

DOCUMENT RESUME

ED 207 037

CS 006 293

AUTHOR Kieras, David E.
 TITLE Modelling Reading Times in Different Reading Tasks with a Simulation Model of Comprehension. Technical Report, May 1, 1978 through March 30, 1979.
 INSTITUTION Arizona Univ., Tucson, Dept. of Psychology.
 SPONS AGENCY National Inst. of Mental Health (DHEW), Rockville, Md.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
 REPORT NO ONR-TR-2
 PUB DATE 30 Mar 79
 CONTRACT ONR-N00014-78-C-0509
 GRANT NIMH-1-R03-MH31985-01
 NOTE 79p.

EDRS PRICE MF01/PC04 Plus Postage.
 DESCRIPTORS *Cognitive Processes; Higher Education; Language Processing; Models; *Multiple Regression Analysis; *Reading Comprehension; Reading Processes; *Reading Rate; *Reading Research; *Research Methodology; Simulation

ABSTRACT

Reading times can be a valuable source of data on comprehension processes, especially in the case of recording reading times on individual sentences in a passage. To overcome the methodological problems encountered in other research efforts concerning reading times, a multiple regression method was used to compare an ordinary language processing simulation model to single sentence reading times by fitting a linear equation to the reading times. Approximately 90 undergraduate college students read simple passages one sentence at a time, with the order of sentences in the passage varied from subject to subject. Three different reading tasks were studied: an immediate free recall task, a topic identification task, and a free reading task. The results revealed that (1) the multiple regression method of fitting the simulation model to reading time worked; (2) the reading times on individual sentences could be broken down into the times required for each process in the simulation--parsing, representation building, memorizing for recall, topic identification, and topic pointer maintenance; (3) reading times appeared to be determined mostly by representation building and memorization; and (4) the reading tasks differed substantially and in meaningful ways as to how much and what kind of processing was required on individual sentences. (FL)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED207037

Modelling Reading Times in Different Reading Tasks with a Simulation Model of Comprehension

David E. Kieras
University of Arizona
Tucson, Arizona 85721

U S DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Technical Report No. 2
March 30, 1979

This research was supported by the National Institute for Mental Health, under Grant No. 1-R03-MH31985-01 to the author, and the Personnel and Training Research Programs, Office of Naval Research, under Contract Number N00014-78-C-0509, Contract Authority Identification Number NR 157-423. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for Public Release; Distribution Unlimited.

The University of Arizona is an Equal Employment Opportunity/Affirmative Action Employer. In compliance with Title IX (Educational Amendments of 1972), Title VII (Civil Rights Act of 1964), and Section 504 (Rehabilitation Act of 1973), the University does not discriminate on the basis of sex, race, creed, color, national origin, or handicapping condition in its educational programs or activities, including admissions and employment. Inquiries concerning application of Title IX, Title VII, and Section 504 may be referred to Dr. Jean Raley Kearns, Assistant Executive Vice President, Administration 503, phone (602) 626-3084.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER Technical Report No. 2	2. GOVT ACCESSION NO. ---	3 RECIPIENT'S CATALOG NUMBER ---
4 TITLE (and Subtitle) Modelling Reading Times in Different Reading Tasks with a Simulation Model of Comprehension		5. TYPE OF REPORT & PERIOD COVERED Technical Report (1 May 78 - 30 Mar 79)
		6. PERFORMING ORG. REPORT NUMBER
7 AUTHOR(s) David E. Kieras	8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0509	
9 PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology University of Arizona Tucson, AZ 85721		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N; RR 042-06; RR 042-06-02; NR 157-423
11 CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, VA 22217		2. REPORT DATE 30 Mar 79
		13. NUMBER OF PAGES 70
14 MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16 DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17 DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18 SUPPLEMENTARY NOTES		
19 KEY WORDS (Continue on reverse side if necessary and identify by block number) Reading, Comprehension, Simulation Model		
20 ABSTRACT (Continue on reverse side if necessary and identify by block number) Reading times are a very detailed source of data on comprehension processes, especially in the case of recording reading times on individual sentences in a passage. But conventional verbal or mathematical theories of comprehension fail to fully exploit such data because the comprehension processes are not represented in enough detail. Computer simulation models of language processing are adequately detailed, but there has been difficulty in applying such models to experimental data. This paper presents the results obtained by using a		

multiple-regression method to compare an ordinary language processing simulation model to single-sentence reading times by fitting a linear equation to the reading times. The candidate predictor variables in the regression analysis are variables representing the number of steps in each process performed by the simulation in understanding each sentence. This method was applied to a detailed set of single-sentence reading times. Subjects read simple passages one sentence at a time, with the order of the sentences in the passage being varied. Three different reading tasks were studied: a standard immediate free recall task, a topic identification task, and a free reading task. The results of this simulation analysis show that: (1) The multiple-regression method of fitting the simulation model to reading time data works; that is, it yields useful information on the accuracy of the model. (2) The reading times on individual sentences can be decomposed into the times required for each process in the simulation, such as parsing, representation building, memorizing for recall, topic identification, and topic pointer maintenance. (3) In these simple passages, reading times appear to be mostly determined by representation building and memorization processes, while parsing, memory search, and topic pointer maintenance contributed weakly, although significantly, to the reading times. (4) The reading tasks differ substantially and in theoretically meaningful ways in how much and what kind of processing is required on individual sentences.

Modelling Reading Times in Different Reading Tasks
with a Simulation Model of Comprehension

David E. Kieras
University of Arizona
Tucson, Arizona 85721

Technical Report No. 2

March 30, 1979

This research was supported by the National Institute for Mental Health, under Grant No. 1-RO3-MH31985-01 to the author, and the Personnel and Training Research Programs, Office of Naval Research, under Contract Number N00014-78-C-0509, Contract Authority Identification Number NR 157-423. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for Public Release; Distribution Unlimited.

Abstract

Reading times are a very detailed source of data on comprehension processes, especially in the case of recording reading times on individual sentences in a passage. But conventional verbal or mathematical theories of comprehension fail to fully exploit such data because the comprehension processes are not represented in enough detail. Computer simulation models of language processing are adequately detailed, but there has been difficulty in applying such models to experimental data. This paper presents the results obtained by using a multiple-regression method to compare an ordinary language processing simulation model to single-sentence reading times by fitting a linear equation to the reading times. The candidate predictor variables in the regression analysis are variables representing the number of steps in each process performed by the simulation in understanding each sentence. This method was applied to a detailed set of single-sentence reading times. Subjects read simple passages one sentence at a time, with the order of the sentences in the passage being varied. Three different reading tasks were studied: a standard immediate free recall task, a topic identification task, and a free reading task. The results of this simulation analysis show that: (1) The multiple-regression method of fitting the simulation model to reading time data works; that is, it yields useful information on the accuracy of the model. (2) The reading times on individual sentences can be decomposed into the times required for each process in the simulation, such as parsing, representation building, memorizing for recall, topic

identification, and topic pointer maintenance. (3) In these simple passages, reading times appear to be mostly determined by representation building and memorization processes, while parsing, memory search, and topic pointer maintenance contributed weakly, although significantly, to the reading times. (4) The reading tasks differ substantially and in theoretically meaningful ways in how much and what kind of processing is required on individual sentences.

Modelling Reading Times in Different Reading Tasks
with a Simulation Model of Comprehension

David E. Kieras

University of Arizona

An increasingly popular measure in the study of reading comprehension is that of reading time, or inspection time, the time a reader takes to process a piece of verbal input. The rationale for using a reading time measure is that variables that influence the type or amount of processing should thereby affect the amount of time required to complete the processing. In contrast to the more commonly used recall measure, which assesses the results of comprehension, the reading time measure taps an aspect of the comprehension process as it occurs. For this reason, the study of reading times can contribute uniquely to our growing understanding of comprehension processes.

One of the major uses of reading times has been to measure the time taken to process an entire passage. Although this is an old use of the measure, it has become really valuable theoretically only with the work of Kintsch and others which provides characterizations of passages in terms of independent variables such as the number of propositions or arguments (e.g. Kintsch & Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). At the other extreme, by using a laboratory computer, Aaronson (1972; Aaronson & Scarborough, 1976) has

measured reading times for individual words in isolated sentences. The profile of word reading times followed the phrase structure of the sentence, but, most importantly, only if the subject's task was a recall task. A task involving reading just for a later simple question produced only very small variations in reading times.

Most of the work with reading times has required subjects to make some judgement, such as true or false, in response to a sentence, with some linguistic property of the sentence being manipulated. Often the manipulation involves supplying additional information, such as a picture (e.g., Chase & Clark, 1972) or earlier sentences (e.g., Carpenter & Just, 1975), that change some implicit property of the test sentence. The logical unit of study of such experiments is thus the comprehension of the single sentence. While this procedure is the obvious one to use to explore many issues in comprehension, clearly making a true-false judgement of a sentence is a task that might dramatically alter the processing time compared to some other task.

Other studies have used reading times to study how sentences are integrated. While the time to process a single sentence is still the measure, the interest is directed towards the relations between the sentences. For example, Haviland and Clark (1974) studied the operation of the given-new contract in sentence comprehension. According to this view of integration processes, each sentence specifies some new information about one or more given items or referents. In order to understand a

sentence, the reader must identify the given and new portions of the sentence, locate the representations of the given information in memory, and add the new information to these representations. Haviland and Clark's subjects read pairs of sentences which differed in whether references in the second sentence could be resolved to referents postulated by the first sentence. Subjects simply pressed a button when they thought they understood each sentence. Sentences whose referents were given were processed faster than if they were new. Other examples of integration studies using reading times Haberlandt and Bingham(1978), who showed that verbs can govern the integration of sets of simple sentences, and Mynatt and Smith(1977) who tested a model for integration in extremely simple passages describing linear orderings.

A set of experiments by Kieras(1978), using reading times for individual sentences in a passage, is of special interest here because they are the precursors of the experiment reported in this paper. Integration was studied in simple passages, similar to those of Bransford and Franks(1971), which differed in the goodness of their text structure. The passages consisted of seven simple one-proposition sentences that could logically be presented in any order. The main idea sentence could appear either at the beginning or the end. The sentences could be in a coherent order, in which each sentence contains at least one given referent, allowing it to be immediately integrated. Or, the sentences could be arranged in an incoherent order, such that the first several sentences contain only new referents and so can not be immediately integrated. Both main idea position

and coherence had effects on apparent theme, recall, and reading times. In particular, the reading times were shorter on sentences with given referents than when these same sentences contained only new referents. The incoherent order thus required longer average reading times, particularly on the initially unintegratable sentences. The interpretation was that the initially unintegratable sentences in the incoherent order produced a higher memory load and so slowed down processing. However, the strongest and most consistent reading time differences were between sentences of different structure, such as The ants were hungry and The ants ate the jelly. Furthermore, the pattern of the reading times depended on the task, showing different patterns in an immediate recall and in a theme identification task, and no effects at all in a delayed recall task (see Kieras, Note 1, 1978).

The conclusion to be drawn about these studies of reading time is that measuring this variable is clearly a powerful methodology that can yield a lot of information about comprehension processes. However, as commented above, the pattern of reading times can differ considerably depending on the task. Hence providing a temporal characterization of comprehension processes based on reading times will require that the reading task be taken seriously into account.

Modelling Reading Times

The logic of using a reading time measure is that differences in the amount or type of processing should show up as differences in the amount of time required. However, the

comprehension process is believed to be extremely complex, consisting of several subprocesses such as word perception, syntactic analysis, memory search, integration, memory storage, and other processes, such as schema-using strategies. The problem in devising models for reading time is how to assign differences in reading times to these several subprocesses, and so decompose the total reading time into theoretical components. For example, in the data in Kieras(1978), the longer reading times for the unintegrable sentences could be due to differences in memory search, integration, topic maintenance, or storage processes. Also, the differences in reading time between sentences of different types, such as between the simple one-noun sentences like The ants were hungry and the more complicated two-noun sentences like The ants ate the jelly, can not be interpreted at all since they differ not only with regard to these same processes, but also in the amount of syntactic analysis and the amount of content information as well.

The approach generally taken to ensure that reading times can be interpreted in a theoretically useful way is to use careful manipulations that hopefully affect only the process of interest, and to use either general models that make only directional predictions, for example, Kintsch(1977), or highly specific mathematical models tailored to particular experiments, such as Carpenter and Just(1975). These approaches are required because the extant empirically-based theories of comprehension are not able to describe comprehension processes at a level of detail that would enable the theorist to completely divide up the total reading time for a sentence among the subprocesses.

Because of this, the field is faced with an impasse in that computerized laboratory technology and the systems of propositional analysis of verbal materials now available (e.g., Kintsch, 1974) make it possible to collect extremely detailed data about reading times, but conventional methods of theorizing can not be depended on to exploit this empirical detail to develop general models.

In contrast to conventional theories of comprehension, the computer simulation models of comprehension and the allied efforts in artificial intelligence are committed to describing at a general level and in great detail exactly what must be done in comprehension. Because of this feature, it is possible to use computer simulation as a theoretical tool, to try out different ideas about comprehension to see if they actually work. This has enabled theorists about comprehension to account for a wide variety of comprehension processes, ranging from the basic level of sentence integration (Kieras, 1977), to high-level schema-directed comprehension (e.g., Schank, 1978). However, little has been done with such models as explanations for the usual type of laboratory data. In fact, a major problem is how such models could be tested against data. At a recent national conference on the evaluation of computer simulations in psychology (Ann Arbor, August, 1975) the consensus was that although simulations are clearly one of psychology's most powerful theoretical tools, there is a major problem in evaluating the empirical accuracy of a simulation model. There is simply no equivalent to the simple, elegant, and rigorous techniques familiar in the older mathematical models. Hence, as

interesting and theoretically stimulating as simulation models are, their value in actually accounting for experimental data remains questionable.

Simulation modelers have attempted to cope with the evaluation problem in several ways. Some, such as Newell and Simon(1972), emphasize modelling individual subjects, circumventing the difficult problem of inter-subject variability, but suffering the problem of a lack of generalizability. Others, such as Anderson(1972) and Hintzman(1968) have developed stochastic simulations which produce noisy data similar to that of subjects, and so can be compared directly to the data by ordinary statistical techniques. However, such simulations are difficult to work with and expensive to run. The artificial intelligence technology typically used in computer simulations is much more suited to modelling deterministic processes. Other modellers have, in effect, disclaimed trying to account for experimental data, apparently feeling that the theoretical return on devising a simulation is worth the ambiguity in its empirical status.

An Approach to Simulation Testing

This paper makes use of an approach to testing a simulation model of prose comprehension that enables the model to predict reading times on individual sentences in simple passages like those in Kieras(1978). The extent to which the model succeeds in fitting, and the details of the fit, allow a comprehensive evaluation of the accuracy of the model to be made.

The approach is a larger scale version of that used by Carpenter and Just(1975) to fit simple constituent-testing models of verification time to the observed mean times. Their models contain a variable consisting of the number of constituents to be matched. Under the assumption that each constituent match takes a constant amount of time, the predicted verification time is thus a linear function of the number of constituents. An ordinary regression line is fitted to the data; the slope of the line is the coefficient of the variable, the time per match, and the intercept is the time to perform the other, remaining, components of the task. Finally, the correlation coefficient, or its square, provides a measure of the goodness-of-fit of the model in terms of the percent of variance among the observed mean times accounted for.

The method used here is essentially the same, but in a multivariate case. Again, the reading time on each sentence is the dependent variable. The independent, or predictor, variables are variables whose values are supplied by the simulation model. That is, the simulation model has many different internal processes, each of which consists of a series of steps. The number of steps in the individual processes may vary, depending on exactly what the simulation must do in order to understand an individual sentence in the context of the preceding sentences. The number of steps in each kind of process for each input sentence is a set of values for predictor variables in a multiple regression analysis. The results of the regression analysis are a set of coefficients, one for each variable, that can be interpreted as the number of seconds

required for each step in the corresponding process. As in the simple case, the intercept of the regression line is the amount of time taken for all other processes. And also as before, the square of the correlation coefficient gives the proportion of variance among the mean reading times accounted for by the regression equation, and is thus a measure of the accuracy of the model. By examining where reading times predicted from the regression equation match and mismatch the observed reading times, this overall measure of goodness-of-fit can be augmented by the detailed conditions under which the model is a good fit or a bad fit.

This technique has a major advantage in that a deterministic simulation model, using the available artificial technology, can be used to predict statistically noisy experimental data. However, there are important logical implications of this approach. The resulting modelling situation is that we have not just a simulation model and a set of data, but also a statistical model that provides a bridge between the simulation and the data. Certain assumptions are thus implied. The data are assumed to be essentially deterministic, with the only random component being a simple error term. Hence the psychological processes are assumed to be basically deterministic.

More significantly, the reading times are assumed to be a simple linear function of the variables provided by the simulation model, with no other variables involved. If so, then the equation and the logic of the regression analysis make up

nothing more than a statistical model linking the simulation to the data. However, if in the course of building a good regression equation, variables are included that are not directly provided by the simulation, the result is that the statistical model has substantive content which the simulation fails to include. While such a situation is not necessarily bad in terms of the overall value of the analysis, it clearly means that the simulation alone is not an accurate or complete model of the substantive features of the data.

This method was tried out with a subset of the data from Kieras(1978); the results of this trial were reported in Kieras(Note 2). Two sets of reading times for individual sentences in simple passages were used: one set was from a coherent presentation order, the other an incoherent order. The reading time profiles were rather different for the two presentation orders. Applying this method to fit the simulation model (basically that of Kieras, 1977) to this small set of data showed that substantively different versions of the simulation could be distinguished by their goodness-of-fit to the data. The worst fitting version could account for only 62% of the variance among mean reading times, the best for 83%. However, these results could hardly be considered conclusive, since they were based on only 14 data points, and these were extracted from a more complicated set of reading times. Furthermore, the data were obtained in a recall task; as described above, other tasks could be expected to produce rather different patterns of reading times.

The remainder of the paper is divided into two sections. The first describes an experiment using three different tasks that yielded a detailed set of data on reading times in simple passages, and the standard statistical analysis of these results. The second section of the paper is an analysis of the reading times using the above-described method of fitting a simulation model to the reading times. The simulation is first described, followed by the results of model fitting. Since a relatively complicated experiment, a simulation model, and the model fitting results, all have to be described, this paper is necessarily more complex and extensive than the usual empirical report.

EXPERIMENT ON READING TIMES

The experiments in Kieras(1978) suggested that memory load resulting from incoherence and also the position of the main idea at either the beginning or end of the passage were important to reading time, recall, and apparent theme. Passages in which some of the sentences could not be immediately integrated took longer to read on those same sentences, and demonstrated lower recall, especially on these same unintegrable sentences. However, the results on reading time were obtained only by imposing a heavy processing load on the subjects; they were required to read and integrate the sentences from three simple passages at the same time. This is a rather artificial reading situation. As a follow-up to these results, the present experiment was designed to produce effects on reading time as a

function of presentation order with subjects reading only a single passage. This was achieved by means of longer passages, which would impose a higher memory load, and having subjects produce immediate recall (see Kieras, 1978). Furthermore, rather than contrasting only coherent with incoherent presentation orders, three levels of coherence were used. The level of coherence was determined by following a suggestion in Kieras(1978) and computing the average number of chunks, or unintegrated sentences, being maintained during the reading of each sentence. The coherence factor was crossed with the position of the main idea sentence, which appeared at the beginning or at the end, giving a total of six different presentation orders. In order to assess the effects of different reading tasks, three different task instruction conditions were used. These were a Recall Task of standard immediate gist recall, a Topic Choice Task requiring subjects to choose the topic of the passage, and a Free Reading Task similar to that used by Haviland and Clark(1974).

Experimental Method

Materials. Twenty-four passages were used. These were similar to those used in Kieras(1978) but had additional sentences added for a total of eleven sentences per passage. The passages had a propositional structure shown in Figure 1, consisting of a central proposition labelled A, and two branches elaborating on the subject noun concept(N1) and object noun concept(N2) of this proposition. Logically, proposition A would seem to be the main idea, and so the Type A sentence can be termed the topic sentence. This is also justifiable because the

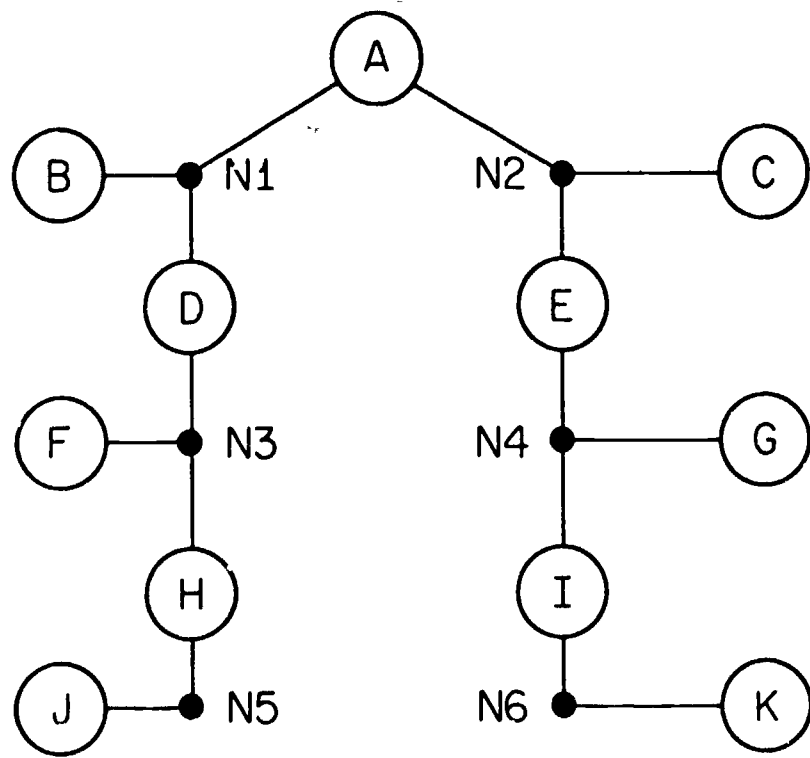


Figure 1. Propositional structure of the passages.

Type A sentence is the most popular choice for the apparent theme in the results in Kieras(1978) and several unpublished experiments on apparent theme in these simple passages.

The propositions B, C, F, G, J, and K were simple predications of one of the noun concepts, and were expressed in a sentence frame having the form The noun was adjective. The propositions A, D, E, H, and I consisted of a relation between two of the nouns, and appeared as the sentence frame The noun was relation the noun. The relation in these sentences was either an -ing form of a verb, a preposition, or sometimes a more complex relation such as was attached to or was in tune with. These more complex forms were used only for the Type H and I sentences, and only when necessary to permit a reasonably natural passage to be constructed. Unlike the passages in Kieras(1978), the proposition A sentence was not different in form from the other two-noun sentences. The verb appeared in the -ing form, which was similar to the other two-noun sentences. Table 1 contains an example passage with the sentences labelled according to Figure 1.

The different passage presentation orders are shown in Table 2. There are three pairs of orders; the members of each pair differ in whether the topic sentence appears at the beginning or at the end. Within each of these pairs, the sentence order is identical except for the position of the Type A sentence. The three pairs differ in the level of coherence. Shown in Table 2 is a measure of coherence, the average number of chunks, or unintegrable sentences that have been previously

Table 1
Example of Passages

- A. The ants were eating the jelly.
 - B. The ants were hungry.
 - C. The jelly was grape.
 - D. The ants were in the kitchen.
 - E. The jelly was on the table.
 - F. The kitchen was spotless.
 - G. The table was wooden.
 - H. The kitchen was equipped with the blender.
 - I. The table was against the stove.
 - J. The blender was white.
 - K. The stove was hot.
-

Table 2

Passage Sentence Presentation Orders
 Showing Chunks Present During Each Sentence

Coherence	Topic	Order	Average Chunks
High	First	A D B F H J C E G I K	
	Chunks	0 1 1 1 1 1 1 1 1 1 1	0.91
High	Last	D B F H J C E G I K A	
	Chunks	0 1 1 1 1 1 2 2 2 2 2	1.36
Medium	First	A B C H F J I G K D E	
	Chunks	0 1 1 1 2 2 2 3 3 3 2	1.82
Medium	Last	B C H F J I G K D E A	
	Chunks	0 1 2 3 3 3 4 4 4 3 2	2.64
Low	First	A J K F G B C H I D E	
	Chunks	0 1 2 3 4 5 5 5 4 3 2	3.09
Low	Last	J K F G B C H I D E A	
	Chunks	0 1 2 3 4 5 6 5 4 3 2	3.18

presented, and thus have to be maintained while reading each sentence. Although the position of the topic sentence in these passages unavoidably affects the coherence to some extent, the members of each pair of orders are fairly close in coherence level.

Design and Subjects. The task factor was between-subjects and the two presentation order factors of Topic Position and Coherence were within-subject. Since the orders were obtained by simply changing the order of presentation of the passage sentences, all 24 passages could appear in any combination of first or last topic sentence position, and high, medium, or low coherence. Each subject saw all of the 24 passages, which were assigned at random to presentation order conditions for each subject under the constraint that each subject saw four passages in each of the six presentation order conditions. Order of appearance of the passages in the experiment was also randomized separately for each subject. Hence the experimental manipulations were all within-passage manipulations, and interactions of individual passages with experimental conditions would be confounded with subject variability.

The Recall Task subjects were run first, followed by the Topic Choice and Free Reading Task subjects about a month later. In the Recall Task portion of the experiment, about 40 subjects were run individually, with serious equipment failures causing partial data loss for all but 29 subjects. Only the data for these 29 subjects was retained for analysis. After the defective equipment was replaced, 30 subjects for each of the

Topic Choice and Recall Task were run in groups of one or two, with alternate groups of subjects run in each task condition.

Subjects were University of Arizona undergraduates of both sexes. The subjects in the Recall Task were introductory psychology students who received extra credit for participating, while the Topic Choice and Free Reading subjects were recruited through campus ads and were paid \$2.00 for participating. There is a definite possibility of a general motivational difference between the credit and paid subjects; the extra credit subjects appeared to be very diligent and highly motivated compared to the paid subjects.

Equipment and Procedure. A Data General MicroNOVA laboratory computer was used to generate the randomized passage sequences and to display the passages one sentence at a time and collect reading times (Kieras, in press). Up to two subjects could be run concurrently. Each subject sat in a booth at an 80 X 24 upper/lower-case Teleray video terminal driven at 9600 Baud. The sentences appeared in the center of the screen. When the subject tapped the space bar, the sentence disappeared and was replaced with the next. When the last sentence of a passage had been read, an instruction appeared for the subject to recall, write a title, or to continue on to the next passage, depending on the task condition. The time each sentence was left on the screen was recorded with a grain of 1.83105 ms, the period of the MicroNOVA's internal built-in real-time clock.

After reading a set of instructions, the subject read and responded to a practice passage presented in an arbitrary sentence order. After being checked to ensure understanding of the task, the subject was allowed to proceed to read and respond to the 24 experimental passages. The subject wrote the recall or titles down by hand on a notepad.

Task Instructions. The instructions for the three tasks were intended to inform subjects as fully as possible about the nature of the task and its requirements. The Recall Task subjects were told that the experiment was "concerned with how people read and remember information from simple stories or paragraphs under conditions in which there can be a lot of information to remember for a short time." They were warned that "Some of the paragraphs will be presented in a way so that the sentences may not fit together easily. The sentences may seem mixed up. Your task is to keep track of the information, and fit the entire story together. Even if the paragraph seems confused, it always fits together. Sometimes this will be easy to do, other times, it will be fairly difficult." The instructions required that the recall be in the form of complete sentences, but verbatim recall was not required. Subjects were urged not to waste time and to avoid pausing while reading time was being recorded.

The Topic Choice subjects were instructed similarly to the recall subjects except that instead of recall, they were asked to produce "a title, or a statement about the topic of the paragraph." This title statement had to obey the following

rules: (1) It must name a person or thing actually mentioned in the paragraph, rather than something inferred or deduced. (2) It must name a single person or thing; (3) It should be a short phrase, not a sentence. Subjects were urged not to waste time as described above.

Because of the lack of the demand in the Free Reading Task, the instructions for this condition were carefully prepared. The Free Reading subjects were told that "the experiment concerned the amount of time it takes to understand each sentence ... For this experiment to work, it is very important that the time you leave each sentence on the screen is an accurate reflection of how long it actually takes you to understand it. ... read and think about each sentence long enough to understand it, and understand how it fits into the complete paragraph. But do not spend any longer on the sentence. Especially do not try to memorize the sentences. This is NOT a memory experiment. You will not be tested on what you remember. ... so as you read each sentence, as soon as you feel that you fully understand its meaning, and how it relates to the other sentences in the paragraph, immediately tap the space bar and go on to the next sentence."

Experimental Results

The data were first subjected to a set of standard statistical analyses to confirm that the experimental manipulations were effective and that the basic data used in the simulation analyses were reliable. The results under these

standard analyses will not be given any detailed discussion here since the simulation analysis is able to provide a more comprehensive and theoretically motivated account.

Recall Task Recall Results. The recall protocols were scored blind for gist reproduction of the presented sentence propositions, as described in Kieras(1978). The recall proportion was computed for each of the eleven proposition types in the six presentation order conditions for each of the 29 subjects, collapsing across passages, thus yielding 66 data points per subject. The mean proportions of recall are shown in Table 3. These data were subjected to a 2 X 3 X 11 within-subjects ANOVA in which Topic Sentence Position, Coherence Level, and Proposition Type were factors. The overall mean proportion of recall was .64. Recall was higher for topic sentence appearing first, .67, than for topic sentence appearing last, .62, ($F(1,28)=17.04$, $p<.01$). Recall declined as coherence declined, being .71, .65, and .58, for high, medium, and low coherence, respectively ($F(2,56)=38.13$, $p<.01$). The proposition types differed in level of recall ($F(10,280)=25.57$, $p<.01$). Proposition A was recalled best, followed by the linking propositions D and E, with all others having roughly similar levels of recall.

Interactions appeared between Coherence and Proposition Type ($F(20,560)=2.13$, $p<.01$), Topic Position and Proposition Type ($F(10,280)=2.08$, $p<.05$), and Coherence, Topic Position, and Proposition Type ($F(20,56)=2.77$, $p<.01$). These interactions, especially the three-way interaction, show that recall for

Table 3

Mean Proportion of Recall for Each Sentence Type

Condition		Two-Noun Sentences					One-Noun Sentences					
Coherence	Topic	A	D	E	H	I	B	C	F	G	J	K
High	First	.86	.82	.84	.72	.74	.64	.76	.72	.66	.71	.78
High	Last	.92	.81	.78	.66	.59	.60	.59	.71	.53	.57	.59
Medium	First	.85	.76	.81	.67	.59	.66	.67	.60	.49	.65	.55
Medium	Last	.82	.76	.66	.68	.57	.70	.56	.57	.41	.60	.53
Low	First	.76	.73	.80	.52	.58	.60	.63	.52	.46	.43	.40
Low	Last	.78	.63	.72	.59	.56	.38	.48	.52	.48	.63	.51

Note. The MSE for these means is .0445.

individual propositions was differentially affected by presentation order, confirming that the presentation order manipulations were effective.

Topic Choice Results. The topic choices were scored blind for verbatim or synonymous reproduction of one of the six nouns from the presented passages. The distribution of choices is shown in Table 4, where the labels N1 through N6 correspond to those in Figure 1. About 20% of the responses were scored as "other" in that they mentioned more than one noun, or nouns that were generalizations, such as a label for a standard situation described in the passage. These responses were essentially schema labels, but will not be discussed further here. As a simple check on whether the pattern of topic choices in Table 4 depends on the presentation order, a chi-square statistic was computed in which the null hypothesis was that Coherence Level, Topic Sentence Position, and Topic Choice were independent factors. The departure from mutual independence was significant ($\chi^2(32)=50.44$, $p<.05$), showing as above, that the subjects' responses were differentially affected by the presentation orders, meaning that the order manipulations were also effective in the Topic Choice Task.

Reading Times. The data of primary interest, the reading times, were collapsed for each subject across the four passages in each condition to yield a set of 11 reading times for each subject, one for each sentence type, in each of the six presentation order conditions. The mean reading times for the three tasks are shown in Tables 5, 6, and 7. To show that the

Table 4
Distribution of Topic Choices of Each Noun

Coherence	Topic	N1	N2	N3	N4	N5	N6	Other
High	First	.41	.23	.08	.08	.00	.00	.20
High	Last	.46	.18	.07	.02	.03	.02	.22
Medium	First	.48	.18	.02	.10	.02	.01	.20
Medium	Last	.53	.13	.06	.07	.03	.02	.17
Low	First	.50	.16	.06	.08	.02	.01	.18
Low	Last	.35	.13	.12	.10	.08	.02	.22

experimental manipulations had effects on reading times, the data from each task were subjected to separate ANOVAs.

The Recall Task reading times had a grand mean of 3.802 seconds, and showed no significant main effects of Coherence ($p > .1$), or Topic Position ($F < 1$). Proposition Type did have a significant main effect ($F(10,280) = 7.49, p < .01$). The presence of interactions between presentation order conditions and Proposition Type indicates that the presentation conditions did influence reading time on individual sentence types. Proposition Type interacted with Coherence ($F(20,560) = 1.97, p < .05$), Topic Position ($F(10,280) = 2.56, p < .01$), and with both Coherence and Topic Sentence Position ($F(20,560) = 1.75, p < .05$).

The Topic Choice reading times had a grand mean of 1.817 seconds, and showed significant effects of Coherence ($F(2,58) = 7.04, p < .01$) in which the lower the coherence, the longer the reading time. The topic sentence appearing first produced shorter reading times than if it appeared last ($F(1,29) = 3.25, p < .05$). Proposition Type was significant ($F(10,280) = 20.42, p < .01$). The three-way interaction was not significant ($p > .1$) but Proposition Type interacted with Coherence ($F(20,580) = 2.71, p < .01$) and with Topic Position ($F(10,290) = 3.68, p < .01$).

The Free Reading Task reading times had a grand mean of 1.525 seconds. Lower coherence was again associated with longer reading times ($F(2,58) = 36.87, p < .01$), but Topic Position had no effect ($F < 1$). Proposition Type was significant ($F(10,290) = 43.20, p < .01$). Coherence and Topic Position interacted

($F(2,58)=4.09$, $p<.05$). Proposition Type interacted with Coherence, ($F(20,580)=6.65$, $p<.01$), with Topic Position ($F(10,290)=5.81$, $p<.01$), and marginally with both Coherence and Topic Position ($F(20,580)=1.547$, $.05<p<.10$).

Discussion of Experimental Results

The standard data analyses are intended only to show that the experimental manipulations produced reasonable effects in the reading times, with additional confirmation provided by the effects on recall and topic choice in two of the tasks. The most general indication that the presentation order manipulations were effective appears in the many significant interactions of Proposition Type with one or both presentation factors. These show that the reading time, apparent topic, and recall of the very same passage sentences differs with the order in which they appear in the passage. The recall and topic choice effects are comparable to the corresponding effects reported in Kieras (Note 1; 1978) and so will not be discussed any further here because the reading times are of primary interest in this paper.

Overall, the Recall Task by far produced the longest and most variable reading times. As would be expected, the Free Reading Task produced the smallest and least variable reading times. The Topic Choice Task reading times were only slightly slower and more variable than the Free Reading times. Hence in terms of overall reading time, the Free Reading and Topic Choice Tasks are very similar, with the Recall Task being markedly

different.

In terms of the reading times for each sentence, the means in Tables 5, 6, and 7 present no easily discernable pattern. While many differences can be pointed out and discussed by piece-meal comparisons, conventional ways of looking at results of such detail and complexity produce no coherent theoretical conclusions. The simulation analysis reported next is able to provide a comprehensive account of the pattern of reading times in the three tasks.

SIMULATION ANALYSIS

The original goal in developing the simulation model was to explore in some detail the processes that would have to be involved in sentence integration. The model would have to understand given and new references correctly, so that, after a series of sentences were read in, the model would have in memory an integrated representation of the passage content. In order to devise the model, theoretical ideas had to be developed on the function of noun phrases, the use of long-term memory information during integration, and how the processing of each sentence should take into account the context supplied by the previous sentences. An important step was assigning the passage topic the specific role of acting as a summary of the passage context. Throughout the simulation, certain popular pre-existing ideas about language processing were used, such as current notions about memory representations, memory search, and sentence parsing. This means that what is being tested

Table 5
 Mean Reading Time of Each Sentence
 in Each Presentation Order in the Free Reading Task

Sentence Type	High Coherence		Medium Coherence		Low Coherence	
	Topic	Topic	Topic	Topic	Topic	Topic
	First	Last	First	Last	First	Last
A	1.873	1.533	1.776	1.658	1.841	1.624
B	1.295	1.220	1.270	1.476	1.454	1.504
C	1.320	1.323	1.261	1.425	1.388	1.585
D	1.607	1.731	1.597	1.674	1.689	1.703
E	1.590	1.477	1.570	1.651	1.585	1.621
F	1.404	1.183	1.261	1.372	1.525	1.690
G	1.350	1.250	1.296	1.337	1.527	1.582
H	1.630	1.529	1.755	1.809	1.707	1.802
I	1.540	1.443	1.770	1.816	1.711	1.787
J	1.380	1.221	1.319	1.352	1.571	1.644
K	1.410	1.310	1.361	1.396	1.592	1.659

Note. The MSE for these means is .0444.

Table 6
 Mean Reading Times for Each Sentence
 in Each Presentation Order in the Topic Choice Task

Sentence Type	High Coherence		Medium Coherence		Low Coherence	
	Topic	Topic	Topic	Topic	Topic	Topic
	First	Last	First	Last	First	Last
A	2.747	2.345	2.411	2.074	2.479	2.301
B	1.393	1.527	1.437	1.840	1.704	1.667
C	1.509	1.648	1.440	1.712	1.523	1.629
D	1.797	2.246	1.994	1.988	2.190	2.277
E	1.778	1.739	2.094	2.076	1.931	2.027
F	1.437	1.601	1.486	1.525	1.654	1.692
G	1.553	1.534	1.477	1.545	1.719	1.611
H	1.771	1.782	2.004	2.088	2.039	2.351
I	1.806	1.787	1.924	2.003	1.960	1.964
J	1.509	1.509	1.508	1.605	1.743	1.949
K	1.622	1.711	1.647	1.664	1.848	1.752

Note. The MSE for these means is .2261.

Table 7
 Mean Reading Times for Each Sentence
 in Each Presentation Order in the Recall Task

Sentence Type	High Coherence		Medium Coherence		Low Coherence	
	Topic	Topic	Topic	Topic	Topic	Topic
	First	Last	First	Last	First	Last
A	4.474	3.740	4.593	3.259	4.780	3.965
B	3.424	2.780	3.290	3.562	3.400	4.289
C	3.409	3.397	3.614	3.388	3.455	4.136
D	3.686	4.263	4.085	3.634	3.841	3.947
E	4.141	3.572	3.818	3.370	3.365	3.938
F	3.974	3.555	3.603	3.237	3.788	4.503
G	3.356	3.218	3.045	3.117	4.015	4.452
H	4.122	4.967	4.720	4.856	4.106	4.483
I	3.672	3.670	4.307	4.380	3.882	4.298
J	3.823	3.876	3.790	3.676	3.456	3.623
K	3.019	3.225	3.490	3.283	3.752	3.948

Note. The MSE for these means is 1.127.

empirically in this work are some of the commonly held notions about how language might be processed.

The simulation model was not originally intended to be used to model experimental data at the quantitative level. Rather it was developed to meet the informal, loose criteria for accuracy and plausibility commonly used in simulation work. That is, the model had to behave in an apparently realistic manner by understanding a variety of sentence and text forms, and operate using psychologically plausible mechanisms. Although this criterion is only loosely stated, in practice it is very hard to meet because the model is required to behave appropriately for a very wide range of linguistic inputs, a much wider range than is usually addressed by conventional psychological theories of language processing. Hence, although the model was not originally intended to be subjected to the formal quantitative test that is the subject of this paper, it nonetheless had to give the appearance of realism by meeting many of the requirements involved in actually processing prose.

Description of the Simulation Model

Rather than indulge in a full presentation of the simulation here, the reader is referred to the description in Kieras(1977). Only some of the most important properties of the simulation will be described here. Since, as mentioned above, the simulation is used to produce numeric variables reflecting the amount of processing, the emphasis in the following description will be on the mechanisms in the simulation relevant to these variables.

In brief, the simulation model is a large computer program, written in the LISP programming language, that accepts connected discourse as input, one sentence at a time. Using a set of text- and sentence-grammar rules, it identifies the given-new structure in the input sentence. It then searches memory for the representations of the given referents, and after finding them, attaches to them the structure representing the new information. A broad set of kinds of given and new referents can be handled by the model.

At the current stage of theory in cognitive psychology and artificial intelligence, any model of language processing will be severely limited. This model is no exception. Some obviously important aspects of language, such as verb phrase structure, were essentially ignored. For the purposes of the work described in this paper, an even simpler model would have sufficed. The simple passages simply do not require much in the way of grammatical or semantic analysis. Thus, much of the machinery in the model described in Kieras (1977) was not actually used in processing the passages for this work. However, there is nothing in the structure of the model itself that would definitely prevent the extension of this work to more complicated sentences and passages.

Memory Representations. The memory representations are basically a modified version of HAM memory structures (Anderson & Bower, 1973), being most similar to HAM2 and ACT representations (Anderson, 1975, 1976). These are semantic network representations, consisting of nodes interconnected with

links A natural way to quantify the amount of memory structure in such a representation is to simply count the number of links in it. This measure is used in the work reported here. The finer details of the network structure are not important in the present context, and so will not be discussed here.

An important property of the representations that can be justified on linguistic grounds (Bach, 1968) is that the main noun class membership of the referent of a noun phrase is represented by a separate proposition, rather than by a special link type as in Anderson and Bower's HAM (see Anderson, 1972, 1976). Hence the sentence The ants were hungry would consist of two propositions, one being X is ants, which defines the main class membership of X, the other X is hungry. Likewise, a sentence such as The ants ate the jelly would consist of three propositions X is ants, Y is jelly, and X ate Y.

This decision on how to represent the content of noun phrases is important, especially in regard to how the given-newness of a referent is to be represented. If a referent is given, then the proposition defining its main class membership does not have to be built; if it is new, then this proposition must be constructed. For example, if The ants is given in The ants are hungry, then only the X is hungry proposition must be constructed. While there are certainly other ways in which the given-new status of a referent might be represented, this one has the advantage that the given-newness is represented in terms of propositions, and is thus homogenous in representation with the rest of the sentence content.

Another important property of the representations has to do with the nature of the propositions. Using the above examples, the representation for the proposition X is hungry is simpler than that of the proposition X ate Y. The first requires only the Subject-Predicate pair of links; the second requires two more links to make up the Relation-Object form of predicate. Hence, this representation system requires more structure to represent subject-verb-object sentences than simple subject-adjective sentences.

This HAM-like system may be contrasted with that of Kintsch(1974). First of all, in Kintsch's representations, referents which appear as arguments in propositions are not separately defined by other propositions (But see Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). Hence, the given-new differences would have to be represented in some other way besides a simple difference in number of propositions. Furthermore, in the Kintsch systems, the "fine structure" of propositions is not defined, as in the semantic network models. Thus, the Subject-Verb-Object and Subject-Adjective sentence forms both appear as on unit proposition, such as (HUNGRY,ANTS) and (ATE, ANTS, JELLY). Here the difference between the two kinds of propositions could be measured in terms of differences in number of arguments; however, this measure would have to be distinguished from the given-newness of each argument. The HAM-like system used in the simulation simply expresses all of these differences in referents and propositional form in terms of the number of links required for the propositional structure.

It must not be thought that some things are possible in the HAM-like representations that are not possible in the Kintsch representations. While there are some problems with the Kintsch representations, due to the fact that they are not a thoroughly formalized system, we know that representation systems can not be distinguished empirically, as long as they can both minimally adequate to the representational demands of the model (Anderson, 1976, 1978; Kieras, Note 3). The HAM-like representations are used in the model because of their greater suitability to a simulation, and because of their property of reducing sentence content and referent definitions to a uniformly quantifiable variable, namely, the number of links that have to be constructed.

Parsing. The next portion of the simulation to be discussed is the parsing mechanism and its system for representing grammatical rules. The parser is a simple form of Augmented Transition Network (ATN) parser. Discussion of ATN parsers in general can be found in Woods(1970), Anderson(1976), Rumelhart(1977), and of the parser used here in Kieras(1977).

In brief, an ATN parser consists of a network representing grammatical parsing rules, and a program which uses the network to analyze a sentence. The ATN consists of nodes, which represent states of the parse, and directed links between the nodes which represent possible transitions between the parsing states. Each link consists of a condition, such as a test on the part of speech of the current word in the sentence, which if satisfied, allows a transition to the next state. An action

associated with the link is performed if the transition is made. This action can consist of activities such as building memory structure, advancing to the next word in the sentence, or saving the meaning of the current word in a list. In processing a sentence, the ATN parser works through the sentence in a left-to-right sequential fashion, and attempts to find a complete path through the parsing network. The actions done along the way result in memory searches for given referents, and the construction of new memory structure to represent the new information in the sentence. In the simulation model, ATNs are used to represent both the rules for parsing sentences and the rules for the structure of the passage as a whole, although so far, the text-grammars have been extremely simple (see Kieras, 1977).

The advantage of the ATN approach to sentence parsing is that not only is it easy to use from a programming standpoint, but it also admits of a simple quantification. Namely, the number of ATN links crossed in parsing the sentence, that is, the number of transitions made, is a simple and natural measure of the amount of processing required to analyze the sentence. Differences in sentence structure show up as differences in the number of ATN transitions. However, differences in memory search or representations do not show up here, because this processing is done by the other parts of the system.

Memory Search. Another portion of the model that requires discussion is the system for memory search. The model uses spreading activation to identify relevant information in memory,

and then does an examination of specific parts of that information to meet a particular need. This approach to memory search is based on the general ideas in Quillian(1968, 1969), Collins and Loftus(1975), and Anderson(1976). A spreading activation search consists of a breadth-first parallel search, starting from a set of activation sources which are nodes in the memory network, and looking for some path in the network which connects all of the sources together. The mechanics of the search consist of a simulated parallel process of tagging nodes with the name of their activation source. A complete round of tagging is called an activation cycle. When all nodes directly connected with the source have been tagged, the next cycle tags all nodes directly connected to the last-tagged nodes, and so forth. Eventually a node gets tagged from all sources, and this node lies on the path of intersection between all the sources. The search is terminated if too many cycles, as specified by a termination criterion, occur without an intersection.

In terms of processing capacity, the activation process can be either unlimited, meaning that all nodes directly connected to currently active nodes are tagged or limited, in which only a certain number of nodes can be tagged in each cycle. The original version (Kieras, 1977) was unlimited in activation capacity. The work described in Kieras(Note 2) led to the development of a limited capacity activation process, which appeared to be more empirically accurate. This process has an additional parameter besides the termination criterion, which is the total activation capacity per cycle. The two parameters are simply set to a convenient value that allows all of the desired

memory searches to be successful.

In terms of quantifiable measures of processing, each cycle can be assumed to take a fixed amount of time, since it is associated with a fixed amount of tagging work. On the other hand, tagging a node could take a fixed amount of time, and so the memory search time could be a function of the number of nodes that had to be tagged. Both the number of cycles and the number of nodes tagged can be easily counted. But, in the work reported here, these two variables turned out not to be very important.

One more quantifiable aspect of the memory search process is the number of activation sources. The more sources competing for the limited activation capacity, the longer it takes to find an intersection. However, this measure of memory search processing is closely connected with assumptions concerning which nodes are chosen for activation sources and how they are kept track of during reading. That is, a central part of the work originally done with the simulation was a study of how the topic of a passage could be used to guide memory search. This was done by simply including the node representing the passage topic in the list of activation sources for all memory searches, so that the search would "automatically" uncover information relevant in the context of the passage topic. Hence, the activation sources were the concepts mentioned in the noun phrase being processed, together with the topic of the passage. Normally, the topic was taken to be the subject of the sentence introducing the passage.

Experiment Version of the Model. In order to enable the simulation model to process the passages in experiments like those in Kieras(1978) and in this paper, certain additions to the simulation's grammar rules had to be made. First of all, most experiments using simple sentences use only the definite determiner, the. The simulation's grammar for definite noun phrases had to be altered to take into account that if a definite reference could not be resolved, it should be treated as a reference to a new object.

A more profound change was made in the text grammar to enable the simulation to respond to incoherent passages. The change was that if an incoherent sentence is encountered, the topic of the sentence, normally assumed to be the subject noun phrase referent, is added to a list of topics. In a coherent passage, all sentences can be connected to the tonic of the first; hence the topic list never contains more than one topic. However, in an incoherent passage, a topic will be added to the list for each sentence that can not be integrated; as the passage is read further, sentences that tie together some of the previously disconnected content result in topics being deleted from the topic list. In the passages used here, and in Kieras(1978), all of the initially unintegrable sentences were tied together by the end of the passage, leaving only one final topic in the list.

Since the passage topic is used as an activation source, incoherent passages, with their multiple topics, result in multiple activation sources being used during sentence

processing. It is assumed, although not explicitly represented in the simulation, that the topic list is kept in the standard short-term memory with limited capacity and short retention time (see discussion in Kieras, 1978).

Model Fitting Results

The first step in fitting the simulation model was to obtain from the simulation the processing variable values for each sentence in each presentation order. In this work, the simulation makes no use of specific word meanings; hence it was not necessary to have the simulation process each of the actual passages used in the experiment, since these differed only in the specific content words used. Instead, a version of the sentences was prepared in which the content words were replaced by symbols representing a prototypical passage. So, for example, the Type A sentences were represented by the sentence THE N1 WAS RELA THE N2, and the Type B sentences by the sentence THE N1 WAS ADJB, in which the symbols correspond to the labels in Figure 1. This set of symbolic prototype sentences was presented to the simulation in the six different presentation orders. The simulation provided a list of values for the processing variables for each sentence in each presentation order. These variables, which were discussed above, are listed and defined for quick reference in Table 8. An example of the simulation output trace during processing a sentence is shown in Figure 2.

Table 8

Predictor Variables Used in the Simulation Analyses

Variables Supplied by the Simulation Model

TRANSITS	Number of ATN links crossed in parsing the sentence.
LINKS	Number of one-way HAM links built to represent the sentence content.
TOPICS	Number of topic pointer being maintained and used for activation sources in all memory searches done while processing the sentence.
CYCLES	Number of cycles of spreading activation required for all memory searches done while processing the sentence.
TAGGED	Number of memory nodes tagged, or activated, during all memory searches required to process the sentence.

Variables Added to the Statistical Model

MTASK	The constant effect of being in the Recall Task.
TTASK	The constant effect of being in the Topic Choice Task.
MLINKS	The number of one-way HAM links being encoded while processing the sentence for later recall in the Recall Task.
FIRSTA	The constant effect of processing a highly thematic sentence in the Topic Choice Task.

```

(INPUT: (C2 H 4.9670000 (THE N3 WAS RELH THE N5)))
(ACTIVATE FROM (G0446 &N3))CCXXC!
(SEARCH WM FROM &N3 WITH ((SPRED *RO &N3) ($NUM *RO 1)))
(INDIV G0448 FOUND NIL REMAINS)
(TOPIC- G0448)
(ACTIVATE FROM (G0446 &N5))CC?
(SEARCH WM FROM &N5 WITH ((SPRED *R2 &N5) ($NUM *R2 1)))
(CREATE THE UNKNOWN OBJECT USING ((SPRED *R2 &N5) ($NUM *R2 1)))
($BUILD ((SPRED *R2 &N5) ($NUM *R2 1)))
(G0466 HAS BEEN CREATED)
($BUILD ((SPRED G0448 *NO) ($REL *NO &RELH) ($ARG *NO G0466)))
(>>> TRANSITS,LTM LINKS,WM LINKS 37 0 13)
(RFS CYCLES, TAGGINGS, HISTORY 3 17 (0 8 6 0 5))
(NUMBER OF TOPICS IS 1)
(INPUT OF (G0446) COHERENT WITH (G0446))
(NEW *TOPICS (G0446))
(INPUT: (C2 J 3.8760000 (THE N5 WAS ADJJ)))
(ACTIVATE FROM (G0446 &N5))CCXC!
(SEARCH WM FROM &N5 WITH ((SPRED *RO &N5) ($NUM *RO 1)))
(INDIV G0466 FOUND NIL REMAINS)
(TOPIC- G0466)
($BUILD ((SPRED G0466 &ADJJ)))
(>>> TRANSITS,LTM LINKS,WM LINKS 20 0 4)
(RFS CYCLES, TAGGINGS, HISTORY 2 15 (0 8 7))
(NUMBER OF TOPICS IS 1)
(INPUT OF (G0446) COHERENT WITH (G0446))
(NEW *TOPICS (G0446))

```

Figure 2. Excerpt from the simulation output trace, showing the processing of two sentences in the High Coherence Topic Last condition. The first sentence shown is sentence H, which required 37 TRANSITS and 13 LINKS to process. The second sentence is sentence J, which required 20 TRANSITS and 4 LINKS.

These variables comprised the basic set of predictor variables in the multiple regression analyses, with certain other variables being added during the building of the statistical model. These additional variables, also shown in Table 8, were not in fact explicitly represented in the simulation, but are substantive features of the statistical model. The dependent variable in the analyses was the mean reading times for each sentence type in each presentation order for the three tasks, collapsed across subjects and passages. These are the means shown above in Tables 5, 6, and 7.

The regression analyses were conducted using the stepwise multiple regression program, BMD02R, in which variables are added to the prediction equation in descending order of their contribution toward accounting for variance in the dependent variable. As with any multiple regression analysis, decisions had to be made about which and how many variables should be allowed into the regression equation. Three criteria were used: (1) Although many variables could be defined, as few variables as possible were tried, consistent with the quality of the fit and the theoretical interpretability of the resulting equations. (2) Each variable allowed in the equation must have made a significant, at the .05 level, contribution to the total variance accounted for. (3) Since some of the simulation processing variables tend to be correlated (see Kieras, Note 2), sometimes suppression (Cohen & Cohen, 1975) can occur, in which one of the variables acts to suppress some of the variability present in another. In the relevant form of suppression, a new variable enters the equation with a negative coefficient.

Because the simulation processing steps must all take positive amounts of time, such a variable can not be permitted in the equation. Under these criteria, the stepwise process was terminated when additional variables either made no significant contribution, or resulted in suppression with negative coefficients. Hence the regression equations to be reported are all theoretically defensible and statistically predictive.

The model fitting analyses were conducted in three steps. The first was to fit the simulation model to the individual reading tasks. The second was to build a good regression equation for all three tasks together. The third was to contrast the simulation variables with a theoretically empty, but empirically meaningful variable, the number of words in the sentence. This step provides a measure of the value of the simulation compared to models that are considerably less detailed in terms of processes.

Single Task Fits. The multiple regression analysis was applied separately to the reading times for each task, but using the same set of values for the simulation processing variables. A summary of the analyses for the three tasks appears in Table 9, showing the order of entry of each variable, R^2 , and the final regression equation. The criteria discussed above governed at which step the stepwise process was terminated. As can be seen, the best fit is on the Free Reading Task, followed by the Topic Choice Task, with the fit on the Recall Task being poorest.

Table 9
Results of Single Task Regression Analyses

Task	Step	Variable	R ²		Final Coefficient
<u>Free Reading Task</u>					
	1.	LINKS	.59	**	.035
	2.	TAGGED	.71	**	.030
	3.	TOPICS	.75	**	.002
		CONSTANT			1.116
<u>Topic Choice Task</u>					
	1.	TRANSITS	.51	**	.007
	2.	LINKS	.60	**	.036
	3.	TAGGED	.62	*	.003
		CONSTANT			1.210
<u>Recall Task</u>					
	1.	LINKS	.51	**	.082
	2.	TOPICS	.55	*	.068
		CONSTANT			2.950

* F-ratio significant at .05, ** at .01.

Some noteworthy features of the equations can be pointed out. Consistent with the different grand means reported above, the constant in the Recall Task is about three times longer than in the other two, where the constants are similar in value. The estimate for the LINKS coefficient is identical in the Free Reading Task and the Topic Choice Task, but is about twice as large in the Recall Task. The other variables do not follow any obvious pattern, and so will not be discussed further in terms of the single task fits. However, LINKS was clearly the most important variable, as it entered the equation first for both the Free Reading and the Recall Tasks, and second after TRANSITS, in the Topic Choice Task. Hence, a key process in the model, representation building, stands out as important in predicting reading times. However, the process differs between the Recall Task and the other two tasks in that there is a larger coefficient value assigned to LINKS in the Recall Task.

Overall Fits. The next step in the simulation analysis was to model the complete set of data from all three tasks with a single regression equation. As shown by the single task fits, the Recall Task is rather different from the other two. A simple way of representing the difference in baselines and LINKS coefficients was tried. A dummy variable (see Cohen & Cohen 1975) was made available to the regression program. This variable, MTASK, was 0 except for sentences in the Recall Task, where it was 1, and would thus cover any overall mean difference between the Recall Task times and the other task times. An interaction variable (see Cohen & Cohen 1975) was also made available. This variable, MLINKS, was the product of LINKS and

MTASK, and so was 0 except for the Recall Task sentences. where it was equal to LINKS. This variable would thus cover the apparent difference in the LINKS coefficient revealed in the single task fits.

Psychologically, MTASK represents the overall effect of being in a Recall Task, while MLINKS represents the number of links that must be processed for purposes of recall that do not have to be processed in the other tasks. That is, MLINKS is a measure of the amount of memory structure that must be memorized, or encoded in LTM, for later recall. It is 0 in the other tasks because these do not require encoding for recall.

Table 10 shows the summary of the regression analysis. While the overall fit is very good, with 94% of the variance among mean reading times accounted for, it must be pointed out that the dummy variable MTASK entered the equation first and by itself accounted for 89% of the variance. Hence the overall R^2 is somewhat misleading because a degenerate model, consisting of only the constant and MTASK, produces most of this overall R^2 .

As a more realistic measure of the goodness of fit of this equation, times predicted by the equation were correlated with the observed times in each task. These values of r^2 were .74 for the Free Reading Task, .54 for the Topic Choice Task, and .53 for the Recall Task. Hence, compared to the individual task fits, this one equation gives almost as good a fit to the separate tasks as the separately fitted equations from the single task fits. Since this fit is the most interesting of those to be presented, it is shown graphically in Figure 3 for

Table 10

Overall Task Fit Using MTASK and MLINKS

Step	Variable	<u>R</u> ²		Final Coefficient
1.	MTASK	.89	**	1.794
2.	LINKS	.93	**	.033
3.	MLINKS	.94	**	.040
4.	TOPICS	.94	**	.038
5.	TRANSITS	.94	*	.006
	CONSTANT			1.130

* F-ratio significant at .05, ** at .01.

the Recall Task, Figure 4 for the Topic Choice Task, and Figure 5 for the Free Reading Task. The figures show the predicted and observed reading times in each presentation order condition, with the abscissa showing the order of the sentences in the passage.

A supplementary regression analysis was done to determine in more detail the role of the MTASK dummy variable. MTASK was forced to remain out of the equation while MLINKS, LINKS, and TRANSITS were forced into the equation. Without MTASK in the equation, the R^2 was .80, but LINKS had a negative coefficient, apparently compensating for a very large coefficient given to MLINKS and a too-large value for the constant. When MTASK finally was allowed to enter the equation, the R^2 rose to the previous value of .94 and the final equation in Table 10 was obtained. Hence the MTASK variable really is required to allow all of the important variables to have positive coefficients, but only 14% of the variance accounted for is uniquely contributed by MTASK.

Considering the final equation in Table 10, the constant has a value similar to that in the individual task equations in Table 9 for the Free Reading and Topic Choice tasks. As represented by MTASK, the Recall Task adds an additional constant of 1.794 seconds across the board. The .033 seconds/link estimate for the LINKS coefficient agrees with that in the Free Reading and Topic Choice single task fits, while the coefficient for MLINKS covers the additional time per link required in the Recall Task for representation building. In

RECALL TASK

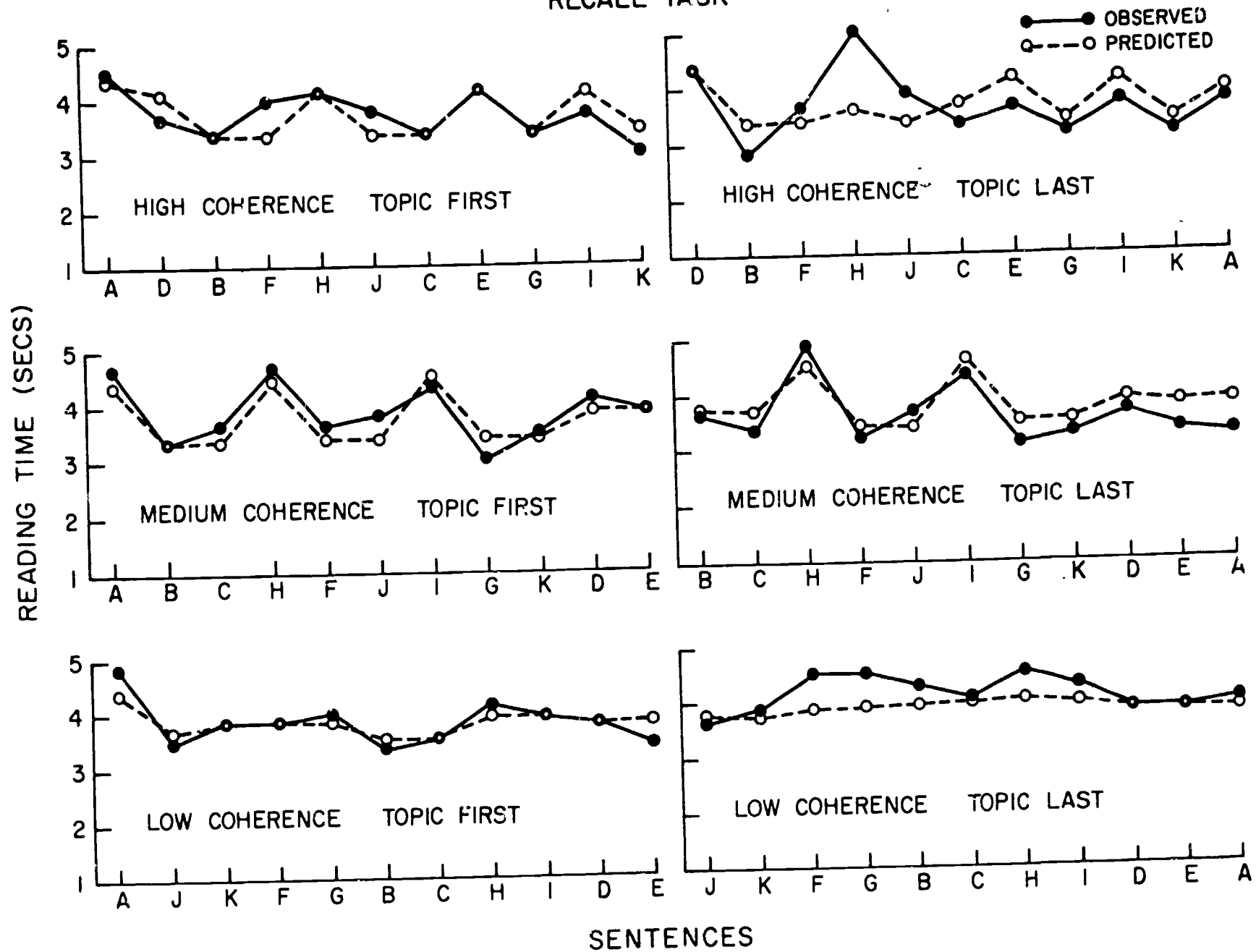


Figure 3. Predicted and observed reading times in the Recall Task.

TOPIC CHOICE TASK

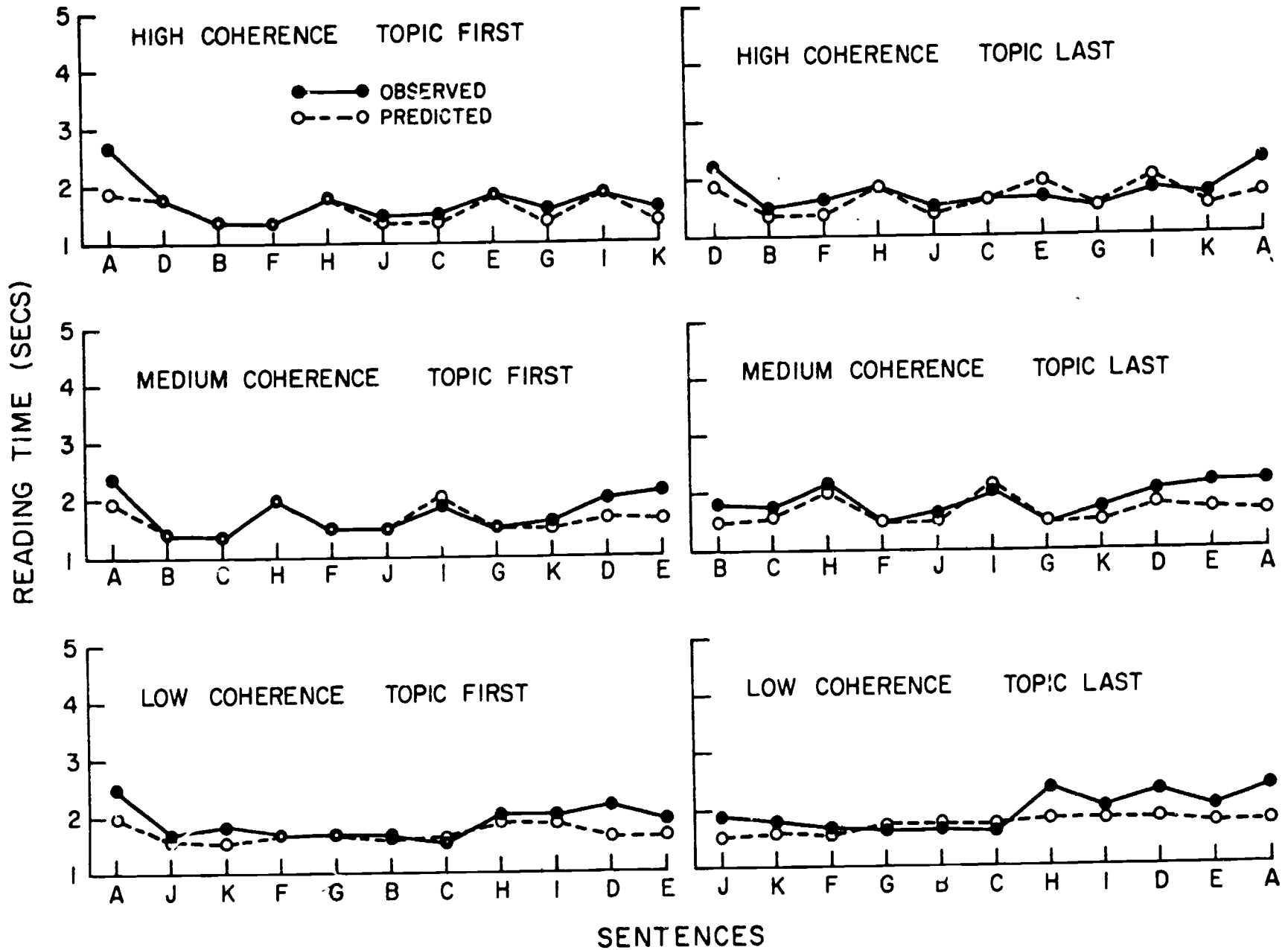


Figure 4. Predicted and observed reading times in the Topic Choice Task.

FREE READING TASK

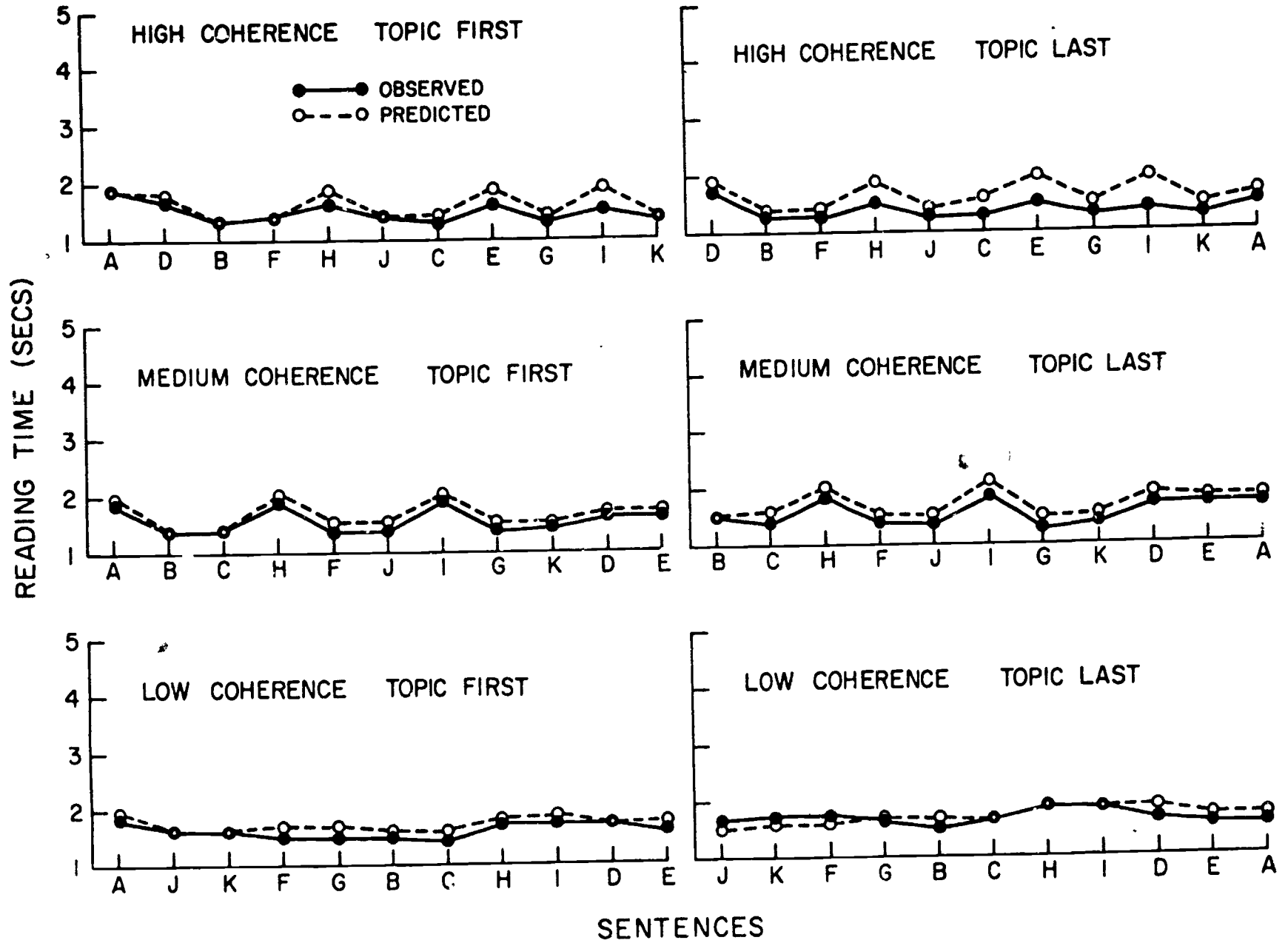


Figure 5. Predicted and observed reading times in the Free Reading Task.

this equation, the parsing variable TRANSITS accounts for a small but significant portion of the variance. Finally, according to the equation, each topic maintained during sentence processing adds an additional .038 seconds.

Despite the overall good fit, examination of the graphs in Figures 3, 4, and 5 reveals some serious discrepancies and some important details of the fit. Most of the presentation order conditions are fit fairly well, except for those in the Recall Task. An interesting detail of where the fit is good appears in the Low-Coherence Topic-Last condition in the Free Reading and Topic Choice Tasks. Both the predicted and observed reading time profiles are relatively flat. This is due to the fact that each of the initial one-noun sentences contains only new information, and so takes relatively long because of the need to build the representation for the new referent. The two-noun sentences that then appear take relatively little time to process since all of their referents are given, meaning that little new representation must be built. Hence both the one-noun and the two-noun sentences require roughly the same amount of processing, and the reading time profiles, both as observed and as predicted, are relatively flat.

The presentation orders in which TOPICS makes the widest excursions are the Low Coherence presentation orders. In these, the number of topics (chunks, or unintegrated propositions being held in working memory) increases with each of the first several sentences (see Table 2). However, for the Recall Task, as shown in Figure 3, the fit of the model in the Low-Coherence

Topic-Last presentation order is very poor; the reading times after the first three sentences do not continue to increase as TOPICS does. Since TOPICS plays only a weak role with its small coefficient, the simulation predicts an essentially flat profile in this presentation order for the Recall Task, as in the other tasks. However the observed reading times do not follow this pattern; they first increase, then decrease. Clearly, Recall Task readers are engaged in some process in this presentation order condition that the simulation completely fails to capture.

Another discrepancy in the fit provided by the Table 10 equation shows up in the Topic Choice Task. Inspection of Figure 4 shows that the model consistently underpredicts the reading times on the first sentence in the Topic-First presentation orders. Moreover, in the Topic-Last conditions, the reading time on the last sentence is also consistently underpredicted. Some special role for these sentences would be expected from the results in Kieras (Note 1, Note 4, 1978), and several unpublished experiments on the simple passages, which show that the first sentence position is particularly important in what readers consider the main item or idea to be. Also, in these simple passages, the Type A sentence is strongly perceived as the main idea, regardless of its position (although it is chosen as the main idea less often if it appears last). Thus, the extra time spent on the first sentence and on a final Type A sentence could be due to topic identification processing performed by the readers. The first sentence is processed longer because it is considered to be an important cue to the main idea, while if the last sentence is a Type A sentence, its

central position in the passage meaning structure also invokes additional processing as a cue to the main item. It is clear that the variables provided by the simulation fail to capture this additional processing. It should be pointed out that the discrepancies also appear in the Single Task Fit for the Topic Choice Task; in fact, they are somewhat more pronounced.

Another statistical model was constructed to see whether representing the extra topic choice processing in the equation would improve the fit. Another dummy variable, TTASK, was defined; TTASK was 0 except for sentences in the Topic Choice Task, where it was 1. An interaction dummy variable, FIRSTA, was defined that was 0 everywhere except for highly topical sentences in the Topic Choice Task, where it was 1. These sentences were the first sentence or a final Type A sentence in each passage.

Notice that MTASK and TTASK comprise a complete set of dummy variables that discriminate the three tasks from each other. By virtue of their definitions, the Free Reading Task can be considered as a reference or baseline condition, with the other tasks requiring additional processes and time above this condition. This characterization is appealing because the Recall and Topic Choice Tasks can be considered as being simply a Free Reading Task with additional demands on the subject.

The results of the multiple regression are summarized in Table 11. The final R^2 is .96. Both TTASK and FIRSTA account for significant portions of the variance. The constant is smaller than the previous fit, since some of the difference in

3

Table 11
Overall Task Fit Using TTASK,
FIRSTA, MTASK, and MLINKS

Step	Variable	<u>R</u> ²		Final Coefficient
1.	MTASK	.89	**	1.887
2.	LINKS	.93	**	.029
3.	TTASK	.94	**	.232
4.	MLINKS	.94	**	.046
5.	TOPICS	.95	**	.051
6.	FIRSTA	.96	**	.447
7.	TRANSITS	.96	*	.006
	CONSTANT			1.010

* F-ratio significant at .05, ** at .01.

the baseline time between the Topic Choice Task and the Free Reading Task is now represented in TTASK. The coefficient for TTASK shows that an additional .232 seconds is required across the board for the Topic Choice Task, compared to the Free Reading Task. As shown by the coefficient for FIRSA, an additional .447 seconds is used on the first sentence or a final Type A sentence. While the overall fit is only slightly better than that of Table 10, the predicted reading times on the first and final-A sentences in the Topic Choice Task are considerably better, being in error by an average of .133 seconds compared to the previous fit average error of .507 seconds.

Comparison with Number of Words. The use of a complex model is justified if it provides a better account than a simpler one. Since the main feature of the simulation model is a complex set of processes that operate on the surface form of the sentence, a good simple opponent to the simulation model is one based only on the surface form, namely the number of words in the sentence. Since the time to read a sentence is clearly a function of the number of words, and the amount of information in a sentence is closely related to the number of words, this variable, WORDS, should be a reasonably good predictor of reading time. If it is as good or better than the simulation, it could be argued that the lack of parsimony in the simulation is too high a price to pay for its predictive power. On the other hand, if the simulation and statistical model together can account for more variance than WORDS, the use of such a complex model is clearly justified.

The variable WORDS was made available to the multiple regression analysis along with the variables that were available in the analysis shown in Table 11. WORDS was set to a value of 4 for each of the one-noun sentences, and at 6 for each of the two-noun sentences. A useful preview of how WORDS performed as a predictor can be had from the correlations between the candidate predictor variables. WORDS and TRANSITS were very highly correlated, .988, but WORDS correlated only moderately with the other simulation variables. WORDS correlated slightly more with reading time, .187, than did TRANSITS, .169. However, other variables correlated more with reading time than WORDS: LINKS correlated .210, MLINKS, .891, and MTASK, .941. Hence it could be concluded that WORDS could replace TRANSITS in the regression equation, but not the other simulation variables, since in the above analyses the other variables were more important than TRANSITS.

In the regression analysis, WORDS entered the equation only after the other variables. It did account for a significant portion of the variance, and carried a coefficient of .068 seconds/word. The other variables had coefficients very similar to those in Table 11. The final R^2 using WORDS instead of TRANSITS was slightly larger, by .0003. Hence, although WORDS is a better predictor than TRANSITS, it is not better to any important extent. Moreover, this regression analysis confirms that the other variables are more important than either WORDS or TRANSITS.

Discussion of Model Fitting Results

Which Task does the Model Fit Best? Both the single task fits and the overall task fits show that the simulation model corresponds most closely to the Free Reading Task. This is reasonable because the simulation was designed primarily to simply comprehend sentences, that is, parse them and build the corresponding memory representations in working memory. It thus has little or no explicit machinery to either choose topics or memorize for recall. The specific places discussed above in the Recall and Topic Choice Task data where the model fails to fit conform to this conclusion. In the Topic Choice Task, subjects performed additional processing on sentences important to the choice of topic, and in the Recall Task, subjects apparently experienced an increasing memory load in the Low-Coherence Topic-Last order, which they apparently met by sloughing the remaining one-noun sentences. Since the simulation has neither memory overload problems, nor a strategy for dealing with memory overload, it simply cannot predict this pattern of reading times.

Adding the variables MIASK, MLINKS, TTASK, and FIRSA to the statistical model results in a good fit in all conditions except for the Low-Coherence Topic-Last Recall Task. But these variables have no direct and explicitly represented counterparts in the simulation model itself. Hence, the final statistical model includes substantive variables. While this certainly means that the simulation by itself is not an adequate model for

the Topic Choice and Recall Tasks, the specific variables that have to be added to arrive at a good statistical model point out in a clear and definite way what the simulation model lacks. On the other hand, the simulation does do a reasonable job of predicting reading times in the Free Reading Task which is closest to the kind of task the simulation was originally designed to do.

How do the Reading Tasks Differ? The Free Reading Task constitutes a baseline task, which as suggested by the simulation analysis, includes parsing and representation building processes, but lacks topic finding, memorization, and memory overload processes.

The Topic Choice Task clearly demands different and additional processes compared to the Free Reading. Extra amounts of time were required on highly topical sentences, and an overall longer time of .232 seconds/sentence compared to Free Reading was required, perhaps because subjects had to keep track of their current topic choices while reading.

The Recall Task required additional time to be spent on memorizing memory representations, and a large increase in overall time, almost 2 seconds/sentence. But this rather large coefficient for MTASK, a variable of considerable predictive importance, is hard to understand. That is, while MLINKS accounts for additional time based on how much information must be memorized, MTASK indicates nothing more than an all-or-none across the board increase in processing time. It is thus unrelated to any of the possible hypothetical processes that

might be required in a memorization task, all of which would be expected to vary in the course of reading a passage.

It is possible that the importance of MTASK is not due to an actual task difference, but to a difference in subject motivation; as described above, the Recall Task subjects were compensated with apparently badly-needed extra credit, and appeared to be very highly motivated compared to the subjects in the other tasks who were only paid \$2.00. However, this explanation is really not satisfactory either, since it is still not clear why a motivational difference would be associated with a greater baseline reading time that does not vary as a function of processing requirements while reading a passage.

Finally, the Recall Task also involved memory overload, and strategies for dealing with memory overload, which did not appear in the other tasks. Hence the hypothesis that incoherence results in increased memory load (Kieras, 1978) appears to apply only in a recall task. In the other tasks, the subjects are able to process the incoherent sentences without suffering from memory load problems.

Thus, by using the simulation as a theoretically based tool for characterizing reading tasks, the important differences between the reading times are pointed out. Moreover, a theoretical explanation for the task differences is then available in terms of processes represented in either the simulation model or the statistical model.

Decomposing Reading Times. The most useful service offered by the simulation analysis is a decomposition of the reading times into theoretical components. Based on the final fit in Table 11, these components can be summarized. The constant value indicates that in these tasks and materials, and with the equipment used, about 1 second is required to visually perceive the sentence and make the response to advance to the next sentence. The parsing process requires about 6 milliseconds per ATN link transition. Hence for a one-noun sentence, about 120 milliseconds are required for parsing. For the two-noun sentences, the figure is about 220 milliseconds. Hence, the total contribution of parsing to the total reading time is small, which could be expected since it is a very overlearned skill.

In terms of predicting reading times, the process of building memory representations is the most important process represented by the simulation. Note that each HAM structure link actually consists of two one-way links, the total number of which is represented by LINKS. The estimated coefficient for LINKS is about 33 milliseconds per link. This gives a total of 297 milliseconds of representation building time for a completely new one-noun sentence, and about 594 milliseconds for a completely new two-noun sentence. The figure would be less for sentences with one or more given referents. The process of memorizing the sentence representations for later recall requires an additional 40 milliseconds per link of structure. Hence, a completely new one-noun sentence would require 360 milliseconds to encode for recall, and a completely new two-noun

sentence would require 760 milliseconds.

It was hypothesized that comprehending an incoherent passage requires the maintenance of topic pointers in memory, one pointer for each unintegrated portion of input material. The coefficient for TOPICS shows that it takes an additional 51 milliseconds per topic to process each sentence. It was also hypothesized that these topic pointers had to be maintained by rehearsal in short-term memory. The preliminary simulation analysis reported in Kieras (Note 2) supported this hypothesis rather well. However, in the present analysis, TOPICS, the index of this hypothesized load, failed to account for the reading times that showed definite memory load effects in the Recall Task.

Thus the status of the topic pointer memory load hypothesis is questionable. If topic pointers must be maintained by rehearsal, apparently only the Recall Task requires such rehearsal, even though at least the Topic Choice Task would seem to require keeping track of the topics. However, the Recall Task subjects were clearly dealing with the high-memory-load incoherent portions of the passages in a way that was rather different from the subjects in the Topic Choice Task and the Free Reading Task. Perhaps the hypothesized topic pointers do not undergo rehearsal unless the material is being encoded for later recall. Another possibility is that the effects represented by TOPICS are different from the rehearsal-related effects appearing in the Recall Task. The effect of additional topic pointers is to add competing sources of activation in the

memory search; if so, then the magnitude of the effect is compatible with the size of the fan-out effects reported by Anderson(1976). Thus, rehearsal of unintegrated sentence content and maintenance of topics may be separate processes.

The memory search variables, CYCLES and TAGGED, did not contribute any important portion of the total reading time in these data. There are several possible reasons for this. One is that the manipulation of values of these variables was achieved mainly by means that also changed TOPICS. Hence, TOPICS might have masked effects of CYCLES and TAGGED. A more substantive possibility is that the simulation misrepresents the process of resolving definite references in these sentences. That is, the sentences contained only a rather simple form of reference. Each referent was always referred to with the same simple noun phrase, such as The ants... Such references might be processed in a very efficient fashion, without the complications of the spreading activation memory search as described in Kieras(1977). Modelling reading times for sentences using more complicated references might resolve the problem. Where TAGGED was important, in the Free Reading and Topic Choice Single Task Fits (see Table 9), it carried a coefficient of 2-3 milliseconds per node activated. The implied effect is in the right direction: As more nodes must be activated, it takes longer to find each sentence referent. Demonstrating conclusive fan-out effects in prose comprehension, and modelling them with a simulation model, is a task for more specific experiments.

Finally, the different reading tasks apparently require some additional task-specific processes that take a constant amount of time for each sentence, and so do not vary depending on the processing requirements on each sentence. Whether this is a reasonable hypothesis, and characterizing the possible nature of these processes, could be determined by explicit modelling in a simulation.

CONCLUSION

The above discussion is a detailed analysis of a set of reading time data based on a simulation model that incorporates many currently popular ideas about cognitive processes, such as ATN parsers, semantic memory networks, and spreading activation memory searches. Hence, it was possible to apply some of our most powerful theoretical ideas to one of our most interesting problems, that of reading, with the result being that a complex set of data could be analyzed and interpreted with the aid of a detailed process model. The success of this venture was indicated not just quantitatively, in which five variables (see Table 10) accounted for a reasonably high proportion of the variance among 198 data points, but also qualitatively, by being able to point out specific sentences in the passages whose reading times, as interpreted with the model, revealed differences in the reading tasks, and where the model was accurate and where it failed. Thus, this initial trial of the simulation model and the multiple-regression method of simulation analysis has been successful. But certainly the

ultimate value of such simulation models and this analysis method will be determined by further attempts to apply them to more varied and interesting reading situations.

Reference Notes

1. Kieras, D. E. Effects of paragraph sentence order on apparent theme. Paper presented at the Midwestern Psychological Association Meetings, Chicago, May, 1976.
2. Kieras, D. E. Comparing a simulation to data using multiple regression. Presented at the Tenth Annual Mathematical Psychology Meetings, University of California, Los Angeles, August, 1977.
3. Kieras, D. E. The "New Mathematics" of knowledge representation in cognitive psychology. Paper presented in the symposium "Frontiers of Behavioral Mathematics" at the Annual Meeting of the American Association for the Advancement of Science, Houston, January, 1979.
4. Kieras, D. E. How readers identify topics in technical prose. Presented at the Psychonomic Society Meetings, San Antonio, November, 1978.

References

- Aaronson, D. Performance theories for sentence coding: Some qualitative evidence. Journal of Experimental Psychology: Human Perception and Performance, 1972, 2, 42-55.
- Aaronson, D., & Scarborough, H. S. Performance theories for sentence coding: Some quantitative evidence. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 56-70.
- Anderson, J. R. FRAN: A simulation model of free recall. In G. H. Bower (Ed.), The psychology of learning and motivation. Vol. 5. New York: Academic Press, 1972
- Anderson, J. R. Computer simulation of a language-acquisition system. In R. L. Solso (Ed.), Information processing and cognition: The Loyola Symposium. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975.
- Anderson, J. R. Language, Memory, and Thought. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976.
- Anderson, J. R. Arguments concerning representations for mental imagery. Psychological Review, 1978, 85, 249-277.
- Anderson, J. R., & Bower, G. H. Human associative memory. Washington, D. C.: Winston, 1973.
- Bach, E. Nouns and noun phrases. In E. Bach and R. T. Harms (Eds.) Universals in linguistic theory. New York: Holt, Rinehart, & Winston, 1968.
- Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.
- Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.
- Chase, W. G. & Clark, H. H. Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.) Cognition in learning and memory. New York: Wiley, 1972.
- Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975.

- Collins, A. M., & Loftus, E. F. A spreading activation theory of semantic processing. Psychological Review, 1975, 82, 407-428.
- Haberlandt, K., & Bingham, G. Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 419-426.
- Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 512-521.
- Hintzman, D. L. Explorations with a discrimination net model for paired-associate learning. Journal of Mathematical Psychology, 1968, 5, 123-162.
- Kieras, D. E. Problems of reference in text comprehension. In P. Carpenter and M. Just (Eds.), Cognitive processes in comprehension. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.
- Kieras, D. E. Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recall. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 13-28.
- Kieras, D. E. Doing it the vendor's way: Running multiple subjects in reading experiments using Data General's Diskette Operating System. Behavior Research Methods and Instrumentation, in press.
- Kintsch, W. The representation of meaning in memory. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.
- Kintsch, W. On recalling stories. In P. Carpenter & M. Just (Eds.), Cognitive processes in comprehension. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.
- Kintsch, W., & Keenan, J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.
- Kintsch, W., Kozminsky, E., Strebby, W. J., McKoon, G., & Keenan, J. M. Comprehension and recall of text as a function of a content variable. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 196-214.
- Mynatt, B. T., & Smith, K. H. Constructive processes in linear order problems revealed by sentence study times. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 357-374.
- Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.
- Quillian, M. R. Semantic Memory. In M. Minsky (Ed.) Semantic information processing. Cambridge: M.I.T., 1968.
- Quillian, M. R. The teachable language comprehender. Communications of the ACM, 1969, 12, 459-476.
- Rumelhart, D. E. Introduction to human information processing. New York: Wiley, 1977.
- Schank, R. C. Computer understanding of natural language. Behavior Research Methods and Instrumentation, 1978, 10, 132-138.
- Woods, W. A. Transition network grammars for natural language analysis. Communications of the ACM, 1970, 13, 591-606.