts, which standard readability measures failed to capture. In other words, one text was found to be more difficult than the other only when processing characteristics of the reader in terms of efficiency in propositional encoding and integration were considered. Because of the additional effort required, L2 readers may be able to buffer and encode fewer propositions at a time than L1 readers.

Only one variable, FREQ, was expected to show sensitivity of the reader at the level of phonological decoding or lexical access. However, neither L1 nor L2 readers had sentence reading times which correlated significantly with FREQ. In other words, reading time was not affected by the familiarity of the word as measured by its log frequency in a large English-language corpus (Kucera & Francis, 1967). One possible reason for this may be that mean log frequency per sentence was not a sensitive enough measure. Indeed, there was very little variance across sentences on this measure ($M$ = 6.218, $SD$ =

.600). An experimental paradigm measuring *word* reading time, such as the moving window display or eye movements, might better capture processing differences reflected by word frequency and other word-level features (see, for example, Haberlandt, 1984; Haberlandt & Graesser, 1985).

SYNTAX, the final within-sentence predictor, was associated with a significant decrease in reading time for L1 readers, but not L2. It seems as though L1 readers were better able to take advantage of simple syntactic structures to facilitate their processing of the text, than were L2 readers. A plausible explanation is that L1 readers had better syntactic knowledge, or more resources available to take advantage of this knowledge through top-down interactive processing to help them assign relationships more quickly in propositional assembly. However, although the difference between the means appears large, it is not statistically significant, probably due to the large within-group variability.

L1 and L2 readers had similar patterns of results for the between-sentence variables of NANS and COREFS. Even once the effect of the PROPS was partialled out, both L1 and L2 readers showed a significant mean increase of 648 and 742 msec respectively for each new argument noun introduced and of 519 and 598 msec respectively for each between-sentence coreferent. Both these variables represent part of the text work associated with propositional assembly and integration and would, therefore, be expected to contribute significantly to reading time. That no significant difference was found between the groups, however, suggests that the distinguishing difference between L1 and L2 readers is at the lower level of propositional assembly, that of actually transforming the text from surface-level structure to a more abstract propositional level. Keeping in mind that these L2 readers were highly skilled in their own L1, it may be that these higher level processes operate to some extent independently of language. That is, once the propositions have been extracted from language-specific text, reading skills acquired in L1 may begin to make a larger contribution to reading in L2. This would be an interesting issue to examine more closely in future work.

The final between-sentence variable, DIST, did not have a significant effect for either the L1 or L2 readers. Failure to find an effect may have been due to the imprecision of this measure. It was a rough measure designed to reflect the maximum processing demands per sentence associated with searching for and reinstating antecedants. Surface distance, the number of sentences back to the antecedant, was chosen to avoid making assumptions regarding the reader's mental representation of the text. However, studies using contrived texts which contrast surface and logical distance have demonstrated that the reader performs this search within a mental representation of the text, not on the basis of surface structure (McKoon & Ratcliff, 1980). Errors in prediction of reading time with this measure may have, therefore, arisen for at least two reasons. First, depending on the importance of a given concept, some readers may not have bothered in all cases to undertake a search for the antecedant, particularly a distant, resource-costly one. Second, depending on the underlying logical structure of a text, an antecedant far in surface distance may actually have been much closer in psychological distance. This suggests that examination of the relative role of reinstatement searches within this experimental paradigm would necessitate building a tentative processing model of the text using a procedure such as that outlined by Kintsch and van Dijk (1978). However, given the complexity of these natural texts and individual differences in processing characteristics, this constructed model may be subject to important measurement errors also.

Both L1 and L2 readers were sensitive to the importance level of a sentence as evidenced by increased reading times as sentences increased in the value of IMPORT. That L2 readers showed a significant effect of importance suggests that they were not focusing solely on local processing; rather they were devoting attentional resources to the global structure of the text and were able to distinguish the relative importance of the information on the basis of this structure. The size of the mean regression coefficients (L1, $M$ = 1489.556; L2, $M$ = 1036.633) suggests that L1 readers may have allocated more resources to this purpose than L2 readers; however, there are large individual variations in the size of the coefficients, particularly among the L2 readers such that the difference between the means is not significant.

Mean narrativity ratings of the paragraphs, NARR, had no effect on the reading times of either group, despite the fact that previous studies have found narrative passages are read more quickly than expository (Graesser, Hoffman, & Clark, 1980; Graesser & Riha, 1984; Haberlandt & Graesser, 1985). The reason for this may be that none of the texts used in the present study were purely narrative or expository, but were a mixture of the two (see the following section of this discussion for a fuller explanation of this issue). An added problem is that the underlying processes implicated in this factor are not clear; further research is needed to illuminate the differences in processing of narrative versus expository texts before the role of this text variable within a cognitive process model of reading can be assessed.

The two predictors which assessed the effect of shifts in subsections within the passages, BEGIN and END, both had unexpected results. Sentences which concluded a section were associated with a *decrease* in reading time for both the L1 and L2 readers, and although the effect for BEGIN was significant only in L1, both groups also evidenced a *decrease* in reading time for sentences which introduced a new subsection. The mode of presentation of the texts may have affected the impact of these factors on the participants' reading times. Recall that sentences appeared on the video monitor one at a time. Consequently, there were no layout cues such as paragraphs to alert the readers to the fact that a section just began or was about to end.

It was anticipated that a reader sensitive to the global organization of a text would devote more attention to introductory sentences in order to activate or construct a schema to facilitate the integration of upcoming information. In the absence of paragraph cues, however, the reader may have failed to recognize these shifts in topic, particularly if the shift was not dramatic. Furthermore, since sentences beginning a subsection also contained more new argument nouns ($r_{BEGIN,NANS}$ = .27), it may be that NANS partialled out much of the anticipated effect of BEGIN.

It was also anticipated that sentences ending a section would be associated with increased reading time as the reader would direct resources to wrapping up or consolidating what had just been read before going on to the next section. Again, without paragraph cues, the reader may have failed to take this into account. Indeed, these sentences may have been relatively easily processed on a local level since they are likely to contain redundant or contextually primed information. This may have accounted for the observed decrease in reading time associated with END.

The last predictor, serial position of the sentence (POSIT), had a small, but significant, effect for L2 readers, once the effect of other related factors such as NANS, IMPORT, and COREFS had been partialled out. The difference between L1 and L2 readers, however, was not significant. It may be that initially L2 readers took longer to 'get into' the text, activate appropriate schema, or accomodate to the writer's style, and that processing of sentences later in the text was facilitated through contextual priming. L1 readers, presumably operating at a high-level of efficiency, would not show as strong an effect since initial sentences would not add substantial burden, and contextual priming in later sentences would not significantly effect processes which are already highly-skilled.

The general picture which emerges from the findings, then, is that propositional text work accounts for the largest proportion of variance in reading time of both groups of readers. This includes propositional encoding, as reflected by PROPS, and propositional integration, as reflected by NANS and COREFS. In the global analyses these three variables accounted for the largest unique proportions of predicted variance. Their importance was further supported by the large mean beta weights associated with them in the individual analyses. It is lower-level propositional encoding, however, not higher level integration, that distinguishes the L1 and L2 readers. L2 readers appear to read more slowly, in part, due to less efficient local processing of propositions (they may also be deficient in processes prerequisite to propositional encoding since these were not adequately partialled out in this study's design). Propositional integration across the text, which takes place in working memory within the reader's abstracted, mental model of the text, does not put differentially larger demands on the L2 reader. This suggests that skill in higher level text processes developed in the reader's L1 may transfer to L2 reading situations, whereas lower-level skills which are more dependent on the reader's ability and practice in the specific language of the text, may be what limits the efficiency of reading in a second language.

*Methodological Issues*

Clearly, interpretation of reading processes based on the size and significance of regression coefficients is exploratory and tentative. Further research is needed to test the interpretations forwarded here. Some of the findings for the L1 readers differ dramatically from those in previous studies. A closer examination of some limitations of multiple regression and factors which may have influenced the results of these and similar analyses offers potential explanations for this, and further alerts the researcher to several methodological issues which need to be considered when employing multiple regression.

One powerful influence on the size and significance of the coefficients is the degree of variability in the independent and dependent variables (Pedhazur, 1982). Of particular importance here is the effect of the variabilities in the independent, or predictor variables. The larger the variability of the values of a given predictor, the greater will appear to be its effect on the dependent variable. Conversely, if the range of the values is restricted, or controlled, its effect will be attenuated. This factor may be partially responsible for the discrepant findings of the present study with those of Graesser and his colleagues. In the present study authentic texts were used in which sentence length was left to vary naturally. The consequent variation in predictors related to sentence length was very large, particularly for number of words (mean = 21.6, SD = 11.1) and number of propositions (mean = 10.6, SD

= 6.2). This contrasts greatly with texts used by Graesser and his colleagues (Graesser, Hoffman, & Clark, 1980; Graesser & Riha, 1984). Their texts were segmented into sentences of between 10 and 16 words, resulting in a restricted variation in related predictors, such as number of words (mean = 11.9, SD = 1.1) and number of propositions (mean = 5.1, SD = 1.3). It is not surprising, therefore, that measures affected by sentence length proved more important predictors in the present study than in those of Graesser and colleagues. Indeed, the one measure which proved most predictive of reading time in Graesser, Hoffman, and Clark (1980) was passage narrativity. This variable was assigned values based on narrativity ratings of the 12 texts used in these studies, six of which had been specifically chosen as highly narrative and six as highly expository. Each of the sentences in a given text were assigned the narrativity rating of the entire text. On a seven-point scale, the ratings of the six narrative passages ranged from 5.75 to 6.75, whereas the ratings of the six expository passages ranged from 1.58 to 3.25 (Graesser, Hauft-Smith, Cohen, & Pyles, 1980). Such a distribution of extreme scores will greatly increase the size and significance of a predictor's coefficient, and this along with the range restriction on sentence length likely played an important role in the results of the analysis in Graesser, Hoffman, and Clark (1980).

Another factor influencing the size and significance of the coefficients is the degree of multicollinearity among the predictor variables. In standard multiple regression analysis the effect of each variable is estimated after the effects of all the other predictors in the equation have been partialled out. This means that the size of the coefficients of independent variables which are highly intercorrelated will be drastically reduced since all explained variance they share will have been removed. Furthermore, the chance of finding significant coefficients in the presence of multicollinearity is also reduced since the larger the correlation among predictors, the larger the standard errors of the coefficients. This can make it difficult to assess the relative importance of a given variable.

In the present experiment, there was a fairly high degree of multicollinearity, even once number of words was removed as a predictor. Consequently, although together the predictors accounted for approximately 85% of the total variance in reading time of both L1 and L2 readers, most of this prediction was shared by the independent variables. Even number of propositions, the strongest predictor, uniquely accounted for only 7.2 % of the total variance in reading time, though it had extremely high zero-order correlations with L1 and L2 mean sentence reading times (L1, r=.86; L2, r=.88). It is perhaps not surprising that a large amount of reading time variance is shared by multiple predictors in this type of study. Many of the text variables are closely related and covary with sentence length. Many of the underlying processes as well interact and are difficult to assess in isolation. Even a design such as the one used in the present study cannot fully disentangle the complex interplay of component processes in reading. Graesser, Hoffman, and Clark

(1980) managed to reduce the degree of multicollinearity among predictors only at the expense of controlling sentence length. By restricting the range, for example, of number of words, the correlation between this and other variables such as number of propositions, would be attenuated. Having done this, however, it is no longer possible to examine the relative role of reading processes within a complete context. Controlling for number of words and using texts which yielded extreme values for narrativity has a profound influence not only on the coefficients of these two variables, but also on all other predictors in the equation. Restricting sentence length precludes accurate assessment of what Perfetti (1985) claims is a large part of the text work of a passage, namely, propositional assembly and integration. The size of the regression slopes associated with many of the predictors of the Graesser, Hoffman, and Clark (1980) study are difficult to interpret since the texts had been manipulated for experimental purposes. In the present experiment, where texts were not significantly altered, text features which reflected processes at the level of propositions accounted for the largest proportion of unique variance. As discussed above, however, caution must be taken in interpreting the size of the regression slopes in the present experiment as well since intercollinearity partials out most of the predictive power of the individual variables and most of the predicted variance remains shared.

A third factor affecting the outcome of any study employing multiple regression concerns the potential consequences of omitting relevant variables from the equation on the interpretation of the regression coefficients of the variables included. This type of specification error can be committed quite inadvertently if, for example, the researcher decides to limit the number of variables to consider or theory is not yet sufficiently advanced to reveal all the relevant variables. The consequences of omitting relevant variables are not too serious if they are not correlated with any of the existant predictors; total predicted variance will be reduced and the significance of the coefficients reduced due to increased standard errors. However, if they are correlated and they are not included in the equation, any predicted variance which is shared will be attributed uniquely to the remaining predictors, thus biasing their estimation. In an area as complex as reading behaviour the risk of inadvertently omitting relevant variables is high. Since many of these relevant variables are likely to be somewhat intercorrelated, caution must be exercised in interpreting and generalizing the results from a single study. The possibility of re-analysing the existing reading time data in light of new theory is an important advantage offered by the paradigm used in this study since the impact of an omitted variable can be subsequently assessed within an otherwise identical data set.

In addition to accidental omission, relevant variables are sometimes intentionally removed from the equation to eliminate sources of multicollinearity. This is an easy, but in light of the consequences of omitting relevant variables, a potentially dangerous way to reduce multicollinearity.

Unless the goal of the analysis is purely to increase the predictive power of a set of variables, elimination of variables must be based on theoretical arguments since the entire pattern of the results may be altered by this omission. In removing number of words in the present study, it was argued that this measure was to a large extent repetitive of number of propositions. The bias introduced by multicollinearity was judged to be greater than the bias resulting from eliminating number of words from the analysis. This decision and its potential consequences need to be considered when interpreting the coefficients of the remaining variables. In particular, number of propositions must be considered a proxy variable not just of those processes directly involving propositional processing, but also of any low-level processes implicated in the processing of small, local units of information. Difficulties with any prerequisite processes to propositional assembly, such as phonological recoding, lexical access, and semantic encoding, have been only partially removed through the mean word frequency variable. Consequently, differences between L1 and L2 readers in the coefficient of number of propositions may be due the quality of several low-level processes, not only propositional encoding.

All of the independent variables in these studies are, in fact, proxy variables used as indicators of underlying processes. Care must be taken to minimize interpretational errors due to misrepresenting their conceptual basis. The stronger the theoretical link to a specific underlying processes, the more informative is the coefficient associated with a given independent variable. Measures such as narrativity, or familiarity (Graesser, Hoffman, & Clark, 1980), are less informative than more specific variables since they are not reflective of any particular underlying process. At the same time, as illustrated above, variables which appear to be highly specific, such as number of propositions, have to be interpreted within the entire context of the regression analysis and reading theory, keeping in mind those processes which are represented by the remaining variables and especially those which are not.

Despite the obvious differences between this study and that of Graesser, Hoffman, and Clark (1980) in the characteristics of the texts used and the consequent influences on the regression coefficients, both these studies ultimately point in a similar direction. Even though Graesser, Hoffman, and Clark (1980) controlled the variation in number of words, it was the second largest predictor of sentence reading time after narrativity (the effect of which it was argued was overestimated). Secondly, when comparing fast and slow L1 readers of English, it was the low-level variables which differentiated between them. Both these observations point to the importance of efficient use of low-level processing skills in skilled reading. In contrast to Graesser, Hoffman, and Clark (1980), the present study, by not controlling number of words, found a much stronger effect for variables related to sentence length, both for variables reflecting low-level, within-sentence processing and higher level, between-sentence processing. However, like Graesser, Hoffman, and Clark (1980), the distinguishing characteristic between skilled L1 readers and less-skilled, in this case L2, readers of English was at the local processing level.

*Recommendations for Future Research*

As evidenced by this and the Graesser, Hoffman, and Clark (1980) study, better understanding of reading processes can only come through multiple studies, using a wide range of texts and a variety of research designs. Experimental designs using artificially constructed texts which isolate a particular process through stringent controls have the advantage of determining whether this process is significantly different in two groups of readers, but cannot address the relative importance of this process in the entire reading situation. Less controlled studies employing multiple regression to address the question of relative importance also encounter limitations in assessing this important issue. While no single method can provide a complete picture, triangulation of findings with careful attention to each design's strengths and weaknesses will contribute to the refinement of current and future models of reading.

Thus far, there is considerable evidence indicating low-level processing deficiencies in less-skilled L1 and L2 readers of English. L2 readers, even those considered skilled bilinguals, have been found to be less efficient in the use of orthographic redundancies in L2 (Favreau, Komoda, & Segalowitz, 1980), to exhibit slow, controlled processing in L2 semantic activation (Favreau & Segalowitz, 1983), and to be less skilled in manipulating phonological codes in working memory in their L2 (Segalowitz & Hébert, 1990). The importance of these deficiencies relative to higher level processes was supported by the results of the present study. While the design of this study is limited in its ability to pinpoint specific underlying processes, when combined with other evidence a more complete view of L2 reading begins to emerge.

Future research is expected to lead to the identification of additional predictor variables and to more accurate measurement of these variables. This will improve the regression model and the stability of the regression coefficients. Studies employing a larger variety of text genres and difficulty will test the generalizability of the present findings across a wider spectrum of reading materials. Other methods of collecting reading times on individual words, such as eye movements and the moving window display, could be usefully applied to multiple regression designs such as the one here to examine in greater depth the nature of local processing factors in L2 reading (e.g., Haberlandt, 1984). The present study used advanced L2 readers of English and fairly difficult texts. A study using simpler texts would allow one to compare L2 readers across different ability levels to examine whether there are any significant differences in resource allocation associated with general L2 ability, or to conduct a longitudinal study examining change in resource allocation within the same subject as skill in the target language increases.

Once the issues of replicability and generalizability of the regression coefficients have been clarified, there are a number of applied studies to which this method of assessing reading skill might be usefully employed. One study may be to explore the effect of pre-reading activities, a form of advance organizers, on subsequent resource allocation. Another study might assess the effects of a trair procedure using a pretest- posttest design in which the first step would be to establish stable estimates of the predictor regression coefficients across a selection of texts. This would be followed by a period of training focusing on a specific aspect of reading skill, after which the regression coefficients would again be estimated to test for a change in the reader's distribution of cognitive resources. A longterm objective for this line of research would be to develop an assessment instrument to identify deficient component skills within a given individual and to design training materials which would focus on the improvement of these components.

*Implications for L2 Reading Instruction*

L2 readers are often criticized for focussing excessively on word- or sentence-level aspects of the text, the implication being that the teacher must direct their attention to more global text processing. The results of the present study do not support this view. The difficulty in processing of local text features exhibited by the L2 readers here suggests that even fairly advanced L2 readers of English may need to improve low-level processing components of reading skill in order to reach reading abilities comparable to skilled L1 readers.

It may well be that leaving them to get 'bogged down' in local processing will not lead to enhanced efficiency of processing. However, reading programs which focus primarily on higher level 'comprehension' skills by having students spend considerable time completing exercises dealing with larger text-level units (e.g., main idea identification, drawing inferences) may not be providing the kind of training appropriate for even advanced L2 readers. While not excluding these higher level components, increased effort needs to be devoted to designing instructional materials which develop greater fluency in low-level processing components, those which have the potential of becoming fast and automatic. The results of this study cannot speak to the form these materials should take. Whether component skills can be practiced in isolation or whether this training is only effective when embedded within a meaningful reading context is an important question. Consideration of this and other issues relevant to improving low-level skills in L2 readers is given in Segalowitz, Poulsen, and Komoda (1991).

REFERENCES

Bever, T.G. (1970). The cognitive basis for linguistic structures. In J.R. Hayes (Ed.), *Cognition and the development of language.* New York, NY: John Wiley.

Cirilo, R.K. (1981). Referential coherence and text structure in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 20,* 358-367.

Cirilo, R.K., & Foss, D.J. (1980). Text structure and reading times for sentences. *Journal of Verbal Learning and Verbal Behavior, 19,* 96-109.

Clark, H.H., & Sengul, C.J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition, 7,* 35-41.

Collins, A.M., & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Cziko, G.A. (1978). Differences in first- and second-language reading: The use of syntactic, semantic, and discourse constraints. *Canadian Modern Language Review, 34,* 473-489.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1-18.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Educational Testing Service (1984). *Test of English as a foreign language.* Princeton, NJ: Author.

Educational Testing Service (1987). *TOEFL Test and Score Manual.* Princeton, NJ: Author.

Favreau, M., Komoda, M.K., and Segalowitz, N. (1980). Second language reading: Implications of the word superiority effect in skilled bilinguals. *Canadian Journal of Psychology, 4,* 370-381.

Favreau, M., & Segalowitz, N. (1982) Second language reading in fluent bilinguals. *Applied Psycholinguistics, 3,* 329-341.

Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory & Cognition, 11,* 565-574.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Fodor, J.A. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales and E. Walker (Eds.), *New Approaches to Language Mechanisms.* Amsterdam: North-Holland.

Goodman, K.S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6,* 126-135.

Gough, P.B. (1972). One second of reading. In J.F. Kavanagh and I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, MA: MIT Press.

Gough, P.B. (1983). Context, form, and interaction. In K. Rayner (Ed.), *Eye movements in reading*. New York: Academic Press.

Graesser, A.C., Hauft-Smith, K., Cohen, A.D., & Pyles, L.D. (1980). Advanced outlines, familiarity, and text genre on retention of prose. *Journal of Experimental Education, 48*, 281-290.

Graesser, A.C., Hoffman, N.L., & Clark, L.F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior, 19*, 135-151.

Graesser, A.C., & Riha, J.R. (1984). An application of multiple regression techniques to sentence reading times. In D.E. Kieras and M.A. Just (Eds.), *New methods in reading comprehension research* (pp. 183-218). Hillsdale, NJ: Lawrence Erlbaum Associates.

Haberlandt, K.F. (1984). Components of sentence and word reading times. In D.E. Kieras and M.A. Just (Eds.), *New methods in reading comprehension research* (pp. 219-251). Hillsdale, NJ: Lawrence Erlbaum Associates.

Haberlandt, K.F., Berian, C., & Sandson, J. (1980). The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior, 19*, 635-650.

Haberlandt, K.F., & Graesser, A.C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114,* 357-374.

Haberlandt, K.F., Graesser, A.C., & Schneider, N.J. (1989). Reading strategies of fast and slow readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 815-823.

Haynes, M. (1984). Patterns and perils of guessing in second language reading. In J. Handscombe, R.A. Orem, and B.P. Taylor (Eds.), *On TESOL '83*. Washington, D.C.: TESOL.

Huggins, A.W.F., & Adams, M.J. (1980). Syntactic aspects of reading comprehension. In R.J. Spiro, B.C. Bruce, and W.F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Just, A.M., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Just, A.M., & Carpenter, P.A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.

Kaplan, R.M. (1975). On process models for sentence analysis. In D.A. Norman, and D.E. Rumelhart, (Eds.), *Explorations in cognition*. San Francisco: Freeman.

Katz, D. (1982, September). Keeping camels down on the farm. *Science82, 3*(7), 79-80.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Kintsch, W., & Keenan, J.M. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology, 5,* 257-274.

Kintsch, W., Kozminsky, E., Streby, W.J., McKoon, G., & Keenan, J.M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior, 14,* 196-214.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L.G. Nilsson (Ed.), *Perspectives on memory research: Essays in honor of Uppsala University's 500th anniversary.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Knight, G.P. (1984). A survey of some important techniques and issues in multiple regression. In D.E. Kieras and M.A. Just (Eds.), *New methods in reading comprehension research.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Langone, J. (1984, October). 'Test tube' skin for burns. *Discover, 5*(10), 12-13.

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), 7-20.

Lesgold, A.M., Roth, S.F., & Curtis, M.E. (1979). Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 18*, 291-308.

Lorch, R.F., Lorch, E.P., & Matthews, P.D. (1985). On-line processing of the topic structure of a text. *Journal of memory and language, 24*, 350-362.

Lorch, R.F., Lorch, E.P., & Mogan, A.M. (1987). Task effects and individual differences in on-line processing of the topic structure of a text. *Discourse Processes, 10*, 63-80.

Lorch, R.F., & Myers, J.L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149-157.

McKoon, G., & Ratcliff, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior, 20*, 358-367.

Meyer, B.J.F., & Rice, G.E. (1984). The structure of text. In P.D. Pearson (Ed.), *Handbook of reading research: Vol. I* (pp. 319-351). New York: Longman.

Mitchell, D.C. (1987). Reading and syntactic analysis. In J.R. Beech and A.M. Colley (Eds.), *Cognitive approaches to reading.* New York: John Wiley.

Mitchell, D.C., & Green, D.W. (1978). The effects of content on immediate processing in reading. *Quarterly Journal of Experimental Psychology, 30,* 609-636.

Morrow, D.G. (1985). Prepositions and verb aspect in narrative understanding. *Journal of Memory and Language, 24,* 390-404.

Neely, J. (1977). Semantic priming and retrieval from lexical memory: Roles of the inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254.

O'Brien, E.J., Shank, D.M., Myers, J.L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 410-420.

Olson, S. (1984, November). Year of the tornado. *Science84, 5*(9), 174-179.

Pedhazur, E.J. (1982). *Multiple regression in behavioral research (2nd Edition).* New York: Holt, Rinehart & Winston.

Perfetti, C.A. (1985). *Reading Ability.* New York: Oxford University Press.

Perfetti, C.A., Goldman, S.R., & Hogaboam, T.W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition, 7,* 273-282.

Quirk, R.S., Greenbaum, R., Leech, G., and Svartvik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

Rumelhart, D.E. (1977). Toward an interactive model of reading. In S. Dornic and P.M.A. Rabbitt (Eds.), *Attention and performance VI.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Segalowitz, N. (1986). Second language reading. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Segalowitz, N., & Hébert, M. (1990). Phonological recoding in the first and second language reading of skilled bilinguals. *Language Learning, 40,* 503-538.

Segalowitz, N., Poulsen, C., and Komoda, M. (1991). Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. In J. H. Hulstijn & J. F. Matter (Eds.), Reading in Two Languages. *AILA Review 8,* 15-30.

Seifert, C.M., Robertson, S.P., & Black, J.B. (1985). Types of inferences generated during reading. *Journal of Memory and Language, 24,* 405-422.

Smith, F. (1973). *Psycholinguistics and reading.* New York: Holt, Rinehart, & Winston.

Stanovich, K.E. (1980). Toward and interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16,* 32-71.

Stanovich, K.E. (1991). Word recognition: Changing perspectives. In R. Barr, M.L. Kamil, P.B. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research, volume II*. White Plains, NY: Longman.

Stevens, A., & Rumelhart, D.E. (1975). Errors in reading: Analyses using an augmented transition network grammar. In D.A. Norman, and D.E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman.

Turner, A., & Greene, E. (1977). *The construction of a propositional text base*. Technical report, University of Colorado.

Vasos, H. (1983). *Semantic processing in bilinguals*. Unpublished doctoral dissertation, Concordia University, Montreal.

Whale and farewell. (1985, May). *Discover, 6*(5), 6-7.

When incubator babies were a world-traveling side show. (1985, August). *Discover, 6*(8), 60.

Williams, E. (1984). *Reading in the language classroom*. London: Macmillan.

Yngve, V.H. (1961). The depth hypothesis. *Proceedings of the XII Symposium in Applied Mathematics*. Providence, RI: American Mathematical Society.

# APPENDIX A

*Experimental Texts*

Table A-1

*Whales Text*

| Case | Sentence |
|---|---|
| 1 | Humans haven't exactly been known as the whales' best friends. |
| 2 | Until recently, we have hunted them incessantly, but this winter the Soviets came to their aid by freeing about 3,000 whales from an icy trap in the Arctic where they faced certain death. |
| 3 | Late in December a group of white whales (also known as belugas) followed a school of codfish into the Senyavina Strait, off the Bering Sea. |
| 4 | More concerned with their next meal than the falling temperature, the whales remained in the narrow strait until a cold east wind locked a wall of solid ice eighteen miles long and eleven miles wide into the passage, closing it off and leaving only small holes in which the animals could surface to breathe. |
| 5 | A native Chukchi hunter discovered the trapped belugas and informed his comrades about the jackpot. |
| 6 | But when other hunters arrived from the nearby settlement of Yandrakinot, they realized there could be too much of a good thing: the pools were crowded with thousands of pure white adult whales, up to fifteen feet long and weighing more than a ton, along with their light-grey young. |
| 7 | The villagers decided (probably with some encouragement from local fishing authorities who knew that the trapped whales represented a tenth of the entire beluga population) to feed the whales instead of killing them. |
| 8 | News of the whales' desperate situation travelled all the way to Moscow. |
| 9 | Helicopters carrying marine mammalogists, reporters, and television crews were sent to the scene, while the good people of Yandrakinot fed the belugas frozen fish (which the whales refused to eat). |
| 10 | But winter was getting colder, and the ice began closing in. |
| 11 | Pressed side to side, the exhausted animals were forced to take turns at the air holes. |
| 12 | Adults came to the aid of their young by supporting them while they gasped for air. |
| 13 | In early February the icebreaker Moskva came to the rescue. |
| 14 | Anatoly Kovalenko, its captain, twice ordered his ship to turn back in the face of the huge wall of ice, afraid the vessel would be crushed. |
| 15 | Meanwhile, the state of the whales grew worse, and some died. |
| 16 | Finally, on February 22, the Moskva succeeded in breaking through the twelve-foot-thick ice and opening a channel to the open sea ten miles away. |
| 17 | The whales didn't immediately realize they had been saved. |

| Case | Sentence |
|---|---|
| 18 | Frightened by the noise of the engines, they refused to follow the ship into the channel, which began freezing over. |
| 19 | For four days Kovalenko and his crew tried to habituate them to the ship, slowing the engines and enlarging the pools. |
| 20 | At first the whales hid. |
| 21 | Then they bagan to approach the boat, and rejoiced and leaped like children. |
| 22 | Still they wouldn't follow. |
| 23 | Then one of the experts on board had an idea. |
| 24 | Remembering that whales are musical creatures, he began playing selections– jazz, popular, military, classical– over the loudspeakers on deck, while the ship advanced slowly through the ice. |
| 25 | After several false starts, the belugas, which seemed to have a preference for classical music, gradually started following, until the happy whales were at last directed to freedom. |
| 26 | Only forty whales died during what Soviet television reporters called Operation Beluga, and the Kremlin even got a congratulatory message from Greenpeace, the activist environmental organization that has harassed Soviet and other whale hunters. |

Table A-2

*Incubators Text*

| Case | Sentence |
|---|---|
| 27 | One of the most bizarre stories in the medical treatment of premature babies is that of Martin Couney, the Incubator Doctor, who during the first half of this century charged admission to show premature infants to crowds at fairs, expositions, and amusement parks all over the world. |
| 28 | Couney's history was pieced together after his death in 1950 by William Silverman, a pediatrician who is also a specialist in prematurity. |
| 29 | According to Silverman, it started in 1878, when a Parisian obstetrician named Emile Tarnier saw a chicken incubator at a fair and asked its designer to build him one for premature babies. |
| 30 | At the Paris Maternity Hospital the device became the most important piece of equipment in the world's first premature baby treatment center. |
| 31 | Couney, who after finishing medical school in Leipzig had begun working in the center, was asked to exhibit the incubators at the 1896 World Exposition in Berlin. |
| 32 | He decided that the exhibit needed live babies, which Berlin's Charity Hospital agreed to lend him, since the babies were expected to die anyway. |
| 33 | The resulting six incubators tended by nurses attracted huge crowds, and Couney boasted that he lost none of this little patients. |
| 34 | Soon a London promoter invited Couney to set up a show which he did, although at the last minute he had to transport three wicker baskets full of French babies across the Channel because Englsih doctors refused to lend him local ones. |
| 35 | Again the exhibit was a hit; even the conservative Englsih medical journal Lancet praised it. |
| 36 | But after a few months, when it became apparent that Couney's success was inspiring imitators to set up baby shows beside stinky menageries and dusty bike-racing tracks, the Lancet's editore began to criticize the idea, and demanded to know what connection there was between the serios matter of saving human life and circus attractions such as the bearded woman and performing horses and pigs. |
| 37 | Nevertheless, the London exhibit represented the start of a career for Couney. |
| 38 | Over the next few years, he put on similar shows in Omaha, Paris, and Buffalo. |
| 39 | Since his equipment and his nursing staff were better than anything available in most hospitals, doctors were glad to help him keep his incubators full. |
| 40 | In 1903 he moved to Brooklyn, settling near the amusement park at Coney Island, where he exhibited babies every summer for the next forty years. |
| 41 | Couney also staged shows in Rio de Janeiro, Mexico City, and various places in the United States, including the 1939 World's Fair in New York, and Atlantic City, where there's a plaque commemorating his exhibits on the Boardwalk. |

| Case | Sentence |
|---|---|
| 42 | Even by today's standards, some of Couney's achievements are remarkable. |
| 43 | In 1901 he saved a two-pound nine-ounce infant, at that time thought to be the smallest infant ever born who lived. |
| 44 | And in 1915, before oxygen was widely used, he apparently cared for and sent home a healthy baby boy who had weighed just a pound and a half at birth. |
| 45 | At the 1939 World's Fair, only ten of his 96 patients are reported to have died-- a survival rate approaching those in the most sophisticated neonatal units today. |
| 46 | How did he do it? |
| 47 | The babies in his care, like those today, were kept warm and clean and well fed. |
| 48 | But more important, most of them came to him several days after birth, meaning that they were strong enough to have survived the critical early hours and days. |
| 49 | One thing frustrated Couney's successes, however. |
| 50 | Some parents were unwilling to visit their children in incubators, and when it came time to take home, they didn't want them. |
| 51 | Couney had to work hard to persuade them; parental rejection of premature babies remains something of a problem to this day. |
| 52 | By the 1940s hospitals in New York had set up their own specially equipped centers for premature infants and didn't need Couney's help anymore. |
| 53 | Even to the public, incubator babies were no longer exciting spectacles, and attendance at the Coney Island exhibit had begun to decline. |
| 54 | When Couney at last closed the exhibit, he told his nephew that his work was done-- he had made propaganda for the premature baby. |

Table A-3

*Skingrafts Text*

| Case | Sentence |
|---|---|
| 55 | On a summer day in Casper, Wyoming, a little more than a year ago, two brothers, Jamie Selby, five, and Glen, six, and a six-year-old neighbour playfully splashed themselves with paint. |
| 56 | Then while trying to remove it with solvent, one of them lit a match. |
| 57 | The solvent ignited, and the boys suffered burns-- mostly third degree (the skin completely destroyed)-- over 97 per cent of their bodies. |
| 58 | The neighbour died, and most doctors gave the brothers virtually no chance for survival. |
| 59 | Yet today the two boys are alive and their chances for staying alive seem good. |
| 60 | What saved them was a revolutionary new treatment involving skin grafts grown from their own skin cells. |
| 61 | This "test-tube" skin could save the lives of the 15,000 people in the United States who are severely burned each year. |
| 62 | When the Selbys were taken to Children's Hospital in Denver, doctors realized that the boys would need extensive skin grafts as a barrier against infection. |
| 63 | But where to get the skin? |
| 64 | There was virtually none left on their burned bodies, and the only alternatives, cadaver skin or artificial skin, offered no permanent cure: cadaver skin is eventually rejected; artificial skin at this stage in its development is also temporary and must be covered with a thin layer of natural skin. |
| 65 | As a last resort, the doctors flew the boys to Boston, where the new technique had been developed and used successfully on a few patients-- though with burns far less severe than the Selbys'. |
| 66 | At the Shriners Burns Institute in Boston, plastic surgeon Gregory Gallico and his colleagues from the Massachusetts General Hospital removed pieces of the boys' remaining unburned skin, from the armpits and lower abdomen. |
| 67 | Then they cut the pieces into tiny fragments, mixed in enzymes that separated them into millions of individual cells, and added growth stimulants. |
| 68 | While waiting for the cells to multiply and grow together, they covered the boys' burned areas with cadaver skin grafts, artificial skin, and grafts from the few spots where skin had healed. |
| 69 | In about twenty days the test-tube skin patches reached the size of playing cards. |
| 70 | Each was carefully removed from its container and a piece of Vaseline-covered gauze was placed on it to make it easier to manipulate. |
| 71 | With the skin side down, the gauze was placed on the burned areas and the corners sutured to underlying muscle. |

| Case | Sentence |
|---|---|
| 72 | About a week later the gauze was removed, leaving the new skin in place. |
| 73 | Within twelve weeks the doctors were able to grow about a half a square yard of new skin for each boy. |
| 74 | Because it was grown from the boys' own cells, their bodies did not reject it. |
| 75 | After about nine months Jamie was well enough to go home and return to school. |
| 76 | Glen has remained in Boston but is recovering nicely. |
| 77 | Both boys still face difficulties: they will need plastic surgery and physical therapy. |
| 78 | But they are lucky to be alive. |
| 79 | Gallico says he knows of no other patients with burns this severe over such a large area who have survived. |
| 80 | The doctors emphasize that it is still too early to make any definitive statements about the long-term usefulness of "test-tube" skin because it has been used on so few patients and it consists only of the outer layer of the skin and contains no sweat glands to allow the skin to breathe. |
| 81 | Doctors do not know if the Selbys will grow an inner layer of skin on their own or if that inner layer is even necessary. |
| 82 | Still, the achievement of Gallico and associates is remarkable. |

Table A-4

*Tornadoes Text*

| Case | Sentence |
|---|---|
| 83 | In the spring of 1984 an unusually large number of storm systems passed over the American Midwest and the East, generating the severe thunderstorms that create tornadoes. |
| 84 | By September, the National Weather Service's count of reported tornadoes had reached 1,235, and with still three months to go, had already surpassed the previous record of confirmed tornadoes, which was set in 1973 at 1,102. |
| 85 | The tornadoes in 1984 were also especially violent. |
| 86 | More than 120 people died, after nearly a decade in which fatalities never reached one hundred. |
| 87 | Ed Ferguson of the National Severe Storms Forecast Center in Kansas City warned that the residents couldn't hope to hang on to the low fatality rates they had been experiencing. |
| 88 | The center spent a lot of energy prior to the disastrous 1984 year telling people that they shouldn't think the tornado problem had gone away because the death toll had been so low. |
| 89 | That year their prophecy came true. |
| 90 | The number of deaths caused by tornadoes, like other natural disasters, depends on factors that can seem frighteningly unpredictable. |
| 91 | Tornadoes may measure anywhere from a few yards to a mile wide, and their devastation can be surgically precise. |
| 92 | When a tornado moves through a town, it can demolish one block of homes while leaving the next block virtually untouched. |
| 93 | The time of day also makes a difference. |
| 94 | One reason the Carolina tornadoes were so deadly is that they came after sunset. |
| 95 | With no way to see a tornado coming, many residents had only last-minute warnings from radio and television-- and from their own ears-- to tell them they were in its path. |
| 96 | Even during the day, tornadoes may hit with little warning. |
| 97 | They can move deceptively fast, up to 70 miles per hour. |
| 98 | By the time people see or hear an approaching tornado, they have just a few minutes, or possibly seconds, to find protection. |
| 99 | If those who have survived such an event have one thing in common, it is probably the terror they felt. |
| 100 | A tornado hits with such ferocity that it's more like being in an explosion than in a storm. |
| 101 | Large tornadoes contain the strongest winds on the face of the Earth, as fast as 300 miles per hour. |

| Case | Sentence |
|---|---|
| 102 | Winds of such magnitude can reduce a house to pieces in seconds. |
| 103 | For researchers, a tornado's winds have another, particularly problematic effect: they make tornadoes very difficult to study. |
| 104 | Tornadoes destroy any standard meteorological instruments over which they pass. |
| 105 | So researchers have to rely on indirect measurements and eyewitness reports to check their theories. |
| 106 | However, the overall rotation associated with a tornado may be the key to better tornado warnings. |
| 107 | Most tornadoes-- and all of the worst ones-- form only within thunderstorms that have themselves begun to rotate. |
| 108 | On average, perhaps a few thousand of these dangerous rotating thunderstorms hit the United States each year, most often in the Midwest. |
| 109 | They are extremely hard to predict and detect. |
| 110 | Usually the first time a weather forecaster knows such a storm is in the area is when someone sees a tornado coming. |
| 111 | But in 1988 the weather service is going to begin installing a new generation of radars that will make it much easier to detect these travelling storms. |
| 112 | Conventional radar simply indicates how much rain is falling at a given place. |
| 113 | The new radars, known as Doppler radars, detect how fast the rain is moving toward or away from the radar by using the same principle as police traffic radar. |
| 114 | The full system of Doppler radars will not be complete until well into the next decade, but already forecasters are anxious to begin using it. |
| 115 | In a two-year study at the National Severe Storms Laboratory in Norman, Oklahoma, a team of forecasters using Doppler radar predicted nearly 70 percent of the tornadoes in the area with an average lead time of more than 20 minutes. |
| 116 | Another team of forecasters using everything but Doppler radar predicted almost as many tornadoes, but with a much higher false alarm rate and an average lead time of just two minutes. |
| 117 | The new technique's most precious service will probably be the few extra minutes of advance notice it can provide. |
| 118 | In years to come, at least some people will have more of a warning than the sound of a distant roar in the midst of a thunderstorm. |

# APPENDIX B

*Matrix of Predictor Values*

Table B-1

*Matrix of Predictor Values*

| | Predictor Variables | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sentence | WRD | FREQ | PROP | SYN | REF | DIST | NAN | IMP | BEG | END | POS | NARR |
| 1 | 10 | 5.478 | 5 | 0.465 | 0 | 0 | 3 | 0.30 | 1 | 0 | 1 | 4.60 |
| 2 | 33 | 5.872 | 20 | 0.499 | 2 | 1 | 6 | 0.79 | 0 | 1 | 2 | 4.60 |
| 3 | 25 | 5.892 | 8 | 0.565 | 1 | 1 | 5 | 0.45 | 1 | 0 | 3 | 6.20 |
| 4 | 54 | 6.420 | 34 | 0.556 | 2 | 1 | 5 | 0.85 | 0 | 1 | 4 | 6.20 |
| 5 | 15 | 5.220 | 8 | 0.675 | 1 | 1 | 3 | 0.33 | 1 | 0 | 5 | 5.70 |
| 6 | 49 | 6.416 | 27 | 0.515 | 1 | 2 | 5 | 0.30 | 0 | 0 | 6 | 5.70 |
| 7 | 33 | 5.996 | 16 | 0.616 | 3 | 4 | 3 | 0.73 | 0 | 1 | 7 | 5.70 |
| 8 | 12 | 6.350 | 6 | 0.553 | 2 | 1 | 3 | 0.54 | 1 | 0 | 8 | 5.70 |
| 9 | 30 | 5.560 | 11 | 0.624 | 3 | 2 | 5 | 0.54 | 0 | 1 | 9 | 5.70 |
| 10 | 11 | 6.635 | 4 | 0.509 | 1 | 4 | 1 | 0.21 | 1 | 0 | 10 | 5.70 |
| 11 | 16 | 6.564 | 7 | 0.571 | 2 | 5 | 0 | 0.27 | 0 | 0 | 11 | 5.70 |
| 12 | 16 | 6.906 | 6 | 0.818 | 2 | 6 | 1 | 0.24 | 0 | 1 | 12 | 5.70 |
| 13 | 10 | 6.130 | 4 | 0.699 | 0 | 0 | 3 | 0.61 | 1 | 0 | 13 | 6.60 |
| 14 | 26 | 6.357 | 12 | 0.642 | 3 | 4 | 2 | 0.24 | 0 | 0 | 14 | 6.60 |
| 15 | 11 | 6.659 | 6 | 0.664 | 1 | 3 | 1 | 0.36 | 0 | 0 | 15 | 6.60 |
| 16 | 24 | 5.962 | 10 | 0.581 | 3 | 13 | 2 | 0.97 | 0 | 0 | 16 | 6.60 |
| 17 | 9 | 6.056 | 4 | 0.625 | 1 | 2 | 0 | 0.15 | 0 | 0 | 17 | 6.20 |
| 18 | 20 | 6.943 | 7 | 0.718 | 3 | 2 | 2 | 0.54 | 0 | 0 | 18 | 6.20 |
| 19 | 21 | 6.345 | 10 | 0.628 | 5 | 8 | 2 | 0.58 | 0 | 0 | 19 | 6.20 |
| 20 | 5 | 5.751 | 2 | 0.750 | 1 | 1 | 0 | 0.15 | 0 | 0 | 20 | 6.20 |
| 21 | 13 | 6.781 | 9 | 0.795 | 2 | 2 | 0 | 0.27 | 0 | 0 | 21 | 6.20 |
| 22 | 4 | 6.073 | 3 | 0.788 | 1 | 1 | 0 | 0.18 | 0 | 1 | 22 | 6.20 |
| 23 | 10 | 7.696 | 4 | 0.635 | 0 | 0 | 2 | 0.30 | 1 | 0 | 23 | 6.30 |
| 24 | 26 | 5.407 | 15 | 0.615 | 4 | 22 | 4 | 0.73 | 0 | 0 | 24 | 6.30 |
| 25 | 28 | 6.161 | 13 | 0.510 | 2 | 3 | 3 | 0.94 | 0 | 1 | 25 | 6.30 |
| 26 | 34 | 5.117 | 20 | 0.434 | 1 | 1 | 7 | 0.73 | 1 | 0 | 26 | 4.40 |

| | Predictor Variables | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sentence | WRD | FREQ | PROP | SYN | REF | DIST | NAN | IMP | BEG | END | POS | NARR |
| 27 | 47 | 6.259 | 22 | 0.569 | 0 | 0 | 12 | 0.91 | 1 | 0 | 1 | 4.50 |
| 28 | 22 | 5.554 | 10 | 0.670 | 1 | 1 | 7 | 0.09 | 1 | 0 | 2 | 4.50 |
| 29 | 32 | 5.528 | 10 | 0.619 | 3 | 2 | 7 | 0.58 | 0 | 0 | 3 | 4.50 |
| 30 | 22 | 6.423 | 10 | 0.524 | 2 | 1 | 3 | 0.54 | 0 | 1 | 4 | 4.50 |
| 31 | 27 | 6.088 | 14 | 0.591 | 3 | [illegible] | 4 | 0.73 | 1 | 0 | 5 | 5.30 |
| 32 | 24 | 5.990 | 11 | 0.539 | 3 | 2 | 1 | 0.61 | 0 | 0 | 6 | 5.30 |
| 33 | 21 | 5.387 | 12 | 0.491 | 3 | 2 | 2 | 0.58 | 0 | 1 | 7 | 5.30 |
| 34 | 42 | 6.088 | 20 | 0.537 | 1 | 1 | 7 | 0.42 | 1 | 0 | 8 | 5 60 |
| 35 | 15 | 6.141 | 9 | 0.481 | 1 | 1 | 2 | 0.21 | 0 | 0 | 9 | 5 60 |
| 36 | 64 | 5.963 | 38 | 0.521 | 2 | 2 | 14 | 0.58 | 0 | 0 | 10 | 5 60 |
| 37 | 12 | 6.275 | 3 | 0.678 | 2 | 2 | 1 | 0.24 | 0 | 0 | 11 | 5 30 |
| 38 | 15 | 6.605 | 9 | 0.629 | 2 | 2 | 4 | 0.21 | 0 | 0 | 12 | 5 30 |
| 39 | 25 | 6.346 | 15 | 0.474 | 2 | 6 | 4 | 0.76 | 0 | 1 | 13 | 5 30 |
| 40 | 25 | 5.864 | 12 | 0.599 | 2 | 8 | 6 | 0.54 | 1 | 0 | 14 | 5 30 |
| 41 | 38 | 5.640 | 18 | 0.582 | 2 | 3 | 10 | 0.42 | 0 | 1 | 15 | 5 30 |
| 42 | 10 | 5.671 | 8 | 0.505 | 1 | 1 | 2 | 0.33 | 1 | 0 | 16 | 4 80 |
| 43 | 21 | 5.948 | 14 | 0.547 | 1 | 1 | 3 | 0.64 | 0 | 0 | 17 | 4 80 |
| 44 | 30 | 6.632 | 15 | 0.508 | 1 | 1 | 5 | 0.70 | 0 | 0 | 18 | 4.80 |
| 45 | 28 | 6.060 | 16 | 0.510 | 2 | 4 | 3 | 0.70 | 0 | 1 | 19 | 4 80 |
| 46 | 5 | 7.827 | 2 | 0.773 | 2 | 4 | 0 | 0.06 | 1 | 0 | 20 | 4 80 |
| 47 | 16 | 6.872 | 10 | 0.461 | 2 | 7 | 1 | 0.54 | 0 | 0 | 21 | 4 80 |
| 48 | 28 | 7.141 | 19 | 0.410 | 2 | 1 | 3 | 0.73 | 0 | 1 | 22 | 4 80 |
| 49 | 6 | 4.269 | 5 | 0.293 | 2 | 7 | 1 | 0.03 | 1 | 0 | 23 | 4 30 |
| 50 | 23 | 6.926 | 12 | 0.636 | 1 | 11 | 4 | 0.45 | 0 | 0 | 24 | 4.30 |
| 51 | 21 | 6.210 | 14 | 0.593 | 3 | 4 | 2 | 0.70 | 0 | 1 | 25 | 4 30 |
| 52 | 24 | 5.534 | 12 | 0.566 | 2 | 1 | 5 | 0.82 | 1 | 0 | 26 | 4 80 |
| 53 | 22 | 5.932 | 12 | 0.530 | 2 | 13 | 3 | 0.27 | 0 | 0 | 27 | 4 80 |
| 54 | 24 | 6.730 | 9 | 0.622 | 3 | 2 | 3 | 0.61 | 0 | 1 | 28 | 4.80 |
| 55 | 31 | 5.577 | 15 | 0.511 | 0 | 0 | 9 | 0.42 | 1 | 0 | 1 | 6 40 |
| 56 | 14 | 6.811 | 8 | 0.608 | 2 | 1 | 2 | 0.39 | 0 | 0 | 2 | 6 40 |

| | Predictor Variables | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sentence | WRD | FREQ | PROP | SYN | REF | DIST | NAN | IMP | BEG | END | POS | NARR |
| 57 | 22 | 5.633 | 20 | 0.609 | 2 | 1 | 3 | 0.76 | 0 | 0 | 3 | 6.40 |
| 58 | 14 | 6.315 | 8 | 0.615 | 2 | 3 | 3 | 0.45 | 0 | 0 | 4 | 5.00 |
| 59 | 15 | 6.466 | 6 | 0.565 | 2 | 1 | 1 | 0.27 | 0 | 1 | 5 | 5.00 |
| 60 | 17 | 5.624 | 10 | 0.494 | 1 | 1 | 3 | 0.82 | 1 | 0 | 6 | 5.00 |
| 61 | 21 | 6.543 | 9 | 0.582 | 1 | 1 | 4 | 0.30 | 0 | 1 | 7 | 5.00 |
| 62 | 25 | 5.900 | 8 | 0.651 | 1 | 2 | 6 | 0.54 | 1 | 0 | 8 | 5.00 |
| 63 | 6 | 7.838 | 3 | 0.600 | 1 | 1 | 0 | 0.09 | 0 | 0 | 9 | 5.00 |
| 64 | 49 | 6.008 | 29 | 0.505 | 3 | 7 | 7 | 0.64 | 0 | 0 | 10 | 5.00 |
| 65 | 34 | 6.727 | 14 | 0.591 | 4 | 8 | 4 | 0.42 | 0 | 1 | 11 | 5.00 |
| 66 | 33 | 5.463 | 13 | 0.526 | 2 | 1 | 8 | 0.70 | 1 | 0 | 12 | 6.10 |
| 67 | 23 | 6.034 | 11 | 0.527 | 2 | 1 | 4 | 0.82 | 0 | 0 | 13 | 6.10 |
| 68 | 33 | 5.778 | 12 | 0.532 | 7 | 4 | 2 | 0.61 | 0 | 0 | 14 | 6.10 |
| 69 | 14 | 5.760 | 7 | 0.496 | 1 | 8 | 4 | 0.54 | 0 | 0 | 15 | 5.80 |
| 70 | 23 | 6.626 | 11 | 0.628 | 1 | 1 | 2 | 0.48 | 0 | 0 | 16 | 5.80 |
| 71 | 20 | 6.608 | 6 | 0.613 | 2 | 3 | 3 | 0.58 | 0 | 0 | 17 | 5.80 |
| 72 | 14 | 6.930 | 8 | 0.703 | 2 | 2 | 1 | 0.61 | 0 | 0 | 18 | 5.80 |
| 73 | 21 | 6.760 | 9 | 0.644 | 3 | 5 | 1 | 0.48 | 0 | 0 | 19 | 5.80 |
| 74 | 15 | 6.621 | 4 | 0.552 | 4 | 10 | 0 | 0.67 | 0 | 1 | 20 | 5.80 |
| 75 | 15 | 6.728 | 8 | 0.631 | 1 | 1 | 3 | 0.27 | 1 | 0 | 21 | 5.00 |
| 76 | 9 | 5.519 | 4 | 0.580 | 2 | 10 | 0 | 0.30 | 0 | 0 | 22 | 5.00 |
| 77 | 13 | 5.590 | 8 | 0.477 | 1 | 2 | 3 | 0.67 | 0 | 0 | 23 | 5.00 |
| 78 | 7 | 7.284 | 3 | 0.636 | 1 | 1 | 0 | 0.03 | 0 | 1 | 24 | 5.00 |
| 79 | 20 | 6.358 | 10 | 0.518 | 3 | 22 | 2 | 0.42 | 1 | 0 | 25 | 5.00 |
| 80 | 52 | 6.716 | 23 | 0.572 | 4 | 17 | 4 | 0.76 | 0 | 0 | 26 | 4.00 |
| 81 | 25 | 6.488 | 9 | 0.621 | 3 | 3 | 1 | 0.54 | 0 | 0 | 27 | 4.00 |
| 82 | 9 | 6.505 | 3 | 0.477 | 1 | 1 | 1 | 0.42 | 0 | 1 | 28 | 4.00 |
| 83 | 27 | 6.166 | 11 | 0.518 | 0 | 0 | 7 | 0.36 | 1 | 0 | 1 | 3.60 |
| 84 | 36 | 5.880 | 15 | 0.456 | 0 | 0 | 8 | 0.48 | 0 | 0 | 2 | 3.60 |
| 85 | 8 | 5.596 | 5 | 0.557 | 2 | 2 | 0 | 0.70 | 0 | 0 | 3 | 4.40 |
| 86 | 16 | 5.997 | 9 | 0.438 | 0 | 0 | 3 | 0.50 | 0 | 0 | 4 | 4.40 |

| | Predictor Variables | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sentence | WRD | FREQ | PROP | SYN | REF | DIST | NAN | IMP | BEG | END | POS | NARR |
| 87 | 30 | 6.052 | 10 | 0.589 | 0 | 0 | 5 | 0.30 | 0 | 0 | 5 | 4.40 |
| 88 | 33 | 6.523 | 13 | 0.529 | 4 | 3 | 2 | 0.21 | 0 | 0 | 6 | 4 40 |
| 89 | 6 | 6.189 | 4 | 0.460 | 2 | 1 | 0 | 0.06 | 0 | 1 | 7 | 4.40 |
| 90 | 19 | 5.487 | 12 | 0.460 | 1 | 7 | 4 | 0.45 | 1 | 0 | 8 | 2.40 |
| 91 | 19 | 5.891 | 8 | 0.622 | 1 | 1 | 1 | 0.67 | 0 | 0 | 9 | 2.40 |
| 92 | 21 | 5.971 | 11 | 0.610 | 1 | 1 | 2 | 0.42 | 0 | 1 | 10 | 2.40 |
| 93 | 8 | 7.842 | 3 | 0.671 | 0 | 0 | 1 | 0.45 | 1 | 0 | 11 | 2.30 |
| 94 | 14 | 6.380 | 7 | 0.726 | 0 | 0 | 3 | 0.30 | 0 | 0 | 12 | 2.30 |
| 95 | 31 | 6.751 | 15 | 0.517 | 1 | 1 | 7 | 0.52 | 0 | 0 | 13 | 2.30 |
| 96 | 10 | 6.251 | 5 | 0.588 | 1 | 4 | 2 | 0.24 | 0 | 0 | 14 | 2.80 |
| 97 | 11 | 5.357 | 7 | 0.710 | 1 | 1 | 0 | 0.61 | 0 | 0 | 15 | 2.80 |
| 98 | 22 | 6.659 | 13 | 0.571 | 1 | 1 | 4 | 0.70 | 0 | 1 | 16 | 2.80 |
| 99 | 20 | 7.137 | 7 | 0.513 | 2 | 1 | 2 | 0.21 | 1 | 0 | 17 | 2.40 |
| 100 | 18 | 6.555 | 8 | 0.574 | 1 | 1 | 3 | 0 33 | 0 | 0 | 18 | 2.40 |
| 101 | 19 | 6.191 | 6 | 0.638 | 1 | 1 | 2 | 0.64 | 0 | 0 | 19 | 2.40 |
| 102 | 12 | 6.674 | 3 | 0.673 | 2 | 1 | 2 | 0.18 | 0 | 1 | 20 | 2.40 |
| 103 | 17 | 5.424 | 10 | 0.581 | 3 | 2 | 2 | 0.61 | 1 | 0 | 21 | 2.10 |
| 104 | 10 | 4.811 | 5 | 0.423 | 1 | 1 | 1 | 0.73 | 0 | 0 | 22 | 2.10 |
| 105 | 15 | 5.785 | 7 | 0.549 | 1 | 2 | 3 | 0.58 | 0 | 0 | 23 | 2.10 |
| 106 | 16 | 5.965 | 8 | 0.468 | 1 | 2 | 3 | 0.54 | 0 | 0 | 24 | 2.30 |
| 107 | 18 | 6.249 | 9 | 0.331 | 1 | 1 | 1 | 0.73 | 0 | 0 | 25 | 2 30 |
| 108 | 22 | 6.489 | 11 | 0.358 | 2 | 25 | 1 | 0.36 | 0 | 0 | 26 | 2 30 |
| 109 | 8 | 6.328 | 6 | 0.700 | 1 | 1 | 0 | 0.36 | 0 | 0 | 27 | 2.30 |
| 110 | 22 | 6.676 | 8 | 0.510 | 2 | 3 | 4 | 0.39 | 0 | 1 | 28 | 2 30 |
| 111 | 27 | 6.274 | 12 | 0.583 | 2 | 27 | 3 | 0.88 | 1 | 0 | 29 | 3.00 |
| 112 | 13 | 5.950 | 4 | 0.607 | 0 | 0 | 3 | 0.21 | 0 | 0 | 30 | 3.00 |
| 113 | 29 | 6.069 | 15 | 0.589 | 2 | 2 | 2 | 0.76 | 0 | 0 | 31 | 3.00 |
| 114 | 25 | 6.361 | 13 | 0.470 | 2 | 4 | 2 | 0.33 | 0 | 1 | 32 | 3.20 |
| 115 | 40 | 5.898 | 17 | 0.588 | 1 | 1 | 7 | 0.85 | 1 | 0 | 33 | 3.20 |
| 116 | 31 | 5.938 | 19 | 0.441 | 3 | 1 | 2 | 0.70 | 0 | 0 | 34 | 3 20 |

| | Predictor Variables | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sentence | WRD | FREQ | PROP | SYN | REF | DIST | NAN | IMP | BEG | END | POS | NARR |
| 117 | 19 | 6.533 | 13 | 0.422 | 2 | 1 | 2 | 0.64 | 0 | 0 | 35 | 2.30 |
| 118 | 27 | 7.606 | 9 | 0.589 | 2 | 12 | 3 | 0.42 | 0 | 1 | 36 | 2.30 |

# APPENDIX C

*Illustrative Example of Analysis for SYNTAX Predictor*

Table C-1

*Analysis of a Sample Participant's Predictions for the First Sentence of the Whales Text*

| Actual Text | | Participant's prediction | | |
|---|---|---|---|---|
| Word | Category | Word | Category | Score |
| Humans | noun | | | |
| haven't | auxiliary | beings | noun | X |
| exactly | adverb | the | determiner | X |
| been | auxiliary | the | determiner | X |
| known | verb | ready | modifier | X |
| as | preposition | for | preposition | ✓ |
| the | determiner | the | determiner | ✓ |
| whales' | modifier | first | ordinal | X |
| best | modifier | equal | noun | X |
| friends. | noun | friend | noun | X |

Table C-2

*Sample of Tally Sheet of Syntax Predictions for the First Sentence of the Whales Text*

| Actual Text | | Participants' prediction scores | | | | | | | | | | | prop. correct[a] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Word | Category | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
| Humans | noun | - | - | - | - | - | - | - | - | - | - | - | - |
| haven't | aux | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | X | ✓ | ✓ | ✓ | ✓ | .909 |
| exactly | adv | X | X | ✓ | ✓ | X | X | X | X | ✓ | X | ✓ | .364 |
| been | aux | X | X | X | X | X | X | X | X | X | X | X | .000 |
| known | verb | X | X | X | X | X | ✓ | X | X | X | X | ✓ | .182 |
| as | prep | X | X | X | X | ✓ | X | ✓ | X | X | X | X | .182 |
| the | det | X | ✓ | ✓ | X | ✓ | ✓ | ✓ | X | ✓ | ✓ | ✓ | .727 |
| whales' | mod | X | X | X | ✓ | ✓ | X | X | ✓ | ✓ | X | X | .364 |
| best | mod | ✓ | X | ✓ | ✓ | X | ✓ | X | X | ✓ | X | X | .455 |
| friends. | noun | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 1.000 |

[a]*M* = .465, and is the SYNTAX value entered for sentence 1 in predictors matrix

# APPENDIX D

*Illustrative Example of Propositional Analysis*

*Propositional Analysis of the First Two Sentences of the Whales Text*

Humans haven't exactly been known as the whales' best friends.

1. (REFERENCE, HUMAN, 4)
2. (NEGATE, 1)
3. (EXACTLY, 2)
4. (POSSESS, WHALE, FRIEND)
5. (BEST, FRIEND)

Until recently, we have hunted them incessantly, but this winter the Soviets came to their aid by freeing about 3,000 whles from an icy trap in the Arctic where they faced certain death.

6. (TIME: UNTIL, 8, RECENTLY)
7. (HUNT, A: HUMAN, O: WHALE)
8. (INCESSANTLY, 7)
>9. (CONDITION: IF, 8, 2)
>10. (HELP, A: HUMAN, O: WHALE)
>11. (NEGATE, >10)
>12. (CONDITION: IF, 2, >11)
>13. (ISA, SOVIET, HUMAN)
14. (CONCESSION: BUT, 2, 16)
15. (TIME: THIS, 16, WINTER)
16. (HELP, A: SOVIET, O: WHALE)
17. (MANNER: BY, 16, 18)
18. (FREE, A: SOVIET, O: WHALE, S: TRAP)
19. (ABOUT, 20)
20. (NUMBER OF, WHALE, 3000)
21. (ICY, TRAP)
22. (LOCATION: IN, TRAP, ARCTIC)
23. (LOCATION: WHERE, 24, TRAP)
24. (FACE, E: WHALE, O: DEATH)
25. (CERTAIN, DEATH)

*Note.* Numbers in parentheses refer to embedded propositions.
The > indicates an inferred proposition required to maintain a coherent textbase.

APPENDIX E

*True-False Statements for Experimental Texts*

Table E-1

*True-False Verification Statements for Whales Text*

| Answer[a] | Sentence |
|---|---|
| NEI | Up until the rescue operation, Soviets had been considered the worst offenders of whale hunting quotas. |
| F | The death of the trapped whales would not have seriously affected the global beluga population, but the numbers in Soviet waters would have been drastically reduced. |
| F | The whales could have freed themselves if they had been left in peace. |
| T | Marine experts were sent from Moscow to supervise the rescue. |
| F | The belugas managed to survive on the frozen fish the Soviets fed them. |
| T | The whales' access to the Bering Sea was cut off by the ice. |
| T | The Soviet who discovered the trapped whales was not very concerned about the welfare of the belugas. |
| F | One of the most striking characteristics of young Belugas is their pure white colour. |
| F | The Soviets responded to pressure from Greenpeace to save the whales. |
| T | The whales eventually overcame their fear of the boat, but then wouldn't follow it out of the strait. |
| T | A major threat to the survival of the whales was a lack of air. |
| F | The whales seemed to respond best to jazz music. |
| F | The whales immediately responded to the music and followed the ship out of the strait. |
| T | The Soviets were quick to recognize the seriousness of the belugas' predicament. |
| T | The local fishing authorities believed that the whales should not be killed. |
| T | Greenpeace congratulated the Soviets on their efforts to save the belugas. |
| F | The whales were originally lured into the strait by the Soviet hunters. |
| T | The whales became accustomed to the ship within a few days. |
| NEI | Because of the concern the Soviets displayed in rescuing the whales, Greenpeace has since relaxed its harrassment of Soviet whale hunters. |
| F | The Soviets managed to save only half of the trapped whales. |
| NEI | The Soviets have twice been called upon to supervise similar rescue operations since the successful freeing of the belugas from Senyavina Strait. |

| Answer[a] | Sentence |
|---|---|
| F | The adult whales pushed the young out of the way in order to reach the air holes. |
| T | The whales were attracted into the strait by codfish. |

[a] T = true, F = false, NEI = Not enough information.

Table E-2

*True-False Verification Statements for Incubators Text*

| Answer[a] | Sentence |
|---|---|
| T | Couney treated almost one hundred infants at the 1939 World's Fair and lost only about ten percent of them. |
| F | The first incubator for premature babies was designed by Martin Couney. |
| F | Martin Couney quit medical school to exhibit incubators at fairs. |
| NEI | For several decades, Couney's exhibit was banned in Italy by the Vatican. |
| F | The first exhibit of baby incubators was not very successful, but interest quickly grew soon after. |
| T | The exhibit in London attracted huge crowds. |
| F | Couney was mainly criticized in the 'Lancet' for claiming to save babies that would have been strong enough to survive without his treatment. |
| T | Couney would have had more deaths if he had received the premature babies more quickly. |
| T | London docotrs would not agree to exhibiting English babies in Couney's incubator exhibitions. |
| T | London's medical community became increasingly critical of displaying premature babies alongside freak-show exhibits. |
| T | Couney felt that he had performed an important survice by displaying incubator babies for fifty years. |
| NEI | Couney continued his medical practice during the winter and exhibited the incubators only in summer. |
| F | Couney was unable to convince Berlin's doctors to lend him babies from nearby hospitals so he brought them from France. |
| F | The new parents were proud to have their babies in the famous Doctor Couney's incubators, but they all anxiously awaited the day they could take them home. |
| T | Couney maintained better standards in premature baby treatment than most hospitals could offer. |
| NEI | The diary Martin Couney kept helped Silverman piece together his story. |
| T | Couney's first exhibit of incubator babies was held in Berlin. |
| T | The smallest baby Couney saved weighed less than two pounds. |
| T | Human incubators were developed from techniques used for hatching chicken eggs. |

| Answer[a] | Sentence |
|---|---|
| F | Couney initially refused to exhibit the incubators at commercial fairs and expositions. |
| F | Premature French babies were cared for in incubators as they crossed the channel to England. |
| F | As late as the 1940s, interest in Couney's incubator exhibits was still as high as ever. |
| F | The first exhibit used prematurely-born animals as a substitute for human babies. |

[a] T = true, F = false, NEI = Not enough information.

Table E-3

*True-False Verification Statements for Skingrafts Text*

| Answer[a] | Sentence |
|---|---|
| NEI | The Selbys received generous donations to help finance the huge cost of the operations. |
| F | Jamie and Glen Selby were saved thanks to the superiority of artificial skin over test-tube skin. |
| F | Approximately half the skin on the young boys' bodies was badly burned. |
| T | The doctors progressively replaced the artificial and cadaver skin grafts with the boys' own skin. |
| T | Boston doctors are pioneers in the field of burn treatments. |
| | The new skin patches were placed on the burned areas and then covered with Vaseline. |
| F | Within nine months both boys were back in school. |
| F | Doctors performed this first experiment in test-tube skin grafts on the Selbys before considering other possible treatments. |
| T | It was easier to work with the skin patches if Vaseline-covered bandages were placed on them first. |
| T | Doctors don't know if the two brothers will need an inner layer of skin. |
| NEI | All but one severely burned patient receiving this new type of skin graft have survived. |
| T | The skin grafts the boys received do not have any sweat glands. |
| F | After the successful operation on the Selbys, doctors were sure that they had finally found a permanent solution for burn victims. |
| T | Doctors finally decided to try a new treatment involving skin grafts grown form the boys' own skin cells. |
| F | The doctors were not able to remove the gauze without pulling off the thin layer of skin attached to it. |
| F | The aerosol paint can exploded and caught on fire when one of the boys heated it with matches. |
| T | The corners of the skin grafts were stitched to muscles. |
| F | Doctors were hopeful because several other patients with burns as severe as the Selbys' are known to have survived after treatments with test-tube skin grafts. |
| F | The six-year-old neighbour escaped serious injury, but the two brothers were severely burned. |
| T | Doctors knew that the artificial skin would serve as only a temporary measure. |

| Answer[a] | Sentence |
|---|---|
| NEI | In the future, doctors hope to develop methods of implanting sweat glands to allow the test-tube skin to breathe naturally. |
| T | Skin from dead bodies is sometimes used as a temporary source for skin grafts. |
| T | It took the doctors over two weeks to grow the first test-tube skin patches. |

[a] T = true, F = false, NEI = Not enough information.

Table E-4

*True-False Verification Statements for Tornadoes Text*

| Answer[a] | Sentence |
|---|---|
| NEI | Tornadoes are more numerous in years of heavy rainfall. |
| F | Although there were many more tornadoes in 1984, they were smaller and caused less damage than usual. |
| F | Small tornadoes are more likely to develop from rotating thunderstorms than are large ones. |
| T | Doppler radar can detect the development of tornadoes much earlier than conventional radar. |
| T | Researchers have not been able to install meteorological equipment strong enough to resist the force of tornadoes' winds. |
| T | During the late seventies, many people began to underestimate the severe threat of tornadoes. |
| F | Doppler radar probably won't save many lives, but it is the first step toward a better understanding of how tornadoes develop. |
| F | Rotating thunderstorms are relatively easy to detect. |
| T | Doppler radar relies on rain to detect tornadoes. |
| T | By the time a weather forecaster knows a tornado is coming, it's often too late to warn anyone. |
| F | By 1984, forecasters had finished installing only about half the Doppler radar system needed for coordinated, accurate prediction of tornadoes. |
| T | Doppler radar will be able to detect the development of rotating thunderstorms. |
| F | The author claims the often-made comparison of a tornado to an explosion is inappropriate. |
| F | Tornadoes cause over a hundred deaths in the United States every year. |
| NEI | Residents were particularly fortunate when one tornado in 1984 narrowly missed a densely populated area. |
| T | With Doppler radar, researchers do not seem able to predict more tornadoes than conventional methods, but the predictions are more reliable. |
| T | Using Doppler radar, forecasters are able to predict more than two-thirds of tornadoes. |
| F | Conventional radar detects the direction the rain drops are moving in. |
| T | The number of fatalities due to tornadoes is likely to be lower during the day than at night. |

| Answer[a] | Sentence |
|---|---|
| F | Because of its extremely strong winds, a tornado often causes considerable damage to areas adjacent to its primary path as well. |
| NEI | Some politicians believe that the high cost of installing the complete Doppler radar system is not worth the few extra minutes of warning it will provide. |
| T | Doppler radar is an adapted form of radar used to detect speeders on highways. |
| F | The complete system of Doppler radar will be ready before the end of this decade. |

[a] T = true, F = false, NEI = Not enough information.

# APPENDIX F

*Participant Consent Form*

*Participant Consent Form*

You are invited to participate in a study from which we hope to learn more about the reading process.

If you decide to participate you will be asked to read up to five short passages, rate your familiarity with each of the passage topics, retell each passage in your own words, and verify a series of true/false statements concerning each passage. Following this you may be asked to return for an additional session to complete a standardized test of English reading comprehension. The procedure does not involve any risk or deception and is painless. We hope you find it enjoyable.

Your individual responses will be confidential and will not be disclosed without your permission. Only group data will be reported. Your performance in no way reflects your intellectual abilities or personality.

If you decide to participate we will ask you to come for one to three sessions, each lasting between thirty minutes and an hour. You will be paid $5.00 per hour for your participation in this study. You will receive the payment in full only upon completion of all required sessions.

If you have any questions, please ask me. If you have any additional questions later, Dr. Norman Segalowitz will be happy to answer them (Concordia University, 7141 Sherbrooke West, Room DS-413-5, Montreal, Quebec, H4B 1R6 / 848-2239). Also, if you wish to receive a short report summarizing the findings of this study when the research is completed, he will be happy to arrange for a copy to be sent to you.

If you decide to participate you will be offered a copy of this form to keep.

---

You are making a decision whether or not to participate. Your signature indicates that you have read the information provided above and have decided to participate. You may withdraw at any time after signing this form if you choose to discontinue participation in this study.

____________________ ________________________ ____________

print name signature date

________________________

signature of investigator

---

Payment of $__.00 has been received in full for participation in research conducted by Dr. N. Segalowitz.

____________________________ __________________

signature date

## APPENDIX G

*Instructions for Experimental Tasks*

*Reading Instructions*

In this first part of the experiment you will be asked to read a text in English presented one sentence at a time on a computer video screen. You will progress through the text at your own speed advancing to each subsequent sentence in the following manner:

> When you are ready, press a button marked "next" and the first sentence of the passage will appear on the screen. As soon as you have read and understood the first sentence, press the button again, and the second sentence will appear. Continue to proceed through the passage in this manner advancing to the next sentence once you have finished reading and understood the current sentence on the screen.

Please decide BEFORE you begin which button and hand is most comfortable for you to use to advance through the text. Then, keep just one hand on the box and do NOT change hands or buttons during the reading sections of the experiment.

Once you have begun the sentence presentations, please do not take time out to rearrange yourself as <u>we are measuring the time you take to read and understand each sentence</u>; make sure you are comfortable BEFORE you begin.

Since we are timing your reading of the text, please read as quickly as possible without sacrificing your comprehension.

After you have finished reading the text, you will be given two tasks to assess how well you have understood and remembered the information in the text.

Please note that it is not possible to return to an earlier sentence.

When you have finished reading all of the sentences in the text, "THANK YOU" will appear in the center of the screen and the computer will beep. Please notify the experimenter that you have finished at this point.

If you have any questions, please feel free to ask them now.

*Retell Instructions for L1 Participants*

For this part of the experiment, we would like you to imagine you are going to help prepare a participant who has not read the story to answer a series of true/false statements concerning the story which you have just read. Please retell, in your own words, the story you have just read, being as complete as possible. Remember that you must provide him/her with as much information as is necessary to verify the true/false statements.

If you have any questions, please feel free to ask them now.

*Retell Instructions for L2 Participants*

For this part of the experiment, we would like you to imagine you are going to help prepare a participant who has not read the story to answer a series of true/false statements concerning the story which you have just read. Please retell, in your own words, the story you have just read, being as complete as possible. The next participant is also bilingual, so please feel free to use either French or English -- whichever is easiest or most comfortable for you. You may also switch back and forth freely between these two languages. What is important is to provide him/her with as much information as is necessary to verify the true/false statements.

If you have any questions, please feel free to ask them now.

*T/F Statement Verification Instructions*

In this part of the experiment, a series of statements related to the passage you've just read will be presented on the video screen. Some of these statements are true, some are false, and some contain information that is not discussed in the text at all. Based on your reading of the passage we would like you to verify the truthfulness of these statements. There are four possible responses available for you to choose from:

1. True

2. False

3. I don't remember.

4. Not enough information in text to decide.

Please note that for a statement to be considered TRUE, all the facts contained in that statement must be true. If part of the statement is false, then the statement as a whole must be considered false.

Furthermore, not all the TRUE statements necessarily concern information which is directly stated in the text. Some of the statements can be verified on the basis of indirect information implied by the text, and so should be considered TRUE.

Please reserve the last option (not enough information in text to decide) for statements which simply cannot be verified as 'true' or 'false' because the text did not give you enough information, either directly or indirectly, to verify it.

To select your response, first highlight the desired option by pressing the button marked SELECT. (If you pass the option you wanted keep pressing SELECT until you reach it again.) After you have made your selection, press the two buttons marked ENTER simultaneously. Your response will be recorded and the next statement will appear on the screen.

Please respond to ALL the statements.

If you have any questions, please feel free to ask them now.

# APPENDIX H

*Mean Unstandardized Coefficients from Analyses of Individual L1 and L2 Sentence Reading Times*

Table H-1

*Mean Unstandardized Coefficients from Analyses of Individual L1 and L2 Sentence Reading Times*

| | Unstandardized Coefficient | | | | |
|---|---|---|---|---|---|
| | L1[a] | | L2[b] | | |
| Predictor | Mean[c] | (*SD*) | Mean[d] | (*SD*) | *t*[e] |
| FREQ | -66.355 | (284.170) | -113.865 | (770.392) | .224 |
| PROPS | 260.650** | (61.498) | 428.824** | (140.821) | -4.239*** |
| SYNTAX | -1083.793* | (1898.694) | -432.140 | (2342.270) | -.837 |
| NANS | 647.921** | (195.229) | 741.585** | (464.375) | -.720 |
| COREFS | 519.254** | (186.012) | 598.316** | (385.510) | -.715 |
| DIST | -16.354 | (30.050) | -6.172 | (36.238) | -.838 |
| IMPORT | 1489.556** | (1162.960) | 1036.633* | (1644.617) | .871 |
| BEGIN | -361.414* | (586.353) | -579.285 | (1238.249) | .616 |
| END | -485.370** | (460.246) | -667.554* | (939.937) | .674 |
| NARR | 84.565 | (222.421) | -4.947 | (246.068) | 1.045 |
| POSIT | -20.702* | (37.351) | -25.851** | (14.468) | .498 |

[a]$n$ = 15. [b]$n$ = 15. [c]Asterisks denote significance of mean from 0. [d]Asterisks denote significance of mean from 0. [e]Comparison of L1 and L2 group means.
*$p$<.05. **$p$<.01. ***$p$<.001.