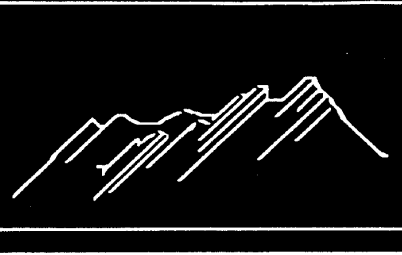


INSTITUTE OF COGNITIVE SCIENCE



Technical Report

University of Colorado, Boulder

The Role of Semantic Relations in Text-Sentence Recognition

by

Thomas Guthke
University of Colorado

Reinhard Beyer
Humboldt-Universität zu Berlin

Institute of Cognitive Science
University of Colorado
Boulder, Colorado 80309-0344

ICS Technical Report #95-03

The role of semantic relations in text-sentence recognition

Thomas Guthke	Reinhard Beyer
University of Colorado	Humboldt-Universität zu Berlin

running head: SEMANTIC RELATIONS

Please address correspondence to:

Until August 15, 1995:	Thomas Guthke Institute of Cognitive Science Campus Box 344 University of Colorado Boulder, CO 80309
After August 16, 1995:	Thomas Guthke Max-Planck-Institut für neuropsychologische Forschung - Neurologische Abteilung - Inselstr. 22-26 04103 Leipzig Germany

Abstract

The article presents an analysis of interactive processes during text comprehension. In the main experiment we used a sentence recognition task (Kintsch, Welsch, Schmalhofer, & Zimny, 1990) in which subjects read texts and afterwards decided whether or not a test sentence had been part of the text. We varied the type of the test sentence (old, paraphrase, inference) and the semantic relation between text and test sentence (e.g. actor, instrument, purpose).

The results showed an effect of both conditions. The relation effect suggests differences in the availability of knowledge during processing, specifically, a distinction between event defining and event referencing semantic relations. To pursue this question, we conducted several computer simulations with different assumptions about the influence and structure of the background knowledge using the construction- integration model of Kintsch (1988, 1992).

The results are discussed with regard to theories about the on-line activation of inferences (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992).

The role of semantic relations in
text-sentence recognition

Reading comprehension is the product of interacting characteristics of text and reader. The effects of readability factors, such as vocabulary, sentence structure, text organization, and coherence have been well documented in the literature, as have various reader characteristics, such as differences in reading ability, in prior knowledge, goals, motivation, interest, and so on. However, research seeking to understand the interaction of these factors is only in the beginning stages. Especially lacking are studies exploring such interactions during on-line processing. In the present study we are concerned with how the reader's prior knowledge and text information becomes integrated into an internal memory representation during comprehension.

We have decided to approach this problem methodologically by investigating integration processes in the context of inferences which occur during on-line comprehension. We assume that inferences drawn during text comprehension will be a good indicator of the type and structure of knowledge that the reader accesses during comprehension. Text comprehension is not possible without drawing inferences. First, inferences are necessary to provide missing links between concepts expressed in the surface structure of the text and thus enable the reader to establish within-text coherence. Second, inferences can establish connections between the content of the text and prior knowledge structures of the reader and therefore activate information about events, facts, and topics that are not mentioned in the text.

These two functions of inferences are generally accepted in the literature (Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992; van den Broek, 1994). On the other hand, there is some controversy about the processing characteristics associated with inferences. Some authors question whether the methods used in the inferencing research induce inferences that would not be otherwise drawn, or whether they lead to wrong conclusions about the time-course of such comprehension processes (Graesser et al., 1994; Guthke & Beyer, 1992; Kintsch, 1993; McKoon & Ratcliff, 1992, van den Broek, 1994).

Our work has also been concerned with establishing an experimentally induced taxonomy of inferences (Guthke & Beyer, 1992, Kintsch, 1993). We think that these questions can only be resolved on the basis of formalized assumptions about the representation of knowledge and its influence on processes of text comprehension (Klix, 1990, 1992; van der Meer, 1987).

In the present study we will first discuss previous studies of test word recognition after reading a sentence. We found an effect of the type of semantic relation between sentence and test word. In the present experiment we used a sentence recognition task in which subjects had to differentiate between old and several types of new test sentences (paraphrases, inferences). The results suggest differences in the availability of knowledge during processing, presumably the result of differences in the amount of cognitive effort needed to activate specific knowledge elements. In an effort to give a precise theoretical account of these data, we apply the construction - integration computer simulation model (Kintsch, 1988, 1992; Kintsch & Welsch, 1991; Mross & Roberts, 1992) to this experiment.

Effects of semantic relations in sentence-word recognition

In our previous studies (Beyer, Artz, & Guthke, 1990; Guthke & Beyer, 1992) we used a sentence-word recognition task for the on-line investigation of inference processes (Mansfield, 1977; McKoon & Ratcliff, 1986). Sentences were presented on a screen to 24 subjects. Each sentence was followed by a test word. Subjects were required to decide whether or not the test word had been part of the sentence. Each subject had to process 60 positive sentence/test word pairs (correct answer: "Yes") and 60 negative sentence/test word pairs (correct answer: "No"). For the negative items we varied the existence of a semantic relation between sentence and test word. In Condition A there is a semantic relation (Sentence: Anne is cleaning the room, Test Word: broom), and in Condition B there is no semantic relation (Sentence: Mary is reading the book, Test Word: water). We compared the rejection time for the same test word

between the semantic relation condition and the control condition (Figure 1).

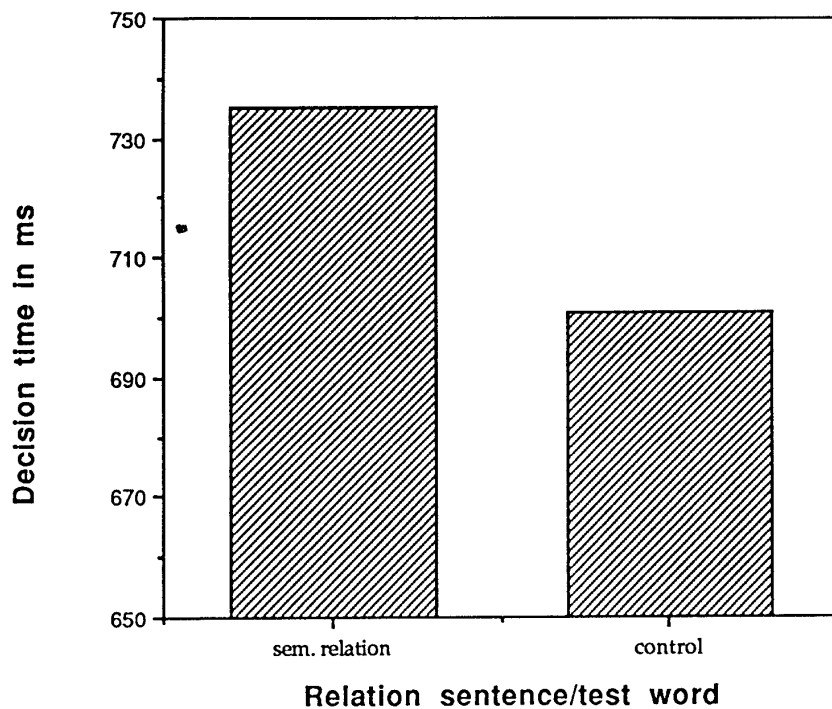


Figure 1. Decision time for rejection of test words for Condition A (semantic relation) and Condition B (control).

We observed significantly longer reaction times when a semantic relation exists between sentence and test word. This delay in the rejection times is due to interference with information activated during the comprehension of the sentence. During the presentation of the sentence activation of prior knowledge takes place. Therefore, since in Condition A the test word is compatible with the activated knowledge, more cognitive

effort is required to reject the test word which had not occurred in the presented sentences.

Why this should be the case can be explained within the theoretical framework of the construction - integration (CI) model of Kintsch (1988, 1992). Accordingly, during the construction phase associations in the semantic net arising from the explicit text contents are activated. Unfortunately, there are no statements which would allow us to differentiate activation strengths of the activated prior knowledge. A critical point in the model is the assumption of equal activation of all associations. A consequence is the assignment of equal connection strengths to all text content/association pairs. An analysis of the word material used by Kintsch (1988) shows that it is possible to differentiate between associations with respect to the type of semantic relation between text content and association (e.g., TO WEED - IN GARDEN: location vs. TO WEED-TO HELP: purpose relation).

In a second phase of our previous experiment (Beyer et al., 1990; Guthke & Beyer, 1992) we varied the type of relationship between sentence and test word (Actor, Patient, Instrument, Purpose, Causal). Figure 2 displays the differences in reaction time ("No" response) for Condition A (semantic relation) and B (control) for the different semantic relations between sentence and test word.

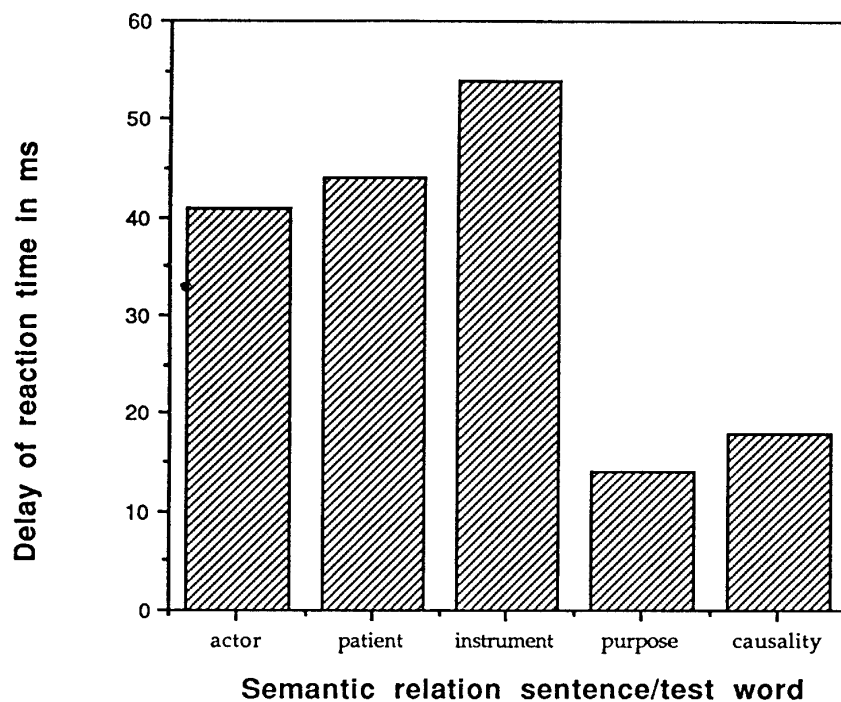


Figure 2. Difference between conditions (A= semantic relation, B=control) in rejection times as a function of the type of semantic relation.

The manipulation of the type of relationship between sentence and test word resulted in a significant modification of the general delay effect, in that the delay of rejection depended on the type of semantic relation between sentence and test word. This result can be explained in terms of differences in relational complexity among concepts. Depending on the degree of relational complexity many or few additional concepts must be activated in order to establish a semantic relationship between two concepts if only a part of the whole structure is given. Various

investigations show, for instance, that an actor relation is accepted as a sensible fact by itself, whereas the decision to accept as reasonable a purpose relation requires the activation of at least two further relations (e.g. an actor and an object (Klix, van der Meer, & Preuß, 1984; Küchler & Preuß 1992, Schmieschek, 1988, van der Meer, 1987, 1994). Therefore, the activation of a complex relation requires greater cognitive effort and presupposes the activation of other implicit knowledge, for example, actor and instrument information. Thus, the probability of activation of implicit actor or instrument information is higher than, say, of purpose information, and this explains why the amount of delay for these relations is lower.

This on-line effect of activation of prior knowledge during sentence processing and the semantic relation effect proved to be stable in various replication experiments using different test materials, different groups of subject, and various paradigms (Beyer et al., 1990; Beyer, Guthke, & Ankert, 1994; Beyer, Guthke, & Thiele, 1995; Guthke & Beyer, 1992).

Our results do not agree with the minimalist hypothesis which McKoon and Ratcliff (1992) proposed. They assume that the only inferences encoded automatically during reading are those necessary to establishing local coherence or which are based on quickly and easily available information. Easily available information is assumed to come from explicit information in the text or from well-known background knowledge. The inferences which we investigated do not fit these categories. The minimalist view predicts that inferences such as instrument inferences are not automatically encoded during the processing of the text. To resolve this contradiction we need to think about processes occurring at the level of the situation model. Assumptions about the situation model can help us determine what information is easily accessible and would therefore be elaborated during processing of the text. We think that our results regarding a differentiation in processing characteristics of inferences can provide a way to approach this question. A critical point will be to test if our results obtained with a simple sentence-word recognition paradigm (Beyer et al., 1990; Guthke & Beyer, 1992) are also valid for comprehending longer texts. For this reason we

have chosen a test-sentence recognition task (Zimny, 1987; Kintsch, Welsch, Schmalhofer, & Zimny, 1990) in order to examine the role of differential processing characteristics of inferences during text comprehension.

Effects of semantic relations in test-sentence recognition

Zimny (1987; Kintsch et al., 1990) have provided evidence of different levels of text representation. Their sentence recognition task is also appropriate to test the influence of prior knowledge during text comprehension. In her experiment (Zimny, 1987) subjects had to read texts about 200 words long and afterwards to decide whether or not a test sentence had been part of the text. Test sentences included an old verbatim test sentence, a paraphrase (minimal word order change or single word change or active/passive change), and an inference sentence. The inference sentence was a sentence which was not mentioned in the text but was a part of the action sequence described by the text. Also two novel sentences were presented. One sentence was contextually appropriate and one test sentence was unrelated to the theme and the content of the text. Their results showed a main effect of sentence type. This effect can be explained with regard to different levels of text representation. For example, the rank position of the acceptance rate of inference statements among the various test sentence types is predicted correctly by the hypothesis that inference statements differ from the text at the surface and propositional level but that they are part of the same situation model.

The relative high percentage of wrong "Yes" responses to the inferences showed the strong influence of prior knowledge. The subjects used their prior knowledge to construct the situation model which also included actions not explicitly mentioned in the text. Immediately after reading the text the representation of the surface features of the text is still very strong, but the data showed an increase in "Yes" responses to inference sentences following delay of the recognition task. This finding demonstrates the appropriateness of this paradigm for investigating the influence of prior knowledge.

Theoretical considerations

In our experiment, like Zimny, we first varied the test sentence type, but in addition, we varied the type of semantic relation between text and test sentence. The text was presented on a computer screen. After the reading of the text the subjects had to decide as quickly and accurately as possible whether or not they had read that exact same sentence earlier while reading the text.

Before describing details concerning the method we would like to present in more depth some theoretical background concerning our choice of independent variables. Our manipulations consisted in varying the test sentence type (verbatim old sentence, paraphrase, inference, appropriate and inappropriate new test sentence) and the semantic relations between text and test sentence (low and high complexity).

Test sentence type effect. We expect to replicate the sentence type effect obtained by Zimny (1987) with our German texts. Our hypothesis is based on the assumption of van Dijk & Kintsch (1983) regarding different levels of text representation. They distinguish between three levels of text representation, the linguistic surface structure, the propositional textbase, and the situation model (also Kintsch, 1988; Schmalhofer & Glavanov, 1986; Fletcher and Chrysler, 1990). These different levels of representation are derived from different sources of information.

The linguistic surface structure of a text contains the text words and syntactic structure. The propositional representation is derived from the semantic content and organization of the text, and the situation model is derived from the reader's knowledge about situational relations in the world which are not expressed directly in the text itself.

We expect the proportion of "Yes" answers to be dependent on the type of test sentences. In the case of an old test sentence the correct response is "Yes". For all other test sentences the correct answer is "No". Nevertheless, we expect a high percentage of "Yes"-answers for paraphrases, followed by inferences, appropriate, and inappropriate test sentences.

Old verbatim sentences are represented at all three levels of representation. Paraphrases differ in terms of the surface structure from the text sentence, but not at the textbase and situation model level. Inference statements differ both in terms of their surface structure and propositional content, but they are part of the same situation model. The appropriate test sentences are only vaguely related to the context and differ at all three text levels. There is no relationship between inappropriate test sentences and the text. The inference statements are of particular interest with respect to the influence of prior knowledge. An acceptance of inference statements is only predictable if prior knowledge is included in the text representation. Inference statements are not mentioned in the text, and the correct answer is "No".

Semantic relation effect. The second variation in our experiment is the type of semantic relationship between text and test sentence. We found in our previous studies (Beyer et al., 1990; Guthke & Beyer, 1992) that the degree of availability of knowledge is dependent on the semantic relationship between sentence and prior knowledge. Our hypothesis about the influence of prior knowledge is derived from models of knowledge representation. Knowledge important for text comprehension can be described as an event script (Nelson, 1986) or an event concept (Klix, 1990, 1992; van der Meer, 1987).

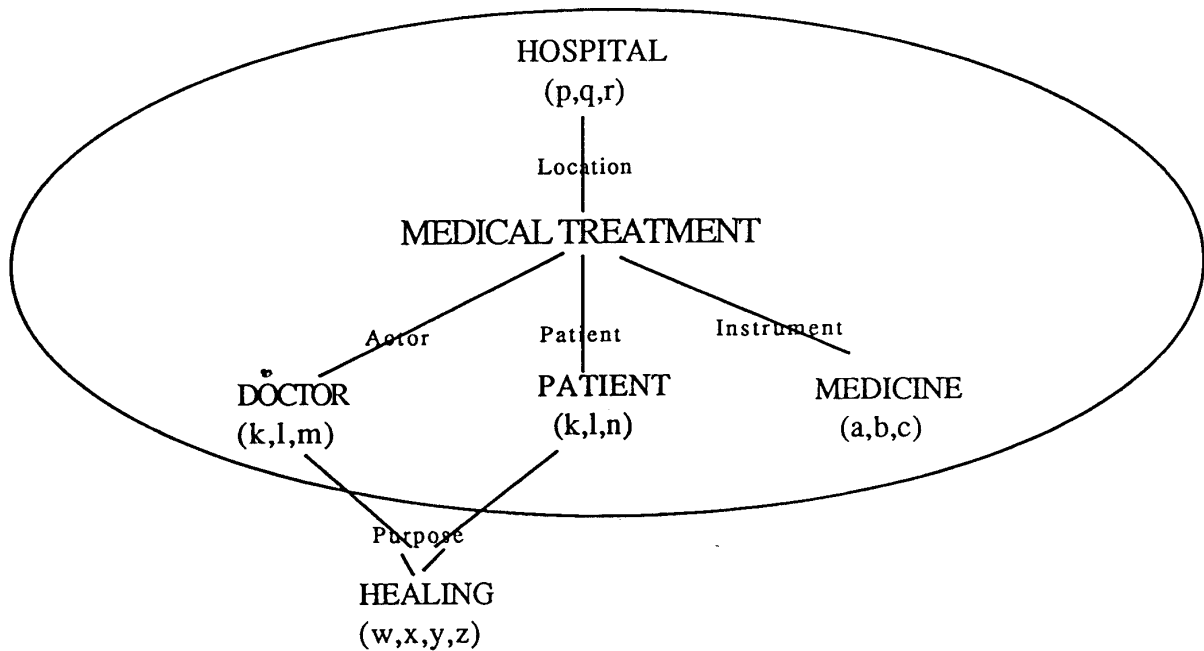


Figure 3. Structure of an event concept (after Klix & van der Meer, 1990).

Figure 3 contains a schematic description of an example of an event concept (medical treatment). Klix and his colleagues at the Humboldt University at Berlin assume that concepts constituting event related knowledge enter into definite relations and are fixed in typical structures. The central concept of this structure is the semantic core (TREATMENT). From this semantic core different relations proceed outwards to other concepts. Examples of the relations are actor, recipient, instrument, purpose, and location. It is important to note that this structure is a structure of concepts, not of words. The letters symbolize the feature sets of the concepts. Thus, "to treat a prisoner" activates a different set of concepts than "to treat a patient".

Experiments which were described previously have shown that the cognitive effort involved in establishing such relations depends on relational complexity (Klix et al., 1984; Küchler & Preuß 1992, Schmieschek, 1988, van der Meer, 1987). These theoretical assumptions and empirical results, together with our findings with the sentence-word recognition task, support the idea of differentiating between semantic relations of low and high complexity.

An additional feature important for distinguishing these semantic relations concerns differences which have been observed in the function of these types of semantic relations (Ankert & Beyer, 1987; Guthke & Beyer, 1992). Relations with low complexity are important for the internal structure of an event representation similar to the way in which a set of features characterizes an object concept. In contrast, relations such as purpose and cause are important for connecting concepts in a chain of events. To illustrate these arguments we refer to the relations with low complexity (actor, recipient, and instrument) as event defining and the relations with high complexity (cause or purpose) as event referencing.

These considerations can then be used to further specify the semantic relationships between the text and the test sentences. Thus, in our experiment we distinguish between two types of semantic relations (event defining vs. event referencing) for each type of test sentence (old, paraphrase, inference).

In general, we expect to find a higher acceptance rate for test sentences with a semantic relation of low complexity. For the inferences, we assume that while reading about a part of an event, other parts of this representation are also activated, especially knowledge which is connected through semantic, event defining relations.

Method

Materials. We constructed seven texts with different content (1 training text, 6 experimental texts). Table 1 shows an example of a text and a set of test sentences.

Anja is a pretty young girl.
 She is concerned about her professional training.
 She is excited about many possible options.
 That is why the decision is very difficult for her.
 So far she likes one possibility the best.
 She would like to work at a hospital.
 Many hard working doctors work there.
 The sick people are treated with great care.
 Actually this is an interesting job.
 Only the irregular working hours would bother her.
 Certainly you often have to get up very early.
 Getting up early ruins the whole day for her.
 Anja is a real lazy bones.
 Thus there are various pros and cons.
 Perhaps she will discuss it again with her parents.

test sentences

sem. relation

1.) old sentences

Many hard working doctors work there.
 That is why the decision is very difficult for her.

actor

2.) paraphrases

The patients are treated with great care.
 Anja is a real sleepy head.

patient

3.) inferences

They also have good equipment there.
 This benefits the healing.

instrument
purpose

4.) new appropriate

She is interested in the new opportunity.

5.) new inappropriate

He reads the book for a few months.

Table1. Experimental text and the critical test sentences. Actor, patient, and instrument are event defining relations and purpose is an event referencing relation

(The original text and test sentences are in German. The length of the test sentences in German is equal.)

Each text consisted of an introduction and conclusion surrounding a description of an event (concert, medical treatment, shopping, lecture, strike, sport, crime case). The texts were about 110 words long. The sentences, describing the event, contained statements about the semantic relations (actor, recipient, instrument, purpose). Fifteen test sentences accompanied each text, 8 of which served our experimental manipulation of the variables. Thus, there were 2 old verbatim test sentences, 2 paraphrases, and 2 inference sentences in order to vary the semantic relation type per test sentence type. In addition, there was one appropriate new test sentence and one inappropriate new test sentence. Seven additional test sentences were necessary to balance the number of old (correct answer is "Yes") and new test sentences (correct answer is "No"). In order to control for word type, formulations, and sentence length effects we constructed a complete permutation of the test sentences. Overall, these adjustments resulted in 12 different text variants per event.

Procedure, subjects. Each subject was asked to read the training text and six experimental texts of the same type. Each text was presented a page at a time on a computer screen and subjects were allowed to read at their own pace. Following each text the 15 test sentences were presented sequentially. The subjects were asked to decide as quickly and accurately as possible whether or not that exact same sentence had occurred in the text. We recorded the subjects' reading times of the texts, their reaction times and their responses. The texts and the test sentences were presented in random order. The experiment took between 35 to 40 minutes to complete. Our subjects were 24 undergraduate students from the Humboldt-University in Berlin between 20 and 30 years of age.

Results

16

Test sentence type. Only the results for the "Yes" responses are described because these are the data used in our simulations. The reaction time data are discussed in another article (Beyer, Guthke, & Pekrul, in press).

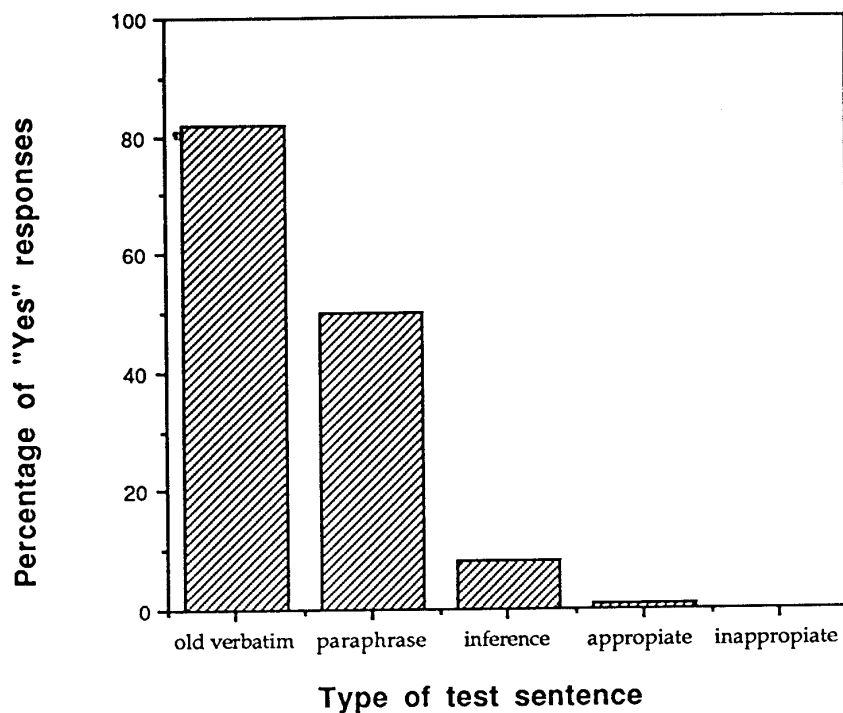


Figure 4. Percentage of "Yes" responses as a function of type of test sentence.

Figure 4 shows the percent of "Yes" responses subjects gave to the different types of test sentences. There is a higher proportion of "Yes"-responses for old sentences, followed by paraphrases, inferences, and appropriate test sentences. The effect of the test sentence type was significant statistically ($F(4,92) = 559.54$; $p < .01$). The results are in good

agreement with the theoretical assumptions of van Dijk & Kintsch (1983) regarding three levels of text representation as well as Zimny's results (1987). In her experiment using English materials subjects responded "Yes" to old verbatim test sentences more often than to paraphrase sentences (73% vs. 52%, respectively) and to paraphrase sentences more often than to inference sentences (21%). The percentage of "Yes"-responses for appropriate new test sentences was approximately two percent and for inappropriate new test sentences one percent. The higher percentage of "Yes" responses for inferences in the Zimny experiment compared with our result (21% vs. 8%, respectively) can be explained by the overall higher error rate in the Zimny experiment and by the fact that she used longer texts (200 words vs. 110 words here). More important is the different type of inferences used in the two experiments: inferences necessary to establish coherence in Zimny whereas we investigated elaborative inferences.

Semantic relation. The second hypothesis concerned whether the type of semantic relation between text and test sentence influences the probability of accepting the sentence as part of the text. These results are shown in Figure 5.

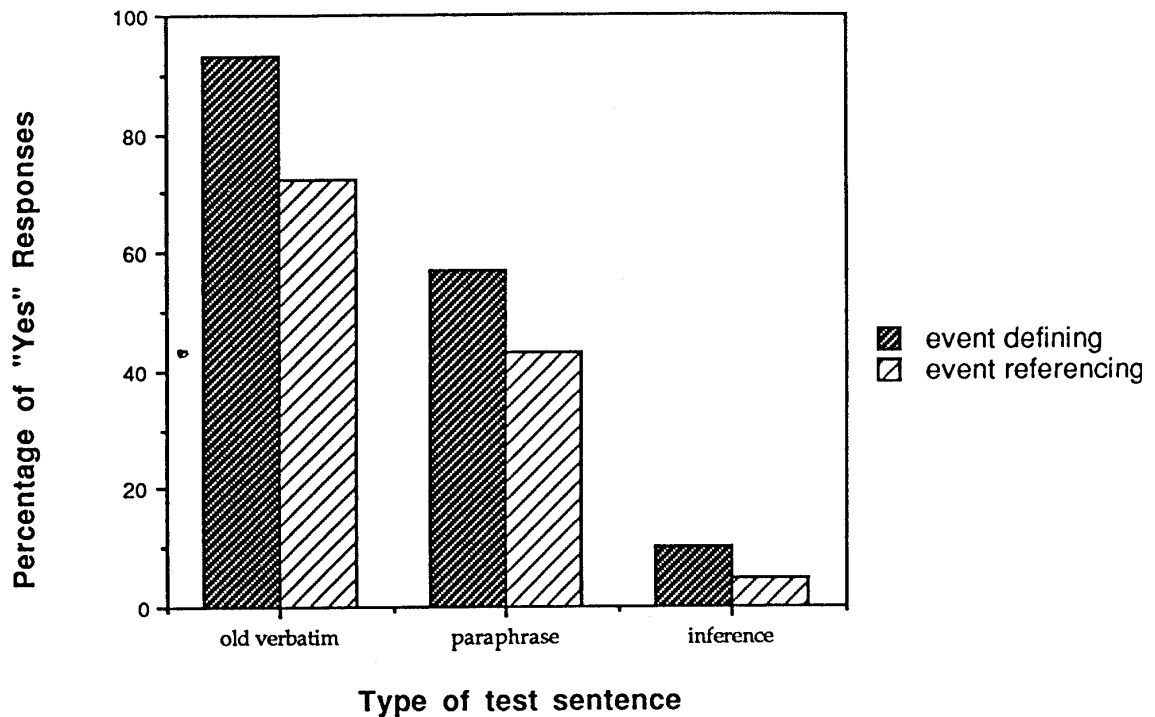


Figure 5. Percentage of "Yes" responses as function of type of test sentence and type of semantic relation between text and test sentence.

As predicted, semantic relations did influence the percentage of "Yes" responses to the test sentences. Overall, the percentage of "Yes" responses is higher for actor, recipient, and instrument relations (event defining) than for purpose relation (event referencing) $F(4,92) = 10.17$; $p < .05$. That is, if the relation between text and test sentence is established by means of an event defining relation, then there is a greater probability that the subject will decide that the test sentence occurred in the text. The wrong acceptance of inference statements reflects the integration of event related knowledge into the text representation. The probability of accepting event

defining inferences is higher than for event referencing inferences ($t(92) = 1.78$; $p < .1$). There is obviously a higher probability that event defining knowledge will be integrated into the text representation.

Discussion

Multi level representation of text. Our data support the idea that at least three levels of text representation are generated during the comprehension of the text. The differences in the probability of the "Yes" responses corresponded to the theoretical assumptions of three levels of text representation since they contribute differentially to specific test item types. Following van Dijk & Kintsch (1983) these levels can be described as the surface level, propositional level, and situation model. The situation model is of particular interest for the investigation of inference processes. The situation model contains information which is not explicitly mentioned in the text, which explains why, both here and in Zimny's experiment, inference statements were wrongly accepted as verbatim parts of the text. One problem with the recognition task is that the subjects are forced to use a strategy of looking for the exact, verbatim impression of the text. Our subjects were able to retain the surface level of the text quite well, which is reflected in the low rate of errors. With this strategy the influence of prior knowledge is limited, resulting in the low rate with which inference statements were accepted as verbatim parts of the text.

On-line activation of inferences. In other experiments, in which we made it more difficult for the subjects to retain the surface level information, we found an increased rate of inference acceptance (Beyer et al., in press). For example, the critical manipulation in one experiment was achieved by modification a single word. In the case of our present experiment we used a test sentence, and all the additional words in the test sentence could be used to determine if the test sentence was part of the text. For this reason, in a subsequent experiment we used only the critical word as a test word instead of the whole sentence. For example, instead of the test sentence: "The patients are treated with great care" we only presented the test word: "patients" for the recognition task. With this

paradigm we were able to replicate the test sentence effect ($F(4,92) = 113.7$; $p < .01$) as well as the relation effect ($F(4,92) = 8.79$; $P < .05$).

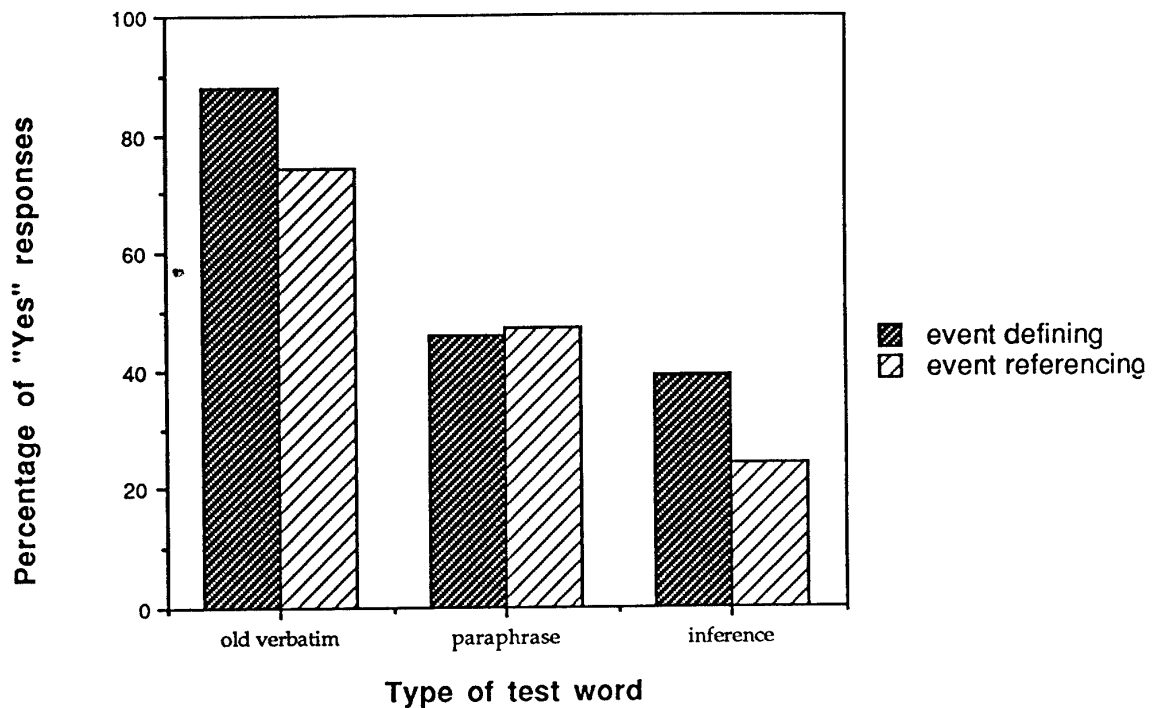


Figure 6. Percentage of "Yes" responses as function of type of test word and type of semantic relation between text and test word.

Additionally, as shown in Figure 6, we observed a much higher acceptance of inference words, since without the additional information it is much harder for the subjects to reject the inference. Here, the difference between event defining and event referencing relations is very strong.

In general, our results support the idea that event related knowledge is integrated into the text representation constructed during reading. Moreover, they shed light on the question about which aspects of a situation are more easily available. We found that event defining relations

(e.g. actor, instrument, patient) are more likely to be activated than event referencing relations (e.g. purpose). This is support for the idea that such information is automatically available during reading and thus our findings argue against the minimalist view of McKoon & Ratcliff (1992).

Instrument inferences. We would like to focus on instrument inferences as an example in discussing our results regarding characteristics of the inference process. Instrument inferences in our experiment can be classified as forward elaborative inferences. McKoon and Ratcliff (1992) assume that instrumental inferences are activated during reading only if the subjects are engaged in special strategies or if the availability of such knowledge is increased by mentioning the instrument explicitly. Neither is the case in our experiment. The subjects had no particular reason to generate these inferences because this would increase the probability of errors. Apparently the semantic constraints alone are sufficient to generate these inferences (Graesser et. al. 1994; McKoon & Ratcliff, 1989; van den Broek, 1994). We have derived this prediction from assumptions about event related knowledge representation (Klix 1990; van der Meer, 1987) and would claim that such semantic information is included in a mental representation of discourse.

To explain the variability seen in the experimental results from research on such inferences we should consider that an inference is activated to a certain degree. The degree of encoding an inference varies along a continuum from not encoded at all, to minimally encoded, to explicitly instantiated (McKoon & Ratcliff, 1989). Many authors have interpreted their experimental results regarding the generation of instrument inferences as an indication that they are not computed on-line (Corbet & Doshier, 1978; Lucas, Tanenhaus, & Carlson, 1990; Singer 1979, 1980). However, other studies provide evidence of on-line activation of instruments (Garrod & Sanford, 1981; Cotter, 1984). The first authors compared the inference condition with a condition in which the instrument was explicitly mentioned and found a difference between these conditions. They explained this difference by asserting that there was no on-line activation of the instrument when the action sentence without the

instrument was processed. However, just because the activation level of inferred information measured by using a certain method is lower than that of information explicitly mentioned in the text does not necessarily mean that the inference is not made. Instead, it may be the case that the inference is encoded in the representation, but does not have the same status in memory that the explicitly mentioned information has. By comparing the implicit condition with a control condition (no semantic relation) we found differences in reading times, decision times, recognition rates between these conditions (Beyer et al., 1990, 1994; Guthke & Beyer, 1992). These differences can be explained in terms of on-line activation of event defining knowledge. However, the degree of this activation is probably not as high as if the information were explicitly mentioned. In the last section of this paper we will describe a computational analysis using the construction-integration architecture of Kintsch (1988; 1992; Kintsch & Welsch, 1991) in which the assumptions about the nature of the situation model and the inference process are varied.

Simulation

We conducted several simulations, varying assumptions about the situation model. These simulations are based on the construction-integration model (Kintsch, 1988; 1992; Kintsch & Welsch, 1991) and use a slightly modified version of the implementation developed by Mross & Roberts (1992). We followed the main principles of the simulation conducted by Kintsch et al. (1990) with some modifications which are described later.

The Construction-Integration Model

The construction-integration (CI) theory is a general theory of comprehension which provides an architecture for the simulation of language processing. Explicit models can be formulated for specific experimental situations. It has been applied to the recall (Kintsch, 1988) and recognition (Kintsch et al., 1990) of text, and to learning from text (Kintsch, 1994) as well as to priming (Kintsch & Welsch, 1991) and question answering (Singer, 1994).

The CI theory is a further development of earlier models of text comprehension (Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). The CI theory combines features of a symbolic system and connectionist principles. In the construction phase a rule based system constructs a network representation of the text and the knowledge that has been activated, and in the integration phase a constraint satisfaction mechanism generates a consistent interpretation of the text).

The purpose of our simulations is to provide an account of the "Yes" responses in our experiment. We are mostly interested in the inference process and believe that we can capture some important features of this process by varying certain assumptions about the situation model. Specifically, we hope to distinguish between a minimalist view and a theory which assumes that event related knowledge is activated on-line.

Method

In this article we will report two simulations. First, we describe the general features of the two simulations and then the differences between them. Two important methodological features of our simulations are that they are based on the whole text and that the analysis includes all of the texts. In previous simulations the authors often chose only a part of the experimental text and an example of the material (e.g., Kintsch et al., 1990). We think that a comparison between empirical data which are averaged over the material and the simulation is more valid if the simulation is also based on all the texts used. This allows us to look at effects of the materials and avoid artifacts which arise from the selection of examples. The disadvantage is the immense increase in the amount of work.

Construction phase. For the two simulations we derived a coherence matrix and an activation vector in accordance with previous simulations conducted by Kintsch et al. (1990). The text is processed in cycles. Each cycle of processing consists of a construction phase and an integration phase. The construction is guided by weak symbolic rules which construct a coherence matrix. The coherence matrix consists of elements from the surface structure, text base, and situation model level.

At the surface level we included the words that make up the text and a sentence node. We did not derive the hierarchical structure of sentence constituents (see Kintsch et al., 1990) because their effects do not play a role in our analysis. At the propositional level we analyzed the sentences into complex propositions (main proposition plus modifiers) according to van Dijk & Kintsch (1983). For the first simulation the situation model consists of a few, quite abstract concepts which overlap entirely with the macrostructure of the text.

The links among the nodes from the different levels of representation are determined by the relationships among these elements specified in the text. In our texts there were links between the sentence node and the word nodes, the words and the corresponding propositions, between propositions (argument overlap, main proposition, and modifiers), between propositions and corresponding units of the situation model and links between elements within the situation model. We assumed an activation value and link strength of 1 following other analyses which were conducted previously (Kintsch et al., 1990; Kintsch & Welsch, 1991; Singer, 1994).

Integration phase. During the integration phase activation settles according to connectionist principles of constraint satisfaction. The integration computation involves the postmultiplication of the active part of the coherence matrix by an activation vector until the criterion (change in activation vector less than .0001) is reached. For each cycle the active elements are the sentence which is being processed and the element with the highest activation from the previous circle, as well as the already constructed nodes of the situation model. The result is a differentiation and weighting of activation values and link strengths. The final activation vector modifies the original coherence matrix to produce a long-term representation of the text.

Test Sentence Recognition. The episodic memory trace is the basis for text based behavior, in our case the recognition task. The recognition task involves connecting the test sentence with the text representation, and the

critical measure is the amount of activation which flows into the representation of the test item.

First, the test sentences are processed and then the links between the test sentence and the text representation are established. The old verbatim test sentence has connections at all three levels of text representation, and all test sentence nodes have a link to the text representation. In the case of paraphrases one link between one word node of the test sentences and the surface level of the text representation is missing. In contrast, inference statements and appropriate new test sentences have links only to the situation model. There is no link between inappropriate new test sentences and the text representation.

The starting activation value of the test sentence nodes is 0, and the link strength is 1. The next step is the integration of the text with the sentence. We clamped the activation value of the text units, resetting them to their initial value after each integration cycle. The integration was completed when the change in the activation vector was less than .0001. We then calculated the average activation value for each test sentence.

Simulation 1

The main point of the first simulations was to test the minimalist view (McKoon & Ratcliff, 1992), that the only knowledge included in a representation is what was explicitly mentioned in the text. The minimalist view does not predict on-line activation of the kind of inferences which were investigated in our experiment and which are the focus of the simulation. Thus, for this simulation the situation model is composed of few, abstract concepts, that are equivalent with the macrostructure of the text.

Results of Simulation 1. Figure 7 shows the activation value for each test sentence. The activation score is the average activation per word of the sentence. We see here the same rank order of old sentences, paraphrases, and inferences as in the empirical data.

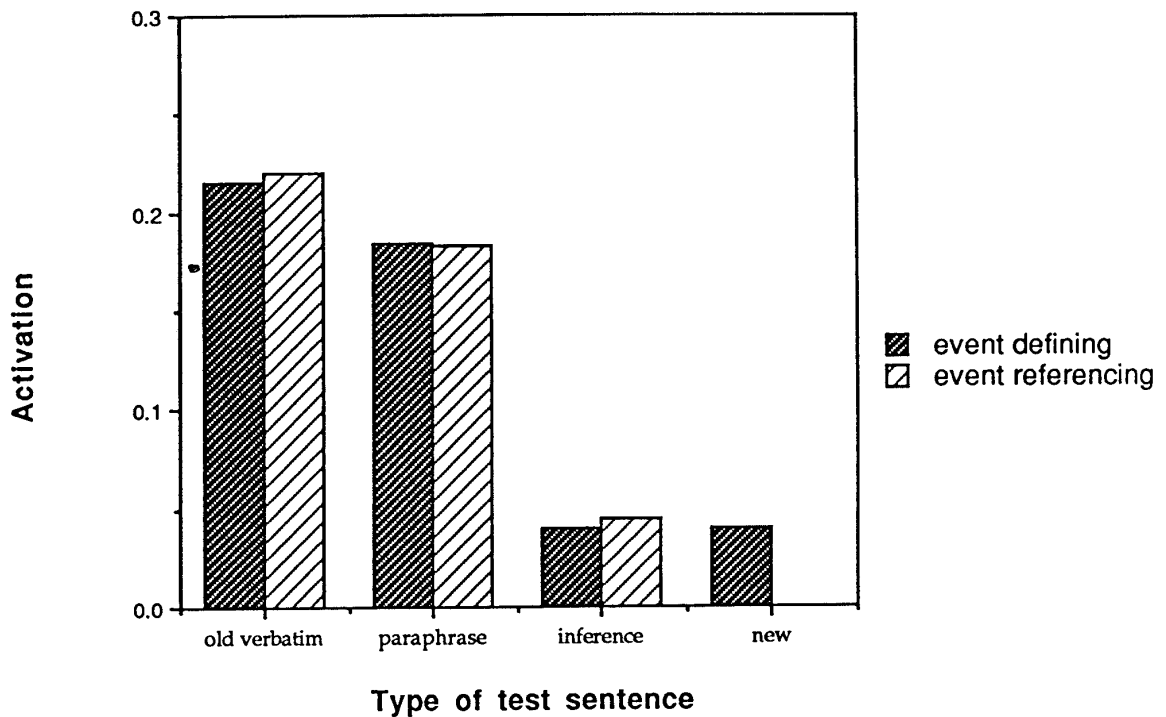


Figure 7. Activation value of the first simulation (average activation per word) as a function of sentence type and semantic relation (no variation in type of semantic relation for new appropriate and inappropriate test sentences; activation for inappropriate = 0).

The further evaluation of the simulation is based on the assumption that high activation of a test sentence promotes "Yes" responses. In Figure 8 the "Yes" responses and the activation value are plotted on a single graph. If we look at the relation between the degree of activation and the tendency of the subject to respond "Yes" then we see that there is good correspondence to the empirical data. Although the quantitative fit

between the two sets of data is quite high ($r^2 = .9$), there are also two major qualitative differences. First, there is no difference between inferences and new appropriate test sentences, and second, there is no effect of semantic relation between test sentence and text.

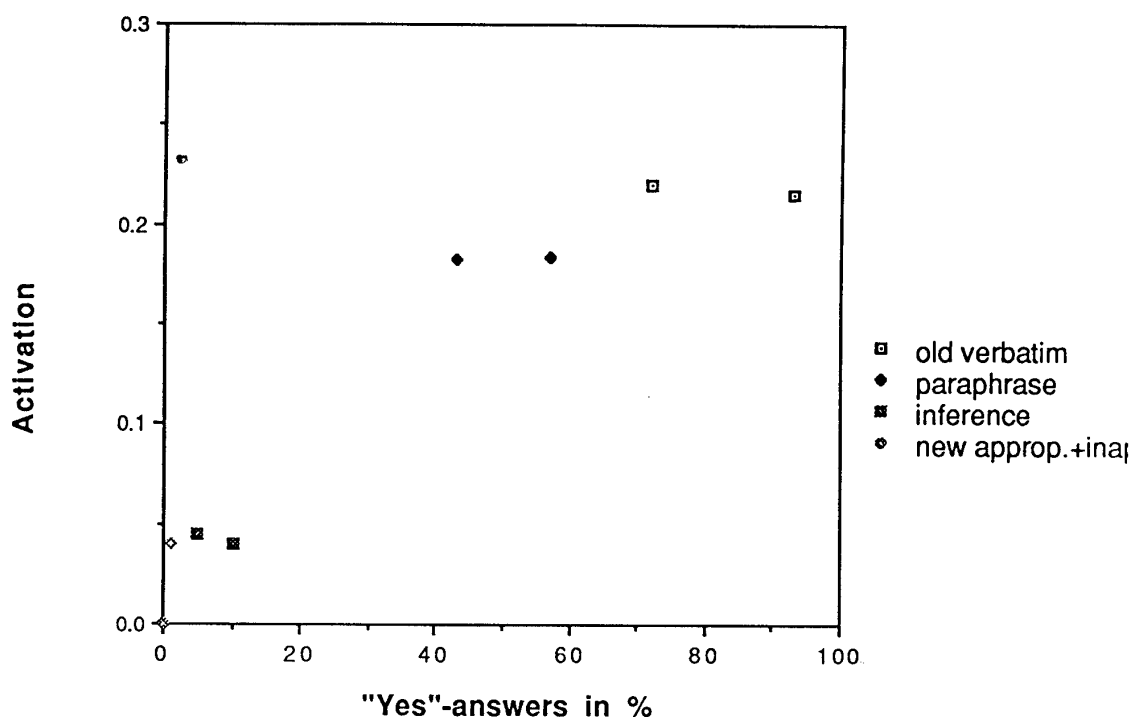


Figure 8. Comparison between empirical data and the first simulation (coefficient of determination $r^2 = .9$). For old, paraphrase, and inference sentences. first plot = event defining, second plot = event referencing).

Both inference sentences and appropriate new test sentences have only one link to the situation model, so there is no basis for differences in their activation value. If we look at the influence of the type of the

semantic relation between text and test sentence, we see that there is no difference between event defining and event referencing relations (Figure 7). Our empirical results showed a higher probability of responding "Yes" if the semantic relation between test sentence and text is an event defining relation. Thus, our first simulation does not explain the difference between inference statements and appropriate new test sentences nor the semantic relation effect. Apparently it is not enough to include only explicit knowledge and to have a situation model in the form of a macrostructure in order to explain the empirical data.

Simulation 2

To pursue these questions we conducted a further simulation, using the assumption of an event concept (Klix, 1990; van der Meer, 1987) to build the situation model. The result is a more elaborated situation model. The situation model now contains implicit knowledge, which is defined as event defining and event referencing semantic relations.

The situation model is activated if only a part of it is mentioned in the text. The semantic core has direct connections to the units which are connected through such relations as actor, instrument, and location. Complex relations are connected with the semantic core indirectly through the other units which are necessary to establish such a relation. The propositions of the text are connected with the corresponding units of the situation model. For example, if one inference statement contains information about the instrument and the other inference statement information about the purpose and neither the purpose nor the instrument are mentioned in the text, both concepts are nevertheless activated. Because of the structure of the event representation the instrument inference is more strongly activated than the purpose inference. The inference sentence has connections to the situation model (to the semantic core and the instrument or purpose unit). All other features of this simulation are the same as in the first simulation.

Results of Simulation 2. Figure 9 shows the average activation value for the test sentence as a function of type of semantic relation. We observe

now the general effect of test sentence type and additionally the effect of semantic relations corresponding to the empirical data. There is a difference between inference sentences and appropriate new sentences because the inference statement now has several connections to the situation model, whereas the new appropriate test sentence has only a single link to the text representation. The structure of the situation model determines that the activation for inferences which are realized by event defining relations be higher than inferences realized by event referencing relations. As a result, the correspondence between empirical and simulations data is now very high.

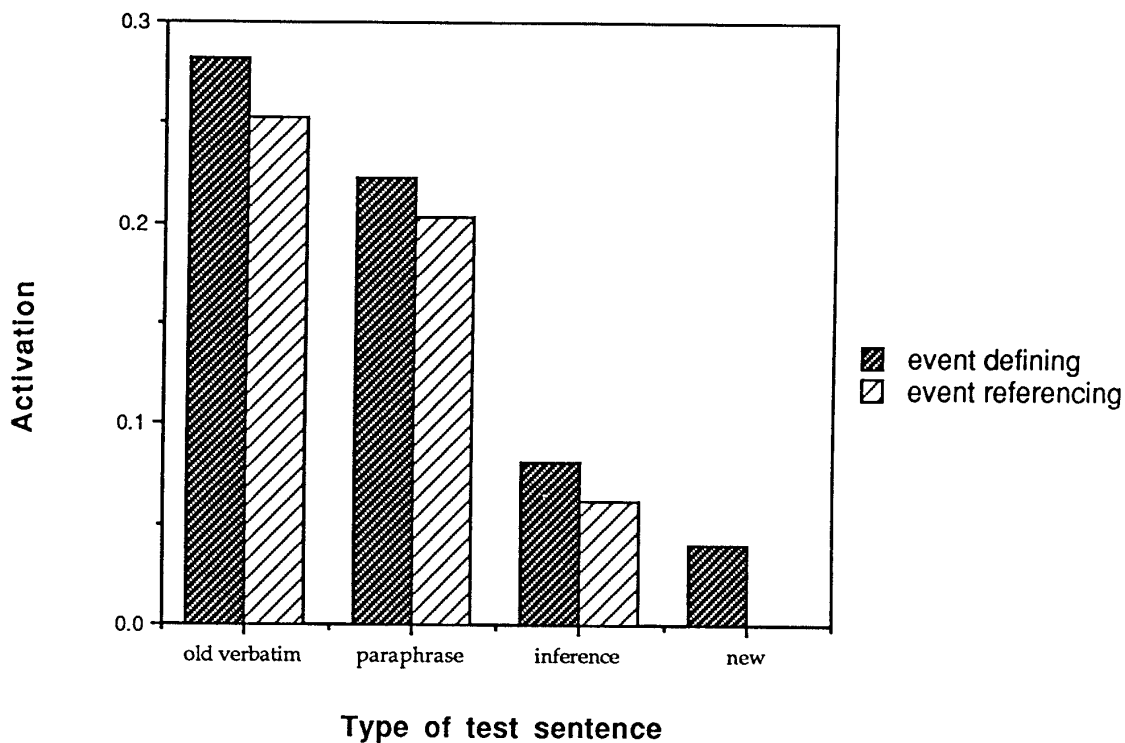


Figure 9. Activation value of the second simulation (average activation per word) as a function of sentence type and semantic relation (no variation of type of semantic relation for new appropriate and inappropriate test sentences; activation for inappropriate = 0).

In Figure 10 the "Yes"-responses and the activation values of the second simulation are plotted on a single graph. As can be seen from the comparison between observed and predicted data, the application of the CI model reproduces the experimental effects very well. The coefficient of determination between empirical and simulation data is .95. There are no qualitative differences between empirical and simulation data.

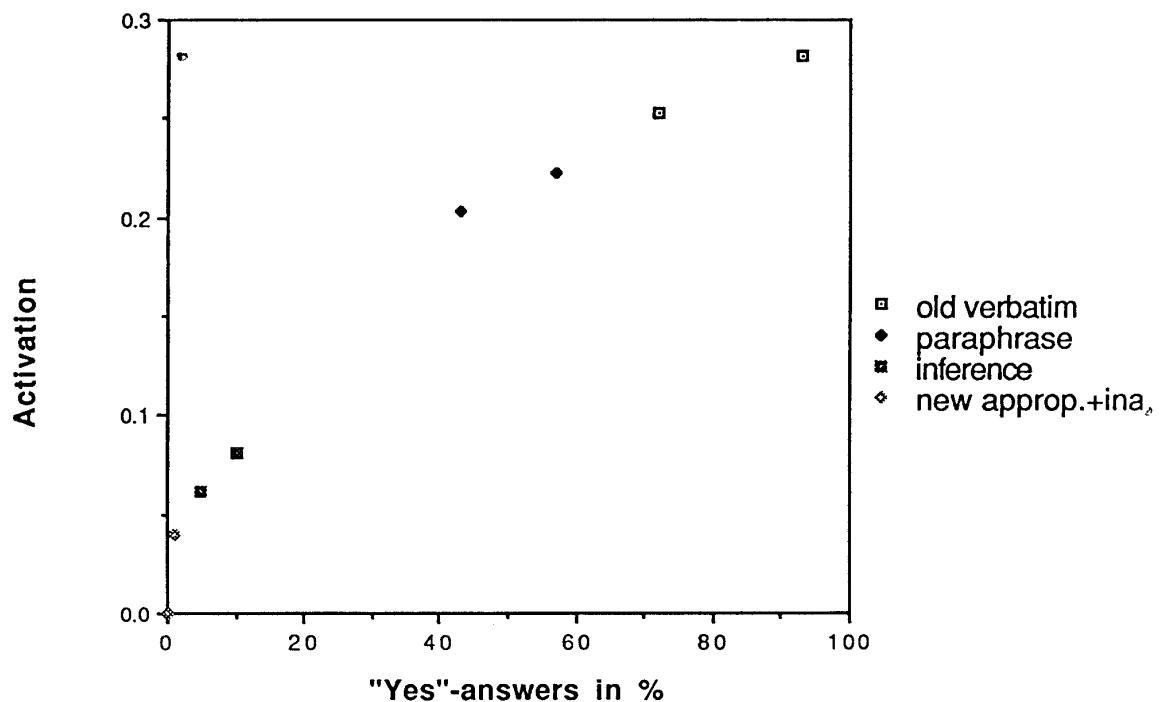


Figure 10. Comparison between empirical data and the second simulation (coefficient of determination $r^2 = .95$). For old, paraphrase, and inference sentences: first plot = event defining, second plot = event referencing).

Discussion

The high correspondence between empirical and simulation data shows first of all, that the CI model is an appropriate description of the comprehension process and second, that assumptions about the differential role of semantic relations can improve the correspondence between empirical and simulation data.

The validity of our results from the simulation is reinforced by the fact that the simulations are based on the entire text and not only on selected parts on it and further that the analysis includes all texts. The CI model can explain the differing probabilities of acceptance of the different types of test sentence because it assumes a multi-level representation of the text. In particular, differences in the acceptance of inference statements can be explained by the type of semantic relation between text and inference. We have made a distinction between event defining relations and event referencing relations. Because the cognitive effort required to establish an event referencing relation is greater than to establish an event defining relation, the probability of activation of event defining relations is higher and thus also the probability of falsely accepting such inferences as part of the text. In our simulation we were able to show that appropriate assumptions about the structure of the situation model, which is constructed during text comprehension through the interaction of text cues and prior knowledge, can help to explain differences in how these kinds of inferences are processed.

Both in our experiment and in our simulations the probability of activation, for example, of the instrument relation was higher than for relations such as purpose. However, this result is not contradictory to assumptions about the special role of causal relations in text comprehension (Graesser 1981; Myers & Duffy, 1990; Trabasso & Suh, 1993; Trabasso & van den Broek, 1985; Singer, Halldorson, Lear, & Andrusiak, 1992). Our experimental texts consist of the description of a single event, thus it is not necessary to connect several events in order to understand the text. However, the main function of event referencing relations, such as purpose, is to connect events., which is why they are so important for text comprehension in general. However, due to the

structure of our experimental texts, such relations were not crucial for understanding the texts in this experiment.

In the present article we are concerned with inferences which afford little cognitive effort, and our experiments were conducted under restrictive laboratory conditions. Currently we are working on an experiment to investigate more complex inferences under more natural conditions. In this experiment the subjects read a text which contains either more or less explicit information about a complex problem-solving task. After reading the text the subjects are asked to build and govern their own city (the computer game SimCity). We are especially interested here in the role of inferences in the interaction of text comprehension and problem solving and hope that this experiment will provide further results for an empirically derived taxonomy of inferences (Guthke & Beyer, 1992; Kintsch, 1993).

- Ankert, H.-G., & Beyer, R. (1987). Untersuchungen zur Textverarbeitung unter besonderer Berücksichtigung von Vorwissenseinflüssen. Zeitschrift für Psychologie, 195(4), 385-399.
- Beyer, R., Artz, E., & Guthke, T. (1990). Zur Differenzierung des kognitiven Aufwandes bei der Anregung von Vorwissen. Zeitschrift für Psychologie, 198, 9-33.
- Beyer, R., Guthke, T., & Ankert, H. (1994). Reader's knowledge and text comprehension. In F. D. Jong & B. V. Hout-Wolters (Eds.), Process-oriented instruction and learning from text Amsterdam: VU University Press.
- Beyer, R., Guthke, T., & Pekrul, U. (in press). Repräsentation von Textwissen im menschlichen Gedächtnis. Zeitschrift für Psychologie.
- Beyer, R., Guthke, T., & Thiele, B. (1995). Integration von Vorwissen beim Sprachverstehen in Abhängigkeit vom Lebensalter. Zeitschrift für Psychologie, 203, 361-378.
- Corbett, A. T., & Doshier, B. A. (1978). Instrument inferences in sentence encoding. Journal of Verbal Learning and Verbal Behavior, 22, 121-132.
- Cotter, C. A. (1984). Inferring indirect objects in sentences: Some implication for the semantics of verbs. Language and Speech, 27(Part 1), 25-45.
- Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. Discourse Processes, 13, 175-190.

- Garrod, S. C., & Sanford, A. J. (1981). Bridging inference and the extended domain of references. In J. Long & A. Baddeley (Eds.), Attention and performance IX Hillsdale, NJ: Erlbaum.
- Graesser, A. C. (1981). Prose comprehension beyond the word. New York: Springer Verlag.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101(3), 371-395.
- Guthke, T., & Beyer, R. (1992). Inferenzen beim Satz- und Textverstehen. Zeitschrift für Psychologie, 200, 321-344.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95(2), 163-182.
- Kintsch, W. (1992). A cognitive architecture for comprehension. In H. L. Pick, P. van den Broek, & D. C. Knill (Eds.), Cognition: Conceptual and methodological issues (pp. 143-164). Washington, D.C.: American Psychological Association.
- Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. Discourse Processes, 16, 193-202.
- Kintsch, W. (1994). Text comprehension, memory, and learning. American Psychologist, 49(4), 294-303.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 5, 363-394.

- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. Journal of Memory and Language, 29, 133-159.
- Kintsch, W., & Welsch, D. M. (1991). The construction-integration model: A framework for studying memory for text. In W. E. Hockley & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (pp. 367-386). Hillsdale, NJ: Erlbaum.
- Klix, F. (1990). Wissensrepräsentation und geistige Leistungsfähigkeit im Lichte neuer Forschungsergebnisse der kognitiven Psychologie. Zeitschrift für Psychologie, 198(2), 165-187.
- Klix, F. (1992). Die Natur des Verstandes. Göttingen: Hogrefe.
- Klix, F., van der Meer, E., & Preuß, M. (1984). Semantische Relationen: Erkennungsaufwand und psychophysiologische Reaktionstendenzen. In F. Klix (Ed.), Gedächtnis, Wissen, Wissensnutzung (pp. 156-172). Berlin: Deutscher Verlag der Wissenschaften.
- Kühler, E., & Preuß, M. (1992). Die Prozesse der Wissensbereitstellung und Wissensnutzung bei elementaren Erkenntnisstrukturen. Zeitschrift für Psychologie, 200, 19-35.
- Lucas, M. M., Tanenhaus, M. K., & Carlson, G. N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. Memory & Cognition, 18(6), 611-631.

- Mansfield, A. F. (1977). Semantic organization in the young child: Evidence for the development of semantic feature system. Journal of Experimental Child Psychology, 23, 57-77.
- McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. Journal of Experimental Psychology: Learning, Memory and Cognition, 12(1), 82-91.
- McKoon, G., & Ratcliff, R. (1989). Semantic associations and elaborative inference. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(2), 326-338.
- McKoon, G., & Ratcliff, R. (1992). Inferences during reading. Psychological Review, 99(3), 440-466.
- Mross, E., & Roberts, J. (1992). The construction-integration model: A program and manual (Tech. Rep. No. 92-14). Boulder: University of Colorado, Institute of Cognitive Science.
- Myers, J. L., & Duffy, S. A. (1990). Causal inferences and text memory. In A. C. Graesser & G. H. Bower (Eds.), The psychology of learning and motivation: Inferences and text comprehension (pp. 159-173). San Diego, CA: Academic Press.
- Nelson, K. (1986). Event knowledge: Structure and function in development. Hillsdale, NJ: Erlbaum.
- Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. Journal of Memory and Language, 25, 279-294.

- Schmieschek, M. (1988) Strukturbeschreibung von ereignisbestimmtem Wissensbesitz. Unpublished doctoral dissertation, Humboldt-Universität zu Berlin.
- Singer, M. (1979). Processes of inference in sentence encoding. Memory & Cognition, 7, 192-200.
- Singer, M. (1980). The role of case-filling inferences in the coherence of brief passages. Discourse Processes, 3, 185-201.
- Singer, M. (1994). Comprehending consistent and inconsistent casual text sequences: A construction-integration analysis. (Tech. Rep. No. 94-04). Boulder: University of Colorado, Institute of Cognitive Science.
- Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. Journal of Memory and Language, 31, 507-524.
- Trabasso, T., & Suh, S. Y. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. Discourse Processes, 16, 3-34.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. Journal of Memory and Language, 24, 612-630.
- van den Broek, P. W. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), Handbook of Psycholinguistics San Diego, CA: Academic Press

- Schmieschek, M. (1988) Strukturbeschreibung von ereignisbestimmtem Wissensbesitz. Unpublished doctoral dissertation, Humboldt-Universität zu Berlin.
- Singer, M. (1979). Processes of inference in sentence encoding. Memory & Cognition, 7, 192-200.
- Singer, M. (1980). The role of case-filling inferences in the coherence of brief passages. Discourse Processes, 3, 185-201.
- Singer, M. (1994). Comprehending consistent and inconsistent casual text sequences: A construction-integration analysis. (Tech. Rep. No. 94-04). Boulder: University of Colorado, Institute of Cognitive Science.
- Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. Journal of Memory and Language, 31, 507-524.
- Trabasso, T., & Suh, S. Y. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. Discourse Processes, 16, 3-34.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. Journal of Memory and Language, 24, 612-630.
- van den Broek, P. W. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), Handbook of Psycholinguistics San Diego, CA: Academic Press

- van der Meer, E. (1987). Mental representations of events. In E. van der Meer & J. Hoffmann (Eds.), Knowledge aided information processing Amsterdam: North Holland.
- van der Meer, E. (1994). Gedächtnis und Inferenzen. In D. Dörner & E. v. d. Meer (Eds.), Das Gedächtnis: Probleme-Trends-Perspektiven Göttingen: Hogrefe.
- van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.
- Zimny, S. T. (1987) Recognition memory for sentences from a discourse. Unpublished doctoral dissertation, University of Colorado, Boulder.

Author Notes

The work presented in this article was supported by a postdoctoral grant to the first author by the German Academic Exchange Service. We would like to thank Eileen and Walter Kintsch very much for providing a stimulating theoretical framework, for their comments to the manuscript and for their help in making our English understandable.

