ADVANCES IN PSYCHOLOGY

79

# Text and Text Processing

Guy Denhière
Jean-Pierre Rossi
Editors

North-Holland

# TEXT AND
# TEXT PROCESSING

# ADVANCES IN PSYCHOLOGY

# 79

*Editors:*

G. E. STELMACH

P. A. VROON

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

# TEXT AND TEXT PROCESSING


*Edited by*

Guy DENHIERE

*Groupe Textima*
*Université de Paris VIII*
*Saint-Denis, France*

Jean-Pierre ROSSI

*Laboratoire de Psychologie du Langage*
*Université de Poitiers*
*Poitiers, France*



1991

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

**NORTH-HOLLAND**
ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the United States and Canada:
ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
655 Avenue of the Americas
New York, N.Y. 10010, U.S.A.

ISBN: 0 444 88484 X



Printed in The Netherlands

# FOREWORD

This volume presents a comprehensive overview of current research on the cognitive and linguistic aspects of text processing during comprehension, knowledge acquisition and discourse production. It brings together psychologists and linguists whose joint concern is finegrained description of text structure levels as they relate to readers' ongoing activities. Part of the originality of this volume lies in its systematic efforts to compare and contrast theoretical models and empirical data. Equally important is its international scope: the contributors are among the best known European and American specialists in the field.

This book is divided into four parts. Part One analyses the ways in which representation construction processes are related to text type and cognitive architectures. Part Two is devoted to knowledge acquisition from text and presents the first model inspired by connectionist theory. Part Three deals with lexical and syntactic tools that convey text coherence and examines their role in the construction of representations. Part Four covers cognitive processes in discourse production.

In the first part of this book, Rossi lists and describes the components of information processing in text comprehension. Each of these processes is characterized in terms of goals, the features of the input components these processes operate on, and the nature of the output forming the basis for the next level of processing. This description serves as a backdrop for a critical examination of current concepts in the field, and an inventory of the major conceptual and functional issues for each level of processing.
Frederiksen and Donin describe recent developments in semantic representation theory for knowledge acquisition through discourse. They present a theoretical approach to semantic discourse processing in which semantic representations are generated at different levels by rule application. They show that these structure-generation and transformation occur on multiple levels. A formal semantic representation is presented for each of these levels.

Flammer and Lüthi reports a series of data on relational between text format and mental representation formats. Findings show that the mental representations subjects form for purposes of selective recall are not isomorphic to the structure of the text they have read.
Fayol examines the psychological validity of text typologies and presents evidence indicating that for certain well-defined categories, different text types can be processed in a variety of ways and recall performance can vary considerably. It is argued that each type of text may be associated with the specific surface structure which is the most appropriate to it. Nevertheless, descriptive criteria need to be better defined since they are less sharp than those used for problem solving taxonomies.
Kurtz analyses three metacognitive aspects of text processing in metalinguistics, reading strategies, and comprehension monitoring. Three types of metalinguistic skills are described: phonological skills, lexical skills and syntactic skills. Knowledge about reading strategies is differentiated from instructions to use certain strategies. One fundamental conclusion is that much more research is needed to bridge these three areas of study to test competing theories.

In chapter 2, Kintsch presents a construction-integration model of discourse comprehension, so termed to highlight its most salient features. The model combines a construction process --in which a text base is built up from the linguistic input and the comprehender's knowledge base-- and an integration phase where this text base is integrated into a coherent whole. The model links production systems and connectionist theory.
Baudet and Denhière present a broad conceptual framework for the acquisition of knowledge from text. The model incorporates relational, transformational and teleological (functional and intentional) systems which associate domain representations and learners' cognitive representations. Positive empirical support is presented for the assumption that learners gradually build up a system-organized representation.
Ballstaedt and Mandl define knowledge as an interconnected structure of cognitive elements. Their chapter examines knowledge change during reading through the Heidelberger-Struktur-Lege-Technik (SLT). Three different methods of assessment (word association, card sorting and diagram making) are used to examine concept modifications in a segmented expository that embeds contextual dissonance.
Le Ny models the acquisition of concepts as the addition of new branches to an existing semantic network. The parsing of each chunk of information is defined

as a sequence of operations yielding a partial semantic representation which is then added to the current overall representation. The concept of spontaneous recovery is analyzed on four levels of conceptualization, all of which contain increasingly more abstract schemata.

François describes a hypothetical sequence of cognitive processes enabling experienced news editor to acquire, update, and organize pieces of information (temporally, causally, teleologically) about a macro-event from a series of wires, in such a way that the outcome of the reading process is a mental or situational model of the information in the wires. A series of wires on hijacking is used to illustrate these cognitive processes.

The two chapters by Charolles and Ehrlich present an extensive overview of text continuity in linguistic and psychology, and aim at sparking debate on the potential value of combined linguistic and psycholinguistic approaches. Studies in linguistics are mainly aimed at identifying the nature and roles of markers of text continuity and connectedness in text composition in different languages. Charolles and Ehrlich show that cohesion and conexity markers, which guide the interpretation and construction of text coherence, can also result in text processing difficulties. The use of these markers is governed by specific rules in each language and the formal description of these rules is the basic task in linguistics. The basic issue in psychology is how individuals construct a coherent and organized mental representation of discourse. The authors survey the main text processing models in psycholinguistics.

Garrod's paper presents a cognitive perspective on pronouns which integrates constraints on human attentional resources and short term memory. The fundamental issue in cognitive psychology is to define the nature of the reader or listener's transient mental representation of text which can serve as a basis for anaphor interpretation. The data suggest that a special type of cognitive linkage, termed thematic subjecthood, may be implemented, and that is also a major feature of pronoun interpretation.

Trabasso analyses strategies for the construction of a coherent interpretation of narratives by local inferences based on the reader's knowledge of intentional action. A causal network model for discourse analysis is used to account for the considerable variation in structure and interpretation of content accross subjects' representations of narratives. From a developmental perspective, story comprehension takes place in a sequence in which within-episode integration (through inference) precedes between-episode integration. Changes in both these types of integration are observed with age.

In their review of text production research in psychology, Esperet and Piolat analyze three interrelated features of text production: production conditions, texts and text characteristics, and output in pre-specified conditions. The overview centers on the functional aspects of production (processes), and emphasizes planning (pauses, errors and self-correction) and control processes. The authors show that real time data obtained from computer-assisted composition make step-by-step tracing of individual activity possible.

Breuleux's paper has three objectives: a critique of the issue involved in the use of the think-aloud technique in writing tasks, a presentation of ways overcoming drawbacks, and a description of computer-based tools to apply a new protocol technique. He shows that representations and processes can be inferred during analysis of the think-aloud protocol if appropriate tools for the analysis and representation of knowledge from discourse are employed. He outlines a semantic (ATN) grammar for planning in which grammar constitutes an explicit encoding schema for planning processes in writer's think-aloud protocols.

In an ingenious experimental production task devised by Coirier, subjects select statements and then arrange them in order of their choice. As a result, the statements that the "author" of an argumentative text feel to be the most important have three specific text functions: they express value judgments, connect arguments, and are the justified rather than the justifying elements of the text.

Quasthoff outlines a linguistic model of discourse production which fulfills the following criteria: empirical testability (can be falsified), integrates the cognitive and interactive constituents of discourse planning, integrates global and local procedures. Data are reported for a "formal" and an "informal" condition. In the "informal" condition, the adult listener guides conversation with a child such that retelling the story can be predictable. These data provide insights into the "interactive achievement of reportability", i.e. the sequential implicatures which require a narrative discourse unit within the specific replay discourse pattern.

Zammuner reviews a series of artificial intelligence language generation models and programs which have potential for a theory of discourse production and may have psychological relevance. Her main conclusion is that the range of available strategies, their interaction with the problem structure, circumstances affecting their use, the influence of prior experience or motivational factors, and the possibility of strategy transfer across tasks or domains are all open issues in poorly structured tasks.

As this overview has shown, the chapters presented in this volume are addressed to researchers and students in psychology and linguistics with an interest in the growing field of cognitive science, as well as to teachers and educators who are called upon to deal with problems of language processing and knowledge acquisition.

Guy Denhière Jean-Pierre Rossi

This Page Intentionally Left Blank

# CONTENTS

# TEXT REPRESENTATION: STRUCTURES AND PROCESSES

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# INPUT-OUTPUT: PROCESSING AND REPRESENTATION

Jean-Pierre ROSSI

Laboratoire de Psychologie du Langage
Université de Poitiers
Poitiers, France

During the reading of connected discourse, the subject performs processes to construct a representation of meaning. The aim of research in cognitive science is to track the components of information processing or at least to describe the major ones. The present articles gives a list of these processes, by specifying their goals, the features of the input components they act upon, and the characteristics of the output which form the basis for the next level of processing. This description, which focuses on a critical analysis of current concepts in the field, also inventories the major conceptual and functional issues that arise at each level of processing. The scope and the difficulty of these problems clearly point to the necessity of analyzing the preconditions for the construction of a text processing model.

## 1. PROLEGOMENA TO A PROCESSING MODEL

### 1.1. Description of input and output

The researcher interested in describing cognitive processing can be compared to an observer in front of a factory gate who watches parts go in and cars come out. To comprehend what takes place inside the factory he has two types of products available for scrutiny: the products going in (input) and the finished products coming out (output). Comparison of input and output can serve to hypothesize relations between them which cannot be observed directly; this metaphor suggests why test descriptions are as varied as they are numerous.

Text models can be classified broadly into three major types: concept maps, inter- and intra- propositional grammatical analyses, and text typologies. The studies of Trabasso and Sperry (1985) belong to the first category. These authors describe the causal network connecting different elements in the text. Similarly, Ballstaedt and Mand (1985) constructed a conceptual map of a text and characterized the reader's knowledge before and after reading, so as to track conceptual organization.

Although related to this theoretical perspective, the propositional analysis (van Dijk and Kintsch, 1983; Frederiksen, 1985...) differs slightly in approach in that the basic unit is the proposition. Within this framework, a distinction must be made between grammars which specify relationships between the predicate and the arguments (one or several), and grammars which enumerate the various links between propositions. Frederiksen's (1986) interpropositional grammar is an offshoot of intra grammar; only the processing unit has changed.

In contrast, the third approach differs greatly from the first two in that the aim is to constitute a text typology: description, exposition, narration... This approach aims at identifying the structural constants of each type of text and thus constructing a schema modelling each category. Each category can be divided into subcategories. Mosenthal (1985), for example, ranks expository texts according to their structural complexity, from simple description to hypthetico-deductive reasoning.

These three types of models are described in detail elsewhere in this volume. They differ as to units of processing and to points of view. Drawing a hand is different from drawing the whole body although the hand is part of the entire silhouette. Similarly, a brick is a unit in a dividing wall, but the room is the unit of measure used to describe an apartment -- in other words, each level of processing has its own processing unit. Letters and words are identification units whereas the propositon is the unit which serves to describe sentences. There are as many units as there are types of processing. The nesting of these units is never a simple matter; text comprehension is more than the comprehension of each of the words which compose it. The definition of a processing unit or units associated with each level of analysis is thus one of the researcher's major concerns.

Aside from the scope of the processing unit, the description varies as a function of theoretical perspective. Regardless of the talent of fame of an artist, a portrait will always be different from the description of the anatomy of the face. A given object will receive different descriptions depending upon the angle from which it is seen. Descriptions of processing units found in the current literature only have an interest for the psychologist in so far as they can evidence a subject's behavior, and by extension can lead to an analysis of the operations a subject implements to understand a texte s/he is reading. Thus the distinction between narration/exposition is useful only if it can be shown that subjects process these two types of texts differently, or that there is a narrative schema whose mastery facilitates the understanding of the story. The issue is thus the evaluation of the psychological validity of descriptions. A description, in our opinion, is relevant only if it operates from the viewpoint of what the subject is doing and thus concerns, the processes he must perform.

## 1.2. The task

A task defines the purpose, or the goal, and by extension the operations to be accomplished. In this sense a task structures a subject's activity. An article will be read differently if the goal is to summarize it versus to memorize it. The task defines the content of representations and thus the operations which must be performed by the subject. Even when the task is specified after a text has been read, it can modify the content of the representation. This shows that any study of subject's behavior must be founded on a task model, i.e., one which reflects task constraints and specificities. The complexity of the model increases as a function of the nature of operations to be performed at any point in the text, and the nature of previously encountered information. Certain forms of processing tend to be invariant from the beginning to the end of the text, whereas others arise as a function of specific needs. Similarly, information will be processed differently as a function of the reader's linguistic, thematic and procedural knowledge. As a general rule, the type of processing strategy used by a reader depends upon the structure of the text, the reader's knowledge and the nature of the task. Any study of information processing should take these knowledge parameters into account.

The definition of a task model should incorporate the reading goal, the required task, and the type of test. The task must be precise and constraining enough to control what the subject is doing at least partially if it is to provide access to his/her processing strategies. It is totally unrealistic to describe the processes implemented by a subject in situations as vague and nonconstraining as tasks where the subject is asked to "read for the sake of reading".

In defining a task model, the nature of the task presented in instructions to the subject must be distinguished from the type of test used to measure performance. The three main tasks generally assigned to subjects are comprehension, retention and summarizing. These tasks are combined with testing, which is normally in the form of comprehension questions, measures of recall, and summarizing. Although this list suggests there is a one-to-one relationship between task and test, there are numerous examples in the literature where a recall test is associated with a comprehension task or the contrary, where comprehension is measured through recall, despite the fact that the operations implemented by a subject during retention of a text clearly differ from those employed in text comprehension. This gap between task and test may be due to a prevailing interest in describing subjects' knowledge, rather than in the way this knowledge has been organized or constructed. It goes without saying that this imbalance can only be a further source of difficulty when attempting to model information processing.

Recall, summarizing and question answering are normally used to assess subjects' knowledge after reading of a text. However they generally provide little information on what the subject does while reading. In contrast, on-line studies using indirect measures of subjects' activities provide a means of analyzing these processes. Although it is possible to access subjects' knowledge without exploring the underlying operations, it is not interesting to study these operations without taking their outcomes into consideration.

Models of processing strategies, in that they apprehend individual processing strategies, cannot disregard the differential approach. This implies forming groups of subjects whose performance is equivalent (Ehrlich and Rossi, 1986). Although equivalent performance may be achieved via different strategies, it seems even more intuitively logical that different performances can arise from different strategies. Unfortunately, neither is the case since a given strategy may result in high performance in one subject, and low performance in another.

Task model are confronted with two other obstacles: the first is that the task assigned to the subject through explicit instructions may differ from the task the subject assigns to him/herself. Instructions and design should be planned in such a way as to make the two coincide as closely as possible. The second obstacle concerns levels of processing. Subjects may exhibit identical results measured by whatever test is used, yet their similar responses have resulted from processing at very different levels. This points to the necessity of appropriate assessment of levels of processing during the course of reading.

Bearing these considerations in mind, the model presented below attempts to apprehend information processing by moving step by step from input to output. The description of each step will indicate: a) the type of ongoing process; b) the operations performed; c) the processing units involved; d) the permanent structures involved, e) the type of knowledge activated. Issues related to each level of processing will also be examined.

## 2. PERCEPTUAL IDENTIFICATION AND PREDICATIVE ACTIVITY

Perceptual identification refers to processing of word features and the activation of word meaning. The subject perceives forms, or graphemes, which trigger feature detectors (Hubel and Wiesel, 1962). The characteristic features will in turn activate the schemas stored in the graphemic repertory. Letters and frequent letter strings are recognized, combined, and probably assigned a phonological interpretation. These operations result in word recognition and the activation of meanings. Identification is preceded and at least partially constrained and preactivated by the context. This activation of meanings involves a predicative analysis so that the identification of a predicate leads to searching the arguments that must be associated to it.

During the identification stage, the processing units are characteristic features, letters, frequent letter stings (bigrams, trigrams...) and words. The structure activated is obviously the lexicon. The main issues on this level are related to the organization of this structure. Is the distinction between word features and word meanings relevant here? If the answer is in the affirmative, how many repertories of form are necessary: can it be shown that a graphemic repertory containing the graphemic representations, a phonological repertory constructed from phonological representations, an articulatory repertory constituted by the articulatory motor programs are all necessary and sufficient? What can be said about the content of these repertories and their organization? What can be said about the content of these repertories and their organization? Do they have an equivalent status or should we, for example, consider that the phonological repertory alone governs acces to meanings? How are meanings organized, classified and associated? Are there linking or tradeoff systems between these different repertories?

Although these issues have been explored experimentally (Rossi, 1985) we have no clearcut answer, although there is evidence that there are numerous and varied forms of interconnections between the different systems. Whereas there is general agreement that the context can help to prepare and preactivate elements of the repertories and thus facilitate reading and instantiation of the relevant meaning, little is known about how it operates. What, for instance is the role of redundancy and different knowledge bases (linguistic, lexical...) in the so-called preactivation or preparation for later processing? The notion of automaticity of coding and ractivation of all these systems warrants further investigation.

Equally fundamental is the connection between the lexicon and the subject's knowledge bases. Are knowledge bases distinct from the lexicon, whereas others are automatically integrated so that all forms of this new knowledge are assigned to each notion (knowledge of a domain, linguistic knowledge, procedural knowledge...)?

Categorization in predicates and arguments most likely takes place at this level. Is this categorization part of the activation of meaning or is it distinct, thus requiring a supplementary operation? For the moment there is no answer. It seems difficult to isolate categorization from activation of meanings, since each item has a specific function. This categorization or "predicative activity" is the basis for the construction of propositions.

## 3. CONSTRUCTION OF PROPOSITIONS

Following Kintsch and van Dijk (1978), the proposition is generally assumed to be the fundamental unit for the construction of representation. A proposition is made up of a predicate and arguments. It is the smallest unit having a truth-value. This definition is rarely if ever challenged, but researchers tend to adapt it to their own systems. Frederiksen (1986), for example, uses the term "proposition" and the notions of predicate and argument, but enlarges their scope, in that he defines the proposition as a "chunk of semantic information in working memory". Each item in this definition is important, but is it possible to be more precise? Is it possible to define what is meant by semantic information? If one admits the idea that a proposition contains several items of information, is it possible to specify its range? Is a chunk equivalent to a proposition? Can a proposition be composed of several predicates and arguments? What characterizes a chunk? Is it simply a quantity of information, or should it be considered as a semantic unit, Is it a set of information corresponding to a topic, a subtopic, a concept, a predicate... Can it be defined as a packet of information associated with a predicate? On what basis are the chunks constructed? Which cues allow the subject to form these packets of information? What is the status of this unit: a linguistic unit, a logical unit or a cognitive one? Are the relationships described in these grammars connected to each concept or on the contrary, are they open and concept-independent?

The classification of the proposition also varies. Turner and Greene (1977), who were the first to develop a detailed propositional analysis, distinguish among predicate propositions expressing a relationship of action or state, modifier propositions used to quantify, qualify, or negate, and connective propositions which express various forms of restriction or limitation. Frederiksen's BNF grammar (1985), which is based on the same principles, has 88 different propositions. Given such a variety, does it make sense to presuppose the existence of a single construction procedure, or should different forms be envisaged?

Another issue is related to the relationships between the whole and its parts. If a sentence can be separated into a sequence of propositions, is the global meaning of the sentence the sum of the meaning of each of these units? The issue is the same for the entire text, whose meaning is greater than the sum of the meanings of each sentence. In general, and regardless of the basic unit, its contribution to the whole cannot be the result of a simple addition. How may one account for this type of phenomenon?

Note that Turner and Greene's original objective (1977) was to develop a method to determine the semantic content of texts. This method was applied to the text base. Its use in the analysis of the representation of meaning in memory raises a series of questions: what is the format of representational units: concepts or propositions? What is the status of propositions: cognitive units, logical units or linguistic units? Does propositional analysis have psychological validity, or is it on the contrary a purely theoretical construct far removed from psychological reality? Are propositional analyses able to account for "all the meaning"; are they exhaustive; which propositional analysis is the best?

Information processing models must be able to show whether text comprehension is based on propositional analysis, and if so, account for how the reader builds propositions, and assigns status to them. It is legitimate to inquire in this respect whether the reader can construct propositions independently of the reading task, the goal, or prior knowledge. The reader who is only interested in how penicillin was discovered will go directly to that part of the text without spending time on anecdotes concerning the personal life of Sir A. Fleming. However, to decide whether a piece of information is important, its meaning must first be activated. Even though there may be cases where one has the impression that what follows is of lesser importance careful reading will nevertheless require the activation of all meanings.

A recapitulation of processing sequences will help to illustrate propositional analysis in the study of individual processing models. At the beginning the reader is confronted with the printed word. The identification of graphemic features leads to the activation of meanings (see section on identification). These units of meaning are retrieved and combined into another unit termed (for lack of a better word) a proposition. This new processing unit is thus a chunk of meanings whose function can be identified by predicate analysis. Linguistic, semantic, as well as logical and cognitive information help the subject to organize these new units, i.e., to identify the meanings forming the argument and those forming the predicate, and to establish a connection between them. The construction of these relationships may draw on inferences that require activation of the different knowledge bases available to the reader. Recent studies have introduced the idea that propositions are not constructed independently of the task, the reader's knowledge, text content, and moment of proposition formation: the characteristics of the units are not the same during topicalization as during argument search... Since it is likely that each new proposition is composed with reference to preceding ones, the effects of immediate dependency should not be disregarded. This is why we are led to hypothesize that an executive structure intervenes at this level, which adjusts the different levels of knowledge. At this point, it is necessary to specify when and how these different knowledge bases intervene. Before exploring this question, we will take a look at the operations which are thought to be performed on propositions.

## 4. PROCESSING OF PROPOSITIONS

For Kintsch and van Dijk (1978), processing of propositions consists mainly of selecting, constructing, and generalizing: propositions which add nothing to the interpretation of subsequent propositions are deleted, and those which condition, or are the consequence of an action described in another proposition are selected. Sequences of propositions are replaced by more general propositions which encompass them; macropropositions replace those propositions which compose, modify, or are the direct consequences of the original propositions.

The construction of a proposition results in the elaboration of a local meaning which may require inference. The decision as to the future of this new unit presupposes matching, i.e., comparing the information stored in memory against the new unit to

decide whether it is useful for the interpretation of previously stored information, on the basis of whether it modifies or is the consequence of information in memory. If no matching procedure can be instigated, but the unit appears to have links with subsequent text, it can be kept in the buffer until recycling. This is generally what happens when the reader encounters the interrogative form:
"What happened 65 million years ago?"
This sentence generates expectations which are not resolved until later in the text:
"At the end of the Mesozoic era, dinosaurs vanished from the face of the Earth". In this case, it is clear that the first information should be kept in the buffer since the interrogative form indicates the necessity of temporary activation. Interrogatives often cue topic shifts or new information. Temporary activation generated by the question-form is stored in working memory before being modified into statement form. This example provides evidence that each proposition has a definite function in building the macrostructure. Some announce macrostructural information, others modify it, enhance it, or on the contrary, restrict it, etc... The processing performed on these propositions depends on their functions. A typology of these functions would be vitally necessary to understand the operations performed by the reader. The difficulty of such a projet lies in the fact that in most cases, the processing unit would be suprapropositional. it could be a predicate proposition and all the propositions modifying the arguments.

If the subject reads to produce a summary, his activity will be planned so as to select relevant information and condense information which is judged redundant. Summarizing a text in fact consists of identifying the sequences to be included in the summary, characterizing them, and condensing the information they contain. This activity can be conceptualized in two ways: through schemata, or, on a more functional level, through identification of topics and topic shifts non-integrated in a previously acquired text schema. The notion of schema was originally introduced by Bartlett as early as 1932, and was first adapted to the study of connected discourse by Rumelhart and Ortony (1977), who defined a schema as a data structure which represents generic concepts stored in memory. According to Rumelhart and Norman (1983), a schema is composed of constants and variables: the schema for HEAD contains the one for EYE(s). The schema represents knowledge rather than definitions. Schema construction includes an internal evaluation feature which tests the schema for relevancy agains other schema. The schema represents knowledge on all levels, from the graphemic features to text content. In this framework, the text schema is defined as a mental surperstructure describing internal regularities of each category of text (narrative, expository, ...). This superstructure, integrated by the subjects, helps the comprehender to understand texts and if necessary, to write a summary. A text schema should be distinguished from a schema of a text, which corresponds to the representation of the subjects' knowledge after reading a text.

In this framework, it is assumed that a good adult reader will have a variety of schemas available in memory representing different types of previously encountered

texts. Using the narrative as an example, Denhière (1986) points out that "the partisans of story grammars share the following basic assumptions:
(1) There is a canonical narrative structure which can be described in the form of a constituent grammar with a finite numer of rewriting rules (Denhière, 1984);
(2) This conventional structure is gradually internalized through repeated exposure to narratives (listening and reading);
(3) This cognitive structure is organized as a hierarchical structure with, at the lower levels, empty slots -variables- which will be assigned a content during the hearing/reading of narratives;
(4) This cognitive schema fulfills a certain number of functions, the main ones being to guide comprehension and govern the retrieval of information stored in memory when the individual is called upon to recall or summarize the story".
This cognitive schema is part of the knowledge required of all good adult comprehenders of written prose.

The activation of a schema in this model reduces the reader's activity to that of inserting variables into a fixed framework. In other words, after variables have been selected, they are assigned a function: a given proposition is important and should be assigned to a fixed part of the schema. This formulation presupposes that it is possible to construct a text grammar, and that any text can be reduced to a text-type. In chapter 3 of this volume, Adam demonstrates the illusory nature of such an enterprise. A scientific article may contain descriptions about scientific methodology or the story of how a discovery took place within the section describing the scientific issues involved. The reader may take an interest in the scientific procedure, the argumentation, or the events leading up to the discovery. Thus it is not unusual to find several schemata coexisting in a given text. The subject goes from one schemata to another, choosing the one which best correspnods to his/her goal. This is why, in contrast to a schema theory based on reading using text typology it seems more appropriate to characterize a processing strategy model based on identification of sequences. The basic principales of this model are founded upon the idea that a text is composed of a chain of sequences. The reader identifies the topic of each sequence. Each new proposition is matched against the topic, and the reader locates topic boundaries, or topic shifts. The propositions which best correspond to be topic will make up the summary. These kernel propositions (which we term N propositions) form the basis for comparisons. Any new propositions will be ranked (ordered) as a function of its distance from the kernel proposition. Propositions N-1 complete, restrict, enlarge or announce propostions N-1. The good interpreter will focus his attention on the N propositions around which he will structure all his activity. This selection is evidenced by longer reading times on the important i.e. kernel sentences (Rossi and Bert-Erboul, 1987). The description of the three levels (N, N-1, N-2) is connected to the typology of functions since propositions N-1 serve some functions in relation to N, whereas N-2 are organized around N and N-1.

The operations which serve to identify a topic of a sequence and the features used to detect important propositions, topics and topic shifts also require attention. This segmenting of the text is related to issues of coherence and cohesion which are dealt with in chapter 3. What is important to note here is the fundamental role of punctuation, connectives and lexical markers which serve to orient the subject.

The selection of a new proposition may lead the reader to complete, modify, recycle or even delete all or part of the contents of memory. To illustrate these different operations, let us return to the first sentence in the example from the dinosaur text. "What happened 65 million years ago? That something happened, everybody agrees. We know the outcome: at the end of the Mesozoic era the dinosaurs vanished from the face of the Earth". The first sentence generates expectation: "what happened 65 million years ago". The second sentence reinforces the first, and the third contains information which leads the reader to delete everything so far in memory and replace it by "at the end of the Mesozoic era, the dinosaurs vanished..." Elimination of expectation results in a reformulation of the latter propositions by condensing, since the reader will only retain from this paragraph that the dinosaurs disappeared at the end of the Mesozoic era. The reader is thus led to reformulate: by condensing, adding a new piece of information, modulating (restricting, enlarging, specifying, ...) given information, or employing concatenations relying upon temporal, spatial or causal links. Thus the third sentence "all perished and 75% of the animal world was hit by this terrible crisis", leads to a reformulation which could take the following form "at the end of the Mesozoic, the dinosaurs and 75% of the animal world disappeared". Thus the text is composed both of propositions which confirm or reinforce certain ideas already present in the summary and others which cause the buffer to be emptied and another version to be stored. As the previous examples have shown, the selection of a proposition which itself requires examination of the contents of working memory leads to recycling and reorganizing these contents. One of the obstacles at this level is related to modes of processing propositions. Are they processed as soon as they are encountered, or are they processed in packets? According to the first hypothesis, each proposition is processed as soon as it has been constructed, with this construction itself dependent upon the contents of working memory. One of the major drawbacks to this hypothesis is that processing is slow and it is difficult to isolate a proposition without information as to what directly precedes or follows it. The second hypothesis requires defining the size and the nature of the packets. Criteria can be quantitative, qualitative, or both at the same time. The quantitative criterion is the following: the number of propositions forming a packet is predetermined and depends upon the processing capacity of the system. The capacity and if necessary the different factors which are likely to cause it to vary must be specified. The qualitative criterion states that propositions are grouped according to common topic. In this case it is necessary to describe the features of this semantic unit. Quantitative and qualitative criterion: the topic dominates but the number of units making up a chunk is constrained. The theoretical options presented here must be discussed. Because the issues of the plausibility of text grammars and their efficiency during comprehension are developed in the next chapters, we will only mention here the problems raised by this outlined model.

A first series of quetions concerns the markers which allow the reader to assess the importance and relevance of information with regard to reading goals. These markers and the procedures implemented are such that the selection of important information can occur from the first reading (Rossi and Bert-Erboul, 1987) and be quite automated in able adult comprehenders (Rossi and Coulombier, 1987). The representation created after this first processing permits a significant improvement of the selection during the following readings. If this guidance can be easily considered, the markers that allow the identification of the informative functions of propositions and their role in the construction of macrostructure remain to be explored.

The same holds true with redundancy markers and cognitive operations that permit the building of macrostructure. Beyond the description of these operations, the format and the status of knowledge structures are still very hypothetical. There is no evidence that the representation is of a propositional type. However that may be, processing of propositions does not stop when chunks have been constituted; these new units must be reorganized with a hierarchical structure.

## 5. HIERARCHY OF SEQUENCES

The types of processing described above result in the building of macrostructural units which will not remain independent but will have to be organized into a hierarchical structure. To be more precise, this structuration accompanies the operations that generate macrostructure and in most cases it cannot be isolated. Nevertheless, the reader may perform subsequent restructuration whenever the comprehension of the text causes him difficulties, when his knowledge of the topic is as important as relevant or when his reading goal requires it of him.

The organization of macrostructural units can be described either in terms of pre-existing structures or of networks to elaborate. The former refers to text and script grammars. The structure exists in the reader's memory, and must be activated upon reading. This theory, assumes that it is possible to build a text grammar precise enough to be applied to most texts. Critical points are a) the activation of the schema (identification of the markers that permit the selection of the adequate schema; moment when the schema is activated; modalities of non-activation), and b) the association between macro-propositional units and the headings of the schema (markers permitting the identification of the heading to which the macro-propositional unit will be assigned; procedure of assignment).

If it is impossible to build a text grammar with the qualities mentioned above, we must hypothesize that the reader constructs a different structure for each text. This is

Trabasso and Sperry's proposition (1987) when they describe interpropositional relations in terms of a causal network. Following their argument, causal relations would constitute the basic structure of all texts. The logical extension of this hypothesis results in considering any relation as causal and may make meaningless the notion of causality itself.

Between text grammars and causal networks, we find the theory of frames developed by Frederiksen (1987). Frames characterize sequences of information. Thus the procedural frame describes the operations performed to fulfil a given objective. This procedure yields a result. Thus one follows the course of the text by characterizing step by step the information provided. This processing system is flexible and can be used with all texts. It assumes that the reader has in his memory frames and procedures to identify them. Critical issues in this theory are dealing with the enumeration of frames and the description of the procedures. Besides theses issues, one might consider the efficiency of frames and their utility in text comprehension. Can the frame knowledge promote text comprehension? That is the fundamental question. Rossi and Bert-Erboul (1987) demonstrated that poor comprehenders are easily able to identify the sequences of a text, but are unable to answer questions concerning its argumentation or problematics. Thus, the organization of information in sequences is just as important as is the identification of those sequences.

All the models presented above tend to describe the "text base" it does not necessarily correspond to the representation constructed by the subject since it is entirely divorced from her knowledge, attitudes and goals for reading. To illustrate the disparity between text base and subject's representation, we will again refer to the dinosaur example. This text is made up of the six sequences: 1) presentation of problem, 2) description of the world of dinosaurs, 3) proof of their sudden disappearance at the end of the Mesozoic era, 4) discussion of the necessity for interdisciplinary collaboration in scientific research, 5) presentation of a hypothesis to account for this disapearance, 6) arguments supporting this hypothesis. A good summary of this text will include those six elements. By contrast, only sequences 1, 3, 5 and 6 are necessary to understand the reason for the disappearance of the dinosaurs. Non-specialists will concentrate on these four sections, whereas the student of paleontology will only focus on the arguments, using his knowledge to look for weak points, and restructuring the text with the aim of finding its contradictions. Thus one text will be structured differently by readers according to their goals and their previous knowledge.

At this stage, the processing unit is composed of bundless of propositions that must be organized through a hierarchization process which is based on an assessment of the relative importance of sequences and the construction of intersequential relations. These operations are dependent upon the decision to comply with the linear presentation of the text sequences or to modify this sequence. This decision is made as a function of the comprehender's goal, interests, attitudes and knowledge.

The activation of knowledge bases and how knowledge intervenes at each level of processing thus is at the heart of the issue of text comprehension.

## 6. SITUATION MODEL AND TYPES OF KNOWLEDGE

The operations involved in the formation of a propositional text base have been described above. These involve processing of the graphemic component, the semantic component, predication operations, forming of propositions, construction of the macrostructure, and hierarchization of propositions in the form of schemas or frames. Processing of the graphemic component requires knowledge of letters or groups of letters (bigrams, syllabes). This presupposes the existence of graphemic and phonological repertoires in memory. In a similar way, activation of meaning necessitates a prior semantic store. Each processing category thus draws on specific knowledge bases; in fact, systems of knowledge can be characterized by the type of operations in which they engage. Text comprehension thus requires lexical, predicative, linguistic, and schematic knowledge, independent of knowledge of the topic. Furthermore, metalinguistic and metacognitive knowledge are also necessary. This suggests that a knowledge of the procedures involved in performing a task is one of the essential tools behind successful summary writing. Before these procedures can become automatic, a long period of training is necessary.

These different categories of knowledge are not equivalent. Some are composent of repertoires, such as discriminant graphic features which differentiate letters or groups of letters. Other knowledge bases are composed of procedures; for example, the procedure for summary writing. The repertoires themselves also differ. The semantic repertoire differs in content and organization from access repertoires. Switching from one repertoire to another is rule-governed and is currently the focus of theoretical description.

The description of knowledge implementation is further complicated by the fact that bottom-up and top-down processes are present in all operation categories. Word identification is facilitated by the context, whereas construction of propositions is dependent upon lexical knowledge. In other words, different types of knowledge are associated with each type of processing. These interactions are dependent upon the properties of the input, the nature of the task, and the reader's knowledge. Input properties have received the most attention: the role of vocabulary, syntax, propositional and hierarchical analysis. In contrast, the tasks have seldom been explored. This is also the case for the role of prior knowledge, although in this area several formulations have been proposed. It is clear that the story of the discovery of penicillin will be read differently by an historian than by a biologist. The more prior knowledge of the topic is important for reading, the more reading will distance itself from the text frame. Since the role and the impact of each knowledge base are combined in a specific way which is situation-dependent, it has been assumed that these are accessed either by simple activation or through the actions of an executive structure which controls data management.

Activation is clearly more elementary. The reader activates his knowledge of the topic; if his knowledge is poor, the impact will be low, whereas if the subject has broad knowledge of the topic, this knowledge will be crucial. If text information is not compatible with what is stored in memory, the reader will be forced to proceed to extensive processing to test for the existence of contradictions between text and personal information. However, even before devising theoretical interaction systems between different knowledge categories, the specific features of each must be categorized; i.e., what do they do, in what form are they stored, how are they organized, and how are they activated. Models of subject performance are highly dependent upon the form take by this type of description.

This attempt at a description of the fundamental processes implemented during text processing sheds light on the difficulties inherent in such a task. Whether one adheres to a particular theoretical frameword (we have often referred to the one developed by Kintsch and van Dijk), or takes an independent task, the unresolved issues remain as numerous as they are arduous. This is why overviews of current research in the field often resemble lists of questions rather than truly operational models.

## REFERENCES

**Ballstaedt, S.P. and Mandl, H.**, (1985). - **Diagnosis of knowledge structures in text learning**. Technical Report Deutsches Institut fur Fernstudien an der Universitat, Tubingen.

**Bartlett, F.C.**, (1932). - **Remembering**. Cambridge: Cambridge University Press.

**Coulombier, G.**, (1987). - **Chronopsychologie des tâches de compréhension**. Poitiers: Technical Report, Laboratoire de Psychologie.

**Denhière, G.**, (1986). - Connaissances générales et connaissances linguistiques dans la production de textes. Communication au 3ème **Colloque International de Didactique du Français**, Namur.

**Ehrlich, S. and Rossi, J.P.**, (1986). - Exploration oculaire et opérations cognitives de compréhension lors de la lecture d'un texte. **L'Année Psychologique, 86**, 63-82.

**Frederiksen, C.H.**, (1986). - **Constructing and manipulating multi-level semantic representations in comprehending and producting text**. Paper presented at the "Text and text processing", Symposium, Poitiers.

**Hubel, D.E., Wiesel, T.W.**, (1962). - Receptive field binocular interaction and functional architecture in the cat's animal cortex. **Journal of Physiology, 160**, 106-154.

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

**Mosenthal, P.B.**, (1985). - Defining the expository discourse continuum: towards a taxonomy of expository text types, **Poetics, 14**, 387-414.

**Rossi, J.P. and Bert-Erboul, A.**, (1987). - **Strategies of information selection in the reading of scientific texts**. Paper read at the Symposium "Acquisition des Connaissances", Tubingen.

**Rossi, J.P.**, (1985). - Les mécanismes de la lecture Paris: Publications de la Sorbonne.

**Rumelhart, D.E., Norman, D.A.**, (1983). - **Representation in memory.** Technical Report, Center for Human Information Processing, University of California, San Diego.

**Trabasso, T. and Sperry, L.L.**, (1985). - Causal relatedness and importance of story events. **Journal of Memory Language, 24**, 595-611.

**Turner, A., Greene, E.**, (1977). - **The construction and use of propositional text base.** Technical Report, n°63, University of Colorado.

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

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# CONSTRUCTING MULTIPLE SEMANTIC REPRESENTATIONS IN COMPREHENDING AND PRODUCING DISCOURSE*

Carl H. FREDERIKSEN and Janet DONIN

Laboratory of Applied Cognitive Science
Department of Educational Psychology
McGill University
Montréal, Canada

## 1. INTRODUCTION

It is well established that discourse communication involves semantic representations --of the "global" knowledge structure that is to be communicated, and of explicit meanings that are encoded locally in the sentences that comprise a discourse. In comprehension, a reader or listener must "extract" these explicit sentence meanings ("propositions") and use them as a basis for constructing a representation of the knowledge structure that is to be understood. This construction of a knowledge representation from a sequence of propositions encoded in discourse involves a number of steps or "levels." Each of these involves particular types of semantic representations and processes. The specification of these representations and the processes associated with them is a central problem for cognitive theories of discourse communication.

This chapter describes recent developments in semantic representation theory for knowledge communicated by discourse. It also discusses implications of these developments for modeling of cognitive processes that are involved in the generation of semantic representations in text comprehension and production. The developments reported have their origins in two approaches to research on natural language comprehension. The first was concerned with the problem of defining a model for a semantic, i.e., *propositional*, representation of natural language discourse as a component of a model of discourse comprehension and semantic memory representation (Frederiksen, 1975a; Kintsch, 1974). The second was concerned with explor-

---

* This research was supported by grants from the Natural Sciences and Engineering Research Council (Canada), the Ministry of Higher Education and Science (Québec), and the Fonds FCAR (Québec).

ing *constructive* operations in text understanding (Bartlett, 1932; Frederiksen, 1975b; Spiro, 1980). Constructive operations often have been characterized as "synthetic" or "knowledge-based," as opposed to "analytic" or "text-based." They involve the generation of semantic representations by readers or listeners based on a variety of information sources including text, prior knowledge, and situational or other sources. The developments that will be reported may be seen as an attempt to elaborate these two perspectives on discourse processing by, first, providing explicit models of semantic representations at several levels, and second, specifying constructive processing operations for the generation and manipulation of these representations.

First, a theoretical approach to semantic processing of discourse will be outlined in which semantic representations of discourse are generated at different levels by the application of rules. Semantic representations are viewed as expressions in formal symbolic languages that are defined by syntactic rules (just as programming languages are defined by syntactic rules). In comprehending discourse, a reader generates symbolic representations (e.g., propositions, frames) that are expressions in an abstract symbolic representation language by applying rules that define the language, or that transform expressions in the language. These structure-generation and transformation operations are seen as occurring at multiple levels as a reader proceeds from local sentence meanings to a global knowledge structure. In production, the same operations are seen as occurring "in reverse" as global domain knowledge is converted into propositional knowledge which in turn is expressed in natural language discourse. This theoretical approach to discourse comprehension and production may be characterized as a "rule-based" approach, in contrast to a "frame-based" approach (in which natural language expressions are analyzed or produced by retrieving and modifying semantic structures, i.e."frames" or "schemata," that were previously stored in memory).

Second, the problem of defining a *semantic representation language* for a particular level or type of semantic representation of discourse will be addressed. An approach will be advocated in which (1) a semantic representation language is defined by means of a set of recursive rules that constitute a transition network grammar for the language, and (2) *semantic representations* are defined as structures generated by application of rules in such a *semantic grammar*. One advantage of this approach is that rules that define a representation language also provide a model of processes by which propositions may be generated for natural language expressions and by which global semantic representations of discourse may be generated in comprehension. Furthermore, the same rules may be used to describe processes of structure generation at various levels in discourse production. Semantic grammars thus provide a unified theory of structure generation applicable across different situations (e.g., production, comprehension) and levels of symbolic representation.

Semantic grammars will be described at two levels: (1) *propositions* (Frederiksen, 1986), and (2) supra-propositional levels of semantic representation called *frames* (Minsky, 1975). Propositions represent chunks of semantic information in working memory that are normally communicated within a clause or sentence and are manipulatable by logical inference operations; frames are more general semantic network structures that represent connected knowledge structures in long term memory.

The final sections will describe examples of applications of these models in experimental research on discourse comprehension and production. The models also have been applied to other situations of semantic information processing such as knowledge acquisition and integration (Kubes, 1988), and knowledge-based reasoning and problem solving (Patel & Groen, 1986; Joseph & Patel, 1986). The problem in accounting for successful performance and learning in all of these "knowledge rich" domains is to specify what rules subjects (at various levels of knowledge and proficiency) apply to generate semantic representations, express them in language, and apply them to achieve communicative, learning or problem-solving goals.

## 2. RULE-BASED PROCESSES IN TEXT COMPREHENSION

When cognitive research on text comprehension became active in the early 1970's, attention was focused on the propositional content of text and the processes that a reader or listener applies to a text to generate propositions from the text. Influenced by new work in case grammar (Fillmore, 1968) and semantics (e.g., Chafe, 1970; Leech, 1970; Simmons, 1972), a model of comprehension was proposed in which a reader or listener was viewed as generating propositions, i.e., abstract representations of semantic relations among concepts, to represent the meaning of a text in memory. Propositions were not limited to sentence meanings but also included inferred propositions as well (Bransford & Franks, 1971; Frederiksen, 1975b). This *text-based theory* viewed comprehension as a process of generating propositions by: (1) interpreting the literal semantic content of sentences, or (2) making inferences that connect, complete, summarize, re-interpret, or elaborate the literal propositional content of a text. An important component of the theories developed at this time, therefore, was the specification of propositional representations, either as semantic network structures (Frederiksen, 1975a) or (equivalently) predicate-argument structures (Kintsch, 1974). A second component was an attempt to specify types of inferential operations that could be applied to a "propositional text base" to generate inferred propositions that link or elaborate text propositions, or "macro-propositions" that summarize or represent text -level semantic structure (Kintsch & Van Dijk, 1978).

However, it soon became apparent from analyses of readers' inferences that text-based theories alone could not solve the problem of explaining why readers' text recalls contain particular inferences and not others (e.g., Anderson & Ortony, 1975). It was as if a reader's selective inferences were being directed by some "top-down" strategy that specified in advance what types of information the inferences were supposed to generate. A theory of top-down processing was proposed in which text comprehension was viewed as a kind of pattern-recognition process in which text propositions are "fit" to pre-existing relational knowledge structures in memory called "frames" (Minsky, 1975), "schemata" (Rumelhart, 1980) or "scripts" (Schank & Abelson, 1977). Frames are defined as specific semantic network structures in memory which contain variables. In comprehension, frames are retrieved from prior knowledge pertinent to a text's content, and "instantiated" by replacing their variables with particular text-derived information (e.g., concepts, propositions) or default values. The resulting *text frames* are representations of specific declarative knowledge structures that can be used to link a text's propositions into a single coherent semantic structure. Thus, in *canonical frame theory*, a reader is seen as fitting incoming semantic information to pre-existing frames in memory while in the process modifying these frames as necessary to fit the incoming "data".

As a consequence of frame theory, semantic representations now were thought of as occurring at two levels: (1) a *propositional level* representing chunks of semantic information typically encoded in natural language expressions, represented in working memory, having modality and asserted truth-value, and appropriate for potential logical inferences (Frederiksen, 1985; Van Dijk & Kintsch, 1983); and (2) a *frame level* representing connected semantic network structures composed of nodes and relations linking these nodes into meaningful structural networks. The mechanism postulated for generating text frames was to apply pre-existing "canonical frame structures" in memory as templates into which text information could be introduced. This same process also frequently has been proposed as a mechanism for propositional interpretation of sentences (Sowa, 1984).

One problem with canonical frame theory is that frame-instantiation processes require specific content knowledge and hence are applicable only to texts in familiar content domains. If a reader lacks specific prior knowledge relevant to the semantic content of a text (as would be the case for expository texts used as sources of new knowledge), the reader would lack relevant prior knowledge frames to fit to the text and therefore would have to rely on general purpose text-based processing to understand the text. Put differently, canonical frame theory provides an explanation of how prior knowledge may be used to interpret new situations, but it does not provide an account of how frame representations were originally generated or how they are generated in the absence of relevant prior knowledge.

In our laboratory, we have been investigating children's comprehension of unfamiliar stories and expository texts that convey new information on unfamiliar topics (Frederiksen, 1988). Our research suggests that in such situations children do not rely on general text-based processing abilities. Instead, they comprehend unfamiliar expository and story texts by using selective inference strategies to generate frame representations of a text in the absence of relevant prior knowledge. Furthermore, they appear to have specialized competency in understanding particular *types* of frame structures expressed in texts such as: narrative event structures, procedures, problems, and conversational structures. Some children may be consistently better able to comprehend narrative structures, while others may display an ability to understand problems or procedures. Furthermore, this competency appears not to be associated with specific content knowledge. For example, children display differential ability to understand the structure of problems irrespective of whether a problem is a part of a plot line in a story, or the subject of an expository text in science.

To account for such specialized comprehension abilities, we have been investigating a third type of text processing that may be termed "rule-based." In a *rule-based theory* of text processing, a person comprehending a text is viewed as possessing knowledge of rules for forming or transforming particular types of semantic structures expressed in discourse. These rules may be characterized as *semantic grammars*. In our research, semantic grammars are sets of language-independent rules for generating expressions within a particular type of semantic representation language. Comprehension is assumed to consist of parsing a text at different levels by applying the rules of a semantic grammar in top-down fashion to generate semantic representations. This occurs both at a propositional level, and at higher-levels in which text frames are constructed. The rules in our grammars are written in a simple notation taken from computer science known as a "BNF grammar" (Wirth, 1976). Since there is a recursive transition network (RTN) parser for every BNF grammar, our semantic grammars have particular semantic parsing algorithms associated with them. These parsing algorithms may be explored as possible models of rule-based text processing.

Since it captures the ability to generate novel representations, a rule-based model of semantic processing explains what in generative linguistics is called "generative capacity." Just as the problem for psychological theories of language competence is to account for the capacity to generate new or novel utterances, a similar problem for psychological theories of semantic processing is to account for the capacity to generate new or novel semantic representations. A psycholinguistic model of language competence that lacked knowledge of syntactic rules would have to rely on a large store of sentence patterns (a so-called "slot grammar"). In the same way, a theory of semantic representation that lacks generative rules has to rely on a large store of semantic patterns or frames.

This type of theory of discourse processing has its precursors in research on "story grammars" (Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn,1979; Thorndyke, 1977) and in research investigating the construction of particular structures for stories such as event chains (Trabasso & Nichols, 1978) and characters' plans (Bruce, 1980; Newman & Bruce, 1986). There were two main problems with the story grammar models: first, the decision to attach rules of the grammar to text categories such as "stories" (Black & Wilensky, 1979; Brewer & Lichtenstein, 1981; Beaugrande & Colby, 1979); and second, the failure to specify just how rules are applied to a story to "parse" its structure.

In the research which will be described, frame grammars are assumed to represent abstract representation languages and are not attached to particular types of texts. Texts, such as stories, often represent more than one type of frame structure. For example, one story which we studied could be understood in terms of a conversational frame, a narrative event structure of which the conversation was an important part, a problem structure which represented the situation the characters had to resolve, and a procedure frame which represented the plans underlying the characters' talk and actions. All of these semantic representations are necessary to understand the story, and children at different ages and levels of proficiency showed differential ability in understanding these different aspects of the story structure (Frederiksen, Donin, Kormos & Browns, in prep. ).

## 3. SEMANTIC GRAMMARS AND DISCOURSE ANALYSIS

Both propositions and frames are instances of *semantic networks*, or, alternatively, "conceptual graphs" (Sowa, 1984). These are general relational structures for data representation that are composed of *nodes* (that usually contain concepts or pointers to other semantic network structures) and *links* (labeled arcs or arrows pointing from one node to another). Propositions, which frequently are represented by a "predicate + arguments" formalism, are equivalent to semantic networks. As will be seen in the example that follows, they also are equivalent to *parse trees* .

Propositions are well-formed expressions in a *propositional language* , and frames are semantic network structures that are conjunctions of linked well-formed expressions in a *frame representation language*. Frame representation languages may be defined to represent particular types of declarative knowledge in long-term memory. These include: descriptions, narrative-event structures, declarative representations of procedures, problem representations, and dialogue structures. Representation languages are defined by *semantic grammars.*

When semantic grammars are applied to analyze discourse, the resulting analyses reflect hypotheses about semantic representations that may be generated by subjects in comprehending the discourse. Current multi-level models of discourse comprehension assume that representations of the following levels of structure are generated in comprehension (Frederiksen, Bracewell, Breuleux & Renaud, 1989):

a) *linguistic structures:* representations of language structures in discourse including: (1) sequences of words and morphemes, (2) syntactic structures (syntactic parse trees), and (3) syntactic dependency graphs, i.e., representations of relations between constituents of different syntactic tree structures;

b) *propositions:* representations of semantic information (i.e., relations among concepts) explicitly encoded in natural language structures, plus propositions generated by inferential rules of various kinds such as local coherence inferences, propositions derived by means of logical operations, and macro-structure inferences (propositions are thought of as "intermediate representations" that mediate between natural language and conceptual frame information);

c) *semantic network (conceptual frame) structures:* integrative processing operations may be applied to construct descriptive semantic networks from propositions to link semantic information into a connected memory structure, and rules in a frame grammar may be applied to generate semantic networks when particular types of knowledge structures are required to understand discourse.

In comprehension research, texts presented to subjects are analyzed in terms of these different levels of representation. The resulting representations of the text are used to predict measures of subjects' comprehension such as recall of text information, inferences, on-line interpretations of a text, and reading times (Frederiksen and Renaud, 1987, in prep.). Also of interest are interactions of these variables with developmental and individual differences in predicting comprehension performance measures. Thus, discourse analyses reflect hypotheses about the representations and rules that are applied by competent readers in comprehending a text. A well-developed theory of discourse representation and processing leads to a well-defined discourse analysis procedure (and vice-versa).

We will illustrate semantic discourse analysis with a text employed in a study of children's comprehension of a narrative account of a problem and its solution in science (Frederiksen, Donin, Kubes & Browns, in prep.). This text consisted of an account of the discovery of Penicillin by Alexander Fleming.

FLEMING TEXT (first paragraph)

*Alexander Fleming was working in his laboratory at Saint Mary's Hospital in London, England. It was a warm day and the windows were open. A dish of disease germs, needed for an experiment, had been left uncovered. It looked like cloudy soup. As Fleming walked by he glanced at the dish. Something caught his eye. He looked again. (P1.1) There was a patch of blue-green Penicillium mold growing on the dish. But instead of cloudy soup, thick with germs, there was a clear circle around the mold. All of the germs near the mold were dead! Fleming knew that this was something important. (P3.1) He scraped off the bit of mold and (P3.2) put it in a dish of its own with some food. The mold plant spread and the blue-green patch grew larger. (P2.2) When it was big enough, (P3.3) Fleming set to work to (P1.1) find out why the mold had killed the germs. (P3.4) After many weeks of difficult experiments, (P3.5) he finally managed to squeeze from the mold a few drops of brownish fluid. (P2.2) This remarkable fluid was the germ-killer. (P3.6) Fleming named it penicillin.*

*Propositions* are generated for each sentence by applying rules in a propositional grammar in top-down depth-first fashion. This will be illustrated for the first sentence of the text. The analysis is done with the assistance of a computer program (called CODA) which interprets rules in any BNF grammar, presents a user with requests for information and decisions as to which rule to apply, generates a graphic parse tree on the screen representing the rules that were applied, and stores a list representation of this tree in a file. The set of propositions for a text defines the "literal propositional content" of the text.

Propositional analysis with CODA produces a propositional representation in the form of a tree-structure. To illustrate, the complete parse tree for the first proposition generated for sentence 1 of the Fleming text is given in Figure 1. Each non-terminal node in the tree refers to a rule which was applied and "leaves" (terminal nodes) are *concepts* (e.g., "work"), *literals* specified by the grammar (e.g., "PAST"), or *proposition numbers* (e.g., "1.1").

The tree-structure in Figure 1 may be represented more simply by deleting certain nodes and by representing some of the remaining nodes as relations. When this is done, the resulting tree is a semantic network (see Figure 2). Nodes in this network are either simple (e.g., *work*) or complex (e.g., OBJECT); complex nodes decompose into other nodes.

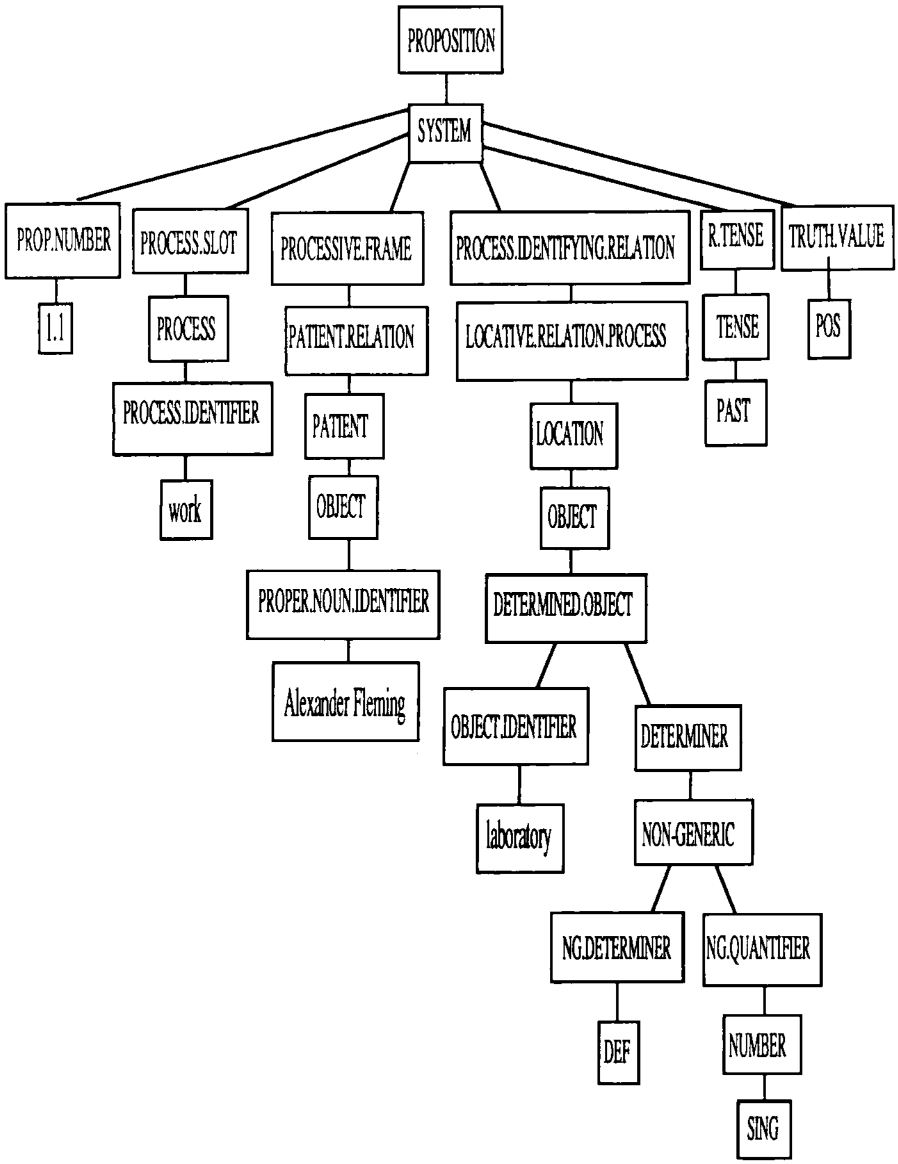



FIGURE 1
Complete parse tree for the first proposition generated for sentence 1 of the Fleming text.

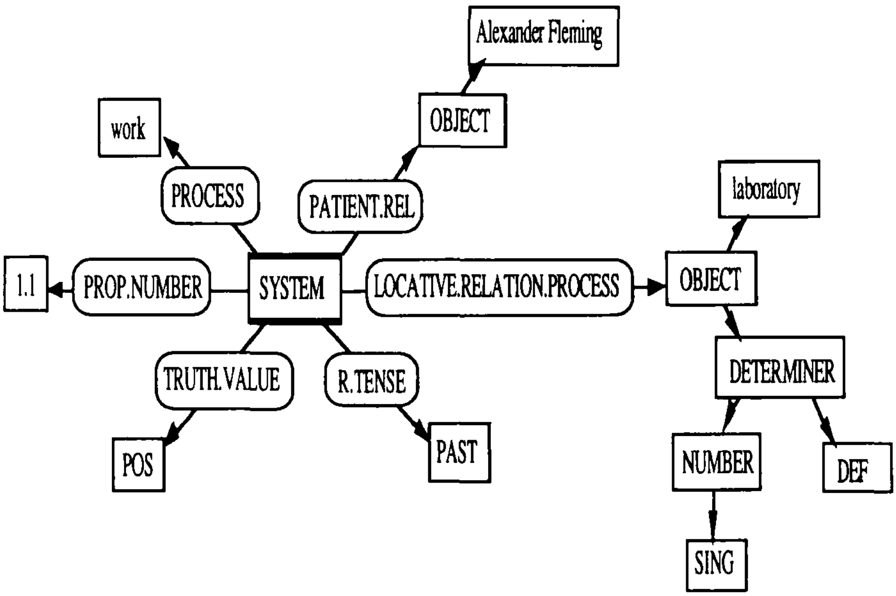



FIGURE 2
Semantic network representation of proposition 1.1 (simplified parse tree for the first proposition generated for sentence 1 of the Fleming text ).

Finally, the semantic network of Figure 2 may be represented in terms of its equivalent list structure as follows (the relations in Figure 2 have been represented as boldface) :

<u>PROPOSITION (SYSTEM (PROP.NUMBER (1.1)</u> **PROCESS** (work) PROCESSIVE.FRAME (**PATIENT.REL** (OBJECT (Alexander Fleming))) PROCESS.IDENTIFYING.RELATION (**LOCATIVE.RELATION.PROCESS** (OBJECT (laboratory DETERMINER (DEF NUMBER (SING))))) **R.TENSE** (PAST) **TRUTH.VALUE** (POS)))

If the underlined part is interpreted as a *predicate*, the remainder of this expression may be interpreted as an *argument structure* in which argument slots are labeled by the boldface relations. This list representation is transformable into a quantified formula in predicate calculus with types ascribed to predicates and arguments. If slots in the list were to contain embedded propositions, a second-order logic would be required. Note that unlike standard predicate-calculus notation, determination and quantification are represented as properties of objects (variables) occupying slots. It has been demonstrated that this form is equivalent to the standard prefix notation for quantification (Simmons, 1972; Sowa, 1984). Since the propositional grammar was developed to represent semantic forms expressed in natural languages, it permits unrestricted embedding of propositions in argument slots and includes complex modal categories of truth value. Therefore, a calculus for handling natural language inferences is much more complex than a first-order two-valued logic (Jackendoff, 1983).

The "hand" generation of propositions (with the assistance of CODA) is in fact "hand" semantic parsing of sentences (analogous to "hand" syntactic parsing). Frederiksen & Décary (1989; Frederiksen, Décary & Hoover, 1988) currently are developing a computer system for automatic semantic parsing of English sentences according to a semantic grammar. Propositions are generated by top-down application of rules in a propositional grammar. When a rule is applied, a series of tests are made on a syntactic parse tree for the sentence (generated automatically by a deterministic parser). Each test corresponds to a syntactic pattern that can encode a semantic category. The system generates both semantic parse trees and simplified predicate-argument representations. The feasibility of the system has been established and rules currently are being written to cover an increasing range of syntactic and semantic structures. This system serves both as an automated system for semantic analysis of discourse and as a computational model of rule-based comprehension processes.

*Descriptive frames* may be generated from the propositional representation of the text by representing propositions as semantic networks (as in Figure 2) and linking these into connected network structures. Before the semantic network can be generated, local inference operations on propositions (e.g., slot filling, replacement of anaphoric elements with their referents) may be necessary. The resulting linked semantic networks provide a highly detailed description of the content of a text. Research on integrative processes in memory has established that such processes are required to generate connected representations of knowledge derived from texts and to link new knowledge to prior knowledge structures in memory (Yekovich & Walker, 1986; Hayes-Roth & Walker, 1979; Clifton & Slowiaczek, 1981).

To analyze texts that represent specialized frame structures other than purely descriptive networks (such as the Fleming text), a text is parsed using frame grammars (cf. Frederiksen, 1986). We have identified the following types of frames in our research on comprehension of expository texts in science: narrative, procedural and problem frames. *Narrative frames* represent temporally, spatially and causally ordered event structures (Frederiksen, Donin-Frederiksen & Bracewell, 1986). *Procedural frames* are declarative semantic representations of the structure of procedures (Frederiksen, 1989). *Problem frames* represent a trace of problem-solving behavior including: (a) descriptions of problem situations (states), (b) procedures that are applied in a problem situation to attempt to solve the problem, and (c) events that occur as problem-solving procedures are planned or applied. Each frame type is defined by a frame grammar that generates components of a particular type of semantic network. A frame is obtained when these components are linked into a connected semantic network.

Frame analysis will be illustrated with reference to the problem frame that was generated for the "Fleming" text (Figure 3).

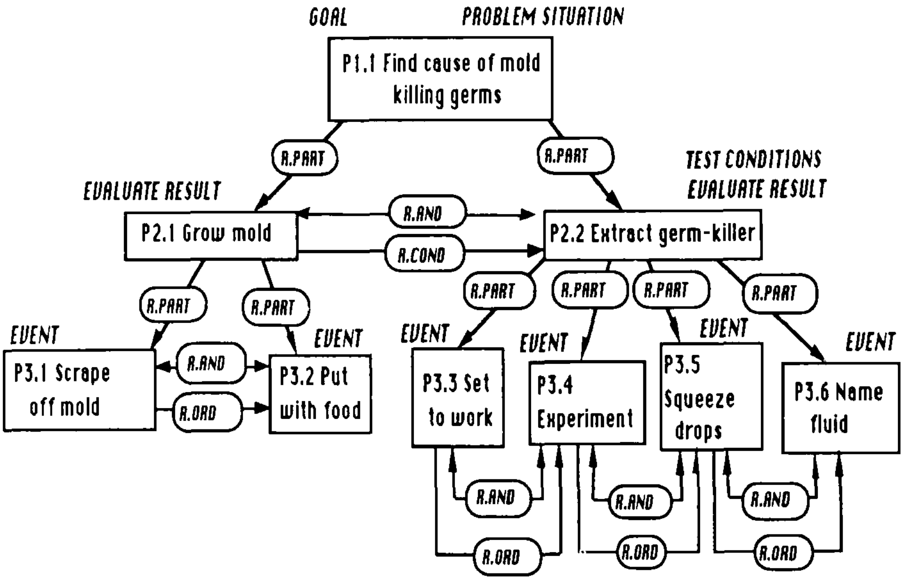



FIGURE 3
Problem frame for first paragraph of the Fleming text.

An important class of texts (of which the Fleming text is a simple example) present descriptions of problems and descriptions of methods for their solution. Problem frame texts may vary in the kind and amount of information about problem-solving that they provide. One kind of problem text, typically found in science or mathematics textbooks, presents only a description of a situation and a goal that is to be attained; the reader has to either apply a known procedure or plan a procedure for obtaining a solution. A second kind of problem text, which we may call "procedural text," presents instructions for solving particular classes of problems in the form of verbal descriptions of procedures (i.e., methods) that may be applied. A third kind of problem text presents descriptions of events that took place in solving a problem. This last type may include as well a description of a problem situation and of procedures applicable to the problem.

The node-link structure in Figure 3 represents a procedural network for the Fleming text. Nodes correspond to procedures (operations that may be executed to achieve goals), and links correspond to relations among procedures. These include: R.PART (a goal is decomposed into a conjunction of sub-goals ), R.CATEGORY (a goal may be attained in more than one way, i.e., it corresponds to disjunctive sub-goals), R.ORD (operations are to be executed in a specific temporal order), and R.COND (one operation is conditional on another being executed). A procedural network represents a hierarchical control structure for carrying out actions (operations) together with temporal and conditional constraints on actions. For example, in the Fleming text, the top-level goal is to find the cause of the mold killing the germs, and this is decomposed into two main sub-procedures (sub-goals), growing a sufficient quantity of the mold and extracting the germ-killing agent. Both of these sub-procedures further decompose into sub-procedures.

A procedural frame also contains semantic information about each procedure node. For each node, the following information is included: (1) the GOAL(s) of the procedure, (2) a description of the SITUATION in which a procedure is to be applied, (3) any DECISION(s) that are required (for disjunctive goals), (4) TESTS to see that preconditions for the application of a procedure are satisfied, (4) EVENT descriptions (including the action, the source state existing prior to an action, the result state existing after the action, the agent, any instruments, etc.), and (5) a description of how the result may be EVALUATED against the goal. These are the main components identified in theories of problem-solving such as the SOAR system (Laird, Newell & Rosenbloom, 1987). In a text describing problem solving, such as the Fleming text, these components of a procedure usually are reflected in different types of events such as: adopting a goal, interpreting a situation, making a decision, carrying out a pretest, enacting or planning a procedural event, or evaluating a result. In the Fleming text, different types of events were described for different nodes. (These event types are marked in Figure 3 in italics adjacent to each node in the procedural network.) For example, the top-level procedure node was represented in the text in terms of goal and situation description events. The middle level nodes were represented in terms of evaluations and, in the case of node P2.2 a pre-test, i.e., "when it was big enough"). Finally, the bottom-level procedure nodes were always represented as enactment events.

## 4. SEMANTIC REPRESENTATION IN TEXT COMPREHENSION

To understand a text such as the "Fleming" text, readers first would have to: (1) generate propositions from sentence structures and lexical information in the text (the literal propositional content of the text), (2) make any inferences necessary to resolve problems of reference (anaphora), elaborate propositions through slot-filling, or generate new propositions (e.g., inferences that connect related propositions, elaborate, or summarize the propositional content of the text); and (3) link propositional meanings into an integrated descriptive frame (i.e., a linked semantic network representation). However, to adequately understand the "Fleming" text, readers also would have to generate a representation of the frame structure that explains the problem Fleming was trying to solve and how he was trying to solve it (Figure 3). The structure of this problem was not previously known to the children. Therefore, to comprehend the problem structure of this text, the children would have to possess knowledge of rules governing the structure of a problem-frame as well as ability to apply these rules to construct the problem-frame representation for this text.

These processes may be studied experimentally by first analyzing an experimental text (such as the Fleming text) and then using the resulting propositional and frame representations of the text to predict subjects' comprehension. Comprehension may be assessed by obtaining subjects' responses to a text such as post-input free recall of text propositions; inferences during reading, recall or responses to questions; on-line "think aloud" responses obtained during reading; and measures of sentence reading times. In the present study, comprehension was assessed using the free recall method.

Twenty-four grade 3 and grade 5 students were asked to read and orally recall the "Fleming" text. Their recall protocols were matched to text propositions and text propositions were coded as: (a) recalled, (b) recalled with slot-filling inferences, and (c) operated on inferentially to generate new propositions. Responses of type (a) reflect literal propositional interpretation of text sentences; responses of type (b) reflect local inferential operations on propositions; and responses of type (c) reflect "deeper" text-level inferential processing. Subjects were grouped by grade and global proficiency level in recalling text propositions (i.e., on the basis of the total number of propositions they recalled). Interest centered on analyzing these response measures as a function of frame variables, grade, proficiency level, and their interactions.

Frame-level semantic processing was studied by carrying out a series of statistical analyses in which properties of the problem frame representation of the text (frame variables) were used to predict responses to text propositions. The assumption was that if structures were being generated at the frame level, then propositions would be selectively processed as a function of their functions within frame structure. If a frame variable was found to produce a main effect, this would provide evidence of selective processing in all groups of children. However, if a frame variable showed a significant interaction with grade, proficiency level, or both, this would provide evidence for developmental or individual differences related to the ability to selectively process text in terms of frame structure.

The results with the "Fleming" passage were as follows.

(1) There was selective processing of problem frame propositions over non-problem frame propositions (i.e., there was a significant main effect of the contrast of frame vs. non-frame propositions). This effect was present in all groups (i.e., there was no significant interaction of the frame vs. non-frame effect with grade or level).

(2) Propositions representing goals and procedural events were more likely to be recalled than problem situation descriptions (i.e., both of these contrasts were significant). However, more high-level inferences were made in generating representations of problem situations than of goals or procedural events (i.e., there was a significant interaction of this frame variable contrast with response type).

(3) There was selective processing of problem frame propositions associated with different sub-trees in the procedural network for the text . These sub-tree effects were as follows.

a) Sub-trees were selectively recalled (i.e., there were significant main effects of sub-tree contrasts).

b) There were differential degrees of inferential processing of sub-trees of the procedural network (i.e., there were significant interactions of sub-tree effects with response type).

c) There were developmental and proficiency level differences in selective processing of sub-trees in which sub-tree effects were larger in groups at higher grade or global proficiency levels (i.e., there were significant interactions of sub-tree effects with level and grade).

These results establish that text propositions are selectively processed. The likelihood that they will be recalled is a function of whether they represent problem-frame information, the type of problem-frame information they represent, and its location within the procedural network. Thus, subjects have been shown to selectively process text propositions in terms of their roles in the underlying problem structure. This selective processing also involves differential amounts of inferential processing. Finally, selective processing of propositions in terms of their frame roles increases developmentally and is more prevalent in children at higher levels of over-all proficiency.

The above results are consistent with those of other studies of children's comprehension of stories and expository science texts (Frederiksen, 1988). These studies also have found selective processing of propositions related to their frame roles. Frame variable effects (a) were specific to particular types of frames, (b) involved differential inference, and (c) were more pronounced in more proficient subjects and subjects who were at a higher grade level. We also have analyzed individual subjects' protocols to study in detail the frame structures subjects generate for a text. This procedure results in a kind of "cognitive diagnosis" of text processing for individual subjects (Frederiksen & Breuleux, 1989). The evaluation of individual subjects' protocols will be illustrated in the next section in reference to discourse production tasks.

In our current research, we are investigating adults' comprehension and acquisition of complex procedures and declarative (i.e., theoretical) knowledge from texts in science and technology, and studying their application to solving problems (Frederiksen & Renaud, 1989). We are interested in the relationship between subjects' prior procedural knowledge, their relevant theoretical knowledge, their comprehension of a problem situation, and the dependence of their solution processes on these representations (Frederiksen & Breuleux, 1989). Since the procedural frame directly reflects a subject's understanding of procedures and their control structures, the quality of this representation ought to be an important determiner of the processes a subject applies in solving a problem (cf., Kintsch & Greeno, 1985).

## 5. GENERATION OF SEMANTIC REPRESENTATIONS IN TEXT PRODUCTION

In text production, we expect the same multi-level semantic representations and language structures to be generated as in text comprehension, but the direction of the process is reversed. We expect the text structures that are produced to be related to semantic structures by means of multiple mappings. These reflect multiple levels of representation and processes a writer uses to generate text to communicate frame-level semantic information (Frederiksen, Donin-Frederiksen & Bracewell, 1986). To do this, a writer has to carry out a number of component processes associated with the generation of various levels of semantic and language structure. These are:

a) *frame generation* : generation of a frame structure based on information retrieved from the writer's store of relevant prior knowledge frames;

b) *chunking and sequencing frame information* : selection of chunks of frame information in a particular sequence (the sequence reflects a particular search strategy);

c) *generation of propositions* : generation of propositions to represent these chunks;

d) *syntactic encoding* : encoding of propositions as sentence structures: choosing lexical and syntactic encodings that are required by the text-level structure;

e) *generation of text structure* : generation of text structure including control of syntactic and lexical encoding to satisfy criteria of local (linear) coherence, cohesion, and topical organization. Topical organization is constrained by local (linear) sentence context and by (global) frame-level semantic content.

The writer is caught between two sets of constraints: the non-linear structure of the frame to be communicated, and the linear structure of a text and its requirement to maintain coherence between propositions (Frederiksen & Dominic, 1982). The reader has the problem of inferring the writer's frame given only its sequential representation in text. Thus we expect the semantic and linguistic structures of texts to be related in ways that reflect the cognitive processes of writers and readers --not by simple mapping rules.

One way in which we have been studying these processes experimentally is by examining students' strategies for retrieving and organizing their declarative knowledge in writing instructions in a delimited domain of knowledge (how to use a word processing program) (Donin, Bracewell, Frederiksen & Dillinger, 1989). There are at least two types of knowledge that one must have in order to produce a set of instructions that another individual can use. The first type of knowledge is of "what to do" --knowledge of the actions required to do the particular task. The second type of knowledge is of how to organize that sequence of actions. In order to describe or carry out a procedure, one must have within one's representation not only knowledge of the *actions* required to accomplish particular task-specified *goals* using the procedure, but also knowledge of how to *organize* that sequence of actions to achieve goals, and *apply* procedural actions to specific situations of application.

To study the students' ability to reflect both types of knowledge in their writing we began by constructing an "expert model" or frame representation of the procedural knowledge required to use those aspects of the word processor for which the students were writing instructions. To do this, we began with a manual that summarized procedures for using the *Bank Street Writer* and modified it to match the procedures required for the computers the students actually were using. From this text we constructed a graphic procedural frame, consisting of three sub-trees: Start-up, Write and Edit, each of which, in turn, have their own sub-procedures. The content of the nodes describe the required actions, and the manner in which the nodes are linked gives the organizational structure of the frame.

A student's text was analyzed by identifying information from the procedural frame present in that text. The sequence in which the nodes and links were produced was recorded, thus allowing us to study both the "expert frame" information that a particular student employed and the linear sequence that was used with respect to that information. We will illustrate these analyses with the following text produced by one student.

BANK STREET WRITER

*Bank Street Writer is a program that allows you to correct mistakes, add more words, delete words, erase words, move paragraphs around in a text that you have typed in, without having to type the text over. (1)/ Bank Street Writer is like a word processing. (2)/ B.S.W. is short for Bank Street Writer. (3)/ When you put the B.S.W. in the disk drive and turn on the computer (4)/ it will ask you for the date, starting from the month, date and year. (5)/ Then it will ask you for the time( 6)/ after you have pressed the return key after everyone of your replies.(7)/ You will see a title cover on the screen (8)/ and then you will be in the writing mode.(9)/ On the top of the screen you will see instructions. (10)/ Before typing in your input you should read the instructions. (11)/ After you have finished typing your text (12)/ you can press ESP (13)/ and that will get you to the editing functions. (14)/ You get to the function that you want by using the arrow key underneath ESP.(15)/ To save you highlight SAVE (16)/ and press return.(17)/ On the screen there should appear questions like what do you want the file to be called and etc.. (18)/ Answer the questions (19)/ and then your input should be saved (20)/ if you have your disk in the second disk drive. (21)/*

An example of the comparison of the organization of semantic information in part of this student's text with that represented in the "expert frame" is given in Figures 4 and 5. The shaded nodes in the "expert frame" are present in the student text. The numbers correspond to the syntactic segments in the student's text in which the coded information is found. Additional procedural information present in the student's text was added to the frame (nodes and links for segments 7, 8, 9, 10, 11, and 12; as well as the top-level node for segment 2) as was any non-procedural (descriptive) information directly linked to procedures (segments 8, 9 and 21)). Thus, the manner in which the writer selected frame information and the linear sequence used are reflected in the Figures.

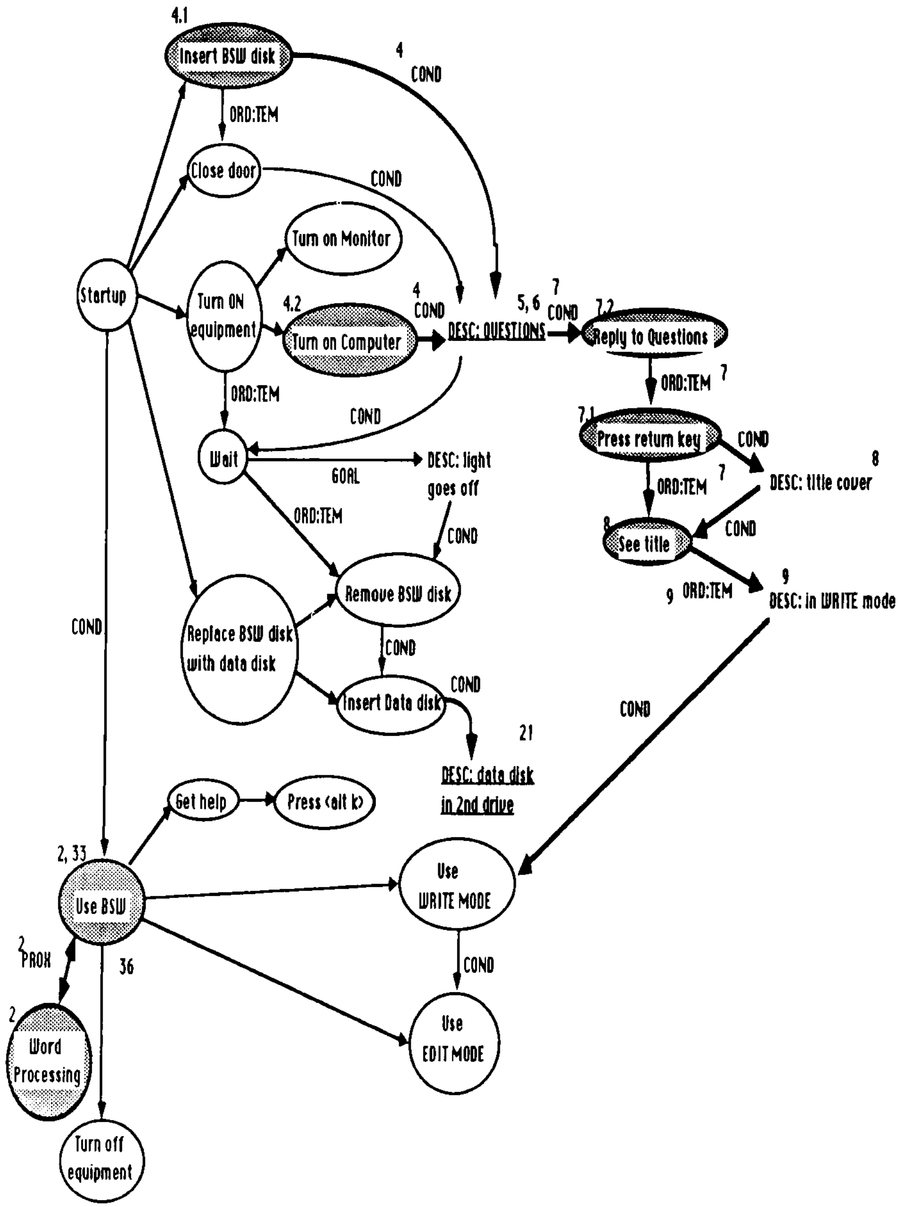



FIGURE 4.
Coded protocol of student for "Startup" procedure.

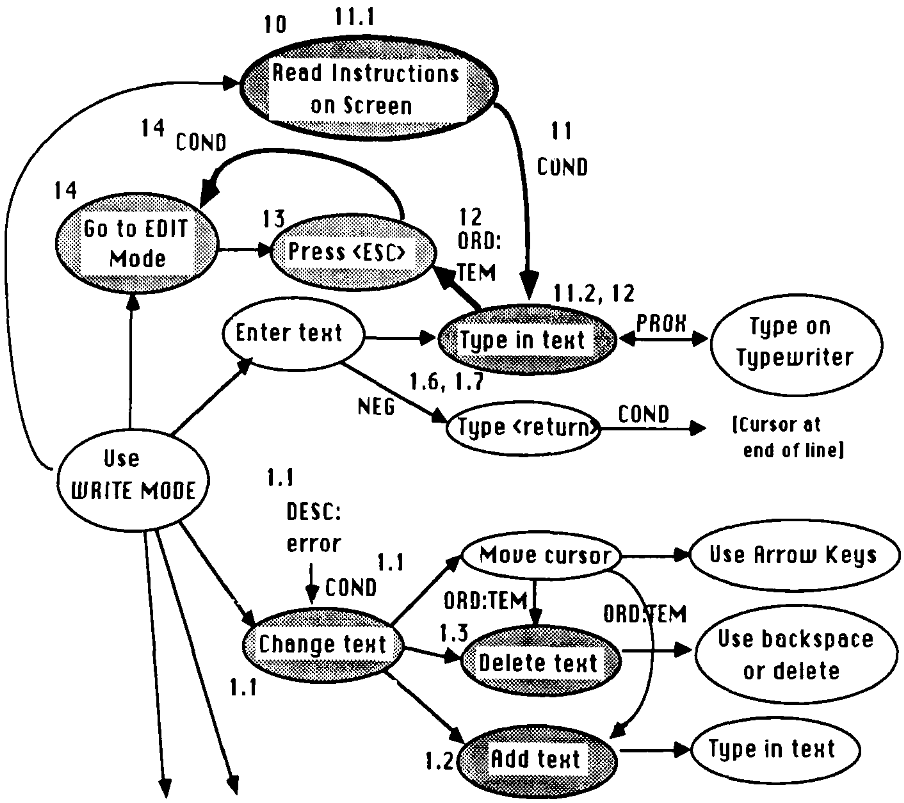



FIGURE 5.
Coded protocol of student for "Write" procedure.

This text was selected as an example of a text in which evidence of the use of the procedural frame was quite good. However, as can be seen, this student's presentation of the procedural frame varied over the different parts of the structure. Within the Start-up procedure the top level node was only implicit. Both procedural information and especially descriptive information pertaining to questions presented on the screen were added along with appropriate links into frame nodes. The information the student presents is not identical to that presented in our "expert frame;" however, it can be seen that it partly takes the form of a procedural hierarchy and partly the form of a temporally-ordered list. Write Mode information is presented in two unconnected parts of the student's text. Part of the Change Text hierarchy is presented in the first segment and this is not in any way linked to the other procedures in Write Mode presented later in the text. These latter procedures, however, are linked both by proximity in the text and by explicit semantic links (such as order in time) to give a well-defined procedural hierarchy.

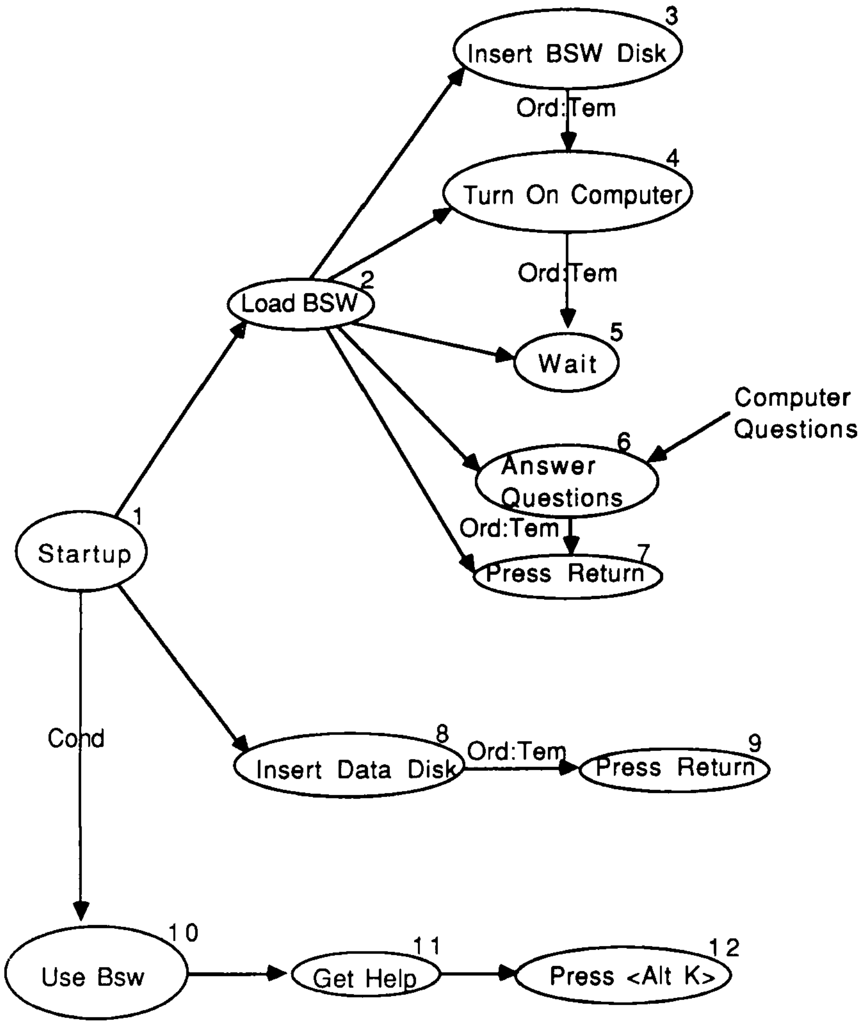



FIGURE 6.
Composite student frame for Bank Street Writer - part 1.

In order to analyze the group data we produced a subset of the "expert frame" that we could call the "expert student." This frame, of which the first part may be seen in Figure 6 gives all the nodes present in at least two student texts. Nodes not appearing in the "expert frame" but included in at least two of the student texts are also included in this diagram; e.g., node #6: answer questions, not considered necessary by the manual writer(s) but definitely considered necessary by the students. The nodes are numbered for the purposes of data analyses.

Our analyses of both individual protocols and group data, using these procedures, provide evidence that, in general, our subjects' writing was driven by their knowledge of the hierarchical organization of the procedural frame. For example, although certain branches in the tree structure were missing, those present were given together and in hierarchical form. In general, the information most often bypassed was that of the "twigs," i.e., the specific, detailed information. Our analyses revealed that our students' texts were generally insufficient with respect to the other type of knowledge, i.e., specific content knowledge. Unfortunately, this study did not allow us to differentiate whether this was the result of these writers assuming too much knowledge on the part of their audience or whether the writers actually did not possess the requisite knowledge. We are currently conducting research in which we are manipulating the information our subjects have about their audience in order to assess these context effects.

Our semantic representation models and analytic techniques based on them allow us to focus on exactly where individual differences in expository writing occur with respect to domain knowledge and its use in written communication. Our results using these tools support the need to study writing from the viewpoint of the knowledge structures being expressed. The product of the writing process is the result of an individual's operating on those knowledge structures to produce a text that is consistent not only with general text cohesion and coherence constraints but also consistent with the structure of knowledge in the domain. The written product, therefore, reflects operations on semantic information to retrieve, select, elaborate, modify, and generate semantic frame representations; "package" this information into propositional units; and organize and encode propositions into a discourse structure. An important goal for future research will be to develop and test explicit models of these operations in discourse production.

## 6. CONCLUSIONS

The techniques that we have presented had their origin in text comprehension research; however, they have been applied successfully not only to studies of text production but also to studies of learning and problem-solving. The general model of comprehension and production within which these analytic techniques have been developed recognizes that these processes are both modular and multilevel. While processing at one level is viewed as operating on the outcomes of processing at other levels, processing is not necessarily sequential. Processes may be operating in parallel with individual processes operating whenever they have data. For example, in comprehending a discourse in a particular situation, a reader or listener constructs a global knowledge representation from propositions derived from lexical and syntactic information. At the same time, knowledge representations that the reader or listener already possesses may be used both in the construction of a text representation and in processing at lower levels. In addition, strategic control may be exerted to organize the comprehension process (VanDijk & Kintsch, 1983).

Our research results have supported this modular multilevel process model. In addition, they have supported the presence of rule-based processing in both comprehension and production. The general model is not restricted to comprehension and production. Learning occurs when new declarative or procedural knowledge is acquired or existing knowledge is transformed or updated. Problem solving occurs in any situation, well-structured or not, in which knowledge is applied to achieve goals. The development of formal models for representing knowledge at multiple levels can allow us to specify more completely the cognitive processes involved in performing complex tasks within and across content domains.

## REFERENCES

Anderson, R.C. & Ortony, A. (1975). On putting apples into bottles --a problem of polysemy. *Cognitive Psychology, 7*, 167-180.

Bartlett, F. (1932). *Remembering*. Cambridge: Cambridge University Press.

Beaugrande, R. de & Colby, B.W. (1979). Narrative models of action and interaction. *Cognitive Science, 3*, 43-66.

Black, J.B. & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science, 3*, 213-230.

Bransford, J. & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 3*, 193-209.

Brewer, W.F. & Lichtenstein,E.H. (1981). Event schemas, story schemas, and story grammar. In A.D. Baddeley & J.D. Long (Eds.), *Attention and performance, IV*. Hillsdale, NJ: Erlbaum.

Bruce, B. (1980). Plans and discourse. *Text, 3*, 253-259.

Chafe, W. (1970). *Meaning and the structure of language.* Chicago: University of Chicago Press.

Clifton, C., & Slowiaczek, M.L. (1981). Integrating new information with old. *Memory and Cognition*, *9*, 142-148.

Donin, J., Bracewell, R. J., Frederiksen, C. H. & Dillinger, M. (1989). *Students' strategies for writing instructions: Organizing semantic information in text* . Manuscript submitted for publication.

Frederiksen, C.H. (1975a). Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology, 7,* 371-485.

Frederiksen, C.H. (1975b). Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology, 7,* 139-166.

Frederiksen, C.H. (1986). Cognitive models and discourse analysis. In C.R. Cooper & S. Greenbaum (Eds.), *Written communication annual: An international survey of research and theory, Vol. I: Studying writing: Linguistic approaches.* Beverly Hills, CA: Sage.

Frederiksen, C.H. (1988). Exploring cognitive processes in text processing within functional contexts. In D. Bloom (Ed.), *Functional literacy activity in education: Cognitive and social approaches.* Norwood, NJ: Ablex.

Frederiksen, C.H. (1989). *The representation of procedures: Acquisition and application of procedural knowledge.* Presented at Colloque Internationale sur l'Informatique Cognitive des Organisations, June 12 - 15, Québec, Canada.

Frederiksen, C. H., Bracewell, R. J. Breuleux, A. & Renaud, A. (1989). The cognitive representation and processing of discourse: Function and dysfunction. In Y. Joanette & H. Brownell (Eds), *Discourse ability and brain damage: Theoretical and empirical perspectives.* New York: Springer-Verlag (in press).

Frederiksen, C.H. & Breuleux, A. (1989). Monitoring cognitive processing in semantically complex domains. In N. Frederiksen, R. Glaser, A. Lesgold & M. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition.* Hillsdale NJ: Erlbaum (in press).

Frederiksen, C.H. & Décary, M. (1989). *Cognitive discourse analysis: Xerox 1186, Version 1.0* [Computer program]. Montréal: McGill University, Laboratory of Applied Cognitive Science.

Frederiksen, C.H, Décary, M. & Hoover, M. ( 1988). Le traitement et la représentation sémantique de la langue naturelle: développement d'un modèle informatisé [The semantic representation and processing of natural language discourse: a computational model]. *Les comptes rendus: Le* *colloque international sur l' informatique et la langue naturelle.* Nantes, France: Université de Nantes.

Frederiksen, C.H. & Dominic, J.F. (1982). Introduction: Perspectives on the activity of writing. In C.H. Frederiksen & J.F. Dominic (Eds.), *Writing: The nature, development, and teaching of written communication.* Hillsdale, NJ: Erlbaum.

Frederiksen, C.H., Donin, J., Kormos, L. & Browns, F. (1989). *Comprehension of multiple frame story texts.* Manuscript in preparation.

Frederiksen, C.H., Donin, J., Browns, F. & Kubes, M. (1989). *Comprehension of different types of expository text structures.* Manuscript in preparation.

Frederiksen, C.H., Donin-Frederiksen, J. & Bracewell, R.J. (1986). Discourse analysis of children's text production. In A. Matsuhashi (Ed.), *Writing in real time.* Norwood, NJ: Ablex.

Frederiksen, C.H. & Renaud, A. (1989). *Representation, acquisition and use of procedural knowledge in chemical engineering.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Hayes-Roth, B. & Walker, C. (1979). Configural effects in human memory: The superiority of memory over external information sources as a basis for inference verification. *Cognitive Science, 3*, 119-140.

Jackendoff, R. (1983). *Semantics and Cognition.* Cambridge, MA: M.I.T. Press.

Joseph, G.-M. & Patel, V. (1986). *Specificity of expertise in clinical reasoning.* Paper presented at the annual meeting of the Cognitive Sciences Society, Amherst, MA.

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

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

Kubes, M. (1988). *Use of prior knowledge in integaration of information from technical materials.* Unpublished doctoral dissertation, McGill University, Montréal.

Laird, J.E., Newell, A., & Rosenbloom, P.S. (1987) SOAR: An architecture for general intelligence. *Artificial Intelligence, 33*, 1-64.

Leech, G. (1972). *Towards a semantic description of English.* Bloomington, IN: Indiana University Press.

Mandler, J.M. & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*, 111-151.

Minsky, M. (1975). A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill.

Newman, D. & Bruce, B.C. (1986). Interpretation and manipulation of human plans. *Discourse Processes, 9*, 167-195.

Patel, V. & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science, 10*, 91-116

Rumelhart, D.E. (1975). Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science.* New York: Academic.

Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, N.J.: Erlbaum.

Schank, R. & Abelson, R.P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Simmons, R.F. (1972). Some semantic structures for representing English meanings. In R. Freedle and J.B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D.C.: V.H. Winston.

Sowa, J. F. (1984). *Conceptual structures: information processes in mind and machine.* Reading, Mass.: Addison-Wesley.

Spiro, R. (1980). Constructive processes in prose comprehension and recall. In R. Spiro, B. Bruce & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension.* Hillsdale, NJ: Erlbaum.

Stein, N. & Glenn, C.G. (1979). An analysis of story comprehension. In R.O. Freedle (Ed.), *New directions in discourse processing, Vol. 2.* Norwood, NJ: Ablex.

Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9*, 77-110.

Trabasso, T. & Nichols, D.W. (1978). Memory and inferences in comprehending narratives. In J. Becker & F. Wilkins (Eds.), *Information integration by children.* Hillsdale, N.J.: Erlbaum.

VanDijk, T. & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic.

Wirth, N. (1976). *Algorithms + data structures = programs.* Englewood Cliffs, N.J.: Prentice-Hall.

Yekovich, F.R., & Walker, C.H. (1986). Retrieval of scripted concepts. *Journal of Memory and Language*, 25, 627-644.

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# THE MENTAL REPRESENTATION OF THE TEXT*

## RELATIONSHIP BETWEEN TEXT CHARACTERISTICS AND CHARACTERISTRICS OF MENTAL REPRESENTATION

August Flammer and Ruth Lüthi

Department of Psychology
University of Berne
Laupenstrasse 4
CH-3008 Berne, Switzerland

## 1. INTRODUCTION: CONTENT OF THE TEXT AND MENTAL REPRESENTATIONS

A text is written by somebody in order to be read by somebody else. It could thus be said that the text is somebody's message addressed to someone.

This might be trivial. And yet: is the message the text or what the text says? The text by itself is clearly of little psychological interest. That is why the behaviorists took it as a stimulus linked or to be linked to a behavior pattern which it provokes or which has provoked it. Since the cognitive revolution, however, a text is not merely regarded as provoking behavior, but first of all as linked to mental, cognitive processes. In Chomsky's ridges (profound or deep structure) psychologists postulated that a text is to be related with a mental representation: it aims at such a representation with the reader and it is constructed from such a representation by the text producer.

(In order to do justice to history two things at least have to be granted: (1) Psychologists have hardly talked of texts before 1965, they rather talked of language; the written form and the complex structure that it assumes as a text was not an explicit object of studies. (2) Well before the psychologists the linguists and philosophers had spoken about the meaning or even of the semantics of language and texts.)

* Parts of the research reported here habe been financed by the Swiss National Science Foundation, projects No. 1.714-0.78 and 1.629-0.82.

The research done during the last 20 years has above all been concerned with the nature of mental representation. Since mental representation cannot directly be inspected, psychologists used mainly models, models borrowed from elsewhere. Some borrowed from logicians (e.g. Collins & Quillan, 1969), others used linguistic concepts (e.g. Anderson & Bower, 1973; Kintsch, 1974) and still others had been inspired by a long tradition according to which our mnesic experiences often take the form of images (e.g. Paivio, 1971; Kosslyn, 1980). Yet, all the experiences made aim at not adopting one of these models in a literal sense and above all at not restricting to one model for all use (see Johnson-Laird, 1983).

According to our position the content of a text can — to get back to our text — be represented mentally in various ways. Does this take place at the same time or is there a typical representation which is supposedly permanent and from which other representations can be constructed for a given duration? Without having final prove we tend to the hypothesis that typically the content of a text finds several mental representations which are supposed to be permanent.

Nevertheless, we should watch out. We have used the term content of text as if it were totally clear what it is and what its totality is. The first is a fundamental problem of epistemology: it is logically impossible to prove that there is identity or only isomorphy between reality as such and perception (or product of perception). The latter follows logically and is as well dictated by practical experience: Since every text contains for every reader what he is capable of understanding and ready to understand in the actual situation, the whole is not the same in every case and there is no definite means to tell when there will be the "total" whole.

Thus, talking of the representation of the text, will in our case not say that we try to solve the problem of "the" mental format of "the" representation and the problem of completeness of such a representation. It rather means to investigate the chosen charateristics of the mental representation or the mental representation of the text, of either the text to be written or the text which has been read. The aim of this contribution consists in gathering a series of experimental facts about the relationship between text format and format of mental representation.

In the first part we shall confront the format of a certain type of text to the format of the mental representation helping to produce this text. The second part will be dedicated to the comparison of a text which has been read and its mental representation being used to recall items chosen selectively. The third part, finally, will relate the format of a text which has been read to the reaction time patterns for the verification of chosen items.

## 2. CONTENT AND SEQUENCE OF QUESTIONS AND FORMAT OF A TEXT TO BE WRITTEN

The first time that our research group was confronted with this problem was during the more or less general fascination of cognitive psychologists about story grammar (see Rumelhardt, 1975; Thorndyke, 1977; Denhière, 1984). We dealt with the processes leading to state questions of information. Starting from the postulate that questions of information derive from mental representations felt as incomplete, we began to predict questions from an exact objective of the experimentation subjects and their supposed mental representation, their aim and present state. More concretely, we asked the subjects to kindly write a police report calling upon witnesses of a traffic accident where the guilty driver had fled. After having given some information about the accident the experimenter interrupted saying that he was not sure about what the subject really needed as information, that he or she could ask all he or she might think suitable and that the experimenter would gladly be ready to answer all questions (what he really did and after a well determined schema).

When the subject had asked all questions that he or she had wanted, he or she wrote the report as agreed. It must be added in advance that we had determined the grammar of such a report according to the rules of our newspapers (table 1). Thirty-seven of our subjects had clearly used this text format; the four others were taken out of our analysis since our interest was to see how the questions followed this grammar of the text.

This is the logic of the experimentation: Our knowledge of the mental representation of the aim and the present state should allow to predict the questions of the subjects. Since we had made sure that the analysed subjects had unterstood their task to ask all necessary information enabling them to write a text of a definite structure, the representation of this text and its content should provide them with the need and the frame for the questions to ask. As much as the text format and the mental representation were isomorphous, the grammar of the report should allow us to predict the content and the sequence of the questions.

Without going too much into details we established (1) that the content postulated by the grammar was indeed a matter of questions; more precisely, the compulsory elements had clearly been more frequently asked than the free elements. (2) The sequence of the questions was however badly predictable on the average. While six subjects behaved exactly as predicted, the 31 others fell into various sequence categories (identifiable except with six subjects). We conclude from this that most subjects "disposed" of the text structure (had it presented faithfully) but orientated themselves according to another representation which should be "representation" with the same right as the knowledge of the text.

What we call text representation can of course not be referred to the text surface as such but rather to the total of its construction rules because the text was still to be written (and if it had already been written it had not been learnt by heart word by word).

A particular demonstration of the multiplicity of the representation lies in the sequence of the elements of Time and Place of the accident. Whereas the (German) grammar prescribes that the Time ranks before the Place (by the way adopted by 33 of the 37 experimental texts), 19 of the 37 subjects asked the Place first and only then the Time.

Table 1
Grammar of the police report calling upon witnesses (according to Flammer, Kaiser & Mueller-Bouquet, 1981, 422)

| | | | |
|---|---|---|---|
| R1 | Police report calling upon witnesses | ——> | Circumstances + event + call |
| R2 | Circumstances | ——> | Time + place + accident |
| R3 | Event | ——> | (Course) + (Damage) + <Reaction of the runaway driver> |
| R4 | Call | ——> | Description of the runaway driver + (Address*) + <Request V Question> + Office to go to |
| R5 | Time | ——> | (Week day ) Date +(Part of the day) + accurate time |
| R6 | Place | ——> | <Region V road> +<Town V District V Road> |
| R7 | Accident | ——> | Accident circumstances$^{1}$ + (vehicle at disposal)$^{3}$ + (runaway vehicle)$^{4}$ |
| R8 | Course | ——> | Vehicle at disposal$^{3}$ ranaway vehicle$^{4}$ + (Person at disposal*)$^{2}$ + accident circumstances |
| R9 | Damage | ——> | (Dead person)$^{5}$ + (Injured person)$^{5}$ + (Material damage$^{5}$ |
| R10 | Description of the runaway driver | ——> | (what is known about the vehicle)$^{4}{}_{2}$ + (What is known about the person)$_{4}$+ (Assumption about the vehicle)$_{2}$ + (Assumption about the person) |
| R11 | Reporting office | ——> | Name + telephone number) |

() voluntary
* may be repeated
< > technical parentheses: composed element
$xx^{n}$ compulsory in one of the $xx^{n}$ elements
(xx) voluntary if effected elsewhere as $(xx)_{n}$ or $xx_{n}$
xxx final element

Just recently our experimentation was repeated by Hans van der Meij (1986) at Leiden. Whereas our subjects were students aged twenty, his subjects attended the 5th class of primary school. It is thus not too much of a surprise that these young pupils did not yet master the canonical format of the police report as required. On this basis it is almost trivial that the sequence of the questions did not correspond either to the text grammar as required, and that the experimentation is therefore not pertinent (there would however have been an interest in comparing the sequence of the questions to the texts as written).— It should only be mentioned incidentally that Van der Meij's work has other merits e.g. the introduction of a variable "interest" ("ask above all what interests you most of all" vs "ask what you need in order to write the text"), which did not lead to significant differences indeed.

To conclude, the knowledge of a text format to be written predicts adequately the content that the subject charged to write the text provides himself with (asking for, e.g.): but this knowledge does not typically guide, item by item, the search for the necessary information.

## 3. TEXT FORMAT AND SELECTIVE RECALL

Let's now look at the mental representation of the text which has been read or more exactly at its content. To what degree does the mental representation of the text content correspond to the organization of the content in the text? In order to investigate this question we can ask subjects to recall a certain text and to write it down from memory. Thorndyke (1977) has been one of the first to do that. The most interesting of his results within our context is not so much that a text organized according to the rules of its species is well recalled (that is to say re-edited) in its original organization, but that a partly disorganized text tends to be re-established when rewritten by the subject. Had the subjects represented the text content themselves according to an organization already different from the text which had been read or had the subjects re-organized what they had read only while they were rewriting it? These are open questions; yet what counts here is that the subjects apparently dispose of (and juggle with) more than only one organization.

Since the knowledge of rules of text writing (as far as its content — or its task — corresponds to a type of established text) can easily compel a subject to "normalization" during the rewriting, it should be possible to ask the subject to recall without implicit constraint towards the good form.

In a series of experimentations, undertaken for a somewhat different aim, this has exactly been effected. The typical course of these experimentations was as follows: The subjects read a text of one and a half pages (or more) and — after an interval of about five minutes filled with another task — they were asked to recall elements chosen from this text. These chosen elements were the elements of a class indicated in the second task.

An example to illustrate the procedure: the subjects had to read a text about the visit someone's house. Besides, the text mentioned pieces of furniture and pieces of clothing the people wore, etc. The second task was to recall and write immediately on a piece of paper all the pieces of furniture mentioned in the text (or all the pieces of clothing, etc.). The research question was whether the seriation of the items would correspond to their sequential position in the text or not. If there was agreement, the conclusion had to be that the seriation of the ideas as expressed in the text had remained at disposal as it had been coded as such or that it had been reconstructed during the recall.

If not: Would there be only disorder, aleatory? We must know that, if subjects are asked to simply write down all the pieces of furniture they know, they do it typically in the decreasing order of their typicality, that is to say that the table typically comes first, then the chair, etc. and only much later the bar or the footstool (Mervis & Rosch, 1981; cf. Arcuri & Girotto, 1985; Cordier, 1980, 1981; Cordier & Dubois, 1981; Flammer, Reisbeck & Stadler, 1985). It may thus be possible that the items in our experimentations should also be given in this order. To explain this one could assume the following processes: The subjects could recall item by item of a certain class to check if it has been in the text or not; if so, it should be written before the next item would be recalled and checked. Surely this strategy would not necessarily imply a different representation from that of the text neither would it imply a multiplicity of mental representation, but a fairly great independence of the given structure, because our subjects had been asked to act immediately.

In order to allow the simultaneous manifestation of the two tendencies we have constructed the texts in such a way that the distribution of the items of each class concerned was increasing in typicality for the first part of the text and decreasing in typicality for the second part of the text; thus the most typical one was the element presented in the middle of the text.

The first experimentation was done with grammar-school students, about 16 years of age. We did not only work with two texts but also with three versions of each text, that is A: critical items in the increasing order of typicality, B: typical items in the increasing order in the first half of the text, and decreasing in the second half of the text. For reasons of recency effects, however, we had been forced to leave out the last two items, and consequently the independence of the correlation coefficients was no longer assured. In table 2 the results of the "non independent" signification tests are put within brackets.

Table 2
Correlations between the order of presentation and the order of recall (Flammer, Grob, Jann & Reisbeck, 1985)

| Text version | r(te) >0 <0 | p | r(ty) >0 <0 | p |
|---|---|---|---|---|
| A: Increasing order of typicality | 104 : 55 | <.01 | 55 : 104 | (<.01) |
| B: Decreasing order of typicality | 90 : 44 | <.01 | 44 : 90 | (<.01) |
| C: Typicality first increasing and then decreasing | 99 : 44 | <.01 | 55 : 88 | (<.01) |

Starting with the contradictions to the predictions: There were no (positive) correlations of the recall order with the decreasing order of typicality. There were correlations but negative ones: Preference of the negative items! On the other hand, the order of the items in the text has clearly had a repercussion on the recall.

In order to avoid the problem of the technical interdependence of the coefficients we repeated the experimentation sticking exclusively to the C version in the first experimentation. Henceforth we constructed longer texts and placed the critical items far away from the beginning and the end of the text to exclude all recency or primacy effects.

Despite a certain variation of the correlation coefficients in table 3 there are two clear coexistent tendencies to notice, one to maintain the order of the text and another one to rearrange the order of the items according to the growing typicality.

Table 3
Correlations between the recall and the order of representation and the recall and the typicality (replications)

| | r(te) | r(ty) |
|---|---|---|
| | >0 <0 | >0 <0 |
| Flammer & Grob (1985): | | |
| grammar-school students | 39 : 33 | 31 : 40 |
| Morger & Flammer (in press) | | |
| subjects at the age of 10 | 11 : 1** | 3 : 10 |
| subjects at the age of 13 | 15 : 6 | 7 : 12 |
| Total | 26 : 7** | 10 : 22 |
| Flammer & Lüthi (1987) | | |
| subjects at the age of 8 | 34 : 20 | 23 : 26 |
| subjects at the age of 10 | 46 : 13 | 16 : 35 |
| Total | 80 : 32 | 39 : 61 |
| Grand total | 145 : 72 | 80 : 123 |

* : $p<.05$
** : $p<.01$

The first result asserts that the mental representation of the text content has at least kept the characteristics of the sequential text order. The second, however, does not allow us to conclude that there is another organization of the mnesic traces but a certain flexibility which the representation permits.

The introspective experience teaches us that, surely for a certain time, we not only keep in mind many details of a text we read which will be lost later on, but that, above all, very superficial characteristics (such as the position of a sentence on a page) will surely be forgotten later on. Thus we are not going to exclude that the characteristics of the sequential organization as manifested in our experiments will be lost after some weeks.

## 4. MENTAL REPRESENTATION AND TIME INTERVAL AFTER READING

To investigate the tenacity in the time of the sequential organization we have tried to replicate our experimentations with a three weeks' interval between the reading and the selective recall. Table 4 gives a comprehensive survey of the results.

The picture has changed a lot: After two weeks the text order is no longer clearly visible in the order of the recalled matter. One cannot say, however, that the influence of the typicality would have prevailed as much.

It is striking, however, that the increasing order of typicality no longer exists. On the contrary, there is a clear tendency towards the decreasing order of typicality, although not always significant. Had the increasing order been due to remaining traces of little duration or to the traces resulting from the fact that the rare items of an item class has much impressed the subjects and has therefore led them to dedicate more time to deal with these items (cf. the paradox of frequency according to Gregg, 1976)?

Table 4
Correlations between the recall and the presentation order and between the recall and the typicality after a three week's interval between the reading and the recall

| | r(te)<br>>0 <0 | r(ty)<br>>0 <0 |
|---|---|---|
| Flammer & Grob (1985): | 28 : 17 | 21 : 22 |
| Morger & Flammer (in press) | | |
| subjects at the age of 8 | 7 : 4 | 9 : 3 |
| subjects at the age of 10 | 3 : 6 | 9 : 3 |
| Total | 10 : 10 | 18 : 6 |
| Grand total | 38 : 27 n.s. | 39 : 28 n.s. |

The results of table 4 are not sufficiently satisfactory because they do not clearly display one of the two strategies. Are we confronted with a mental representation which is not suited to one of these strategies or is there simply a lack of possible strategies? In order to clarify this last question we began slightly to force the subjects to use up all that their representations and their strategic repertory allowed them to do. Thus, we have clearly instructed our subjects to make use of the one or the other of their selective recall strategies. More concretely, we made two groups. The subjects of one of these groups (text group) were instructed to proceed to the recall by following the text mentally and writing immediately, on a piece of paper, the required elements when they "entered their heads". The subjects of the second group (typicality group) were instructed to recall all the items belonging to the given class, to check whether they had been present in the text and, if so, to write them down immediately on a piece of paper. The results of the "text group" are included in table 5 a.

Table 5a
Recall following different instructions: (a) Instruction "text"

| | r(te) >0 <0 | r(ty) >0 <0 |
|---|---|---|
| Flammer & Grob (1985): | | |
| College students | | |
| immediate recall | 46 : 22** | 21 : 45** |
| recall after 3 weeks | 18 : 18 | 15 : 15 |
| Morger & Flammer (in press) | | |
| 10 year olds | | |
| immediate recall | 18 : 8** | 9 : 14 |
| recall after 3 weeks | 2 : 4 | 3 : 3 |
| 13 year olds | | |
| immediate recall | 21 : 1** | 7 : 14 |
| recall after 3 weeks | 3 : 9 | 8 : 4 |
| Total | | |
| immediate recall | 85 : 31** | 37 : 73** |
| recall after 3 weeks | 23 : 31 | 26 : 22 |

* : p<.05
** : p<.01

The subjects seemed to be well capable and ready to follow the instruction "text" when the recall was to be made immediately after the reading, not however, when the recall was asked three weeks after the reading. According to our observations as well as on the basis of the "immediate" results there is no doubt about what the subjects have understood and what they have tried to do when they had been asked to: we have to conclude that the mental representation was such that the temporary organization could no longer exist after the three weeks' interval. And yet, the text content was not (completely) forgotten in the meantime. What was the organization of the representation which stayed on? We do not know. — The correlations with the typicality order are less interesting because they were independent of the instruction. There is, however, a clear tendency towards the increasing typicality order in the immediate recall which will be lost in the recall after three weeks.

Table 5b contains the results of the "typicality group".

Table 5b
Recall order following different instructions: (b) Instruction "typicality"

| | r(te) >0 <0 | r(ty) >0 <0 |
|---|---|---|
| Flammer & Grob (1985): | | |
| College students | | |
| immediate recall | 46 : 31 | 25 : 53** |
| recall after 3 weeks | 27 : 22 | 33 : 17* |
| Morger & Flammer (in press) | | |
| 10 year olds | | |
| immediate recall | 15 : 6 | 11 : 8 |
| recall after 3 weeks | 5 : 6 | 8 : 4 |
| 13 year olds | | |
| immediate recall | 19 : 4** | 7 : 16 |
| recall after 3 weeks | 3 : 4 | 9 : 2 |
| Total | | |
| immediate recall | 80 : 41** | 43 : 27** |
| recall after 3 weeks | 35 : 32 | 52 : 23** |

* : p<.05
** : p<.01

In the immediate recall this instruction had reinforced the increasing order of typicality and asserted the decreasing order in the recall after three weeks. In the immediate recall there remained, nevertheless, traces of the text organization which were completely lost after the three weeks.

These results allowed us to conclude that the mental representation of the text had undergone an important change during the three weeks' interval. The minimal change consists in that the traces due to the special treatment of the rare items have not remained and that the sequential organization of the representation had been lost in the meantime.

It is noteworthy that these differences are not the differences that literature usually attributes to the difference between short-term memory and long-term memory, because the short-term memory could not very well be used to explain the results of the "immediate" selective recall (the text would be too long for this memory, the five minutes' interval too long as well). We have, however, not for the first time been able to confirm the effects of something like a middle-term memory; Flammer and Morger (1985) had found an effect of items activation during a period of ten to twenty minutes, whereas the effects known of the primed recognition are of a length of some tenth of seconds and of seconds.

## 5. TEXT FORMAT, MENTAL REPRESENTATION AND RECOGNITION TIME

Let's turn to the third source of information on the mental organization of the text which has been read: the recognition time. There are the three following basic assumptions underlying our experimentations with the recognition method of primed activation:

(1) The information represented in the memory is vast, and for a determined selection to be recalled it is necessary to activate the representation concerned (cf. Collins & Loftus, 1975).

(2) A punctual activation spreads out through the "adjacent" memory structures (Lorch, 1982; Nelson, McEvoy & Friedrich, 1982). This implies that the activation spreads out according to the organization of the representation. As far as a text which has been read is mentally represented as an independent unity and separated from the rest, its elements are only linked between themselves, and furthermore, if the organization of the representation is isomorphous to the representation of the text, the display patterns correspond to the structure of the text read. We shall call this organization "episodic". We have, however, to assume the possibility that the organization of the text differs greatly from the organization of the mental representation and that, furthermore, the content of the text read is assimilated to already existing complexes, e.g. that the mentioning of a dog in a text is (also) integrated in a structure of various elements of the category "animal".

(3) An activation may derive from the free choice by presenting a stimulus which corresponds to the chosen memory element, e.g. by presenting a certain word.

Up to this day, we have been experimenting twice with the method of the primed activation of recognition. The first time we presented the items to the subjects as soon as he or she had answered the foregoing item (Flammer, Lüthi & Morger, 1985). This procedure led to intervals between the two stimuli (SOA) of some seconds, intervals which are too long.

The second time (Flammer & Lüthi, 1987) we experimented with a SOA interval of 200 milliseconds. The dependent variable was the reaction time in the correct recognitions of certain key words out of the text. The words of activation ("primes") had been chosen according to two independent factors: belonging or not to the same semantic category as the word to be recognized ("target") (semantic relation), and belonging or not to the text vocabulary (episodic relation). The first factor ("category") had two levels: (1a) the word of activation belonged to the same category as the word to be recognized or (1b) did not belong to this category. The second factor ("distance") had three levels: (2a) The word of activation belonged to the same sentence as the word to be recognized, (2b) the word of activation did not belong to the same sentence as the word to be recognized, it was, however, part of the text, (2c) the word of activation did not even belong to the text read. Table 6 shows the reaction time for the six different groups of activation.

Table 6
Recognition time as a function of relationship between the word of activation (WA) and the word to be recognized (WR)

(Only the reaction time for correctly recognized items is taken into account)

| | WA out of the text | | WA not out of the text |
|---|---|---|---|
| | WA and WR from from the same sentence | WA and WR not from the same sentence | |
| WA and WR belong to the same semantic category | 1.84 | 1.87 | 1.38 |
| WA and WR are semantically unrelated | 1.36 | 1.98 | 1.66 |

The analysis of variance shows only one significant main effect, the effect of the second factor. The interaction, however, of the two factors proved to be significant ($F(2,510) = 4,75$: $p<.01$). Thus, we are only discussing the single main effects. There have been two which proved to be significant and were therefore responsible for the interaction. (1) If the word of activation was part of the same sentence as the word to be recognized, recognition was faster in the case of WA and WR belonging to different semantic categories ($F(2,238) = 3.16$; $p<.05$); (2) On the contrary, when the word of activation was not part of the text, recognition was faster in the case of WA and WR belonging to the same sentence ($F(1,220) = 6,28$; $p<.05$).

This is a surprising result; we had expected that either one of the two factors or even the two would, at the same time, simply touch the reality of mental representation, e.g. that 1a and 2a would be superior to 1b and 2b and 2c, respectively. The results may show that the episodic relationship makes recognition easier only if there is no semantic relationship and vice-versa. There seems to be something like interference.

What does interfere? We present the following interpretation which — evidently — needs further clarification: In the case of 1a/2a the WA is strongly linked to two complexes of the mental organization, namely to the complex "part of the text" and the complex "semantic field" of the WA. The activation unfolded is mainly canalized in these two complexes. This leads to a dispersal of the activation (what is called the "fan effect"). In the cases 1a/2c the activation is orientated more exclusively towards the complex "semantic field" and thus reaches the word to be recognized more quickly (no loss through the complex "part of the text") and reaches quickly the word to be recognized because there is no loss through the complex "semantic field".

What does this imply for our problem of mental representation? If our interpretation of the interaction is correct, the representation of the text read has features in common with the text and with the preexisting semantic structures. The structural features in common with the text have to do with the sequential structure of the representation, possibly the same that the correlations between seriation of the recall and seriation in the text have touched. If so, it is to be expected that the interaction of table 6 will be lost after a certain time, as the correlations with the seriation of the text have been lost after an interval of three weeks. This is an assumption, a hypothesis that we have not tested (yet).

## 6. CONCLUSIONS

It is not surprising that the results of our experiments show that the mental representation of what is read of the text is not isomorphous to the text organization. The details, however, deserve to be looked at:

(1) The experimentation with the questions has shown that the knowledge of a format of text to be written (as well as on the part of the subject as on the part of the researcher) is a good predictor for the content that the subject asked to write the text tries to provide himself with, but that this knowledge typically does not guide item by item the search for the necessary information on the part of the subject. The report format is really something that the subject knows, which he contains in his mnesic repertory, but from which he does not depend for every action concerning this format.

The fact that certain subjects had a clear tendency to first ask for all which was of general interest about the accident, indicates that the subjects felt free and capable to produce the report format at the right moment and — meanwhile — organize knowledge in a different way.

(2) The results on the selective organization of recall show that the structure of the text read was well present and important while memory was still fresh, but that it got lost later on, without the most important content of the text having

been forgotten. The comparison of these results with those of the pre-activated recognition allows to assume that immediately after the lecture the content of the text was organized according to at least two principles. Has the text been encoded twice or did the unique representation include the characteristics of different sources?

## REFERENCES

Anderson, J.R., Bower, G.H. (1973) *Human associative memory*. New York: Wiley.

Arcuri, L., Girotto, V. (1985) Norme di tipicità per sei categorie naturali. Uno studio evolutivo. Dipartimento di psychologia dello sviluppo e della socializzazione. Università di Padova.

Collins, A.M., Quillan, M.R. (1969) Retrieval from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247.

Collins, A.M., Loftus, E.A. (1975) A spreading-activation theory of semantic processing. *Psychological Review, 82,* 402-428.

Cordier, F. (1980) Gradients de prototypie pour cinq catégories sémantiques. *Psychologie Française, 25,* 211-219.

Cordier, F. (1981) Catégorisation d'exemplaires et degré de typicalité: étude chez des enfants. *Cahiers de psychologie cognitive, 1,* 75-83.

Cordier, F., Dubois, D. (1981) Typicalité et représentation cognitive. *Cahiers de psychologie cognitive, 1,* 299-334.

Denhière, G. (1984) *Il était une fois...* Lille: Presses univer-sitaires.

Flammer, A., Lüthi, R. (1987) Episodische und kategorielle Aktivierungseffekte auf Wiedererkennungszeiten. *Archiv für Psychologie, 139,* 209-225.

Flammer, A., Lüthi, R., Morger, V. (1985) Strategien des selektiven Erinnerns: Ein Experiment mit gesteuertem Wiedererkennen. *Archiv für Psychologie, 137,* 255-271.

Flammer, A., Morger, V. (1985) Die Wirkung von Voraktivierung und Typikalität auf die Verifikationszeiten von Begriffsexemplaren. *Schweizerische Zeitschrift für Psychologie, 44,* 1-16.

Flammer, A., Reisbeck, C., Stadler, S. (1985) Typikalitätsnormen für dreizehn Begriffe in einer deutschschweizerischen Studentenpopulation. *Sprache und Kognition, 4,* 49-63.

Gregg, V. (1976) Word frequency, recognition and recall. In Brown, J., Ed., *Recall and recognition*. New York: Wiley, 183-216.

Johnson-Laird, P.N. (1983) *Mental models*. Cambridge: Cambridge University Press.

Kintsch, W. (1974) *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum.

Kosslyn, S.M. (1980) *Image and mind*. Cambridge: Harvard University Press.

Lorch, R.F. (1982) Priming and search processes in semantic memory: a test of three models of spreading activation. *Journal of Verbal Learning and Verbal Behavior, 21,* 468-492.

Mervis, C.B., Rosch, E. (1981) Categorization of natural objects. *Annual Review of Psychology, 32,* 89-115.

Nelson, D.L., McEvoy, C.L., Friedrich, M.A. (1982) Extralist cuing and retrieval inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8,* 89-105.
Paivio, A. (1971) Imagery and deep structure in the recall of english nominalizations. *Journal of Verbal Learning and Verbal Behavior, 10,* 1-12.
Rumelhart. D.E. (1975) Notes on a schema for stories. In Bobrow, D.G., Collins, A., Eds., *Representation and understanding*. New York: Academic, 221-236.
Thorndyke, P.W. (1977) Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9,* 77-110.
Van der Meij, H. (1986) *Questioning*. The Hague: SVO.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# TEXT TYPOLOGIES: A COGNITIVE APPROACH

Michel FAYOL

Laboratoire de Psychologie
6 rue Chabot Charny
21000 Dijon, France

In the present paper, I will discuss how cognitive psychology approaches or could approach questions about text typologies. I have postulated that from a linguist's and rhetorician's point of view, different types of texts exist. At least some of these texts are defined relatively precisely, with little disagreement among theorists (cf. Schneuwly, 1986 for a discussion of these problems).

Starting from a typology defined in a pre-established theoretical framework, cognitive psychology is now attempting to assess the "psychological validity" of such a classification (Mandler and Goodman, 1982). The typology will be deemed valid to the extend that different types of texts are identified as such and subjects who are not necessarily specialists in rhetoric manifest different behavior with respect to each text type.

It is commonly accepted that in cognitive psychology information processing is considered to be the result of an interaction between:

(a) a stimulus of varying complexity definable by a certain number of characteristics; (in this case, the stimuli are texts.) and

(b) a subject who has a relatively limited processing capacity and a body of factual and/or procedural knowledge at his or her disposal. In the domain studied here, factual knowledge refers to the themes dealt with in the text, the text's structure, and as the case may be, the linguistic marks systematically associated with it. Procedural knowledge refers to the processing modes corresponding to each type of text.

Consequently, establishing the psychological validity of a typology of texts may imply that are demonstrated the following points. First of all, despite the diversity of the themes (or contents) of various texts, texts pertaining to the same category are processed in a similar and a consistent manner. Second, texts pertaining to different categories are processed differently, even if their "subject content" is identical.

The processing modes themselves may differ in aspect. In cognitive psychology we conventionally distinguish between 1) processing modes that concern comprehension activities (which include memorization) and 2) processing modes that involve production and 3) processing modes that pertain to awareness ("metatextuel" activities and, in particular, classification activities).

## 1. TYPES OF TEXTS AND COMPREHENSION

### 1.1. Comprehension and problem solving

Oddly enough, the clearest and most complete studies dealing with the impact of type of text on processing mode have been conducted in the field of verbal problem solving (usually arithmetic problems). I will not describe this research here in detail, but would like to mention the following points.

"Experts" are in fundamental agreement regarding the typologies of arithmetic problems (cf Carpenter and Moser, 1983; Rilev, Greeno and Heller, 1983; Vergnaud, 1982. Refer, however, to Mayer, 1981, for an overview of the difficulties encountered in classification).

When confronted with different types of problems, both child and adult subjects rely on different problem-solving strategies that usually -- but not always -- correspond to the different categories defined.

Non-expert subjects widely agree upon how to categorize problems (although their taxonomy is not exactly the same as that of experts).

Non-expert subjects who are familiar with the problems involved manage very quickly -- as soon as they start listening or reading -- to categorize the problem, and thus to determine which problem-solving method to adopt (Hinsley, Hayes and Simon, 1977).

Subjects who have identified the type of problem are able to discern the hierarchy of the information contained in the problem the first time they read it. They know right away whether an item is relevant to the solution or not (Robinson and Hayes, 1978).
The identification of problem type and the relative ease with which a problem is solved by activation of the appropriate processing mode strictly depend on how familiar the subjects are with that problem type. The most uncommon problem statements (in books) are the least well identified and the most difficult to solve (Mayer, ibid., 1985).

The above phenomena might be interpreted in light of "schema" theory (Fayol, 1985). The general procedure would be as follows. When faced with a problem statement, subjects systematically attempt to determine to which category it belongs. As soon as they read (or hear) the first sentence, they use a bottom-up processing mode to try to detect those cues likely to enable them to classify the problem (i.e., based on past experience, to select and activate the appropriate "schema" in long-term memory). If they fail to identify such a schema, they continue this bottom-up processing, sentence by sentence, thus creating a heavy load in working memory. If they succed, the "schema" activated allows them, at least in part, to apply a top-down processing mode. This allows subjects to immediately assess the relevance of the information being processed, and to focus their attention on essential elements, and to more easily retrieve the data likely to accelerate interpretation (by anticipation of the incoming information and establishment of pro-active inferences).

Data collected by cognitive psychologists thus reveal the "psychological validity" of problem typologies. Because relatively well-defined problem categories do exist, subjects who are familiar with these categories are able from the start to identify them and to implement the appropriate problem-solving procedure.

The same set-up and principles may apply to text comprehension in general, although the tasks involved are of a quite different nature. Do we have enough empirical information to support this position?

### 1.2. Studies on text comprehension

As Goetz and Armbruster (1980) emphasized, the critical issue is to discover "how variations in text structure influence the way people read, comprehend, and remember text" (p. 201).

#### 1.2.1. Types of texts and comprehension

In light of available empirical data, most researchers agree that the overall structure of a text has a significant effect on how it is understood. Regardless of what the task is -- silent reading, reading out loud, reading and commenting -- the "proper" or "standard" structure facilitates processing and leads to better understanding (assessed by asking questions). Kintsch and Yarbrough (1982) thus showed that "standard" structures facilitate the comprehension of various types of texts (classification, comparison, description, ...)

This global effect leads us to the question of whether text type influences on-line text comprehension. Little is yet known on this topic, since few researchers have tried to compare the progression of the reading process for each of the different types of texts. Nevertheless, Britton, Graesser, Glynn, Hamilton, and Penland (1983), Graesser, Hoffman, and Clark (1980), Haberlandt and Graesser (1985), and Ehrlich and Tardieu (1986) have shown that higher levels of processing (i.e., those dealing with the text's macrostructure or superstructure, Vipond, 1981), have an effect on on-line encoding (see the overview by Denhière, 1985). The results obtained are often contradictory, however (for instance, descriptions are sometimes read more slowly and sometimes more quickly than stories), and theoretical justifications are lacking. The effect of the superstructure on on-line processing has actually been established only for narratives, with adults subjects (cf. frochot, Zagar and Fayol, 1986). Very little is now known about other types of texts.

Nevertheless, Olson, Mack and Duffy (1981) compared both subjects' reading times and comments made when asked to read and comment on stories and essays that were presented either in the standard order or in a very different order. The results showed very clearly that adult subjects explicitly or implicitly refer to certain conventions concerning types of texts, and that they use these conventions to understand the texts (cf. also Meyer and Rice, 1982). This was manifested in two ways. First of all, "properly organized" narrative texts or essays (i.e. those whose surface structure corresponds to the supposed standart structure) were always processed more easily (better comprehension, faster reading speeds). Second, reading times as well as readers' comments, revealed that narrative texts and essays were handled very differently from each other. The narrative texts were processed using a prospective approach that involved formulation then confirmation of hypotheses. The essays were approached in a retrospective manner that involved relating current elements to preceding ones.

Based on the all - too - scarce data currently available, it thus appears clear that relatively well-educated adults (usually university students) have some knowledge about certain types of texts and their structures. They also seem to use this knowledge even during the reading/comprehension processes. However, the findings obtained so far are still too scarce to allow us to draw definite conclusions. Any conclusions which can be made should be based on additional arguments derived essentially from memory research.

### 1.2.2. Text types and memory

Recall tasks have revealed systematic effects associated with various texts types. For instance, it is well established that when the standard order of a story is maintained, the story is more easily understood and memorized (cf. Fayol, 1985, p. 50-58 for a review). Much less information is available, however, concerning other types of texts.

Nevertheless, Kozminsky (1977) showed that the amount of information recalled from a text is a function of the type of text (narrative, report, description). The clearest results on this subject were obtained by Meyer and Freedle (1984). These authors designed four different types of texts by using paragraph C to vary the relationship between paragraphs A and B, which remained constant across texts. They then determined how much information adult subjects retrieved for each type of text. Their results showed very clearly that the amount of information recalled varies significantly according to the text's structure (macro- and/or superstructure).

### 1.2.3. Conclusion

Data from comprehension and memorization tasks have shown that :
At least for certain, well-defined categories, different text types lead to different processing modes during encoding, and to varying performance at recall.
Each type of text studied seems to be associated with a "surface" structure that is more appropriate for it than other structures. This specific "rhetorical" form is considered standard in that it should be isomorphic to the underlying superstructure (i.e., dependent on a pre-linguistic representation level). The more this isomorphism is respected, the better the quality of the comprehension and memorization. It appears, however, that this standardization is more constricting for some types of texts than for others (cf. Garnham, Gakhill and Johnson-Laird, 1982). Now, does the same phenomena occur during production and metatextual activities (classification, etc.)?

## 2. PRODUCTION AND CATEGORIZATION OF TEXTS

### 2.1. Types of texts and production

Here as in other fields, research on production is quite rare. The results of available studies, however, have led to basically the same conclusions as those obtained for comprehension tasks.

First of all, very little research has examoined on-line production (cf. Fayol and Schneuwly, 1986, for a review). Studies on the effect of text type are for that very reason even rarer. Moreover, the results obtained so far are contradictory. For example, Matsuhashi (1981) found that for expert authors, the mean duration of pauses -- assumed to reflect planning activity -- varied according to text category (generalization, report, persuasion). But Matsuhashi and Quin (1984) did not replicate this result when contrasting argument production to report production, although, granted, only one expert subject was involved. Thus, no clear relationship can currently be established between types of texts and the way subjects produce them.

On the other hand, research on certain linguistic marks has led to clear and consistent findings. It has shown, for example, that the syntactic complexity of sentences varies more according to type of text than to the subject's developmental level : t-units are longer for arguments than for descriptions or narratives (cf. crowhurst and Piché, 1979; Perera, 1986; Rubin, 1984).

The surface marks whose presence and functioning have been most widely studied are the connectors. Here again, it has been shown that different text categories are systematically associated either with different connection modes or with different distributions of the same marks (Langer, 1985; Pelligrini, Galda and Rubin, 1984; Schneuwly and Bronckart, in press; Smith and Frawley, 1983).

This consistent association of surface mark patterns to types of texts was the object of some "correlational" reasearch on production (cf. in particular Bronckart, Bain, Schneuwly, Davaud and Pasquier, 1985). It should be noted, however, that other studies have given comparable results.

Brewer and Hay (1984) requested adult subjects to memorize texts that had very characteristic styles : scientific articles, law texts, children's books, etc. The texts used were rewritten ahead of time in order to eliminate certain aspects of style by replacing them with more common expressions. The authors then analyzed the modifications that subjects made in the texts when asked to recall them literally. The majority of the recalled texts were reconstructions; as a rule, the style of the reproduced texts approached that of the original. In other words, subjects re-established the characteristics that the experimenter had eliminated. A strict interpretation of these results would imply that literal memory of the style of a text does not exist and that only the "content" is memorized; during recall, subjects reconstruct the style -- provided they have mastered it -- by activating the appropriate mark and expression pattern, i.e., by style matching.

It thus appears to be reasonably established that there is a systematic correlation between certain types of texts and certain types of linguistic marks. On the other hand, little is known about the function of these marks; connectors, for example, may serve to indicate to the reader/listener the nature of the inter-statement relationship and/or the "text plan" (Bronckart et al.. ibid.). Experimental confirmation of this hypothesis has not yet been provided, however. The functions of other marks are even more vague. Perhaps they serve the purpose of allowing subjects to identify a text's type, so that it might be processed in the appropriate manner (not only according to its thematic content but also to its superstructure and "discursive force"; cf. Brewer, 1980). This brings us to the questions of classification.

### 2.2. Classification of texts and typologies

Research on text classification is even more scarce than that on production but is nonetheless very important. Indeed, as shown in the domain of problem solving, a given text should not be confused with another if it is to be processed correctly. Moreover, if processing is to be facilitated, the text category should be identified at a early stage in the comprehension process. Therefore, it appears essential to find out if, when, and how readers identify text types. Very little information is currently available on this subject, however.

Most research conducted so far has dealt with narratives, usually with children. Stein and Policastro (1984) found that, contrary to earlier hypotheses, children as young as seven years of age may have relatively precise notion of what a story is, a notion which is identical to that of adults. Children are even able to distinguish narratives told in the past tense from procedures expressed in the future tense.

In a quite different perspective, Langer (1985) found that eight- and nine-year olds (3rd graders) were able to differenciate stories from reports, a difficult distinction. However, a finer analysis of the experiment reveals that the stories in question referred to "make-believe" situations, whereas the reports talked about facts. Moreover, the justifications given by the subjects reveal that the classification criteria pertained mostly to content and not to structure. Structure does not appear to be salient until age fourteen or fifteen (9th grade).

Stein's findings, like Langer's, contradict other data. For example, two series of experiments by Esperet (1984) illustrated that five-year olds begin to distinguish between "text" and "non-text" (unrelated sentences placed side by side), however "elementary" this distinction may be, but that this distinction is not clearly established until age nine (cf. also Buss, Yussen, Mathews II, Miller and Rembold, 1983; Englert and Hiebert, 1984, for similar results). Similarly, Mc Conaughy, Fitzhenry-Coor and Howell (1983) showed in their reserach that children do not possess the same "narrative schema" as adults.

It is clear, then, that even for one particular type of text, certainly the most typical one, narratives, the empirical data collected are still not conclusive. It is true that the developmental aspect of the problem may complicate the issue, but we can also expect difficulties in the protocols of adults.

To our knowledge, only one study is available that provides some information about adults' texts classifications. Faigley and Meyer (1983) asked well-educated adults who were not necessarily trained in rhetoric to perform a sorting task. Although certain text categories were subject to general consensus (narratives, for instance), even the most well-informed subjects did not recognize all the different types of text (eight in all). In a second experiment, only five of the text types were used (description, narrative, classification, definition, and procedure), and a new sorting task was proposed containing three texts in each category.

The results reported by Faigley and Meyer (1983), which were analyzed using multidimensional scaling analysis, show that subjects generally agree on text type, independantly of subjects' knowledge of rhetoric. In brief, texts are located along a one-dimensional axis where narratives are opposed to non-narratives. The authors state that text typologies are based at least in part on cognition. They believe in particular that the most prominent dimension considered by subjects deals with tense and aspect -- texts describing processes that lead to a result (durative or not) would be classified as narrative; those containing verbs of being or verbs describing ongoing activities would be classified as non-narrative. The "finer" distinctions (between description, definition, classification, etc.) would be based on differences in structure (not discussed by the author).

In summary, despite a few contradictory results, the existing data seem to show that the ability to identify a text category is acquired late and is easily disrupted, even in expert subjects if the typology is complex. Little is known about this problem, although it is an essential aspect of text processing, as stressed above.

It seems that some of the difficulties encountered here are due to the fact that the criteria used are not properly distinguished. For example, thematic content and rhetorical structure are rarely considered separately. Indeed, descriptions are often associated with spatial configurations, and narratives with imaginary worlds. It is therefore very difficult to determine how subjects classify a given text. Do they base their classification on thematic content, on the text's superstructure, on its discursive function, or on surface marks?

## 3. TYPES OF TEXTS AND COGNITIVE PSYCHOLOGY : PROBLEMS AND PERSPECTIVES

The above research review illustrates that the cognitive approach to the processing of text types has not yet provided as much clear and precise information as has the cognitive approach to problem solving. Indeed, difficulties persist or information is still lacking in the following areas.

First of all, unlike mathematicians and experts on the didactics of mathematics who have more or less agreed on one or two typologies and clearly differenciate underlying semantic structure (s), rhetorical organization, and, surface structure, linguists have failed to reach the slightest consensus. As a result, only a few (if any) classifications established by experts are available, and we are faced with numerous fluctuations in terminology.

Second, it has not yet been proven that on-line processing varies systematically across text types. Although available data lead us to believe that an effect specific to the narrative form as opposed to the general text form does exist, and although arguments in favor of differentiating narratives from other types of texts may be presented, almost nothing is known about the processses involved in their encoding and comprehension. This is a vast and very basic research field which should be extended to include "transfers" from one type of text to another during reading (or production). It is true that a given speech or text is rarely typologically homogeneous, and the impact ot text category changes on reading mode or, more generally, on comprehension processes should therefore be more clearly defined (cf. Townsend's first studies, 1983, and Adam's remarks, 1985).

Third, in cognitive psychology, reference to text type is based on the postulate that the cognitive structure associated with each text type is both unique and common to everyone (provided the person is familiar with it). Yet a perfect correspondence does not always appear to exist between the so-called standard text structure and the text structure possessed by subjects (cf. Shebilske, 1980, for a discussion). This proves to be particularly true for children (cf. Denhière, 1982), but some experimental data has shown that interindividual differences also exist among adults.

For example, using a recall test, Mosenthal (1979) showed that the performance of children between the ages of seven and ten depends not only on the structure of the text itself (defined by whether the theme summary is located at the beginning or the end of the text) but also on the children's expectations concerning the text's structure. The more congruent expectations and structure are, the better the recall.

In a similar vein, Spiro and Tirre (1980) discovered that certain adult subjects had a greater tendency than others to force a "script"-type structure on the stimulus, and thus attained much higher performance levels when the stimulus was, in fact, structured as such.

Granted, one might object to both of the above experiments that the structures in question are rhetorical level rather than underlying, "deep" level structures, which are not affected. But whether considered superficial or not, these structures do influence performance, and their effects are not distributed evenly across subjects. The problem of interindividual variability for so-called "standard" structures must thus be addressed.

Fourth, if as a working hypothesis we admit that each type of text possesses an optimal structure that facilitates processing, then we should try to find out just how the corresponding cognitive "schema" becomes activated. If such a schema is to be implemented during encoding, it must indeed be identified and triggered as soon as the text is received (cf. problem examples in 1.1). However, in the case of problem statements, we do not yet know whether very rapid recognition of the category even exists. Nor do we know what cues the reader/listener uses to achieve this end. Like Faigley and Meyer (ibid.), we might suppose that subjects use semantic criteria (temporo-aspectual), and that facilitation follows the rapid grasping of certain surface-mark patterns (cf. Brewer et al., ibid.; Bronckart et al., ibid.).

Fifth, if it is true that text type has a differential effect on comprehension and memorization, then we should find out why this is so. The answer to this question requires that we agree that the chronologico-causal structure -- inherent to narratives -- is "in essence" a facilitating factor. This may of course be possible, but Mayer's works (ibid.) lead us to another hypothesis. Are we not dealing here with an effect that is due to subjects' increased familiarity with the structure of narratives? If that is the case, as with problem statements, the rarest text types would also be the most difficult to identify and to process, and the information therein would be less well encoded, stored, and retrieved. Therefore, it is evident that the solution would not lie in favoring the narrative form, but rather developing our knowledge of the other text categories.

To conclude, the cognitive approach to text typologies is only at a very preliminary stage.

## REFERENCES

ADAM, J.M. Reflexion linguistique sur les types de textes et de compétences en lecture. *L'orientation Scolaire et Professionnelle*, 1985, 14 (4), 293-304.

DE BEAUGRANDE, R.. *Text production. Toward a science of composition.* Norwood: Ablex, 1984.

BREWER, W.F. Literary theory, rhetoric and stylistics: implications for psychology. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.) *Theorical issues in reading comprehension.* Hillsdale: Erlbaum, 1980.

BREWER, W.F., & HAY, A.E. Reconstructive recall of linguistic style. *Journal of Verbal Learning and Verbal Behavior*, 1984, 23, 237-249.

BRITTON, B.K., GRAESSER, A.C., GLYNN, S.M., HAMILTON, T., & PENLAND, M. Use of cognitive capacity in reading. *Discourse Processes,* 1983, 6 (1), 39-57.

BRONCKART, J.P., BAIN, D., SCHNEUWLY, B., DAVAUD, C., & PASQUIER, A. *Le fonctionnement des discours;* Neuchâtel, Paris: Delachaux et Niestlé, 1985.

BUSS, R.R., YUSSEN, S.R., MATHEWS II, S.R., MULLER, G.E., & REMBOLD, K.L. Developement of children's use of a story schema to retrieve information. *Developmental Psychology,* 1983, 9 (1), 22-28.

CARPENTER, T.P., & MOSER, J.M. The acquisition of addition and subtraction. In R. Lesh, & M. Landau (Eds.) *Acquisition of mathematics concepts and processes.* New York; Academic Press, 1983.

CROWHURST, M., & PICHE, G.L. Audience and mode of discourse effects on syntactic complexity in writing at two levels. *Research in the Teaching of English,* 1979, 13 (2), 101-109.

DENHIERE, G. Schéma(s)? Vous avez dit schéma(s)? *Bulletin de Psychologie,* 1982, 35 (356), 717-731.

DENHIERE, G. De la compréhension à la lecture. *L'Orientation Scolaire et Professionnelle,* 1985, 14 (4), 305-329.

EHRLICH, M.F., & TARDIEU, H. Le rôle du titre sur le temps de lecture et le rappel de trois types de textes. *Bulletin de Psychologie,* 1986, 39 (375), 397-406.

ENGLERT, C.S., & HIEBERT, E.H. Children's developing awareness of text structures in expository materials. *Journal of Educational Psychology,* 1984, 76 (1), 65-74.

ESPERET, E. Processus de production: génèse et rôle du schéma narratif dans la conduite de récit. In M. Moscato, & G. Piéraut le Bonniec (Eds.) *L'ontogénèse des processus psycholinguistiques et leur actualisation.* Rouen: P.U.R., 1984.

FAIGLEY, L., & MEYER, P. Rhetorical theory and reader's classifications of text types. *Text,* 1983, 3, 305-326.

FAYOL, M. *Le récit et sa construction: une approche de psychologie cognitive.* Neuchâtel, Paris: Delachaux & Niestlé, 1985.

FAYOL, M., & SCHNEUWLY, B. *La mise en texte et ses problèmes.* Communication au "Troisième Colloque International de Didactique du Français". Namur, Septembre 1986.

GARNHAM, A., OAKHILL, J., JOHNSON-LAIRD, P.N. Referential continuity and the coherence of discourse. *Cognition*, 1982, 11, 29-46.

GOETZ, E.T., & ARMBRUSTER, B.B. Psychological correlates of text structure. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.) *Theorical issues in reading comprehension*. Hillsdale: Erlbaum, 1980.

GRAESSER, A.C., HOFFMAN, N.L., & CLARK, L.E. Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 135-151.

HABERLANDT, K.F., & GRAESSER, A.C. Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General*, 1985, 114 (3), 357-374.

HINSLEY, D.A., HAYES, J.R., & SIMON, H.A. From words to equations. In M.A. Just, & T.A. Carpenter (Eds.) *Cognitive processes in comprehension*. Hillsdale: Erlbaum, 1977.

KINTSCH, W., & YARBROUGH, J.C; Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, 1982, 74 (6), 828-834.

KOZMINSKY, E. Altering comprehension: the effect of biaising titles on text comprehension. Memory and Cognition, 1977, 5, 482-490.

LANGER, J.A. Children's sense of genre. *Written communication*, 1985, 2 (2), 157-187.

MANDLER, J.M., & GOODMAN, M.S. On the psychological validity of story structure. *Journal of Verbal Learning and Verbal Behavior*, 1982, 21, 507-523.

MATSUHASHI, A. Pausing and planning: the tempo of written discourse production. *Research in the Teaching of English*, 1981, 15 (2), 113-134.

MATSUHASHI, A., & QUINN, K. Cognitive questions from discourse analysis. Written Communication, 1984, 1 (3), 307-339.

MAYER, R.E. Frequency norms and structural analysis of algebra story problems into family categories and templates. *Instructional Science*, 1981, 10, 135-175.

MAYER, R.E. Mathematical ability. IN R.J. Sternberg (Ed.) *Human abilities*. New York: Freeman, 1985.

Mc CONAUGHY, S.M., FITZHENRY-COOR, I., & HOWELL, D.C. Developmental differences in schemata for story comprehension. In K.E. Nelson (Ed.) *Children's language*, (Vol.4). Hillsdale: Erlbaum, 1983.

MEYER, B.J.F., & FREEDLE, R.O. Effects of discourse type on recall. *American Educational Research Journal*, 1984, 21 (1), 121-143.

MEYER, B.J.F., & RICE, E. The interaction of reader strategies and the organization of texts. *Texts* 1982, 2, 155-192.

MOSENTHAL, P. Three types of schemata in children's recall of cohesive and non cohesive texts. *Journal of Experimental Child Psychology*, 1979, 27, 129-142.

MOSENTHAL, P., DAVIDSON-POSENTHAL, R., & KRIEGER, V. How fourth graders develop points of view in classroom writing. *Research in the Teaching of English*, 1981, 15 (3), 197-214.

OLSON, G.M., MACK, R.L., & DUFFY, S.A. Cognitive aspects of genre. *Poetics*, 1981, 10, 283-315.

PELLEGRINI, A.D., GALDA, L., & RUBIN, D.L. Context in text: the development of oral and written language in two genres. *Child Development*, 1984, 55, 1549-1555.

PERERA, K. Language acquisition and writing. In P. Fletcher, & M. Garman (Eds.) *Language acquisition: studies in first language development* (2nd ed.). Cambridge: C.U.P., 1986.

RILEY, M.S., GREENO, J.G., & HELLER, J.I. Development of children's problem-solving ability in arithmetic. In H. Ginsburg (Ed.) *The development of mathematical thinking*. New York: Academic Press, 1982.

ROBINSON, C.S., & HAYES, J.R. Making inferences about relevance in arithmetic word problems. *Journal of Educational Psychology*, 1974, 66, 817-825.

RUBIN, D.L. The influence of communicative context on stylistic variation in writing. In A. Pelligrini, & T. Yawkey (Eds.) *The development of oral and written language in social contexts*. Norwood: Ablex, 1984.

SCHNEUWLY, B. *Typologie des typologies de textes*. Communication au "Troisième Congrès International de Didactique du Français", Namur (Belgique), septembre 1986.

SCHNEUWLY, B., & BRONCKART, J.P. Connexion et cohésion dans quatre types de textes d'enfants. *Cahiers de Linguistique Française*, sous-presse.

SHEBILSKE, W.L. Structuring an internal representation of text: a basis of literacy. In P.A. Kolers, M.E. Wrolstad, & H. Bouma (Eds.) *Processing of visible language* (Vol.2). New York: Plenum Press.

SMITH, R.N., & FRAWLEY, W.J. Conjunctive cohesion in four English genres. *Text*, 1983, 3 (4), 347-374.

SPIRO, R.J., & TIRRE, W.C. Individual differences in schema utilization during discourse processing. *Journal of Educational Psychology*, 1980, 72 (2), 204-208.

STEIN, N.L., & POLICASTRO, M. The concept of a story; a comparison between children's and teachers' viewpoint. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.) *Learning and comprehension of text*. Hillsdale: Erlbaum, 1984.

TOWNSEND, M.A.R. Schema shifting: children's cognitive monitoring of the prose-schema interaction in comprehension. *Journal of Experimental Child Psychology*, 1983, 36, 139-149.

VERGNAUD, G. A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T.P. Carpenter, J.M. Moser, & T.A. Romberg (Eds.) *Addition and subtraction: a cognitive perspective*. Hillsdale: Erlbaum, 1982.

VIPOND, D. Micro- and macroprocesses in text comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 276-296.

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# COGNITIVE AND METACOGNITIVE ASPECTS OF TEXT PROCESSING

Beth E. Kurtz
Université de Paris VIII
U.R.A. au C.N.R.S. D1297
Equipe Textima
2, rue de la Liberté
93526 Saint Denis, France

## 1. INTRODUCTION

Metacognition has enjoyed an increasingly prominent position in theories of cognitive development during the past two decades. Although most of the early investigations and theories focused on metamemory, recent work has examined metacognition in relation to other cognitive domains, including reading, comprehension, and attention. Despite the growing body of research supporting the position of metacognition in cognitive theories, there continues to be disagreement surrounding its definition. The most widely accepted definition has two facets. Namely, metacognition is (1) knowledge about, and (2) regulation of cognitive states and processes.

Investigations of the knowledge component of metacognition have typically studied verbal reports about memory and problem-solving strategies. Following Flavell and Wellman's (1977) founding work on metamemory, this knowledge includes the cognitive characteristics of people in general and oneself in particular (e.g., I know that I know the way to the bus stop, even though I am not able to name any of the streets I cross on the way); how task characteristics influence performance (e.g., a written multiple choice test is usually less anxiety-provoking than an oral examination); and information about strategies (e.g., I remember my appointments better if I write them down). Thus this aspect of metacognition is a subset of the general knowledge base, including both procedural knowledge about "how to"(e.g., how my task-approach differs in writing a theoretical paper versus a research report) and declarative knowledge "that" (e.g., my performance on yesterday's exam was impaired because I was tired).

The regulatory component of metacognition is responsible for implementing strategies, monitoring their usefulness, and modifying them when they are judged inappropriate. This aspect of metacognition is viewed as a central processor or executive, capable of performing intelligent evaluations of its own operations (Brown, Bransford, Ferrara, & Campione, 1983). Therefore, this part of metacognition is a key link in the acquisition of new metacognitive information (Pressley, Borkowski, & O'Sullivan, 1985). For example, as the learner, reader, or memorizer monitors performance, she learns that certain strategies are appropriate for some tasks but not others, and modifies old strategies to create new ones, thus adding to her metacognitive knowledge about strategies.

As Brown et al. (1983) have shown, the two components of metacognition complicate the problem of definition. That is, while the knowledge component of metacognition is primarily statable, stable, late-developing, and fallible, the regulatory component, in contrast, is not necessarily statable, somewhat unstable, relatively age independent, and task and situation dependent. Nonetheless, this two-pronged definition by far dominates the field (Brown et al., 1983; Schneider & Pressley, 1988; Yussen, 1985). Further, the definition spans the metacognitive literature. For instance, metalinguistics has been defined in relation to the two components analyzed knowledge and cognitive control (Bialystok & Ryan, 1985). However metacognition is defined, a critical area is its relationship to and distinction from cognitive processes which are not metacognitive. The construct metacognition will be useful in the long run only to the extent that its relationship to cognitive processes such as strategy use and the non-metacognitive knowledge base is defined.

The purpose of this chapter was to examine investigations of metacognition and text processing. This literature has been classified into three domains: metalinguistics, reading strategies, and comprehension monitoring. Metalinguistics is primarily concerned with the reader's objective knowledge of language characteristics, and the conscious manipulation of language independent of its meaning (e.g., Bialystok, 1986; Bialystok & Ryan, 1985; Ryan & Ledger, 1980). The section "metacognition and reading strategies" discusses studies of knowledge about and regulation of strategies used to facilitate comprehension of and memory for what one has read (e.g., Paris & Jacobs, 1984; Short & Ryan, 1984). Finally, comprehension monitoring, which was an object of experimental research for several decades before metacognition was invented, concerns readers' and listeners' awareness of and reaction to comprehension failures (e.g., Flavell, Speer, Green, & August, 1981; Markman, 1979; Wagoner, 1983).

These three categories are by no means independent of one another, and several studies reported here fit into more than one category. However, this format proved useful in beginning to organize the immense volume of research related to metacognitive processes and reading. The goal of this chapter was not to give an extensive review of the extant literature; the interested reader is referred to reviews published elsewhere (e.g., Baker &

Brown, 1982; Forrest-Pressley & Gillies, 1983; Garner, 1987; Hakes, 1980; Wagoner, 1983). Rather, the aim was to provide an overview characterizing state-of-the-art research on metacognition and text processing, and especially to evaluate in this research the relationships among metacognitive and cognitive processes.

## 2. METALINGUISTIC AWARENESS

The term "metalinguistics" has been used to describe our conscious knowledge about language, and regulation of its use. Because language use is a cognitive process, a separate, higher-order understanding of language use may thus be termed "metacognitive". When the language in question is written (versus spoken), then it is relevant to issues of text processing, and is the concern here.

Ryan and Ledger (1984) have defined three types of metalinguistic skills: (1) phonological skills require knowledge of sound-spelling correspondences in written text;
(2) lexical skills pertain to knowledge of the identity and use of words; and (3) syntactic skills entail the ability to utilize sentence information about the order and form of words. Although the research literature does not always neatly conform to this classification system, Ryan and Ledger's rubric will be used as a starting point for this discussion, with one additional category: knowledge about the structure of texts.

### 2. 1. Phonological Skills

Underlying the phonological skills of young children is a basic understanding of the nature of text itself: that is, an understanding that marks on a page, highway signs, and restaurant billboards convey meaning. Thus, metalinguistic phonological skills include the knowledge that (1) letters represent sounds, and (2) those sounds together (and therefore those letters together) are a vehicle of communication.

Early interview studies examining children's understanding of the nature and purpose of reading, and their metalinguistic vocabularies, suggested that conceptual uncertainty regarding the nature of written language contributes to the difficulty of learning to read (Downing, 1969; Reid, 1966). However, more recent, methodologically improved experiments have shown that if the questions are phrased in the right way, very young children demonstrate an understanding of the nature of print. For example, pre-schoolers have some awareness of why people read, and what they do when they read (Hiebert, 1981). Further, before learning to read, young children are knowledgeable about conventions related to reading such as directionality, book handling, and the difference between print and pictures (Day, Day, Spicola, & Griffin, 1981; Sulzby, 1985).

Importantly, children's understanding of the nature of text is related to their reading achievement. For example, children's scores on tests of written language knowledge correlate with later measures of their reading ability (Day et al., 1981; Evans, Taylor, & Blum, 1979). Similarly, children's understanding of the terms "letter", "word", and "sentence" correlates with their reading ability (Mickish, 1974). Tornéus (1984) has shown that spelling abilities are primarily dependent on metaphonological abilities, and that improving the metaphonological skills of dyslexic children improves their spelling.

Taylor, Blum, and Logsdon (1986) used a naturalistic training program to examine the relationships between children's metalinguistic awareness and reading achievement. The target children (kindergartners and preschoolers) were from low-income inner-city homes, where they presumedly had little exposure to print. Taylor et al. trained the children's teachers to develop classroom programs, props, and instructional styles which accentuated the importance and meaning-conveying function of print. In particular, the training program was geared to help children (1) see that written language is functional, (2) understand technical and arbitrary aspects of print --e.g., distinctions among letters, words, and sounds, (3) understand the need to be explicit with written language (writing for an absent audience), and (4) increase awareness of text as a source of information. At the end of the school year, children in implementing classrooms outperformed peers in non-implementing classrooms at letter recognition and matching, ability to segment oral sentences into their component words, and awareness of the technical aspects of written language (Taylor et al., 1986).

A recent and very comprehensive examination of phonological skills in young children distinguished among four types of phonological knowledge: concepts about print, graphic awareness, phonemic awareness, and grapheme-phoneme correspondence (Lomax & McGee, 1987). Lomax and McGee measured 3-, 4-, 5-, and 6-year-olds on the above variables, and also on their ability to read isolated words. Concepts about print included children's awareness of written and oral language units, and the correspondence between those units and meaning; knowledge of the purposes and processes of reading; knowledge of the conventions of reading (e.g., left to right); and the ability to recognize print embedded in an environmental context. Graphic awareness represented awareness of the distinctive features in letters and words (e.g., correctly matching letters from a display). Phonemic awareness was measured as children's ability to segment phonemic elements of words and to note similarities and differences among them. Finally, grapheme-phoneme correspondence knowledge was assessed as children's knowledge of the sound units associated with letters, and their ability to apply that knowledge to blend new words or decode unknown words. Through a series of causal models, Lomax and McGee demonstrated that all four metalinguistic constructs were required as separate theoretical entities in order to best explain the variance in the data, and that the four, via direct or indirect paths, predicted reading ability. Correlations between reading and metalinguistic measures were consistently highly significant, ranging from .40 to .80.

## 2. 2. Lexical Skills

In two experiments, Ryan, McNamara and Kenney (1977) examined the lexical skills of young children using tasks which are characteristic of this domain. Namely, children performed tasks measuring their ability to (1) accurately tap the table once for each word while repeating sentences; (2) distinguish among stimuli ranging from single phonemes to two-syllable words and two word phrases as "word", "not a word", or "two words"; (3) in a sentence comparison task, identify the word(s) which distinguished the second sentence from the first; (4) produce sentences using multifunction words in ways differing from those of the experimenter; and (5) separate run-together sentences into words by inserting slashes. In Experiment 1, first and second graders who were average in reading ability were tested on the metalinguistic measures and the Slosson Reading Test; in Experiment 2, the subjects were third and fourth graders receiving remedial reading instruction. Median splits on reading ability in each group showed that better readers outperformed poorer readers on most of the metalinguistic measures.

These results were consistent with those obtained in other studies in which children's reading ability was significantly related to word discrimination performance (Bowey, Tunmer, & Pratt, 1984), and to flexible and accurate use of nonsense words (Smith & Tager-Flusberg, 1982). Further, a metalinguistics measure which included ability to combine separate phonemes into words and pseudowords, and visually distinguishing words with and without sentence contexts correlated significantly with first and third graders' reading scores, and a measure of metamemory (Borkowski, Ryan, Kurtz, & Reid, 1983).

## 2. 3. Syntactic Skills

In a well-known study of children's understanding of language structure, Gleitman, Gleitman, and Shipley (1972) demonstrated that children as young as two and a half years reject improperly formed sentences. Gleitman et al. presented toddlers with sentences which varied on two dimensions: intonation contour (normal length: "Bring me the ball" versus telegraphic length: "Bring ball") and word order (correct: "Bring ball" versus inverted:

"Ball bring"; "Ball me the bring"). The reversed-order sentences consistently resulted in more judgments of "silly"; however, only one of the three girls rejected the telegraphic sentences.

Gleitman et al. (1972) then gave seven children ranging from 5- to 8-years a series of sentences which varied on their grammaticality. The children's judgments showed a systematic progression in similarity to adult judgments, with 8-year-olds giving the most adult-like responses. While the youngest children were consistently sensitive to semantic deformations (e.g., "Golf play my brother"), only the older children picked up syntactic violations (e.g., "Boy is at the door"). Although Gleitman et al. did not presume to present normative data generalizable to other populations, their results provided an intriguing glimpse of the ideas young children have about language structure.

Isakson and Miller (1976) demonstrated that good comprehenders are more sensitive to syntax than poor comprehenders. Fourth graders who were equated on word recognition ability, but who differed on comprehension ability, read sentences which were normal, had semantic violations, or semantic and syntactic violations. A sample triad of sentences is the following:

A. The old farmer <u>planted</u> the bean seeds in the rich, brown soil.
B. The old farmer <u>paid</u> the bean seeds in the rich, brown soil.
C. The old farmer <u>went</u> the bean seeds in the rich, brown soil. (p. 789, Isakson & Miller, 1976).

The subjects read the sentences aloud, and their spontaneous errors were recorded as a measure of the degree to which they were disrupted by the violations. Poor comprehenders made the same number of errors across sentence types. Good comprehenders, in contrast, made an increasing number of errors with increasing complexity (i.e.,. violations) of the sentence structure. Isakson and Miller concluded that good comprehenders are more sensitive metalinguistically than poor comprehenders.

Developmental studies have shown that older children are more adept than younger children at judging the grammaticality of negative sentences (Scholl & Ryan, 1975, 1980) and correcting syntactic errors in sentences (Ryan & Ledger, 1980). Significant although moderate correlations have been reported between children's reading scores and metalinguistic skills as measured by judgments of sentence grammaticality and error corrections (Borkowski et al., 1983; Ryan & Ledger, 1980).

## 2. 4. Knowledge about the Structure of Texts

In addition to the meaning conveyed by individual words, and the rule-based grammars which govern syntax, another characteristic of written language is the structure of texts--the sequences in which ideas are placed, the placement (or absence) of topics sentences, and so forth. Just as young readers' and pre-readers' metalinguistic understanding of phonemes, words, and syntax is related to their ability to read, so might their verbalizeable knowledge about text structures be related to reading performance.

Several studies have examined children's developing awareness of text structure. Danner (1976) presented topically organized and disorganized passages to children in Grades 2, 4, and 6. He found that awareness of the structural differences increased with age, as well as children's ability to sort sentences by topics, and their selection of review notes on the basis of passage organization. Williams, Taylor, and Ganger (1981) asked children in Grades 4 and 6 to select the best titles for a series of expository texts, and to produce a main-idea sentence for each passage. Sixth graders made more appropriate specific-topic responses than did fourth graders. Garner, Alexander, Slater, Hare, Smith, and Reis (1986) asked average readers from Grades 3, 5, and 7 to construct paragraphs from sentence strips, and in a second task, to place constrained sentences in appropriate positions in pre-structured paragraphs. Garner et al. found that the oldest students were most effective at excluding unrelated information, and were most accurate in pairing cohesively tied sentences.

Differences in awareness of text structure have also been found between good and poor readers. Meyer, Brandt, and Bluth (1980) asked good, average, and poor comprehenders to read and recall well-organized expository passages. When organizing their recalls, most good comprehenders used the same organizational pattern that the text author had initially employed, while most poor comprehenders did not. In addition, students who are aware of the characteristics of text structure are more likely to use a structural strategy when they read than are unaware students (Richgels, McGee, Lomax, & Sheard, 1987), and show improved recall on well-organized texts in comparison to unaware readers (Taylor & Samuels, 1983).

An inherent aspect of text structure is the relative importance of idea units. Readers' assignment of importance level to text units, and their recall as a function of importance level has been called metacognitive behavior (Brown & Smiley, 1977, 1978). While we have cautioned elsewhere about labeling these processes as "metacognitive" (see Kurtz & Schneider, 1988), a brief summary of some key research findings from this area will be included here.

Recall of prose passages by children and adults follows the importance-hierarchy rankings of adults. That is, readers tend to recall the greatest number of units which have been independently rated by adults as "most important", a lesser number of average-important units, and few "least important" units (Brown & Smiley, 1978; Brown, Smiley, Day, Townsend, & Lawton, 1977; Denhière & Legros, 1987; Moore & O'Driscoll, 1983).

Interestingly, even though young children produce stable importance rankings which are different from those of adults (usually, humorous, frightening, or otherwise highly-salient information is ranked as most important), their recall patterns conform to the importance rankings of adults, rather than to those of their own age-mates (Denhière & LeNy, 1980).

With additional study time, adult recall improves as a function of importance level: university students improve their recall of important units, but not of less-important units; this pattern is not found in fifth graders (Brown & Smiley, 1978). Brown and Smiley posited that mature readers' protocols reflect sophisticated metacognition. That is, with additional study time, the older readers focused on important text units, used effective strategies like note-taking and underlining, and consequently showed improved recall of important, but not unimportant, text units.

Consistent with this hypothesis, Rinehart, Stahl, and Erickson (1986) found that summarization instruction improved sixth graders' recall of important, but not unimportant, text units. Path analyses indicated that students' note-taking played a crucial intermediate role in influencing students' recalls, and was more important than amount of study time. Rinehart et al. argued that these results confirm a metacognitive hypothesis; namely, that summarization instruction resulted in a focus on those text units judged to be most important, with a subsequent improvement in recall of those units.

### 2. 5. Discussion: Metalinguistics and Reading

This research area has shown some convincing evidence of reliable relationships between metalinguistic skills and reading performance. Unfortunately, that relationship is blurred by lack of clarity regarding the construct metalinguistics. If metacognitive research as a whole has suffered from a fuzzy definition, this is particularly true of metalinguistics. As Bialystok and Ryan (1985) have pointed out, metalinguistic abilities as they have been defined and measured can be found or not found in any age group studied. The tasks used to measure metalinguistic skills are sufficiently diverse that individual variations in performance are difficult to interpret. In the studies described above, measurement of metalinguistics included such diverse tasks as sound blending (Tornéus, 1984), ability to distinguish print from pictures (Lomax & McGee, 1987), and table tapping to demonstrate recognition of what a "word" is (Ryan et al., 1977). Clearly, this area could profit from a careful re-examination of the definition of metalinguistics, perhaps following a classification system such as that of Ryan and Ledger (1984).

Attempts at redefinition should consider the question of consciousness (e.g., Karmiloff-Smith, 1986). In other words, how much of metalinguistic skill is verbalizeable knowledge of which the reader is conscious, and how much is automatized, unconscious processing? Second, theorists might consider developmental implications of the concept. One result of the current vagueness in terminology is that research in metalinguistics appears age-graded with little cross-over among domains: Phonetic skills are primarily relevant for pre- and beginning readers, lexical skills for beginning readers, while only sophisticated readers possess accurate knowledge about text structure.

A larger concern is whether or not metalinguistic skills are a subset of metacognitive processes. If metacognition is defined as knowledge about and regulation of cognitive states and processes, a critical distinction must be made here between knowledge about language, and knowledge about language use. Knowledge about language use (e.g., recognizing that a focus on important units of text will aid future memory for that text; recognizing that letters, words, and sentences convey meaning for sharers of a common language) is more arguably metacognitive than is knowledge about language (e.g., English sentences always have a verb; the "g" in "gift" is different from the "g" in "enough"). Knowledge about language--whether about phonemes, grammar, or text structure--might be called metalinguistics, but I would argue that most of such knowledge is not metacognitive as metacognition is commonly defined. However, many of the studies cited above did examine knowledge about language use (e.g., Lomax & McGee, 1987; Taylor et al., 1986).

A final point directed to this discussion, but also relevant for metacognitive research as a whole, regards the inference of metacognitive processes underlying cognitive behaviors. It is too easy, as skilled readers, writers, and speakers, to observe less-skilled readers, writers, and speakers behave in a seemingly informed manner (e.g., correcting the grammatical errors of a younger sibling; placing topic sentences at the beginning of paragraphs) and to conclude that these actions are metacognitively-informed. An example of this relates to the "metacognitive" label which many researchers have attached to the finding that with increased study time, adult text recall (or with instruction, children's recall) improves as a function of importance of text units (Brown & Smiley, 1978; Rinehart et al., 1986). Although a deliberate decision to focus on specific units because of their perceived importance is obviously metacognitive, text recall alone is not. In other words, only when it is demonstrated that readers' strategic behaviors are indeed metacognitively-informed actions does the "meta" label apply (Kurtz & Schneider, 1988). Without a convincing demonstration that a reader's text recall is determined by metacognitively directed strategy use, the recall of text units according to importance level must be considered a cognitive, not a metacognitive, phenomenon.

## 3. METACOGNITION AND READING STRATEGIES

Effective strategies accomplish cognitive purposes (Pressley, Forrest-Pressley, & Elliott-Faust, 1987). Thus efficient reading--whether to understand, or to remember--is often characterized by the use of metacognitively-guided strategies (Baker & Brown, 1984). While the first investigations of readers' knowledge about strategies were primarily descriptive, recent work has used experimental manipulations and instructional programs to build theory and to develop practical programs aimed at ameliorating metacognitive deficits in poor readers. This section will first address experimental work which defined good and poor readers' metacognitive knowledge about reading strategies, and then turn to instructional studies. A large body of literature exists concerning study behaviors, much of which is theoretically linked to this examination of metacognition and reading strategies. However, because of space restraints and the enormity of this area, the interested reader is again referred elsewhere (Anderson & Armbruster, 1984).

### 3. 1. Knowledge About Reading Strategies

The first wave of research examining metacognitive knowledge about reading strategies was characterized by more-less comparisons: poor (versus good) or younger (versus older) readers were asked questions about reading strategies or reading itself, and were found to be deficient compared to peers. For example, older children know more than younger children about their limitations in reading ability, how to use context when deducing word meaning, the usefulness of contextual cues when performing a cloze task, and strategies which can be used to resolve comprehension failures (Byrd & Gholson, 1985; Myers & Paris, 1978). Poor readers are not as aware as are good readers of the negative influences of some strategies (Paris & Myers, 1981). Similarly, comparisons of adult good and poor readers have shown differences in awareness of the characteristics of a good reader, and knowledge about strategies useful in the event of comprehension failure (Gambrell & Heathington, 1981). Adult readers, in comparison to children, know more about summarization strategies, and provide better matches between reading strategies and task demands (Aaronson & Ferres, 1986; Brown & Day, 1983).

Readers' strategic behaviors are likely driven by their understanding of the goals of reading. That is, a young child who views reading as "sounding out the words" is not as likely to monitor comprehension, nor to possess a ready repertoire of comprehension-coping strategies as is a child who views reading as a search for meaning. In fact, younger (poorer) readers tend to view reading as a decoding task, while older (better) readers view it as a meaning comprehension task (Gambrell & Heathington, 1981; Johns, 1974; Johns & Ellis, 1976; Myers & Paris, 1978; Paris & Myers, 1981).

As most of these early studies did not include any measures of reading ability or reading strategies, the only inferences we are able to draw about links between cognitive and metacognitive performances are indirect, based on what is known about the contrasting performances of older versus younger, or good versus poor readers. However, some studies have examined these relationships more directly. Lipson, Irwin, and Poth (1986) reported correlations ranging from .30 to .60 between metacognitive knowledge about reading strategies, and both cloze performance and postreading comprehension. Although the correlation values themselves were not reported, Paris and Myers (1981) found that text recall and knowledge about reading strategies were related in a group of fourth grade good and poor readers. In particular, performance and knowledge about strategies were related for precisely those metacognitive items on which the poor readers had been shown to be deficient. In contrast, Byrd and Gholson (1985) found generally low correlations among metareading, reading ability, and memory performance measures. However, it should be noted that the authors did not use maximally powerful statistical procedures (e.g., did not aggregate data; Rushton, Brainerd, & Pressley, 1983), and that their reading measures, subtests from John's Basic Reading Inventory, were measures of reading level rather than strategy use or transfer.

One problem with these early studies is a general coarseness or lack of focus in framing the hypotheses, and a lack of experimental designs. Taking a group that is already deficient in one sense (here, poor readers) and showing them to be deficient in another, related, area (e.g., their ability to describe what actions they might take when encountering an unknown word), is a useful first step in understanding their lack of skill, but little more. It is also regrettable that the reported studies represented a hodgepodge of testing materials and procedures, with very little information provided about test-retest reliability, internal validity, or construct validity of the metareading measures employed. Thus many of these studies suffered from both a lack of a clear theoretical focus, and from poor measurement.

On the other hand, these projects supplied some helpful preliminary evidence that metacognitive knowledge about reading might be an important factor influencing reading performances. The robust finding that poor readers view reading as a decoding task, while better readers view it as a comprehension task is a case in point: this finding alone has prompted a wealth of further investigations. Finally, the correlational data reported provide converging evidence that metareading might indeed be an important causal factor determining readers' deployment of reading strategies, and, therefore, their comprehension of and memory for texts. The strength of that relationship has been more thoroughly examined through training studies, which will be discussed next.

### 3. 2. Instruction of Reading Strategies

As in other areas of metacognitive research, the second generation of investigations of metareading skills has used more complex experimental designs and more sophisticated statistical analyses to examine the impact of metacognitive knowledge about reading on performance. In particular, this area has enjoyed a proliferation of training studies, which have used experimental manipulation to observe how changes in knowledge about reading strategies influence reading performance and knowledge about reading.

Many of these instructional programs occurred in the classroom, and were carried out by the regular classroom teacher. For example, in a study by Sivan, Book, Meloth, and Putnam (1987), 20 third-grade teachers participated in either a training or a control condition. In half of the classrooms, the teachers instructed reading strategies, including explicit instructions regarding flexible and adaptive use of the strategies. Interview probes of pupils' metacognitive awareness of the reading lessons were conducted after every lesson, and at the school-year's end, target pupils responded to questions about the goals of reading and reading strategies. Sivan et al. (1987) found that trained classrooms showed higher metacognitive knowledge at posttest than control classrooms, and that for trained children, pupils' awareness of lesson content was significantly related to end-of-the-year reading knowledge. Unfortunately, Sivan et al. did not report measures of reading behaviors.

Both reading performance and metareading were assessed in a classroom intervention project by Paris and his colleagues (Cross & Paris, 1988; Paris, Cross, & Lipson, 1984; Paris & Jacobs, 1984). 8- and 10-year-olds were pre- and posttested on (1) the Comprehension subtest of the Gates- MacGinitie, (2) error detection, (3) cloze task, (4) awareness about reading, and (5) knowledge about reading strategies. The reading awareness measure inquired about evaluation of the reading task, planning to reach a specified reading goal, and regulating reading through the use of monitoring strategies (Paris & Jacobs, 1984). Children received four months of instruction about reading comprehen-sion strategies, and how, when, and why to use them. After intervention, trained children showed increased reading awareness, and improved performance on the error detection and cloze tasks. No correlations between metareading knowledge and reading performances were reported.

Short and Ryan (1984) examined the interactive effects of reading strategies and attributional beliefs in fourth grade poor readers. All children were measured on baseline reading ability, awareness of the goals of reading, and strategic behavior and recall for narrative texts. Strategy training included a description of the use of "wh" questions (who, what, where, when, why) as an aid to comprehension monitoring and preparation for recall, and other reading strategies (e.g., note question cues in the margins, underline important phrases). In addition, children in the Strategy-Attribution condition received attributional retraining regarding the importance of effortful behavior in determining task outcome, and instruction to give themselves positive self-statements about reading. Children who received Strategy training showed better recall and strategy use after training than did control children, and also demonstrated improved awareness of the goals of reading. Further, children who received both Strategy and Attribution training were superior on a transfer error detection and correction task when compared to Strategy-only and Control children (Short & Ryan, 1984).

Palincsar and Brown (1984) used a reciprocal teaching method to instruct comprehension fostering and comprehension monitoring strategies to poor readers. Through the reciprocal method, students actively interacted with the teacher while developing predicting, questioning, clarifying, and summarizing skills. Training, which included positive attributional messages about students' competence and control, resulted in impressive gains in summarization and comprehension on laboratory tasks, and in improved reading skills in the classroom. In a carefully controlled follow-up to Palincsar and Brown's findings, Lysynchuk, Pressley, and Vye (1988) evaluated the usefulness of the reciprocal teaching method with Grade-4 and -7 comprehenders. Importantly, Lysynchuk et al. included pre- and post-experiment assessments of standardized reading achievement, and compared reciprocal instruction to a control condition which differed from the experimental condition only in that it did not receive strategy training. Greater gains in standardized reading comprehension for the trained students documented the value of the reciprocal instruction method.

Although the above studies showed that children's use of reading strategies improves with instruction, none of them adequately demonstrated that metacognitive instructions are superior to intensive traditional reading instruction. Hasselhorn and Körkel (1986) addressed this issue in German schools, also controlling for readers' prior knowledge about the text topic. Traditional reading instruction was modeled after German exercise books; instructions included emphatic reading of texts, retelling of contents, identification of syntactic and grammatical components of the sentences, and discussion of text content in detail. Metacognitive instructions were designed to increase children's awareness about reading, and their self-regulation skills in text processing. The instructions were especially geared to foster a flexible, goal-adaptive use of the specific reading strategies instructed. The sixth grade participants had been previously identified as soccer experts or novices. Hasselhorn and Körkel found that while both instructional programs were effective in improving comprehension performance and metacognitive knowledge, benefits of the traditional approach were largely restricted to soccer experts, while novices benefitted from metacognitive instructions. Correlations between metacognitive and reading performance measures were fairly high before instruction; these correlations were less after instruction, but still primarily in the .40 to .50 range (Hasselhorn & Körkel, 1986).

### 3. 3. Discussion: Reading Strategies

The above training studies, while only a sampling of recent investigations of instructional programs, provide conclusive evidence that metacognitive instruction results in improved reading skills. Teaching reading strategies, and especially teaching reading strategies in conjunction with metacognitive knowledge about why and when to use them, and instructions to monitor their usefulness, resulted in improved performance on cloze tasks, comprehension tests, error-detection tasks, as well as metacognitive tests about reading and reading strategies. In a few of the studies, correlational data provided converging evidence for the relationship between reading behaviors and metacognitive knowledge.

These studies also speak to the complex interrelationships among children's cognitive, metacognitive, and personality-motivational behaviors and beliefs. Results of the Palincsar and Brown (1984) study showed the value of instructing monitoring behavior simultaneously with task-appropriate strategies. Short and Ryan (1984) showed that instructional effects are enhanced by changing readers' attributional beliefs about why they fail, and the relationship between effortful behavior and successful performance. Transfer data from several of the projects similarly implied the complex interrelationships among these variables: changing motivation resulted in improved monitoring (Short & Ryan, 1984); improved knowledge about strategies was accompanied by an improved understanding of the nature of the reading task itself (Paris & Jacobs, 1984); instructional effects were mediated by readers' prior knowledge related to the text topic (Hasselhorn & Körkel, 1986).

Two notes of caution are in order here. First, as Pressley et al. (1987) have pointed out, there is little empirical evidence for the usefulness of many strategies which are recommended in the reading instructional literature, and which are frequently included in instructional programs (cf. Pressley, Johnson, Symons, McGoldrick, & Kurita, in press). Second, and relatedly, too many strategy instructional programs are complex mismashes of information, integrating the cognitive, metacognitive, and motivational factors as mentioned above, but with the unfortunate result that no clear conclusions may be drawn from the experimental results about the contributions of the various components. Thus, in summary, although this research area has made valuable contributions to our understanding that cognitive and metacognitive instruction about strategies can improve reading, more research is needed to define the specific roles of specific strategies in children's reading behaviors, to identify specific roles of different aspects of metacognition, to determine the direction of causality among cognitive and metacognitive factors, to test competing theories, and finally, to learn the best way to teach strategies.

## 4. COMPREHENSION MONITORING

Comprehension of written (and spoken) messages is an integral part of reading (and listening). When comprehension failures occur, strategic readers stop and reread, reflect, or attempt in other ways to locate the source of the failure. However, sometimes readers believe they have comprehended a text, when in fact they have failed. Readers' (and listeners') regulation of whether or not comprehension has occurred is called comprehension monitoring. The discussion here includes reports of research concerning oral as well as written discourse. A detailed analysis of these areas treated separately may be found in Wagoner (1983).

Most research on comprehension monitoring has measured readers' and listeners' detection of either inconsistencies or inadequate (missing) information in spoken or written messages. For example, in one of the earliest investigations of comprehension monitoring, Markman (1977) gave first and third graders inadequate instructions regarding how to play a card game and how to perform a magic trick. She found that while Grade 3 children were moderately sensitive to gaps embedded in the instructions, first graders rarely noticed them. That is, the younger children reported that they had understood the instructions, and only when they attempted to play the card game or perform the magic trick themselves did they realize that the instructions had been inadequate. In a second series of experiments, Markman (1979) found that third, fifth, and sixth graders were more likely to notice explicit than implicit contradictions in text, but even the oldest children frequently failed to monitor comprehension. Markman suggested that in order to notice inconsistencies, children must encode and store the new information, draw relevant inferences from it, retrieve and maintain the inferred propositions in working memory, and compare them to new information as it is read or heard. Young children do not seem to perform this series of complex cognitive behaviors spontaneously (Markman, 1979).

Developmental studies and comparisons of good and poor readers have found reliable differences in monitoring performances on those dimensions. In general, younger children are less aware of ambiguities in messages (Patterson, Cosgrove, & O'Brien, 1980; Patterson, O'Brien, Kister, Carter, & Kotsonis, 1981), tend to blame the listener rather than an inadequate message for communication failure (Robinson & Robinson, 1976), and are less able to locate and correct specific sources of inconsistencies in written texts (Garner & Taylor, 1982). Similarly, poor readers (i.e., poor comprehenders) are less successful than good readers at detecting inconsistencies in texts (Garner & Kraus, 1981-82), and at spontaneously correcting them (Beebe, 1980). Owings, Petersen, Bransford, Morris, and Stein (1980) found differences in the monitoring behavior of successful and less-successful fifth graders related to their performances on a story remembering task. Owings et al. gave the children stories which varied in the degree to which they made sense. For example, in the easy version of a story, the hungry boy ate a hamburger, while in the difficult version, the hungry boy played baseball. Successful students spontaneously monitored as they read, and were aware of why they had trouble remembering the randomly-constructed stories, in contrast to less-successful students, who were unable to rate story difficulty accurately. When the children were allowed to regulate their learning time, successful students spent more time on the difficult stories, while less-successful students spent roughly equal time studying each story (Owings et al., 1980). Other research has shown that inconsistencies are most easily noted in texts when they are explicit rather than implicit (Markman, 1979), when the information is syntactically marked as new (Glenburg, Wilkinson, & Epstein, 1982), and when the text topic is familiar and the text is well-organized (Raphael, Myers, Tirre, Fritz, & Freebody, 1981).

It has been argued that younger and poorer readers do not monitor comprehension well because they do not make the inferences that adults and better readers make automatically, and because, in general, they do not process texts as actively (Capelli & Markman, 1982; Ryan, Ledger, Short, & Weed, 1982). This difference in processing is likely tied to poor readers' deficient metacognitive knowledge (Ryan et al, 1982). For starters, young children probably do not know the difference between knowing and guessing when encountering inconsistencies. For instance, Flavell et al. (1981) found that young children often resolved ambiguities in messages arbitrarily, and then judged that the messages were unambiguous. Further, because younger and poorer readers have a poorer understanding of reading strategies, and also of the nature of texts themselves, they are less successful at reading purposely and systematically, at recognizing structure in texts which might be integrated into their previous knowledge structures, and finally, at comprehending and at monitoring comprehension. In fact, there are so many reasons why poorer and younger readers might show deficient monitoring skills, that a problem of this research area has been to sort out the various possibilities.

One way researchers have attempted to learn more about the causes and possible ameliorations of poor comprehension monitoring has been through instructional studies. Simple instructions to look for problems increase the probability of error detections, and especially increase the detection of the type of errors which the reader was instructed to seek. For instance, Markman and Gorin (1981) told 8- and 10-year-olds to find the problems in essays, defining a problem as "something that might confuse people or something that people might have trouble understanding" (p. 321). Some of the children were then given examples of falsehoods, one treatment group was given examples of inconsistencies, and one group received only the find-problems instructions. Markman and Gorin found that problem detection corresponded to children's set: Children in the falsehood condition found more falsehoods, while children in the inconsistency condition were more likely to locate inconsistencies. Problem detection in the neutral condition was quite low.

Other instructional studies have demonstrated that children's comprehension monitoring may be improved with imagery instructions (Gambrell & Bales, 1986), hypothesis-testing (Capelli & Markman, 1982), self-questioning (Wong & Jones, 1982), a cloze strategy (Dewitz, Carr, & Patberg, 1987), or general instruction of reading strategies (Paris & Jacobs, 1984). In fact, any procedure which encourages the reader to process the text more actively seems to result in improved comprehension monitoring. One recent instructional study which contrasted the relative benefits of several intervention programs will be described in greater detail.

Elliott-Faust and Pressley (1986) used 8 distinct instructional programs to determine whether comparing different parts of texts increases children's error detections, and if self-control training produces more durable use of the strategy. Third graders in the three comparison processing conditions were taught local comparison only (i.e., analyzing the two most recent sentences), both local and wholistic comparison (i.e., comparing the most recent sentences with each other and with the meaning of the entire passage), or local and wholistic comparison with self-instructions. The self-instructions included knowledge about the value of the comparison strategy, when to use it, and instructions to check its effectiveness periodically. In other words, children in this condition were given additional metacognitive knowledge which presumeably would increase their probability of maintaining the strategy. Children in a "passive" training condition received additional information from the experimenter about the texts, which resembled ideas generated by children in the comparison groups. These four training conditions were evaluated against the three conditions used by Markman and Gorin (1981). Namely, one group was given practice listening to stories ("to see if they make sense", p. 28); one group was given a standard of what constitutes a sensible passage; and one group was given both the standard and practice listening to stories. A final, eighth group received only the posttest measures.

Elliott-Faust and Pressley found that children who received local and wholistic comparison with self-instruction outperformed all other groups on monitoring tasks given immediately after training and one week later. Although children who received both local and wholistic training showed superior performance immediately after training, they did not maintain the strategy as did children in the self-instruction group. This study gives us important evidence that (a) children's monitoring behavior can be improved by teaching them to actively compare different texts components, (b) the comparison of contradictory text components training was more effective than exposure to stories, receiving an appropriate standard of evaluation, or the "passive" training, which gave children additional information about the texts, and (c) training which included metacognitive knowledge about when and why to use the strategy, and instructions to monitor its effectiveness resulted in successful strategy maintenance (Elliott-Faust & Pressley, 1986).

One conclusion that may be drawn from the training projects described above is that comprehension monitoring behaviors can be easily modified. A second possible conclusion is that with only minimal instruction, children might not be understanding the task demands of error detection tasks. Winograd and Johnson (1980) have evaluated the error detection paradigm, identifying a number of reasons why young children might have problems with it, including the lack of an established criteria for adequate comprehension, and the overuse of probes. Although the relative problems inherent in this approach versus its validity are still under debate, it is clear that accurate error detection should not be equated with good comprehension monitoring. That is, comprehension monitoring is a complex process which alternately results in comprehension, noncomprehension, rereading, a conscious decision to "forget" part of the text because it seems unimportant to the reader's needs or inconsistent with the main tenets of the text, and so forth. Obviously, error detection is but one small part of this process. Indeed, when a skilled reader reads a text, it is seldom for the purpose of detecting errors. It might be argued that a skilled reader who, as he reads, is busily incorporating ideas into his existing knowledge schemas and making appropriate inferences would be most unlikely to detect errors embedded in texts. How many of us have required a sixth or seventh reading of one of our own articles to detect the word that was used inappropriately or the negative that completely reversed one of our arguments, simply because during every previous reading, the mind had efficiently made the appropriate inferential leap, causing the eye to ignore the misplaced word?

Another problem related to error detection tasks is the confounding factor of comprehension. If poor comprehenders are inefficient at detecting errors and omissions in text, it may be due to their lack of comprehension, rather than their lack of comprehension monitoring. In fact, there is a danger of circularity in this research area as a whole, which creates methodological difficulties. Namely, whenever poor comprehenders read a text and believe they have understood it while in fact they have not, it is virtually impossible to conclude whether they have failed to monitor comprehension, or have simply failed to comprehend.

Nevertheless, good comprehension monitoring is clearly a key ingredient in good text processing skills, and is also an area where understanding the interface between metacognitive (here, monitoring behavior) and cognitive processes is informative in understanding how people read. As a final note, the training study by Elliott-Faust and Pressley is noteworthy in its rigorous evaluation of separate treatment components, and its demonstration of how the successful integration of cognitive and metacognitive components in an instructional program may result in durable performance gains.

## 5. IMPLICATIONS FOR FUTURE RESEARCH

Several themes emerge from this examination of cognitive and metacognitive aspects of text processing. First, the research described here illustrates not only that metacognitive factors bear an important role in skilled reading, but more broadly, that cognitive, metacognitive, and motivational factors act and interact together to determine readers' comprehension of and memory for texts. Results of the instructional studies in particular illustrate that good readers possess task-appropriate strategies, metacognitive knowledge about where to apply the strategies and how to adapt them, good monitoring skills, and adequate motivation to tackle tasks, including adequate motivation to deploy strategies (Pressley, Borkowski, & O'Sullivan, 1984).

A second, related theme is that poor readers are passive readers (Ryan et al., 1982). Research in the reading strategies section shows that poor readers do not actively deploy strategies. Comprehension monitoring research indicates that poor readers fail to actively process texts, failing to activate prior knowledge to compare with just-read information, and failing to compare segments of texts with each other. As Ryan et al. (1982) have argued, this passive processing of texts likely results both from deficient metacognitive systems, and from motivational problems.

Although we still have much to learn about text processing, we can begin to be concrete in recommendations for intervention programs, based on the findings cited above. Reading performances may be improved by instructions which are geared towards improving the motivational state of the reader, which are sensitive to the reader's prior knowledge of text material and knowledge about language structure and usage, and which include metacognitive information about reading and reading strategies.

Strategy instruction will be most effective when it includes metacognitive information about how to use strategies flexibly, especially with explicit instructions regarding monitoring and strategy modification.

Finally, much more research is needed, both to provide more rigorous tests of hypotheses regarding specific strategy and knowledge components as they relate to particular reading skills, and also research bridging the three domains discussed here, to test competing theories. Research programs need to be more programmatic in order to tease out the effects of individual variables (e.g., Elliott-Faust & Pressley, 1986); at the same time, research needs to be more complex in order to permit simultaneous comparison of the many factors that influence reading performance. Such research will allow us to understand the relative strengths of each component as well as interactions among them as they influence reading performance, thereby enriching theory, and guiding instruction.

* Author's current address: Department of Psychology; Davie Hall, CB 3270; University of North Carolina; Chapel Hill, NC 27599-3270

ACKNOWLEDGEMENTS

I am grateful to Michael Pressley, Wolfgang Schneider, and Martha Carr for their comments on earlier drafts of this manuscript.

## REFERENCES

Aaronson, D., & Ferres, S. (1986). Reading strategies for children and adults: A quantitative model. *Psychological Review*, 93, 89-112.

Anderson, T. H., & Armbruster, B. B. (1984). Studying. In D. Pearson (Ed.), *Handbook of reading research* (pp. 657-679). New York: Longman.

Baker, L., & Brown, A. L. (1984). Metacognitive skills in reading. In D. Pearson (Ed.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Beebe, M. J. (1980). The effect of different types of substitution miscues on reading. *Reading Research Quarterly*, 15, 324-336.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, 57, 498-510.

Bialystok, E., & Ryan, E. B. (1985). Towards a definition of metalinguistic skill. *Merrill-Palmer Quarterly*, 31, 229-251.

Borkowski, J. G., Ryan, E. B., Kurtz, B. E., & Reid, M. K. (1983). Metamemory and metalinguistic development: Correlates of children's intelligence and achievement. *Bulletin of the Psychonomic Society*, 21, 393-396.

Bowey, J. A., Tunmer, W. E. & Pratt, C. (1984). Development of children's understanding of the metalinguistic term Word. *Journal of Educational Psychology*, 76, 500-512.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of child psychology :Cognitive development* (pp. 77-166). New York: Wiley.

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22, 1-14.

Brown, A. L., & Smiley, S. S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 48, 1-8.

Brown, A. L., & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development*, 49, 1076-1088.

Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M., & Lawton, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development*, 48, 1454-1466.

Byrd, D. M., & Gholson, B. (1985). Reading, memory, and metacognition. *Journal of Educational Psychology*, 77, 428-436.

Capelli, C. A., & Markman, E. M. (1982). Suggestions for training comprehension monitoring. *Topics in Learning and Learning Disabilities*, 2, 87-96.

Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, 80, 131-142.

Danner, F. W. (1976). Children's understanding of intersentence organization in the recall of short descriptive passages. *Journal of Educational Psychology*, 68, 174-183.

Day, K. C., Day, H. D., Spicola, R., & Griffin, M. (1981). The development of orthographic linguistic awareness in kindergarten children and the relationship of this awareness to later reading achievement. *Reading Psychology*, 2, 76-87.

Denhière, G., & Legros, D. (1987). L'interaction narration*description dans le récit I. Etude de la mémorisation de differents types de séquences descriptives. *L'Anée Psychologique*, 87, 345-362.

Denhière, G. & LeNy, J. F. (1980). Relative importance of meaningful units in comprehension and recall of narratives by children and adults. *Poetics*, 9, 147-161.

Dewitz, P., Carr, E. M., & Patberg, J. P. (1987). Effects of inference training on comprehension and comprehension monitoring. *Reading Research Quarterly*, 22, 99-119.

Downing, J. (1969). How children think about reading. *Reading Teacher*, 23, 217-230.

Elliott-Faust, D. J., & Pressley, M. (1986). How to teach comparison processing to increase children's short- and long-term listening comprehension monitoring. *Journal of Educational Psychology*, 78, 27-33.

Evans, M., Taylor, N., & Blum, I. (1979). Children's written language awareness and its relation to reading acquisition. *Journal of Reading Behavior*, 11, 331-341.

Flavell, J. H., Speer, J. R., Green, F. L., & August, D. L. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development*, 46, 1-65.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, Jr., & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, N. J.: Erlbaum.

Forrest-Pressley, D. L., & Gillies, L. A. (1983). Children's flexible use of strategies during reading. In M. Pressley and J. R. Levin (Eds.), *Cognitive strategy research* (pp. 133- 156). New York: Springer Verlag.

Gambrell, L. B., & Bales, R. J. (1986). Mental imagery and the comprehension- monitoring performance of fourth- and fifth-grade poor readers. *Reading Research Quarterly*, 21, 454-464.

Gambrell, L. B., & Heathington, B. S. (1981). Adult disabled readers' metacognitive awareness about reading tasks and strategies. *Journal of Reading Behavior*, 13, 215-222.

Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, N. J.: Ablex.

Garner, R., Alexander, P., Slater, W., Hare, V. C., Smith, T., & Reis, R. (1986). Children's knowledge of structural properties of expository text. *Journal of Educational Psychology*, 78, 411-416.

Garner, R., & Kraus, C. (1981-82). Good and poor comprehender differences in knowing and regulating reading behaviors. *Educational Research Quarterly*, 6, 5-12.

Garner, R., & Taylor, N. (1982). Monitoring of understanding: An investigation of attentional assistance needs at different grade and reading proficiency levels. *Reading Psychology*, 3, 1-6.

Gleitman, L. R., Gleitman, H., & Shipley, E. F. (1972). The emergence of the child as grammarian. *Cognition*, 1, 137-164.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition*, 10, 597-602.

Hakes, D. T. (1980). *The development of metalinguistic abilities in children.* Berlin: Springer-Verlag.

Hasselhorn, M., & Körkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text-processing. *Human Learning*, 5, 75-90.

Hiebert, E. (1981). Developmental patterns and interrelationships of preschool children's print awareness. *Reading Research Quarterly*, 16, 236-260.

Isakson, R. L., & Miller, J. W. (1976). Sensitivity to syntactic and semantic cues in good and poor comprehenders. *Journal of Educational Psychology*, 68, 787-792.

Johns, J. (1974). Concepts of reading among good and poor readers. *Education*, 85, 58-60.

Johns, J., & Ellis, D. W. (1976). Reading: Children tell it like it is. *Reading World*, 15, 115-128.

Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition*, 23, 95-147.

Kurtz, B. E., & Schneider, W. (1988). The effects of age, study time, and importance of text units on strategy use and memory for texts. *European Journal of the Psychology of Education*, 3, 191-199.

Lipson, M. Y., Irwin, M., & Poth, L. (1986). The relationship between metacognitive self-reports and strategic reading behavior. In J. A. Niles & R. V. Lalik (Eds.), *Solving problems in literacy: Learners, teachers, and researchers.* Thirty-fifth yearbook of the National Reading Conference.

Lomax, R. G., & McGee, L. M. (1987). Young children's concepts about print and reading: Toward a model of word reading acquisition. *Reading Research Quarterly,* 22, 237-256.

Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1988). Reciprocal instruction improves standardized reading comprehension performance in poor grade-school comprehenders. Manuscript submitted for publication.

Markman, E. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development,* 48, 986-992.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development,* 50, 643-655.

Markman, E. M., & Gorin, L. (1981). Children's ability to adjust their standards for evaluating comprehension. *Journal of Educational Psychology,* 73, 320-325.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly,* 16, 72-103.

Mickish, V. (1974). Children's perceptions of written word boundaries. *Journal of Reading Behavior,* 6, 19-22.

Moore, D. W., & O'Driscoll, M. P. (1983). Strategies for studying texts: A comparison of Papua New Guinean and American research findings. *Journal of Educational Psychology,* 75, 460-464.

Myers, M., & Paris, S. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology,* 70, 680-690.

Owings, R. A., Petersen, G. A., Bransford, J. D., Morris, C. D., & Stein, B. S. (1980). Spontaneous monitoring and regulation of learning: A comparison of successful and less successful fifth graders. *Journal of Educational Psychology,* 72, 250-256.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension fostering and monitoring activity. *Cognition and Instruction,* 1, 117-175.

Paris, S. G., Cross, D. R. & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology,* 76, 1239-1252.

Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development,* 55, 2083-2093.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior*, 13, 5-22.

Patterson, C. J., Cosgrove, J. M., & O'Brien, R. G. (1980). Nonverbal indicants of comprehension and noncomprehension in children. *Developmental Psychology*, 16, 38-48.

Patterson, C. J., O'Brien, C., Kister, M. C., Carter, D. B., & Kotsonis, M. E. (1981). Development of comprehension monitoring as a function of context. *Developmental Psychology*, 17, 379-389.

Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1984). Memory strategy instruction is made of this: Metamemory and durable strategy use. *Educational Psychologist*, 19, 94-107.

Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1985). Children's metamemory and the teaching of memory strategies. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance, Volume 1*, (pp. 111-154). Orlando: Academic Press.

Pressley, M., Forrest-Pressley, D. L., & Elliott-Faust, D. J. (1987). What is strategy instructional enrichment and how to study it: Illustrations from research on children's prose memory and comprehension. In F. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Pressley, M., Johnson, C. J., Symons, S., McGoldrick, J. A., & Kurita, J. A. (In press). Strategies that improve children's memory and comprehension of what is read. *The Elementary School Journal*.

Raphael, T., Myers, A., Tirre, W., Fritz, M., & Freebody, P. (1981). The effects of some known sources of reading difficulty on metacomprehension and comprehension. *Journal of Reading Behavior*, 13, 324-334.

Reid, J. (1966). Learning to think about reading. *Educational Research*, 9, 55-62.

Richgels, D. J., McGee, L.M., Lomax, R. G., & Sheard, C. (1987). Awareness of four text structures: Effects on recall of expository text. *Reading Research Quareterly*, 22, 177-196.

Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly*, 21, 422-436.

Robinson, E. J., & Robinson, W. P. (1976). The young child's understanding of communication. *Developmental Psychology*, 12, 328-333.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, 94, 18-38.

Ryan, E. B., & Ledger, G. W. (1980). Grammaticality judgments, sentence repetitions, and sentence corrections of children learning to read. *International Journal of Psycholinguistics*, 6, 23-40.

Ryan, E. B., & Ledger, G. W. (1984). Learning to attend to sentence structure: Links between metalinguistics development and reading. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 149-172). New York: Springer Verlag.

Ryan, E. B., Ledger, G. W., Short, E. J., & Weed, K. A. (1982). Promoting the use of active comprehension strategies by poor readers. *Topics in Learning and Learning Disabilities*, 53-60.

Ryan, E. B., McNamara, S. R., & Kenny, M. (1977). Linguistic awareness and reading performance among beginning readers. *Journal of Reading Behavior*, 9, 399-400.

Schneider, W., & Pressley, M. (1988). *Memory development between 2 and 20*. New York: Springer Verlag.

Scholl, D. M., & Ryan, E. B. (1980). The development of metalinguistic performance during the early school years. *Language and Speech*, 23, 199-211.

Scholl, D. M., & Ryan, E. B. (1975). Child judgments of sentences varying in grammatical complexity. *Journal of Experimental Child Psychology*, 20, 274-285.

Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attribution training. *Journal of Educational Psychology*, 76, 225-235.

Sivan, E., Book, C., Meloth, M., & Putnam, J. (1987). Students' concepts of reading, their metacognitive awareness, and the explicit instruction of mental processing in reading.

Smith, C. L., & Tager-Flusberg, H. (1982). Metalinguistic awareness and language development. *Journal of Experimental Child Psychology*, 34, 449-468.

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly*, 20, 458-481.

Taylor, N. E., Blum, I. H., & Logsdon, D. M. (1986). The development of written language awareness: Environmental aspects and program characteristics. *Reading Research Quarterly*, 21, 132-149.

Taylor, B. M., & Samuels, S. J. (1983). Children's use of text structure in the recall of expository material. *American Educational Research Journal*, 20, 517-528.

Tornéus, M. (1984). Phonological awareness and reading: A chicken and egg problem? *Journal of Educational Psychology*, 76, 1346-1358.

Wagoner, S. A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, 18, 328-346.

Williams, J. P., Taylor, M. B., & Ganger, S. (1981). Text variations at the level of the individual sentence and the comprehension of simple expository paragraphs. *Journal of Educational Psychology*, 73, 851-865.

Winograd, P., & Johnston, P. (1980). *Comprehension monitoring and the error detection paradigm* (Tech. Rep. No. 153). Urbana: University of Illinois, Center for the Study of Reading.

Wong, B. Y. L., & Jones, W. (1982). Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training. *Learning Disability Quarterly*, 5, 228-240.

Yussen, S. R. (1985). The role of metacognition in theories of cognitive development. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Cognition, metacognition, and human performance, Volume* 1 (pp. 253-283). Orlando: Academic Press.

This Page Intentionally Left Blank

# KNOWLEDGE ACQUISITION FROM TEXT

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# THE ROLE OF KNOWLEDGE IN DISCOURSE COMPREHENSION: A CONSTRUCTION-INTEGRATION MODEL

Walter KINTSCH

**Institute of Cognitive Science**
University of Colorado at Boulder
Muenzinger Building
Boulder,CO 8O309-0345
(U.S.A.)

In contrast to expectation-based, predictive views of discourse comprehension, a model is developed in which the initial processing is strictly bottom-up. Word meanings are activated, propositions are formed, and inferences and elaborations are produced without regard to the discourse context. However, a network of interrelated items is created in this manner, which can be integrated into a coherent structure through a spreading activation process. Data concerning the time course of word identification in a discourse context are examined. A simulation of arithmetic word-problem understanding provides a plausible account for some well-known phenomena in this area.

Discourse comprehension, from the viewpoint of a computational theory, involves constructing a representation of a discourse upon which various computations can be performed, the outcomes of which are commonly taken as evidence for comprehension. Thus, after comprehending a text, one might reasonably expect to be able to answer questions about it, recall or summarize it, verify statements about it, paraphrase it, and so on.

To achieve these goals, current theories use representations with several mutually constraining layers. Thus, there is typically a linguistic level of representation, conceptual levels to represent both the local and global meaning and structure of a text (e.g., the micro- and macrostructure, constituting the *text base* in van Dijk & Kintsch, 1983), and a level at which the text itself has lost its individuality and its information content has become integrated into some larger structure (e.g., van Dijk & Kintsch's situation model).

---

This research was supported by Grant MH 15872 from the National Institute of Mental Health. The work on word arithmetic problems was supported by Grant BNS8741 from the National Science Foundation.

Many different processes are involved in constructing these representations. To mention just a few, there is word identification, where, say, a written word like *bank* must somehow provide access to what we know about banks, money, and overdrafts. There is a parser that turns phrases like *the old men and women* into propositions such as AND [OLD [MEN], OLD [WOMEN]]. There is an inference mechanism that concludes from the phrase *The hikers saw the bear* that they were scared. There are macro-operators that extract the gist of a passage. There are processes that generate spatial imagery from a verbal description of a place.

It is one thing for a theorist to provide some formal description (e.g., a simulation model) for how such processes can occur and for what the computational steps were that led to a particular word identification, inference, or situation model. It is quite another to control construction processes in such a way that at each point in the process exactly the right step is taken. Part of the problem has to do with the characteristic ambiguity of language: How do we make sure that we access the financial meaning of *bank*, and not the meaning of *riverbank*? Why did we parse *the old men and women* as we did--maybe the women were not old at all. Why did we infer that the hikers were scared rather than that they had their eyes open, or a myriad of other irrelevancies? Of all the many ways macro-operators could be applied, how did we get just the right sequence to reach a plausible gist without making the wrong generalizations? The number of possible alternative steps is distressingly large in constructing discourse representations, and without firm guidance, a computational model could not function properly for long. That is where knowledge comes in.

General knowledge about words, syntax, the world, spatial relations --in short, general knowledge about anything-- constrains the construction of discourse representations at all levels. Indeed, this is what makes it possible to construct these representations. There is a striking unanimity among current theories about how this is done.

Our conceptions about knowledge use in discourse comprehension are dominated by the notions of top-down effects and expectation-driven processing. Knowledge provides part of the context within which a discourse is interpreted. The context is thought of as a kind of filter through which people perceive the world. At the level of word recognition and parsing, it lets through only the appropriate meaning of an ambiguous word or phrase and suppresses the inappropriate one. Through semantic priming, the feature counter of the logogen for *bank as a financial institution* will be incremented and will reach its threshold before that of *riverbank* in the right context (Morton, 1969). Parsing a sentence is often thought of as predicting each successive constituent from those already analyzed on the basis of syntactic rules (Winograd, 1983). Scripts, frames, and schemata constrain the inferences an understander makes (as in Schank & Abelson, 1977), thereby preventing the process from being swamped in a flood of irrelevancies and redundancies. Arithmetic strategies generate just the right hypothesis in solving a word problem and preclude the wrong ones (Kintsch & Greeno, 1985). In a word, knowledge makes understanding processes smart: It keeps them on the right track and avoids exploring blind alleys. People understand correctly because they sort of know what is going to come. This program of research is well expressed by the following quotation from Schank (1978, p. 94), which served as a motto for Sharkey's (1986) model of text comprehension:

"We would claim that in natural language understanding, a simple rule is followed. Analysis proceeds in a top-down predictive manner. Understanding is expectation based. It is only when the expectations are useless or wrong that bottom-up processing begins".

Empirically, this position is questionable: Even fluent readers densely sample the words of a text, as indicated by their eye fixations (Just & Carpenter, 1980), making the bottom-up mode appear the rule rather than the exception. Computationally, it is not an easy idea to make work. It is difficult to make a system smart enough so that it will make the right decisions, yet keep it flexible enough so that it will perform well in a broad range of situations. On the one hand, one needs to make sure that exactly the right thing (word meaning, proposition, inference) will be constructed; for that purpose one needs powerful, smart rules that react sensitively to subtle cues. On the other hand, humans comprehend well in ever-changing contexts and adapt easily to new and unforeseen situations; for that purpose one needs robust and general construction rules. Scripts and frames, as they were first conceived, are simply not workable: If they are powerful enough, they are too inflexible, and if they are general enough, they fail in their constraining function. This dilemma has long been recognized (e.g., Schank, 1982; van Dijk & Kintsch, 1983), and efforts have been undertaken to make expectation-driven processes sufficiently flexible (e.g., Schank's memory organization packets, or MOPs). In this article, an alternative solution to this problem will be explored.

## Construction of Discourse Representations

The traditional approach to modeling knowledge use in comprehension has been to design powerful rules to ensure that the right elements are generated in the right context. The problem is that it is very difficult to design a production system powerful enough to yield the right results but flexible enough to work in an environment characterized by almost infinite variability. The approach taken here is to design a much weaker production system that generates a whole set of elements. These rules need to be just powerful enough so that the right element is likely to be among those generated, even though others will also be generated that are irrelevant or outright inappropriate. An integration process will then be used to strengthen the contextually appropriate elements and inhibit unrelated and inappropriate ones. Weak production can operate in many different contexts because they do not have to yield precise outputs; on the other hand, a context-sensitive integration process is then required to select among the outputs generated. The integration phase is the price the model pays for the necessary flexibility in the construction process.

The model proposed here has been termed a *construction-integration* model to emphasize its most salient feature. It combines a construction process in which a text base is constructed from the linguistic input as well as from the comprehender's knowledge base, with an integration phase, in which this text base is integrated into a coherent whole. The knowledge base is conceptualized as an associative network. The construction process is modeled as a production system. Indeed, it is a generalization of the production system used in earlier work, such as the simulation-of-comprehension processes developed by Fletcher (1985) and Dellarosa (1986) after the model of Kintsch and Greeno (1985). The main difference is that instead of precise inference rules, sloppy ones are used, resulting in an incoherent, potentially contradictory output. However, this output structure is itself in the form of an associative net, which can be shaped into a coherent text base via relaxation procedures in the connectionist manner (e.g., Rumelhart & McClelland, 1986). Thus, the model represents a symbiosis of production systems and connectionist approaches.[1]

Certain limitations of the present article are worth noting at this point, for it does not offer a solution to all the problems in discourse understanding. Thus, it is not primarily concerned with the specific strategies (or rules) for the construction of text propositions or inferencing. Instead, it relies in this respect on what is available in the literature as well as on whatever future researchers will be able to come up with. The only point it makes is that whatever these strategies or rules are, they will be easier to formulate within the present framework, which allows them to be both weaker and more general. Thus, one need not worry about constructing just the right inference, but can be content with a much sloppier rule. Sometimes, of course, even the latter type of rule may be hard to come by, whereas in other cases (e.g., in the word problems discussed later) promiscuous hypothesis generation is straightforward (while selecting just the right one can be tricky).

## Knowledge Representation

The process of constructing a discourse representation relies heavily on knowledge. To understand how it operates, one must first have an idea of how the to-be-used knowledge is organized. Typically, theorists have tried to create knowledge structures to support smart processes: semantic nets, frames, scripts, and schemata. As has been argued elsewhere (Kintsch, in press), such fixed structures are too inflexible and cannot adapt readily enough to the demands imposed by the ever-changing context of the environment. Instead, a minimally organized knowledge system is assumed here in which structure is not prestored, but generated in the context of the task for which it is needed. An

---

[1]Conceivably, a purer connectionist model might be constructed. In the present model, an associative knowledge net is used to build a textbase net, which is then integrated. McClelland (1985) has put forth the idea of a connection information distributor, which is a subnetwork in which the units are not dedicated and connections are not hardwired. Instead, this subnetwork is programmable by inputs from the central network where the knowledge that controls processing in the subnetwork is stored. One could say that the production rules in the present model have the function of programming such a subnetwork.

associative net with positive as well as negative interconnections serves this purpose.

Knowledge is represented as an associative net, the nodes of which are concepts or propositions.[2] The notes in this net are interconnected. Connections among nodes have a strength value, which may be positive, zero, or negative, ranging from 1 to -1. Nodes consist of a head plus a number of slots for arguments. Thus, the nodes of the knowledge net are formally equivalent to the propositions used to represent texts (e.g., Kintsch, 1974).[3] The slot specifies the nature of the relation between the head and the argument. Slots may represent attributes, parts, cases of verbs, or arguments of functions. They need not be named, but may be named if the relation is a common one (such as the cases of verb frames). The arguments of a proposition are concepts or other propositions. The number of arguments in a proposition may vary from one to some small number. Examples of common types of nodes in the knowledge net are (a) MARY, (b) CAKE, (c) SWEET [CAKE], (d) BAKE [agent: MARY, object: CAKE], (e) CONSEQUENCE [condition: NOT [WATCH [agent: MARY, object: CAKE], effect: BURN [object: CAKE]]. Examples A and B are lexical nodes that have associated with them perceptual procedures that identify certain patterns in the environment --either the objects themselves or the written or spoken words, such as MARY and CAKE, respectively. In the following I shall not deal with these perceptual procedures explicitly. The semantic and associative relations into which MARY and CAKE enter, which constitute a part of the general knowledge net, are the focus of interest here. MARY and CAKE appear as arguments in Examples C through E in various roles (the agent and object slots, etc.).

There are two ways of looking at the list of propositions in Examples A through E. On the one hand, it could be considered simply as a portion of a general knowledge network, whereas on the other hand, it could be considered the propositional base of a (brief) discourse, in which a particular Mary bakes and burns a particular cake.[4] Thus, the elements of which knowledge nets and text bases are formed by selecting, modifying, and rearranging propositional elements from the knowledge net. However, text bases are not part of the knowledge net, but separate structures with their own properties.

Concepts are not defined in a knowledge net, but their meaning can be constructed from their position in the net. The immediate associates and

---

[2]Formally, concepts and propositions can be treated alike (e.g., Anderson, 1980).
[3]This use of the term *proposition* differs from the standard one in logic. Furthermore, not all authors who use comparable semantic units in their analyses use the same term. For instance, Dijk (1980) talked about *predicates* and *terms* combining to form *predications.* Wilensky (1986) used *relation* and *aspectuals.* In spite of this terminological disarray and the possibility of confusion with the meaning of proposition in logic, *proposition* appears to be the most widely accepted term and will be retained here.
[4]The extreme informality of this notation is chosen for each of exposition. Frequently, of course, a more precise formalism is needed. It is fairly straightforward to elaborate the present informal notation whenever that is the case. For example, in the computer simulation of word problem solving by Dellarosa (1986), the LOOPS language provides a ready-made type-token distinction. There seems to be no reason, however, to burden a general discussion like the present one with a cumbersome, formal notation when it is not needed.

semantic neighbors of a node constitute its core meaning, however, can be obtained only by exploring its relations to all the other nodes in the net. Meaning must be created. As a first step one could add all propositions in the net directly related to a node to obtain what Mudersbach (1982) termed the first level of meaning; then all propositions directly related to the propositions at the first level can be added to form a second level, and so on, until the whole knowledge net is involved. Note, however, that such a construction is a theoretical exercise without direct psychological correspondence. It is not possible to deal with the whole, huge knowledge net at once. Instead, at any moment only a tiny fraction of the net can be activated, and only those propositions of the net that are actually activated can affect the meaning of a given concept. Thus, the meaning of concept is always situation specific and context dependent. It is necessarily incomplete and unstable: Additional nodes could always be added to the activated subnet constituting the momentary meaning of a concept, but at the cost of losing some of the already activated nodes.[5]

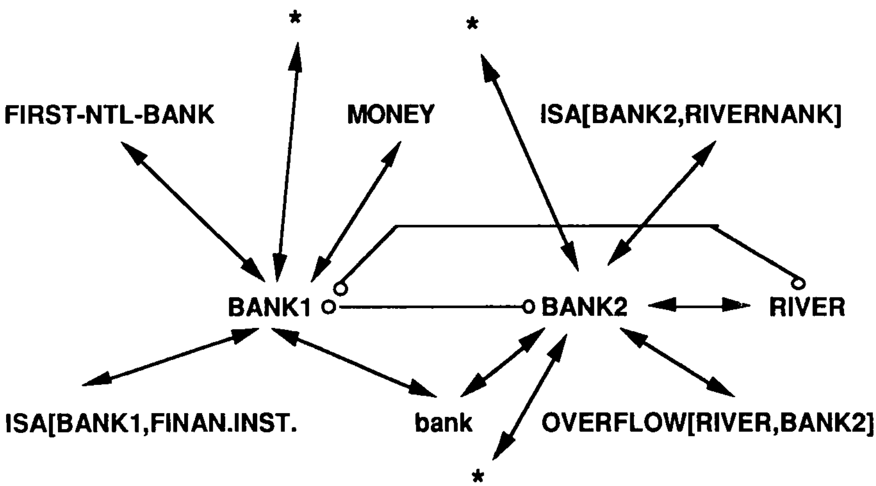



**Figure 1. A fragment of the associative net for BANK. (Positive connections are indicated by arrows, negative ones by circles; asteriks indicate further, unnamed nodes.)**

[5]As with *proposition*, this is a nonstandard use of the term's meaning. Meaning is used here as shorthand for the momentary, subject- and situation-specific activated semantic and experiential context of a concept. Clearly, this is not what many people have in mind when they speak about the meaning of a word--though it is a conception of meaning quite appropriate for a psychological processing model.

The notion of an associative net is not unfamiliar, but it is usually thought of as relating concepts only, not propositional nodes. Two extremely simple examples will illustrate the nature of such an associative net. First, consider the representation of the homonym BANK in an associative net. Positive connections are indicated by arrows, negative ones by circles. Asterisks indicate further, unspecified nodes. Of course, each of the concepts and propositions shown in Figure 1 participate in the general knowledge net beyond the single connection shown here. As a second example, consider the proposition BAKE [agent: PERSON, object: CAKE] (see Figure 2). Once again, only a fragment of the complete network is shown, just to illustrate certain types of connections.

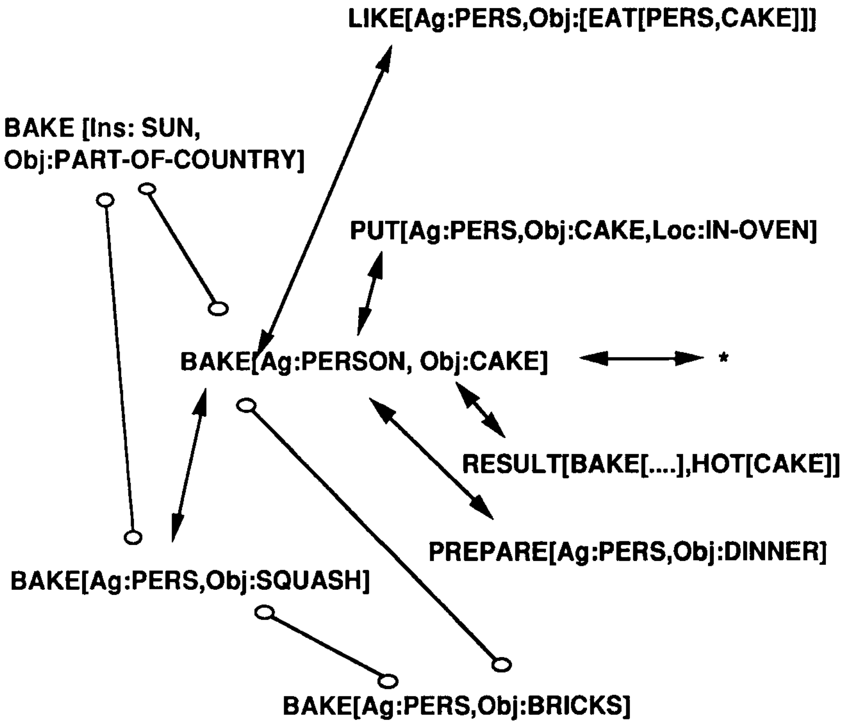



**Figure 2. A fragment of the associative net for BAKE.**

Representing knowledge in a propositional network has several advantages. Primarily, it provides a common format for the knowledge base and for the mental representation of discourse. Furthermore, we have by now considerable experience working with propositional structures, whereas other forms of representation are less well understood (e.g., the spatial-imagery and linear structures of Anderson, 1983; the mental models of Johnson-Laird, 1983; or whatever the appropriate representation in the affective system might be, as in Zajonc, 1980). However, the decision to use a propositional representation does not imply that all other forms of knowledge are to be considered unimportant or nonexistent. It would be desirable to expand the model to include nonpropositional representations, but one would first have to learn how to operate with such forms of representation.

## Construction Processes

The steps in constructing a text base according to the construction-integration model involve: (a) forming the concepts and propositions directly corresponding to the linguistic input; (b) elaborating each of these elements by selecting a small number of its most closely associated neighbors from the general knowledge net; (c) inferring certain additional propositions; and (d) assigning connection strengths to all pairs of elements that have been created.

The result is an initial, enriched, but incoherent and possibly contradictory text base, which is then subjected to an integration process to form a coherent structure.

In Step A of this process, a propositional representation of the text is constructed from a parsed linguistic input, such as the words of a text with suitable syntactic annotations, and from a knowledge system as envisioned earlier. Note that the parser itself is not a part of the present model. The basic process of proposition building has been described in van Dijk and Kintsch (1983, chapter 4) and Kintsch (1985). I will illustrate it here with some simple examples. Consider the sentence *Mary bakes a cake.* The parser output needed is *Mary* (agent of BAKE) *bakes* (predicate) *a cake* (object of BAKE). *Mary, bake,* and *cake* activate their corresponding lexical nodes, and MARY and CAKE are assigned the roles of agent and object in the BAKE proposition. As was suggested in Figure 2, BAKE requires a PERSON as agent, hence a test is made whether MARY is a person. This may either involve a search through the knowledge net for the proposition ISA [MARY, PERSON] or, should that search prove unsuccessful, an attempt to infer this proposition (e.g., the net may contain only propositions to the effect that MARY is a name and that persons have names; exactly how such problem-solving activity occurs within an associative net will not be considered here).

The present model, however, differs in a significant way from my earlier conceptions: It does not require that the right, and only the right, proposition always be formed. Instead, the construction rules for building propositions can be weakened, allowing for the formation of incomplete or "wrong" propositions. Proposition building is on-line, and frequently, all the relevant information for building just the right one is not available on-line, leading to false starts or incompleted attempts. In the aforementioned example, this has no interesting consequences; for example, if in response to the phrase *Mary bakes* . . . the proposition BAKE [MARY, $] --the dollar sign indicates an unfilled slot-- is formed, it will simply be replaced by the complete proposition when the rest of the sentence is processed. However, consider an example discussed by Frazier and Rayner (1982); *The linguists knew the solution of the problem would not be easy.* Here, the on-line construction of propositions is not so simple. First, the proposition KNOW [LINGUISTS, $] is formed. Then, by the strategy of minimal attachment, the subsequent noun phrase is interpreted as the object of KNOW, yielding KNOW [LINGUISTS, SOLUTION]. The final verb phrase, however, requires a subject, so [NOT [EASY] SOLUTION]] is constructed. As Frazier and Rayner pointed out, this does not involve a reinterpretation of the sentence. Subjects do not go back, noting in some way that *solution of the problem* had been attached to the wrong proposition, and repair this error. Instead, the incorrectly formed KNOW proposition somehow just disappears; the description of the integration process that follows shows how.

A third example of proposition building, involving pronoun identification, will be discussed here. There exists good psychological evidence that pronouns may activate more than one possible referent (e.g., Frederiksen, 1981). Thus, in *The lawyer discussed the case with the judge. He said "I shall send the defendant to prison."* the following propositions would be formed: DISCUSS [LAWYER, JUDGE, CASE]: SAY [LAWYER, [SEND [LAWYER, DEFENDANT, PRISON]]]; and SAY [JUDGE, [SEND [JUDGE, DEFENDANT, PRISON]]]. Eventually, of course, the right interpretation comes to dominate the wrong one, as will be shown shortly.

In Step B of the construction process, each concept or proposition that has been formed in Step A serves as a cue for the retrieval of associated nodes in the knowledge net. The retrieval process itself is modeled after well-known theories that have been developed and tested in the memory literature (Raaijmakers & Shiffrin, 1981). Suppose that node *i* in the knowledge net is positively associated with *n* other nodes in the net Let *s(ij)* be the associative strength between nodes *i* and *j*. Then the probability that the retrieval cue *j* will retrieve node *j* is

$$P(j\backslash i) = \frac{S(i,j)}{{}_{h=1}^{n}\; s(i,h)}$$

Note that each concept or proposition in the text base serves as an independent retrieval cue, hence the particularly simple form of the retrieval process. (An intersection search would be required if the items in the text base acted as a compound cue.) On each retrieval attempt, an item among the associates of *i* is selected according to Equation 1. A sampling-with-replacement process is assumed so that dominant associates may be retrieved more than once. The number of retrieval attempts with item *i* as the cue is assumed to be fixed and is a parameter of the model, *k*, was chosen to be 2 or 3, mostly to reduce the complexity of these examples. However, one may speculate that the most realistic value of *k* would not be much higher, perhaps between 5 and 7.

Consider some simple examples.

1. Suppose the word *bank* is presented as part of a text. It will activate the lexical nodes BANK1 (financial institution) as well as BANK2 (riverbank), plus some of their associates; for example, the construction process might pick from Figure 1: BANK1, MONEY, FIRST-NATIONAL-BANK, BANK2, RIVER, OVERFLOW [RIVER, BANK2].

2. Suppose the sentence *Lucy persuaded Mary to bake a cake* is presented as part of a text. The parser should provide a phrase structure tree as output, from which the proposition PERSUADE [LUCY, MARY [MARY, CAKE]] is constructed. Each text proposition activates propositions closely related to it in the general knowledge net, regardless of the discourse context. For instance, in the case of BAKE [MARY, CAKE] we might thus obtain LIKE [MARY, EAT [MARY, CAKE]]. PUT [MARY, CAKE, IN-OVEN], RESULT [BAKE [MARY, CAKE], HOT [CAKE]]. PREPARE [MARY, DINNER]. These propositions are all closely associated with baking a cake (Figure 2). Note, however, that elaborating the text base in this way is not just a question of retrieving associated propositions from the knowledge net. The arguments of these retrieved propositions must be treated as variables that are to be bound to the values specified by the retrieval cue. Thus, because MARY is the agent of the text proposition, MARY is made the agent in the knowledge propositions it brings into the text representation, instead of PERSON in Figure 2. Similarly, although the informality of the present notion hides this, CAKE now is the particular one MARY bakes, not the generic one in Figure 2. These knowledge propositions function as potential inferences. Out of context there is no way of determining which of them are relevant: Maybe Mary really likes to eat cake, but perhaps she is in the process of cooking dinner, in which case PREPARE [MARY, DINNER] might become a macroproposition (what van Dijk, 1980, calls a *construction*). But it is also possible that next she will burn her fingers when she takes the cake out of the oven, making HOT, which plays no role at all in the other contexts, the relevant inference. At this point, the construction process lacks guidance and intelligence; it simply produces potential inferences in the hope that some of them might turn out to be useful.

3. In the third example, if the proposition SEND [LAWYER, DEFENDANT, PRISON] has been formed, the knowledge net contributes nothing, because one presumably does not know anything about lawyers sending defendants to prison. (Of course, LAWYER, DEFENDANT, and PRISON would each be associatively elaborated separately.). If, however, JUDGE rather than LAWYER were the agent of SEND, the elaboration process would contribute the information that this implies that the judge is sentencing the defendant and so forth.

Step C in the construction process, the generation of additional inferences, is necessary because not all inferences that are required for comprehension will, in general, be obtained by the random elaboration mechanism described earlier. In some cases more focused problem-solving activity is necessary to generate the desired inferences. Exactly how this is to be done is, however, beyond the scope of this article. I merely wish to point out here that in addition to the undirected elaboration which results from Step B of the construction process, there is still a need for controlled, specific inferences. Two types of such inferences are of particular importance in comprehension. Bridging inferences (Haviland & Clark, 1974; Kintsch, 1974) are necessary whenever the text base being constructed is incoherent (i.e., whenever either the original text base itself or the elaborated text base remains incoherent by the criteria discussed in van Dijk and Kintsch, 1983, chapter 5). Second, macropropositions have to be inferred (as discussed in general terms in chapter 6 of van Dijk & Kintsch, 1983, and operationalized as a production system by Turner, McCutchen, & Kintsch, 1986). Macropropositions are also elaborated associatively, as described in Step B for micropropositions.

What has been constructed so far is a set of propositions containing the (micro)propositions directly derived from the text, a randomly selected set of associates for each of these, the macropropositions generated from the text, and their associates. The final Step D of the construction process involves the specification of the interconnections between all of these elements. There are two ways in which elements are interconnected. (a) The propositions directly derived from the text (hence referred to as "text propositions") are positively interconnected with strength values proportional to their proximity in the text base. Specific realizations of this principle are described in the discussion of Figure 4. (b) If propositions $i$ and $j$ are connected in the general knowledge net with the strength value $s(i,j)$, -- < $s(i,j) < 1$, and if $i$ and $j$ become members of a text base, the strength of their connection in the text base is $s(i,j)$. In other words, propositions in the text base inherit their interconnections from the general knowledge net. Strength values are additive, up to a maximum of 1, in those cases in which an inherited strength value combines with a text-base-determined connection.

Consider, for instance, the portion of a network that is generated when the word *bank* activates both BANK1 and BANK2, as well as the associations MONEY and RIVER. A possible pattern of connections is shown in Figure 3, where for simplicity, connection strengths have been limited to $\pm.5$ or 1. Alternatively, the graph shown in Figure 3 can be expressed in matrix form as shown in Table 1. BANK1 is associated with MONEY, BANK2 with RIVER, but inhibitory connections exist between MONEY and BANK2 and between RIVER and BANK1.

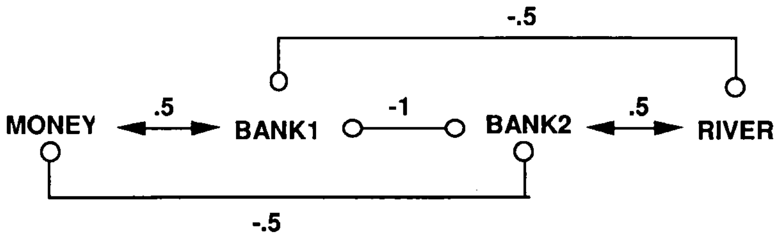



**Figure 3. Connections between BANK1 and BANK2 and their associates.**

**Table 1. Connectivity matrix for the graph shown in Figure 3.**

| Proposition | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. MONEY | ---- | 0.5 | -0.5 | 0.0 |
| 2. BANK1 | 0.5 | ---- | -1.0 | -0.5 |
| 3. BANK2 | -0.5 | -1.0 | ---- | 0.5 |
| 4. RIVER | 0.0 | -0.5 | 0.5 | ---- |

An example of text propositions that are interconnected via their positions in the text base is shown in Figure 4. LUCY is connected most strongly to WEED [LUCY, GARDEN], and least strongly to VEGETABLE [GARDEN]. Although there are many possible ways to assign numerical connection strengths to express this pattern of connectivity, the one chosen here results in the matrix shown in Table 2.

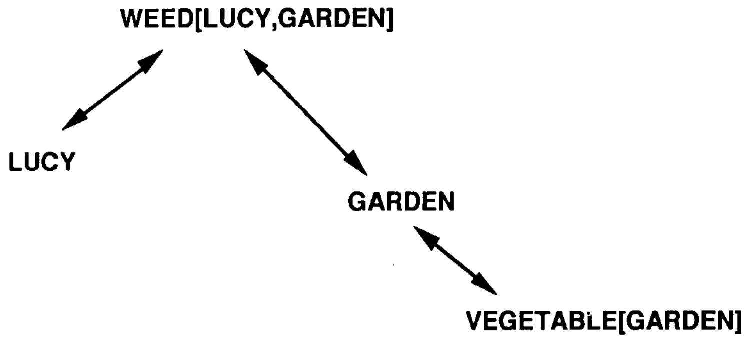



**Figure 4. The text base for *Lucy weeded the vegetable garden.***

**Table 2. Connectivity matrix for the graph shown in Figure 4.**

| Proposition | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. LUCY | ---- | 0.9 | 0.7 | 0.4 |
| 2. WEED | 0.9 | ---- | 0.9 | 0.7 |
| 3. GARDEN | 0.7 | 0.9 | ---- | 0.9 |
| 4. VEGETABLE | 0.4 | 0.7 | 0.9 | ---- |

Inferences inherit positive and negative interconnections from the general knowledge net, as seen in Figure 5. The result of the construction process is, therefore, a network expressable as a connectivity matrix, consisting of all the lexical nodes accessed, all the propositions that have been formed, plus all the inferences and elaborations that were made at both the local and global level and their interconnections.

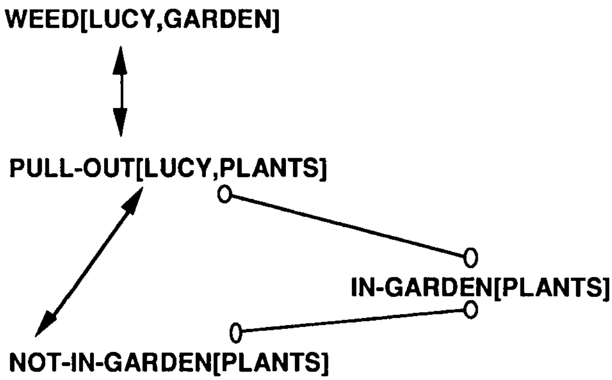



**Figure 5. Inferences generated from WEED[LUCY,GARDEN] and their interconnections.**

## Integration

The network that has been constructed so far is not yet a suitable text representation. It was carelessly constructed and is therefore incoherent and inconsistent. At all levels of the representation, components associated with the text elements were included without regard to the discourse context, and many of them are inappropriate. An integration process in the connectionist manner can be used to exclude these unwanted elements from the text representation (e.g., see Rumelhart & McClelland, 1986, and Waltz & Pollack, 1985, for discourse).

Text comprehension is assumed to be organized in cycles, roughly corresponding to short sentences or phrases (for further detail, see Kintsch & van Dijk, 1978; Miller & Kintsch, 1980). In each cycle a new net is constructed, including whatever is carried over in the short-term buffer from the previous cycle.[6] Once the net is constructed, the integration process takes over: Activation is spread around until the system stabilizes. More specific, an activation vector representing the initial activation values of all nodes in the net is postmultiplied repeatedly with the connectivity matrix. After each

[6]That integration occurs at the end of each processing cycle is proposed here merely as a simplifying assumption. Although there is clearly something going on at the end of sentences (e.g., Aaronson & Scarborough, 1977), integration does not need to wait for a sentence boundary (see the evidence for the "immediacy assumption"; Just & Carpenter, 1980; Sanford & Garrod, 1981). It would be quite possible to apply the relaxation procedure outlined here repeatedly in each cycle, as propositions are being constructed. This would allow for the disambiguation of word senses before the end of a cycle. Because inferences and macropropositions are usually not available before the end of a processing cycle, end-of-cycle integration plays an especially important role.

multiplication the activation values are renormalized: Negative values are set to zero, and each of the positive activation values is divided by the sum of all activation values, so that the total activation on each cycles remains at a value of one (e.g., Rumelhart & McClelland, 1986). Usually, the system finds a stable state fairly rapidly; if the integration process fails, however, new constructions are added to the net, and integration is attempted again. Thus, there is a basic, automatic construction-plus-integration process that normally is sufficient for comprehension. This process is more like perception than problem solving, but when it fails, rather extensive problem-solving activity might be required to bring it back on track. These processes will not be considered further here.

The result of the integration process is a new activation vector, indicating high activation values for some of the nodes in the net and low or zero values for many others. The highly activated nodes constitute the discourse representation formed on each processing cycle. In principle, it includes information at many levels: lexical nodes, text propositions, knowledge-based elaborations (i.e., various types of inferences), as well as macropropositions.

A few simple examples will illustrate what is at issue here. Consider *Lucy persuaded Mary to bake a cake*, which was discussed earlier. The PERSUADE proposition will pull in related knowledge items, just as was shown for BAKE. However, out of context the integration process will not yield any striking results. In the context of *Lucy made tomato soup and sauteed some porkchops with herbs. She set the table and persuaded Mary to bake a cake*, the integration process has very different results: PREPARE [LUCY, DINNER] emerges as the dominant proposition (macroproposition) because most of the other propositions in the text base contribute to its activation value. That the cake was hot, or that she put it into the oven, disappears from the representation with activation values around zero.

Next, consider the example just discussed, where a perfectly good propositional strategy led to a wrong result. For *The linguists knew the solution of the problem would not be easy*, the text base that was constructed is shown in Figure 6. It corresponds to the connectivity matrix exhibited in Table 3 if connection strengths are assigned as in Table 2. (KNOW [SOLUTION] and NOT [EASY] are connected positively via KNOW [$] but negatively via EASY, which adds up to 0.). The activation vector (.25, .25, .25, .25) corresponding to the assumption that all text propositions are equally activated initially is repeatedly multiplied with this matrix, renormalizing the obtained activation values after each multiplication as described earlier. To decide when the activation vector has stabilized, the following criterion was established: A stable state is reached when the average change in the activation values after a multiplication is less than .001. Although this is an arbitrary criterion, even large changes (by one order of magnitude in either direction) make only minor differences in the final activation values obtained in this and many other cases. In the present case, this criterion is reached after 10 operations, yielding the final activation vector (.325, .000, .325, .350) --that is, the wrong KNOW [LINGUISTS, SOLUTION], which does not fit into the text base, has been deactivated. The integration process similarly resolves the problem of multiple pronoun referents. For *The lawyer* *discussed the case with the judge. He said "I shall send the defendant to prison."* propositions were constructed for both lawyer and judge as referents of *he*. However, the process of associative elaboration generated some additional information for SEND [JUDGE, DEFENDANT, PRISON], but not for SEND [LAWYER, DEFENDANT, PRISON]. The resulting text base is shown in Figure 7. To obtain the corresponding connectivity matrix (see Table 4), connection strengths among text base propositions were assigned as in Table 2, and among associates as in Table 3 (other assignments result in different numerical values for the final activation vector, but its pattern remains the same as long as the essential features of the matrix are preserved --for example, which connections are positive, negative, and zero). Assume an initial activation vector of (.25, .25, .25, .25, .25, 0, 0), reflecting the fact that only the text propositions themselves are activated initially. After 19 multiplications with the connectivity matrix, the two propositions in which *he* had been identified as the *lawyer* have activation values of 0, whereas the corresponding *judge* propositions have activation values of .261 and .283, respectively. Just a little knowledge was enough to choose the correct referent.

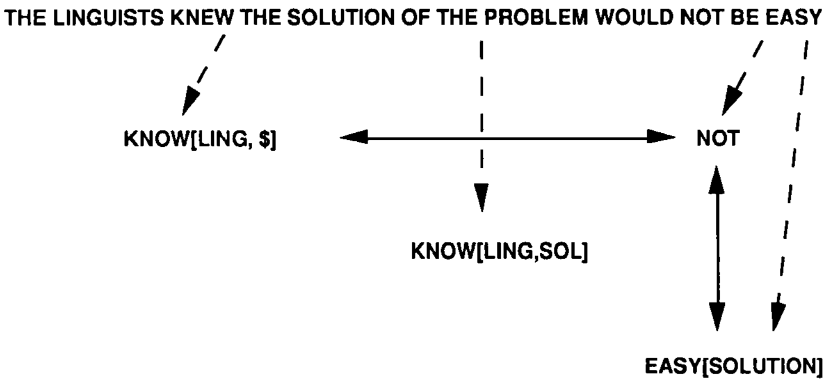



**Figure 6. The strategic construction of a text base: SOLUTION-OF-THE-PROBLEM is first assigned to KNOW, then to EASY. (The dollar sign is a placeholder).**

Table 3. Connectivity matrix for the graph shown in Figure 6.

| Proposition | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. KNOW[$] | ---- | 0.9 | 0.7 | 0.9 |
| 2. KNOW[SOL] | 0.9 | ---- | 1.0 | 0.0 |
| 3. EASY | 0.7 | -1.0 | ---- | 0.9 |
| 4. NOT | 0.9 | 0.0 | 0.91 | ---- |

After this general description of the construction-plus-activation model, two specific applications will be discussed in more detail: how words are identified in a discourse context, and how a propositional text base and situation model are constructed when comprehension depends heavily on activating a rich knowledge set. For that purpose, arithmetic word problems were chosen as the example, because the knowledge that needs to be activated is particularly well defined in that domain, and unambiguous criteria of understanding exist--a solution is either right or wrong. The purpose of these examples is twofold: to show how the general framework proposed can be elaborated into specific models in these experimental situations, and to compare the performance of these models with empirical observations and experimental results as a first test of the psychological adequacy of these models.

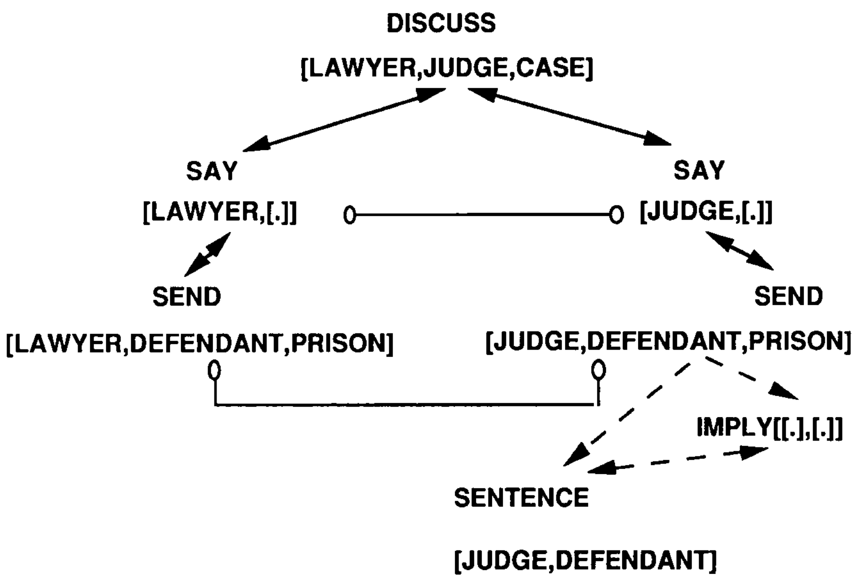



**Figure 7. The strategic construction of a text base: The pronoun *he* is identified with two potential, mutually exclusive referents. (Instead of writing out whole propositions, the abbreviation [.] is used for the arguments of a proposition when they can be readily inferred).**

**Table 4. Connectivity matrix for the graph shown in Figure 7.**

| Proposition | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. DISC | ---- | 0.9 | 0.9 | 0.7 | .0.7 | 0.0 | 0.0 |
| 2. SAY[LAWYER] | 0.9 | ---- | -1.0 | 0.9 | 0.0 | 0.0 | 0.0 |
| 3. SAY[JUDGE] | 0.9 | -1.0 | ---- | 0.0 | 0.9 | 0.0 | 0.0 |
| 4. SEND[LAWYER] | 0.7 | 0.9 | 0.0 | ---- | -1.0 | 0.0 | 0.0 |
| 5. SEND[JUDGE] | 0.7 | 0.0 | 0.9 | -1.0 | ---- | 0.5 | 0.5 |
| 6. IMPLY | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | ---- | 0.5 |
| 7. SENT | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | ---- |

**Word Identification in Discourse**

The first problem to be considered in detail is how knowledge is used in understanding the meaning of words in a discourse. The previously sketched model implies that word meanings have to be created anew in each context, that this is initially strictly a bottom-up process with context having its effects in the integration phase, and that this construction-plus-integration process takes time, with different factors influencing successive phases of the process.

Context effects in word recognition are ubiquitous in the experimental literature, and the explanation of these context effects has been a primary goal of theories of word recognition. Typically, it is taken for granted in these theories that because context influences word recognition, contextual factors interact with the perceptual processes. Context effects are said to be top-down and expectation driven and are said to facilitate (or sometimes interfere with) the perceptual analysis. Similar ideas were once current in related fields, such as the "New Look" in perception (Bruner & Postman, 1949) and the filter theory of attention (Broadbent, 1958). People perceive what they expect or want, attention filters out the irrelevant. Some words are recognized because the context favors them; others are not because the context inhibits them. How these top-down effects of context are realized differs in detail among theories, but all the most influential current theories postulate interactive processes in which contextually expected words are favored. In the logogen model (Morton, 1969, 1979), context primes semantic features that enter into the feature counter of a logogen and therefore strengthen that logogen. In Forster's search model (Forster, 1976), perceptual analysis defines a candidate set that is then searched by semantic relations or by word frequency. In Becker's verification model (Becker, 1976), both a set of sensory candidates and a set of semantic candidates are created, with the latter being verified first. In the cohort model of Marslen-Wilson and Welsh (1978), context is used to exclude members of the cohort from the very beginning. Norris (1986) has recently reviewed these models and pointed out that they all involve some sort of priming mechanism through which context effects are mediated.

The model of how knowledge is used in discourse suggests a fundamentally different approach. Following earlier work by Kintsch and Mross (1985) and Norris (1986), the present model is neither interactive, nor does it involve priming. As these authors have argued, word identification is not simply a matter of lexical access. Rather, it is a complex process that responds to different influences at different stages. These stages, however, are merely convenient verbal labels. In fact, processing is continuous, and there is significant temporal overlap between the different subprocesses defining these stages. In the first stage (which was termed *sense activation* by Kintsch & Mross (1985), the number of word candidates consistent with the perceptual input is progressively reduced through perceptual feature analysis. As in Forster or Becker, a set of sensory candidates is created through perceptual analysis, but its size decreases as the analysis progresses. This process rapidly reduces the number of word candidates to some manageable number, but not necessarily to one. At this point (probably at about 50 ms, see Fischler & Goodman (1978), the semantic context comes into play. Some small number of lexical nodes has now been selected, each one activating a few of its strongest semantic or associative neighbors in the knowledge network. If there is a node whose associates fit into whatever context is present, it will be taken as the meaning of the to-be-identified word. What fits is determined by the integration process sketched earlier. This is the *sense-selection* stage of Kintsch and Mross.

Note that if the perceptual analysis had been allowed to continue for a sufficient period of time, for most words it would have yielded a result eventually by itself, and probably the same one. It is just that the association check helped to shortcut this process. With homonyms, however, the association check plays a crucial role: Perceptual analysis alone cannot decide which meaning of *bank* to select in any given context.

Sense selection by means of an association check is the very first of a possibly very long series of contextual plausibility checks (Norris's term). It comes first because the associative/semantic context of a lexical node can be computed rapidly. As more information about the context becomes available, the sentence and discourse meaning begin to emerge, and more and deeper plausibility checks can be performed as long as there still is time. This is the *sense-elaboration* phase, in which the meaning of a word is contextually explored and elaborated. However, once a response has been made in a recognition experiment, or once the process moves on in a discourse, elaboration is terminated. Thus, word meanings are usually identified long before complex inferences are made in comprehending a discourse.

At this point, a "meaning" has been constructed for the word in this particular context. It consists of the lexical node that has been activated (the contextually inappropriate nodes that had been activated have by now been deactivated through the various context checks), the associative and semantic neighbors of that node, the sentence and discourse context in which the word participated, and some inferences and elaborations that were produced in the course of the various plausibility checks that explored the role of that word in the given context.

What do we need to make such a model of word identification work? We shall disregard the perceptual analysis and take for granted that a certain number of appropriate lexical nodes has been activated (e.g., multiple semantic nodes for a homonym). We then need to compute the sentences and phrases in which the word in question participates, or more accurately, the propositions in which the corresponding concept token (for which the lexical node serves as the type) plays a role. Finally, we need to construct inferences and elaborations when necessary.

A model of word recognition that thus far is identical with the one favored here has recently been developed by Norris (1986). Norris called it the "checking model" and compares and contrasts it with the other extant models of word recognition in the literature. In Norris's model, the plausibility of word candidates in any given context is evaluated. The recognition criterion for contextually plausible words is lowered and that for implausible words is increased. By manipulating criterion bias in this way, Norris accounted for a wide range of observations from threshold and other types of recognition experiments.

Instead of equating plausibility with criterion bias, a different mechanism --integration-- is used here. This mechanism has the great advantage of being applicable not only at the word-recognition level (which is what Norris was concerned with), but it is equally suited to modeling knowledge integration at higher levels.

When a word is perceived, one or more lexical nodes are accessed, and some of their neighboring nodes that are closely related associatively or semantically are also activated. Similarly, when a proposition is constructed, a number of associatively and semantically related propositions are also constructed. Both related concepts and related propositions serve to determine the plausibility of the core words and propositions. A richly interconnected structure is thus formed, through which activation can spread, so that positively interconnected items strengthen each other, while unrelated items drop out and inconsistent items become inhibited. Or, said differently, implausible items will be suppressed, whereas plausible ones support each other --at the level of word recognition as well as of textual integration.

## Time Course of Activation of Words in a Discourse Context

The model of word recognition just outlined is consistent with a great deal of experimental data. Norris (1986) has reviewed the word recognition literature in great detail and shown that his checking model accounts for the rich empirical findings in that area better than any of its competitors. The construction-integration model is closely related to Norris's model. On the one hand, it is more specific in that it proposes computational procedures by means of which Norris's "plausibility check" could actually be achieved, whereas on the other hand it replaces Norris's shift in criterion bias with the computationally more feasible integration mechanism. It appears likely that the present model can handle all the data the checking model accounts for, in just the same way and for just the same reasons as the checking model. There is, however, another part of the literature on word recognition that is not discussed in Norris (1986): the work on word identification in discourse. The empirical findings in this area are also in good agreement with the construction-integration model.

In a lexical decision task, the subject sees a string of letters and must decide as quickly as possible whether it forms an English word. If a target word is preceded by a closely related word, the response to the target word is speeded up (on the order of 20 to 40 ms) in comparison with unrelated control words. This priming effect has been well documented for some time and is obtained in list contexts (e.g., Meyer & Schvaneveldt, 1971) as well as in discourse contexts (e.g., Swinney, 1979). However, the discourse context is actually irrelevant to the priming effect. What matters is merely the associative relation between the prime word and the target word. As has been shown repeatedly (Kintsch & Mross, 1985; Swinney, 1979; Till, Mross, & Kintsch, in press; also equivalent results obtained with a naming task by Seidenberg, Tanenhaus, Leiman, & Bienkowsky, 1982), homonyms will prime strong associates of both their meanings, irrespective of the discourse context and in spite of the fact that the context-inappropriate meaning of the homonym never enters consciousness. Furthermore, context appropriate inferences that are not associatively related to a priming word are not responded to any faster than unrelated control words. However, all of this depends on the amount of time allowed for the processing of the priming word. If the target word closely follows the priming word, so that the processing of the prime is still in its initial stages, the results are as already described. However, if there is enough time for more complete processing of the priming word in its discourse context, quite different results are observed. In this case, context-appropriate associates are still primed, but inappropriate associates no longer are, whereas context-appropriate inferences now become strongly primed. This time course of knowledge activation can be described in more detail by some illustrative experimental results.

In the study by Till et al. (in press), subjects read sentences like *The townspeople were amazed to find that all the buildings had collapsed except the mint.* After the priming word *mint* they were given a lexical decision task, with the target word being either *money, candy,* or *earthquake.* That is, the target was a context-appropriate associate of the prime (*money*), a context-inappropriate associate (*candy*), or a topical inference word (*earthquake*), respectively. In addition, the interval between the presentation of the prime and the target word (stimulus-onset asynchrony, or SOA) was varied from 200 ms to 1500 ms. In the first case, the prime could only be incompletely processed; with an SOA of 500 ms, a somewhat deeper processing of the prime was possible before a response had to be given to the target word; and with 1,000 ms, extensive processing of both the prime word and its discourse context was possible. The data are shown in Figure 8. To keep this presentation simple, Figure 8 shows the average priming effects observed in the three experiments of Till et al. for SOAs of 200, 300, 400, 500, 1,000, and 1,500 ms. The value shown for associates at 200 ms, for instance is the difference between the mean reaction time for context-inappropriate and context-appropriate associates at that prime-target asynchrony. It is the average of two such values obtained in two different experiments--showing the data separately for each experiment merely complicates the picture without changing its essential features. The value for inferences, similarly, is based on the difference between topic words and unrelated control words. The purpose of Figure 8 is, therefore, merely to give an impression of the over-all shape of the results of this study (for more detailed analyses, the original article must be consulted).

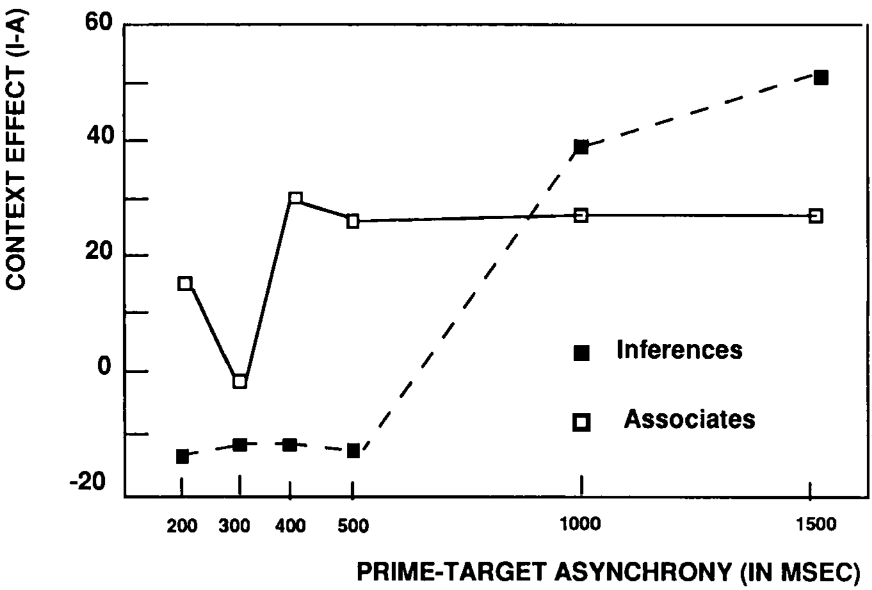



**Figure 8. Context effect as indexed by the reaction time difference to context inappropriate and appropriate associates as a function of processing time, after Till, Mross, and Kintsch (1988).**

Targets that are contextually appropriate associates of the priming word are primed at all four SOAs. Contextually inappropriate targets, however, are primed only when the priming word is still in its initial processing stages; by 400 ms inappropriate associates are no longer activated. Topical inferences are primed only if there is ample time, more than 500 ms, for the processing of the prime and its discourse context. This observation implies that the topic was not inferred immediately as soon as the relevant information became available, but was left for the sentence wrap-up period. Till et al.'s sentences were written in such a way that the topic could have been inferred before the last word in the sentence. This, however, is not what happened: Topics were inferred only after the whole sentence was read, requiring more than 500 ms of processing time. Thus, the full contextual meaning of the prime required about 1 s to emerge.

Data like these suggest that the initial activation of lexical knowledge is independent of the discourse context. What matters is only the (relatively fixed and stable) associate/semantic context of each word by itself. This stage of sense activation, however, is quickly followed by a process of sense selection in which the discourse context becomes effective: By 500 ms, context-inappropriate associates are deactivated (see also Seidenberg et al., 1982, and Swinney, 1979). If given more time, context effects grow even stronger: By 1,000 ms, contextually appropriate inference words are strongly and reliably primed even in the absence of associative connections (similarly for recognition, see McKoon & Ratcliff, 1986).

Clearly, this pattern of results is in excellent agreement qualitatively with the model of knowledge use in discourse presented earlier. Right after a word is perceived, it activates its whole associative neighborhood in a context-independent way, with the consequence that strong associates of a word are likely to be represented in working memory and hence will be primed in a lexical decision task, whether they are context appropriate or not. The knowledge-integration process then results in the deactivation of material that does not fit into the overall discourse context (such as context-inappropriate associates). Note that in order to disambiguate words on-line, the integration phase cannot be delayed until the end of a processing cycle; word senses are disambiguated before that. In the model, therefore, as soon as a text proposition is constructed and its associates have been generated, they will be integrated into whatever context exists at that time in working memory. Thus, each processing cycle involves many integrations, and the single integration operation performed at the end of each cycle in many of the examples discussed here is merely a simplification, adopted whenever one is not concerned with the on-line generation of word meanings. Finally, contextual inferences should require the most time to become activated on the average because although they sometimes result from the initial knowledge sampling, in other cases repeated sampling, or, further, strategic elaboration might be required.

Earlier, an example was given of one of the texts used in the Till et al. (in press) study. The predictions of the model will be illustrated by means of this example. The aforementioned text (*The townspeople were amazed to find that all the buildings had collapsed except the mint*) has the following propositional representation:

1. TOWNSPEOPLE
2. AMAZED [TOWNSPEOPLE, P3]
3. COLLAPSE [P4]
4. ALL-BUT [BUILDING, MINT]
5. BUILDING
6. MINT

Connection strengths of .9, .7, .4, and 0 were assigned to text propositions one, two, three, or more steps apart in the text base (e.g., P1 is two steps away from P3, connected via P2). Next, each text proposition was allowed to access at random two of its neighbors in the long-term associative net. This process was simulated by having an informant provide free associations to phrases based on each of these six propositions. For instance, the phrase *all buildings but the mint* elicited the associations *many buildings* and *mint is a building*. Of course, MONEY and CANDY were chosen as the associates of MINT. Each text proposition was connected by a value of .5 to its associates, yielding an 18 X 18 connectivity matrix. Activation was then allowed to spread from the text propositions to the knowledge elaborations. Specifically, an initial activation vector with 1/6's corresponding to the text propositions and zeros otherwise was multiplied with the connectivity matrix until the pattern of activation stabilized. As a result, text propositions achieved activation values between .0987 and .1612, depending on how closely they were tied into the text base, and the knowledge elaborations had much lower activation values, between .0142 and .0239, with both MONEY and CANDY having a value of .0186. Thus, at this stage of processing, MONEY and CANDY are equally activated.

Activation continues to spread, however, and differences begin to emerge among the activation values for the various knowledge elaborations that have been added to the text base. The reason for this is that the knowledge elaborations are connected not only to the text propositions that had pulled them into the net but also to other text propositions as well as to each other. To approximate these interrelations, a connection value of .5 was assigned to any two propositions sharing a common argument. Because the homophone *mint* contributed associations to the subnet that refers to both of its senses, an inhibiting connection of --.5 was assigned to MINT/CANDY and BUILDING, whereas CANDY and MONEY themselves were connected by a --1. Continued multiplication of the activation vector with this connectivity matrix yielded a stable pattern (average change < .001) after 11 operations. At this point text propositions had activation values ranging between .1091 and .0584. Several of the knowledge elaborations reached values in this range, for example, .0742 for both ISA [MINT, BUILDING] and MONEY and .0708 for KILL [BUILDING, TOWNSPEOPLE], whereas others had faded away by this time, for example, MAN, which entered the subnet as an associate of TOWNSPEOPLE, had an activation value of .0070 and, most significantly, .0000 for CANDY. This stage of processing corresponds to the 400- and 500-ms points in Figure 8: MINT is now clearly embedded in its context as a kind of building, and the inappropriate association CANDY is no longer activated.

The next processing stage involves the construction of a topical inference -- what is the sentence about? While the exact operations involved in the construction of such inferences are beyond the scope of this article, van Dijk and Kintsch (1983, chapter 6) have discussed some of the mechanisms involved, such as a strategy of looking for causal explanations, which is what actual subjects appear to use predominantly in the following case. If given enough time, the modal response of human readers is that the sentence is about an earthquake that destroyed a town. Thus, the (empirically determined) propositions EARTHQUAKE and CAUSE [EARTHQUAKE, P3] were added to the text base and connected with the text-base propositions from which they were derived by a value of .5. The two new propositions were given initial activation values of zero, and the integration process was resumed; that is, activation now spread from the previously stabilized subnet into the newly constructed part of the net. Nine more integration cycles were required before the expanded net stabilized. As one would expect, the two new inferences did not alter the pattern of activation much, but both of them became fairly strongly activated (thereby diminishing activation values in the already existing portion of the net). The topical inferences EARTHQUAKE and CAUSE [EARTHQUAKE, P3] ended up with activation values of .0463 and .0546, respectively, among the most strongly activated inferences in the net. At this point, the process appears to coincide with the time interval between 1.000 and 1.500 ms shown in Figure 8.

The construction-integration model thus accounts for the data in Figure 8 by means of an intricate interplay between construction and integration phases: the construction of the text base and the context-free, associative knowledge elaboration during the first 350 ms of processing; the establishment of a coherent text base, which appears to be complete by 400 ms; and finally, an inference phase, involving new construction and new integration and requiring more than 500 ms of processing under the conditions of the Till et al. study. The model does not account for the time values cited here, but it describes a processing sequence in accordance with the empirically determined time sequence.

In many models of word identification, the problem is thought to be "How do we get from a certain (acoustic or visual) stimulus pattern to the place in the mental lexicon where the meaning of this word is stored?" In the present model, word identification is much more deeply embedded into the process of discourse understanding. The lexical node itself provides just one entry point into the comprehender's long-term memory store of knowledge and experiences, and what eventually becomes activated from that store depends on the discourse context. In conceptions of the lexicon like that of Mudersbach's (1982), the meaning of a word is given by its "neighborhood" in the associative network into which it is embedded. Neighborhoods may be defined narrowly or broadly (nodes one link away vs. nodes several links away). In the present model, the meaning of a word is also given by its neighborhood --narrowly or broadly defined--not in the long-term memory net as a whole, but in the subnet that has been constructed as the mental representation of the discourse of which the word is a part. Because that representation changes as processing proceeds, word meanings change with it.

Figure 9 depicts the changing meaning of MINT in our example. MINT is directly linked to nine propositions in the network; indirectly it is linked to the whole net, of course. If one takes as its contextual meaning only its immediate neighbors, one finds at the beginning of processing mostly closely related propositions from the text base plus three weakly activated knowledge elaborations that in part do not fit into the context at all (CANDY). At the end of the process, however, the context-inappropriate association has dropped out, other inferences have been added, and the activation is more evenly distributed among text propositions and knowledge elaborations. Thus, textual information becomes part of the contextual meaning of a word, in contrast to most traditional conceptions of "meaning."

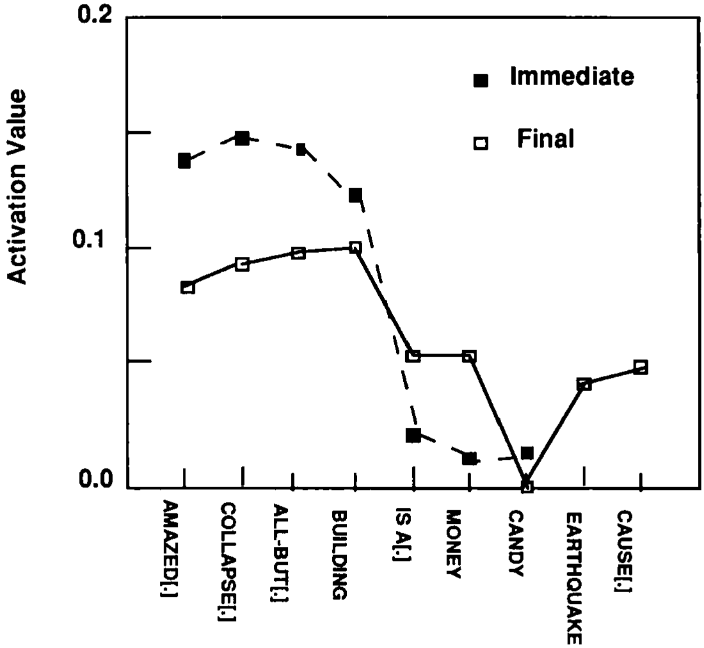



**Figure 9. The changing meaning of MINT. (The activation values of all propositions directly connected to MINT at the beginning and at the end of the process. The [.] notation is used as an abbreviation for the arguments of a proposition).**

This example is, of course, no more than an illustration. Parameters in our calculations could be changed. For example, more than just two associates could be sampled initially in the process of knowledge elaboration. In this case the neighborhood of MINT would contain many more knowledge elaborations than are shown in Figure 9, where there is a strong predominance of text propositions. Not enough is known at present to set some of these parameters with confidence. But Figure 9 does reflect certain aspects of the data correctly: the equal initial activation of MONEY and CANDY, the later emergence of the topical inference EARTHQUAKE. Although much more research is needed to produce a more adequate picture of how the contextual meaning of words is constructed during discourse comprehension, here is a technique that at least may help us to do so.

**Arithmetic Word Problems**

How children understand and solve simple word arithmetic problems provides an excellent domain to try out the construction-plus-integration model. Unlike with many other types of discourse, there are clear-cut criteria for when a problem is solved correctly, and the formal knowledge of arithmetic that is necessary for its solution is easily defined. However, word problems, like all other texts, share the ambiguity and fuzziness of all natural language. Not only formal, arithmetic knowledge is involved in understanding these problems, but all kinds of linguistic and situational knowledge. What makes word problems hard ---and interesting-- are often not their formal properties, but the way a problem is expressed linguistically and the way formal arithmetic relations map into the situations being described. Thus, word problems are ideal from the standpoint of knowledge integration because it is precisely the integration of formal arithmetic knowledge and linguistic and situational understanding that is at issue here.

Another reason for choosing the domain of word problem is that there already exist alternative formal models of how children solve simple word arithmetic problems (Briars & Larkin, 1984; Kintsch & Greeno, 1985). Specifically, the work of Kintsch and Greeno will be taken as a starting point here. Their model represents a union of the work on problem solving in arithmetic by Riley, Greeno, and Heller (1983) on the one hand, and that on discourse understanding by van Dijk and Kintsch (1983) on the other. Kintsch and Greeno (1985) added to the discourse-comprehension strategies of the van Dijk and Kintsch model some special purpose strategies for solving word arithmetic problems, which they named the *arithmetic strategies*. For instance, if the model encounters a quantity proposition, such as "six marbles," it forms a set and tries to fill in the various slots of the set schema: what the objects are, the cardinality of the set, a specification of the objects (e.g,., that the marbles are owned by Fred), and the relation between the present set and other sets in the problem (the six marbles were given to Fred by Tom, which might identify them as a "transfer set"). Thus, the Kintsch and Greeno model for word problems builds a text base in quite the same way as in the van Dijk and Kintsch general theory of text comprehension, but it them forms a very specialized situation or problem model in terms of sets of objects and their interrelations. It solves a problem by recognizing a particular pattern of relations among sets (such as TRANSFER-IN or SUPERSET) and then using a stored-solution procedure appropriate to that case.[7] Thus, in terms of the foregoing discussion about knowledge use in discourse, the Kintsch and Greeno model is a "smart" model: Production rules are formulated in such a way that in each situation exactly the right arithmetic strategy is fired.

The Kintsch and Greeno model of solving arithmetic word problems is useful in several ways. The model identifies different classes of errors, such as errors caused by a lack of arithmetic knowledge, errors caused by linguistic misunderstandings, and errors that do not reflect a lack of knowledge at all but

---

[7]Computer simulations of this model have been developed by Fletcher (1985) and Dellarosa (1986) and are available from the author.

result from resource limitations. Certain formulations of word problems overload the resources of the comprehender, especially short-term memory, leading to a breakdown in processing. As Kintsch and Greeno have shown, within each arithmetic problem type there exists a strong correlation between the frequency of errors made in solving the problem and the memory load imposed by it, even though there are no differences within problem types in either the arithmetic or linguistic knowledge required for solution.

The model distinguishes between linguistic and arithmetic errors and helps us to investigate to what extent errors made by second- and third-grade pupils are caused by a failure to understand properly the text of the word problem, rather than by a faulty knowledge of arithmetic (e.g., Dellarosa, 1986; Dellarosa, Kintsch, Reusser, & Weimer, in press; Kintsch, 1987). If certain linguistic misunderstandings about the meanings of such key words as *have more than*, *have altogether*, or *some* are built into the knowledge base of the model, the model produces a pattern of wrong answers and misrecall of the problem statements that strikingly parallels some of the main types of errors that experimental subjects make. This is a good example of how much can be achieved even with the use of knowledge-poor representations in studies of discourse processing. The Kintsch and Greeno model knows about arithmetic (its arithmetic strategies), and it knows about the meaning of words (its lexicon; a *semantic net* in Dellarosa, 1986). However, it has no general world knowledge that would allow it to understand the situation described in a word problem. It merely picks out the crucial arithmetic information from the discourse and builds a propositional text base for it. This is good enough for some purposes (e.g., the investigation of resource limitations or linguistic factors in understanding as mentioned earlier, or to predict recall, summarization or readability as in Kintsch & van Dijk, 1978, and related work), but it is not good enough for other purposes.

The limits of this approach are illustrated by a well-known observation: If a word problem is embedded into a concrete, familiar situation or action context, it is much easier to solve than when the same problem is expressed abstractly (e.g., Dellarosa et al., in press; Hudson, 1983). Thus, *Five birds saw three worms on the ground, and each bird tried to get a worm. How many birds didn't get a worm?* is easy for first graders, but *There are five red marbles and three green marbles. How many more red marbles are there than green marbles?* is very hard, even though the two problems are equivalent in form.

The Kintsch and Greeno model does not account for this difference. What is needed is a model in which all knowledge relevant to the understanding of a word problem becomes integrated into a representation that is sensitive to arithmetic as well as to situational information. In the model to be described shortly, this is achieved by forming many different hypotheses about the arithmetic relations in the problem, instead of only a single one, and then by looking for information in the text in support of each hypothesis. Thus, situational and arithmetic information can combine in forming the problem interpretation.

## Arithmetic Strategies

Arithmetic knowledge forms a special subset of a person's general knowledge network. Sets of objects can be represented by a propositional schema with the slots object, specification, quantity, and role (i.e., their relation to other sets) -- equivalent to the set schema of Kintsch and Greeno (1985). Superordinate schemata can be similarly defined. Thus, a TRANSFER-IN schema can be set up with slots for a START, TRANSFER-IN, and RESULT SET. With each such superordinate schema, various arithmetic procedures (such as the counting strategies of Kintsch & Greeno, 1985) can be associated.

Arithmetic knowledge is used in the same way as general world knowledge. That is, propositions that represent various hypotheses about the arithmetic structure of a word problem are constructed as the text of the word problem is read and become part of the subnet. Several possible arithmetic hypotheses are constructed at each point, and whatever evidence in the text favors one or the other of these hypotheses is connected positively to it.

The strategies required for solving arithmetic word problems have been described in Kintsch and Greeno (1985) and Kintsch (1984) and have been incorporated into the computer simulation of Dellarosa (1986). However, they are used here in a different way. In the aforementioned works, the intent was to circumscribe the conditions under which each strategy is applied so accurately that only the correct one is fired in each word problem. Here, strategies fire promiscuously whenever they are supported, however weakly, by the text, and it is left for the integration process to weed out what is not wanted, just as all sorts of general knowledge propositions are activated that later turn out to be useless. A problem is solved when the right superordinate schema is more strongly activated than its alternatives, which then triggers the desired arithmetic procedures.

Three forms of arithmetic strategies need to be considered. There are strategies that form arithmetic hypotheses about sets, strategies that determine the nature of the connections between various text propositions and these hypotheses, and strategies that form superordinate arithmetic schemata on the basis of which arithmetic calculations can be performed.

1. Hypotheses about sets are propositions of the form SET [object: X, specification: Y, quantity: Z, role: W], where X refers to an object, such as TULIP; Y is one or more other text propositions, specifying X further --for example, PAST (LOCATION [TULIP, IN-GARDEN]]; Z is a quantity proposition with X as argument --for example, FOURTEEN [TULIP]; and W indicates the role of the set is come superschema, such as WHOLE or PART.

2. Whenever a quantity proposition is encountered in the text base, possible arithmetic hypotheses derivable from it are constructed (e.g., two otherwise identical propositions with the roles WHOLE and PART). Propositions in the text base that provide evidence for any of these alternatives are then connected

positively to it. Key words can be used for this purpose, as in Kintsch and Greeno (1985): Collection terms such as *altogether* indicate WHOLE sets: *give/take, of these,* and *have more/less than* indicate PART sets. In addition, general world knowledge about situations and actions is used to determine what is a WHOLE and what are its PARTS. The strategies involved have been described in Kintsch (1984): restricted subsets, conjunction, and time-ordered possession/location.

**Restricted subsets.** If the specification of one set is more general than that of another, the former is assigned the role of WHOLE and the latter that of PART. Examples are LARGE-WINDOW, SMALL-WINDOW versus WINDOW, or ON-UPPER-SHELF, ON-LOWER-SHELF versus ON-SHELF.

**Conjunction.** If the object or specification of one set consists of the conjunction of the objects or specification of two other sets, the former is the WHOLE and the others the PARTS. This conjunction may be explicit as in YESTERDAY, TODAY, and YESTERDAY & TODAY, or implicit, as in TEDDYBEAR, DOLL, and TOY.

**Time-ordered possession/location.** If the specification slots of three sets contain either HAVE [agent, object] or LOCATION [object, place], or the negations of these propositions, as well as information to establish a temporal order, WHOLE and PART roles can be assigned to the three sets according to the resulting patterns. For instance, if the specifications of three sets are TIME1 [HAVE, JOE, P[MARBLES]]: TIME2 [GIVE, JOE, TOM, Q [MARBLES]], which implies TIME2 [NOT [HAVE, JOE, Q [MARBLES]]: and TIME3 [HAVE [JOE, Z [MARBLES]]. SET1 is indicated as the WHOLE set.

3. The PART-WHOLE schema is the only arithmetic super-schema to be considered in the examples that follow, though various TRANSFER and COMPARISON schemata could have been treated in the same way, as in Kintsch and Greeno (1985). Three hypotheses can be formed about the PART-WHOLE schema, depending on whether the first, second, or third of the sets formed is to be considered the WHOLE set. (Note that the order in which the sets were formed in the word problem, not their true temporal order, is at issue here.). Thus, a proposition with the head PPW, which is simply a mnemonic for PART-PART-WHOLE, expresses the hypothesis that the problem is a PART-WHOLE problem with the third set as the WHOLE: PPW [role [SET1, PART], role [SET2, PART], role [SET3, WHOLE]]. Associated with this schema is the equation Q1 + Q2 = Q3, where Qi is the quantity of the *i*-th set, as well as procedures to solve this equation, depending on which of the quantities happens to be unknown.

### Examples and Issues

Three examples will be analyzed here to show how the model understands, or fails to understand, as the case may be, arithmetic word problems. To see how these examples work, it is necessary to present at least the first one in sufficient detail. This problem is intended simply as an illustration of the basic mechanisms of the model --nothing much of interest happens with respect to the arithmetic, and textually, the only thing of significance is that a simple inference is formed, which, however, is crucial for the understanding of the problem. Two more examples, which will not be presented in as much detail, will serve as illustrations of how the model can account for some well-known facts about word-problem solving that alternative models (Briars & Larkin, 1984; Kintsch & Greeno, 1985) do not handle readily.

### Inferences

Manolita tried to weed her father's garden. "You sure weeded it," said Mr. Mundoza. "There were fourteen tulips in the garden and now there are only six." How many tulips did she pull out by mistake?

This problem, modified from the "Thinking Stories" of Willoughby, Bereiter, Hilton, and Rubinstein (1981), requires for its solution the application of one of the LOCATION strategies: There were so many tulips in the garden, then some were pulled out, and now so many are left. A simple, knowledge-based inference becomes necessary: that the tulips that were pulled out are no longer in the garden. The knowledge-activation mechanism of the present model readily supplies this inference, and the problem will be solved successfully.

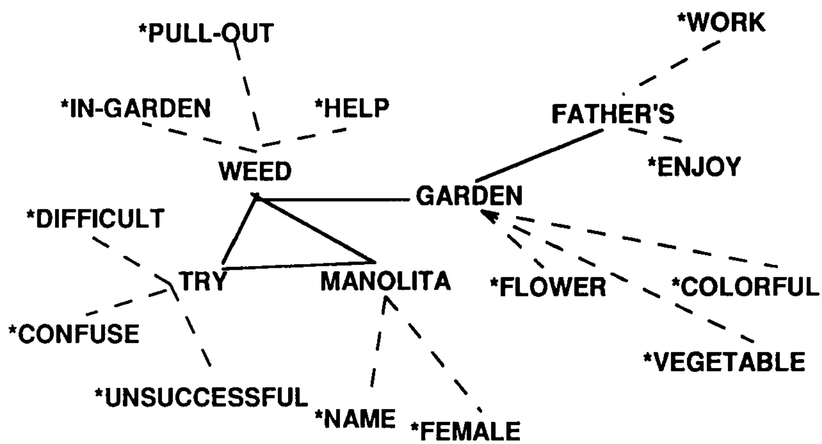



Figure 10. The elaborated text base for the first sentence of the Manolita problem. (Each proposition is indicated by a single word. Text propositions are connected by solid lines, their associates by broken lines. Associates are marked with an asterisk.)

The model processes this problem in three cycles, which includes the first sentence, the statement by Mr. Mundoza, and the question sentence. The first sentence simply sets up a context and is not directly relevant to the arithmetic. In Figure 10, the way the model understands this sentence is indicated, albeit in abbreviated form. The propositions constructed from the sentence itself are (P1) MANOLITA, (P2) GARDEN, (P3) TRY [MANOLITA, P4], (P4) WEED [MANOLITA, GARDEN], (P5) FATHER'S [GARDEN]. Only the first terms of these propositions are shown in Figure 10. Also shown in Figure 10 are the propositions that were added to the text base through the process of associative-knowledge elaboration (they are marked with an asterisk and, once again, abbreviated: *NAME stands for ISA [MANOLITA, NAME], etc). Because no simulation of a general knowledge network is available, or even conceivable, the process of knowledge elaboration must be approximated empirically. An informal procedure was adopted for this purpose: Three persons were asked to provide free associations to phrases corresponding to P1 through P5 (as well as to corresponding phrases from the remaining two sentences of this word problem), and the responses generated by at least two persons were considered as the top associates of each proposition in the general knowledge net (up to a maximum of three associations per proposition).

The text base shown in Figure 10 serves as a basis for deriving a connectivity matrix, using the principles illustrated earlier in Tables 1 and 2: Text propositions are connected depending on their proximity in the text base, each text proposition is connected to its associates by a value of .5, and knowledge derived propositions are interconnected by the same value if they share an argument, or by -.5 if different word senses are involved (this does not occur in the present example).

An initial activation vector consisting of .2 s for the five propositions directly derived from the text, followed by 13 zeros for the propositions generated from the knowledge net, was then repeatedly updated by multiplying it with the connectivity matrix until the activation values stabilized, as in the examples discussed previously. Ins the present case, activation levels stabilize after 10 iterations. The resulting pattern of activation is shown in the first panel of Figure 11, WEED [MANOLITA, GARDEN], whose centrality in the text base is apparent in the graphical representation, has the highest activation value, and the other text-derived propositions also have fairly high activation values. Knowledge-derived propositions are considerably less activated. The four most strongly activated propositions (P1 through P4) are retained in the short-term buffer and enter the second processing cycle.

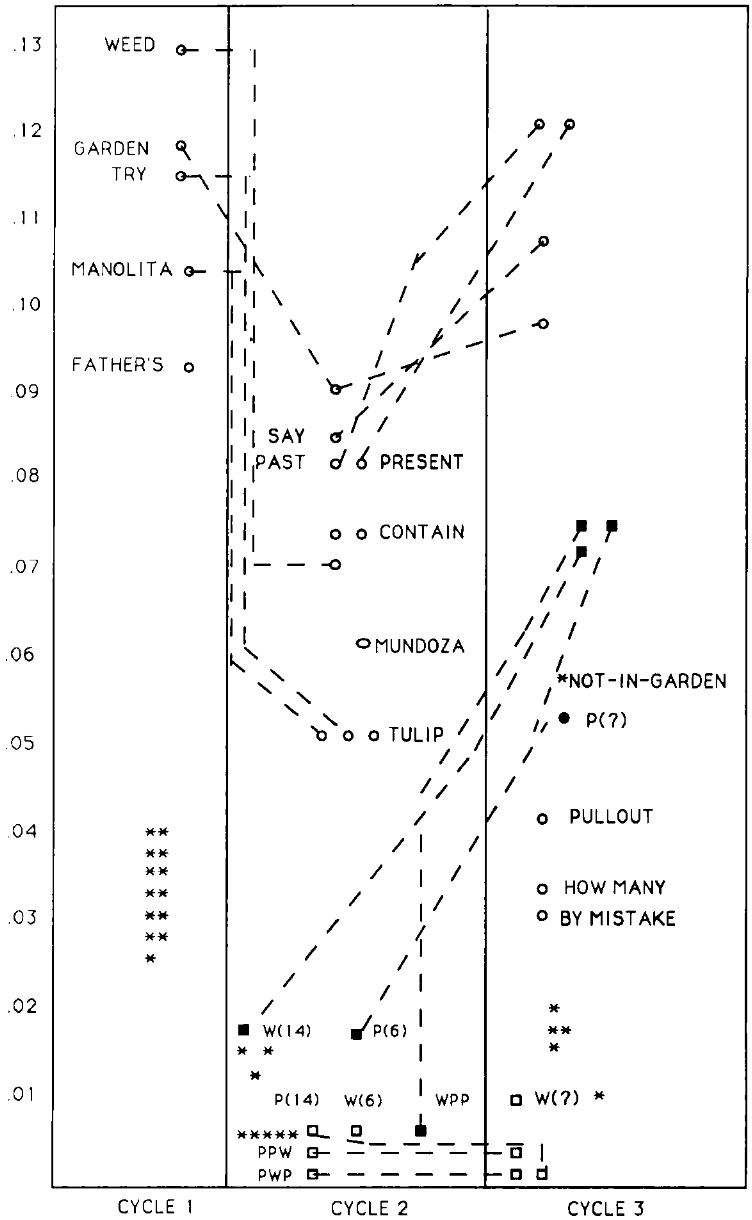



Figure 11. The result of the integration process for the three sentences in the Manolita problem. (Propositions are indicated by single words; inferences are marked by an asterisk; their arrangement in the figure is approximated. The ordinate shows the activation values of each proposition after the process has stabilized. Propositions carried over from one processing cycle to the next are connected by arrows.)

The second processing cycle is shown in Figure 12. The four propositions held over in the short-term memory buffer from Cycle 1 are joined by 9 new text propositions and 11 associated propositions from the knowledge base. (Because of lack of space, the latter are indicated only by asterisks). The quantity propositions FOURTEEN [TULIP] and SIX [TULIP] generate four arithmetic hypotheses: that the 14 tulips that were in the garden in the past are, respectively, a PART or WHOLE set, and that the 6 tulips now in the garden are a PART or WHOLE set. What the reader knows about weeding gardens provides the crucial information that discriminates among these hypotheses: The tulips before the garden was weeded are the WHOLE set, and only a PART is left after the weeding. This knowledge is expressed in the connectivity matrix by connecting PAST with WHOLE [14], and PRESENT with PART [6].

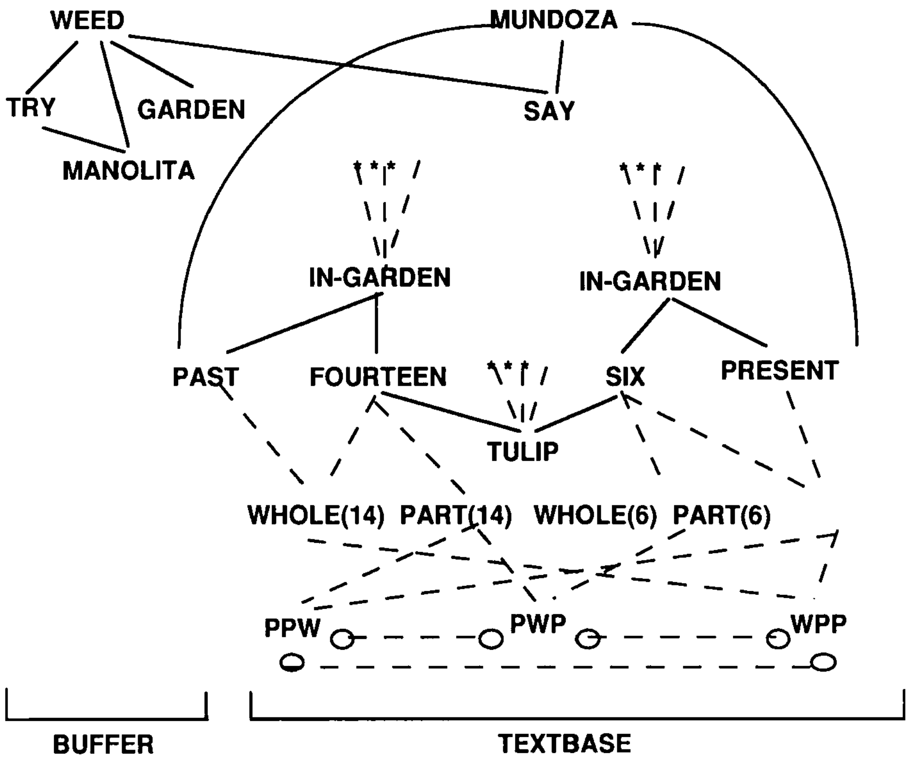



Figure 12. The elaborated text base for the second sentence of the Manolita problem. (Four propositions were carried over from the previous cycle in the short-term memory buffer. Solid lines connect text propositions, broken lines inferences, nonarithmetic inferences are indicated by asterisks only.)

The last three propositions that enter the subnet are the superordinate arithmetic hypotheses PPW, PWP, and WPP. They receive support from their corresponding first-order arithmetic hypotheses. Thus, whatever strength each arithmetic hypothesis gathers from the text is fed into the superordinate arithmetic schemata consistent with it. These schemata are mutually exclusive and inhibit each other with connection values of --1. Note that only at this final level is inhibition among arithmetic hypotheses used: The hypotheses that a particular set of objects plays the role of WHOLE or PART set are also mutually exclusive, but they are not allowed to inhibit each other; they merely collect more or less positive evidence, which they then transmit to the superordinate stage where a selection among alternatives is made.

The resulting connectivity matrix then becomes the multiplier of the activation-state vector for the 28 propositions participating in this second processing cycle. Initially, these activation values are positive for the text-derived propositions, and zero otherwise, except for the propositions carried over in the buffer, which retain the activation values they reached in the last cycle. In this case, the activation vector stabilizes already after seven operations. The results are shown in the second panel of Figure 11. (If the activation process is extended to twice the number of cycles, the activation values for the arithmetic hypotheses, measured to four decimal places, do not change at all). All text-derived propositions remain strongly activated, while none of the textual inferences (e.g., MUNDOZA is a NAME of a MALE, TULIPS are FLOWERS, RED, and GROW-IN-HOLLAND) reach a high level of activation. This is intuitively quite plausible. As far as the arithmetic is concerned, the problem is at this point understood correctly and practically solved: WHOLE [14] is more strongly activated than its alternative, PART [14]. Similarly, PART [6] is stronger than WHOLE [6]. The correct hypothesis, WPP, is the most strongly activated of the three alternative superschemata.

Note that the text propositions and inferences are, in general, much more strongly activated than the arithmetic hypotheses. Therefore, the activation values of the latter must be considered separately, relative to each other, rather than in relation to the text propositions when it comes to selecting propositions to be maintained in the short-term memory buffer. This imbalance is required for the model to work. If the arithmetic hypotheses are weighted more heavily, they draw the activation away from the text itself, and the system cannot stabilize: It will flip-flop between alternative, mutually contradictory arithmetic schemata. The arithmetic hypotheses have to be anchored in a stable text representation.

For the third and final sentence, the short-term memory buffer needs to carry over both text propositions to establish textual coherence and arithmetic hypotheses to take advantage of the understanding of the problem that has been achieved so far. It has been assumed here that the four strongest text propositions as well as the four strongest arithmetic hypotheses are carried over in the buffer, as shown in Figure 13. (There are, of course, other plausible alternatives). The three text propositions generated on the basis of this sentence bring with them into the net six knowledge propositions, one of which is NOT

[CONTAIN [GARDEN, TULIP]], which turns out to be crucial for the solution of the problem. In addition, new hypotheses about the question set are formed, and the schemata PPW and PWP, which were lost after the second cycle, are reconstructed. Because the child knows about weeding gardens, the tulips that were pulled out are identified as a part of those that were in the garden, in the beginning. Hence, a connection that favors the PART hypothesis over the WHOLE hypothesis is formed between the inference NOT [CONTAIN [GARDEN, TULIP]] and PART[?]. It completes the pattern that is the condition for the use of a LOCATION strategy: some tulips at one place in the past, then some not there, now some are left.

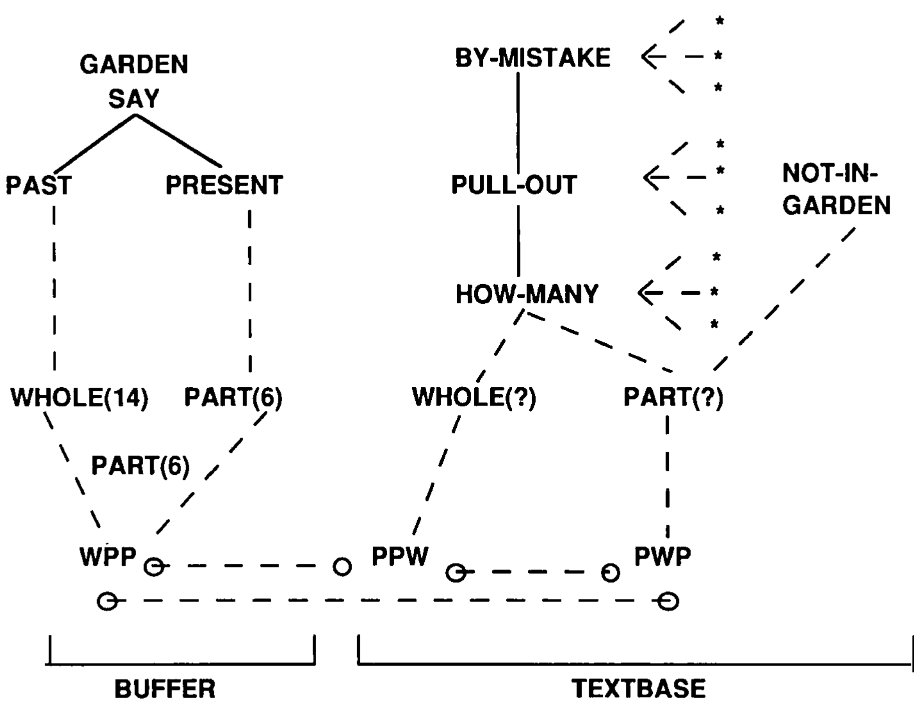



**Figure 13. The elaborated textbase for the third sentence of the Manolita problem.**

The new net requires 43 operations to stabilize. The knowledge-based inference NOT [CONTAIN [GARDEN, TULIP]] achieves an activation level above the range of the text propositions (Figure 11, third panel). The picture is completely clear as far as the arithmetic is concerned: All the correct hypotheses are strongly activated, and all incorrect alternatives have low or zero activation values.

The final steps in the solution of the problem are procedural. From information associated with the WPP pattern the equation 14 = 6 + ? is generated, which is then used to obtain the correct answer. A lot of mountains had to be moved to achieve a very simple result!

The Manolita problem was solved without problem solving. The basic comprehension operations were sufficient; that is, it produced the inference that the pulled-out tulips are not in the garden which was required for the application of the LOCATION strategy. However, this is not always the case. In many, not necessarily difficult, problems, more focused problem-solving operations are required because the random-inference generation process described earlier fails to generate the required inference. Consider the following "thinking problem":

Mrs. Nosho was telling Mark about the two huge aquariums she kept when she was a little girl. "There were 30 fish in one and 40 fish in the other, so you can tell how many fish I had."
How many fish did Mrs. Nosho have?

In a simulation run of this problem the model failed because it did not come up with the transitive inference HAVE [X, Y] & CONTAIN [Y, Z] implies HAVE [X, Z]. At this point, the process needs to go into a problem-solving mode in which the information in the text is elaborated in a more focused manner than is possible with the automatic-comprehension mechanisms discussed here.

### Context Effects

Problems embedded into a familiar situational context are much easier to solve than problems that must be solved without this situational support (e.g., Hudson, 1983). Thus, birds catching worms present a concrete, understandable situation that makes it clear what is the whole and what are the parts, whereas abstract, ill-constrained problems do not. All depends on whether the right arithmetic strategy is used; the situation is of no help.

In the worm-and-bird problem, the text provides a situational constraint for the interpretation of the problem that has very little to do with arithmetic per se. It is the knowledge about birds eating worms that matters. The birds trying to catch the worm are understood as the WHOLE set, with the birds catching worms as one PART, and the birds unable to get a worm as the other PART. This understanding was achieved not because a certain key phrase, like *how many more,* was parsed correctly but on the basis of general world knowledge. If there are birds, some of whom catch and some of whom do not catch a worm, what is the WHOLE set and what are the PARTS is given by general world knowledge that is not specific to arithmetic. The arithmetic can hardly go wrong here because the well-known situation guarantees the right interpretation of the problem. It is this aspect that the present model deals with most effectively.

Context, however, does not always facilitate problem solution, it may also interfere with it. Consider this typical school problem, with its highly impoverished context:

Fred has four Chevies and three Fords. (a) How many cars does he have altogether? (b) How many more Chevies does he have than Fords?

Context is no help with this problem; it must be solved on the basis of specialized arithmetic strategies, on the basis of the key words *have altogether* for Question A and *have more than* for Question B. Of course, children are much more familiar with the former (e.g., Riley et al., 1983), but if the right strategies are available, both problems will be solved. In the model, too, the *altogether* in Question A will be connected with the HOW-MANY/WHOLE hypothesis, and the *have more than* will be connected with the HOW-MANY/PART hypothesis in Question B, and both questions will be answered equally well. After the first sentence, PART and WHOLE hypotheses are established for both the Chevies and the Fords, but there is not much to distinguish them; the superordinate schemata PPW, PWP, and WPP are only weakly activated and hardly differentiated. Question A, on the other hand, correctly activates the PPW hypothesis, and Question B yields the WPP result. Thus, if the arithmetic knowledge is available, it makes very little difference which question follows the problem statement.

In contrast, if the problem is only slightly contextualized, the model can be biased in favor of one of the questions, and actually fails when it gets the wrong one. Suppose, the foregoing problem is changed to read

Fred has a nice collection of antique cars. Four of his cars are Chevies, and three are Fords.

**Collection**, like some, is constructed as a quantity proposition, and hence PART and WHOLE hypotheses for a set of cars with unspecified quantity are established in the first processing cycle. They are both activated equally, however, at this point. This changes dramatically with the second sentence: The *four Chevies* and *three Fords* are both identified as PART sets because of the phrase *of his*. In consequence, the model begins to favor the WPP hypotheses. When it receives Question A, the WPP hypothesis is decisively strengthened, and the problem is solved correctly. On the other hand, if it is given Question B, the model becomes confused between the WPP and PWP hypotheses, which are both equally activated, and fails to solve the problem.

Thus, we have here an example where the problem context interferes with the solution of a problem. It biases the problem in favor of one particular interpretation, so that when another interpretation is required, the whole process fails. It is important, however, to analyze exactly why the model failed to answer Question B correctly: After processing the second sentence, it was so strongly convinced that the *four Chevies* and *three Fords* were both PART sets that it did not carry over the corresponding WHOLE set hypotheses and therefore had no way of using the information in the *have-more-than* question in support of the CHEVIES/WHOLE hypothesis. Thus, rather special circumstances prevented the model from answering Question B. In slightly different circumstances, it could have done so: (a) if the buffer were large enough, the CHEVY/WHOLE hypothesis would not have been lost, or (b) if the model had been allowed to reread the problem statement.

## Question Specificity

The final example illustrates some different aspects of word-problem solving; namely the complex role that redundant specifications of sets may have. On the one hand, overspecifying a set can be helpful because it provides more than one way to refer to it. On the other hand, redundant specifications increase the length of the text and thus the likelihood that some important piece of information is no longer in active memory when it is required. In the following problem, three versions of the question are possible:

Joe had a collection of nine marbles. He started his collection with some beautiful red marbles. Then Lucy added six pink marbles to his collection as a present. (a) How many beautiful red marbles did he start his collection with? (b) How many marbles did he start his collection with? (c) How many beautiful red marbles did he have?

The first processing cycle results in undifferentiated hypotheses about the nine marbles. The set constructed in the second cycle, on the other hand, is clearly a PART set, as is the one constructed in the third cycle. Indeed, at the end of the third cycle, the model understands the problem essentially correctly, with the WPP schema greatly exceeding alternative hypotheses in activation value. To understand what happens next, it is necessary to know which text propositions were maintained in the buffer at the end of the third cycle: Only propositions from the third sentence are carried over, while the propositions from the second sentence are no longer held in active memory at this point. This has nontrivial consequences when the question is asked. In Versions A and B everything is all right, because the question itself identifies the question set as a PART set--*starting a collection* serves this function, just as it did in Sentence 2. Version C of the question, on the other hand, does not yield a correct solution. The question itself does not indicate the role of the question set, and there is no information from the second sentence still available in active memory that would help to identify its role either, because there are already several strong PART hypotheses around, the model tends toward the hypothesis that the question set has the role of a WHOLE; the PWP schema thus becomes more activated than the correct WPP schema.

However, this is far from an unequivocal prediction of failure for Version C of the question. With a slightly larger buffer, or with a little less irrelevant material intervening (*pink marbles, as a present*), the critical information from the second sentence could have been maintained in the buffer and used to solve the problem. Or even more obviously, the problem solver could reread the problem or perform a reinstatement search (Kintsch & van Dijk, 1978; Miller & Kintsch, 1980) to activate the required information from long-term memory. Rather the prediction is that children, like the model, would have more trouble with Question C, and fail more frequently, than with either A or B.

Thus, the more specific the question the better. But how irrelevant or redundant material will affect the difficulty of a word problem is a more complex story. It may be quite harmless, or may even facilitate problem solving, if the question exploits a redundancy in the specification of a set. But it may be a source of difficulty and even a cause of failure when the question is asked in an unhelpful way. The present model has the flexibility to handle these complex effects of context: Many small effects are allowed to add up and pull the model one way or another. The "smart" models of Kintsch and Greeno (1985) and Briars and Larkin (1984) have no ready way to cope with these subtle contextual demands: Either the right strategy is used or not.

**Discussion**

How people recall relevant knowledge when they read a text is reminiscent of another experimental paradigm that has been studied extensively in psychological laboratories: how people recall lists of words. A widely used explanation for the recall of word lists is based on the generation-recognition principle. Some words are recalled directly, perhaps from a short-term memory buffer, and these words are then used to generate other semantically or contextually related, plausible recall candidates. Words that have actually appeared in the to-be-learned list will be recognized among these candidates and recalled, whereas intrusions will tend to be rejected. Generation-recognition theories have had their detractors, and in their most primitive form they are certainly inadequate to account for the more complex phenomena of list recall. However, sophisticated versions of this theory are widely accepted now. Almost every current model of list recall includes a generation/retrieval as well as a recognition/editing stage.

The model of knowledge use in discourse comprehension proposed here has two analogous stages: First, a propositional network must be constructed, and then it must be edited or integrated. The way the construction process is thought of here is a straight extension of previous work on discourse processing. The words and phrases that make up a discourse are the raw material from which a mental representation of the meaning of that discourse is constructed. This mental representation takes the form of a propositional text base. Text bases combine two sources of information: the text itself and knowledge --knowledge about language as well as knowledge about the world. To construct even a single proposition, an appropriate frame must be retrieved from one's store of knowledge, and its slots must be filled in the way indicated by the text. The novel aspect of the present model is that the role of knowledge is greatly expanded in this process. Previously, one could think of the text base --to put it crudely-- as a translation into "propositionaleese" of the sentences in the text. Now, the text base becomes a much richer structure than before. Not only does it contain the propositions directly derivable from the text, but also each of these propositions brings with it a number of other propositions that are closely connected to it in the general knowledge net. Thus, propositions are constructed just as before (e.g., van Dijk & Kintsch, 1983); yet where previously a single proposition was formed, a whole cluster is generated now.

Crucial in the present model is how this cluster of propositions is obtained: by a context-free process of activation of the closest neighbors of the original text-derived proposition in the general knowledge net. Of course, such a process will inevitably activate a lot of material that is irrelevant for any given context and, indeed, inconsistent with it. However, the price that has to be paid for promiscuity is not very high: The resulting text base is a connectionist net in which further spreading activation processes rapidly take care of inconsistencies and irrelevancies. What is gained by this dumb and seemingly wasteful process of random knowledge activation is flexibility and context sensitivity. The research on knowledge activation in psychology, as well as the experience with artificial intelligence systems, suggests that it is very difficult to activate knowledge intelligently. Prediction or expectation-based systems that use frames or scripts do not adapt easily to new contexts; prestructured knowledge hardly ever is exactly in the form that is needed. The construction-integration scheme proposed here may turn out to be more successful in this respect.

The general framework sketched earlier could be extended and elaborated in various ways as more experience with it is obtained. It might prove necessary, for instance, to resort to greater formalization in the propositional notation used here. However, until it becomes quite clear what the gains of greater formalization would be, a robust, easy-to-use system is to be preferred, even at the cost of some imprecision.

Perhaps more important might be elaborations of the knowledge-sampling mechanism. As presented here, each text-derived proposition activates its own strongest associates. It might be worthwhile to explore schemes whereby pairs or clusters of propositions activate their strongest joint associates.

Similarly, other criteria for stabilizing a network might be explored. For instance, networks might be made to maximize some statistic like harmony, as in Smolensky (1986). This might have considerable advantages. For instance, it is not always possible now to compare different networks in terms of how fast they reach equilibrium, because the number of cycles required depends strongly on the number of nodes in the network. In addition, at present there is no really satisfactory way to tell how good an equilibrium a process achieves. In the word arithmetic problems, all one can tell is whether the right hypothesis is more strongly activated than its competitors, but comparisons of the size of that difference across problems are problematic.

Constructive processes other than the ones explored here will need be considered. For word arithmetic problems, the most important constructions involved the arithmetic hypotheses. The construction of macropropositions could be neglected, mostly because the word problems were short ones and their macrostructure played no role in the problem-solving process.[8] For many other types of text, construction rules to form successive layers of abstractions and generalizations, as described by Turner et al. (1986), would be of primary interest. The macrostructure of a text could thus be made an integral part of a text base rather than a separate component, as it is presently treated.

Thus, there are a great many rules necessary to make the construction-integration model work for proposition building, assigning references and coreferences, bridging inferences, forming macrostructures, elaborating knowledge, and so on. Some of these construction rules are reasonably well worked out at this point, others are available within restricted domains, but many problems remain as yet unsolved. Thus, some of the same problems are encountered here as in conventional expectation-driven, top-down models of comprehension--but with one difference: Weaker, more general rules can be used here because these rules need not be fine-tuned to an ever-changing context. Whatever rules are still needed ought to be easier to work out within the construction-integration framework.

In van Dijk and Kintsch (1983), an important distinction was made between text bases and situation models. The former correspond to the propositional representation of a text, both at the level of the micro- and macrostructure. The latter correspond to a representation of the text that is integrated with other knowledge. Thus, in terms of the present model, the integrated text base --after irrelevant and inconsistent information has been deactivated and important knowledge elements have been absorbed-- is a kind of situation model. The qualifying phrase "a Kind of" is needed because text bases, integrated or not, are always propositional, whereas van Dijk and Kintsch specifically left open the possibility that situation models may be nonpropositional (e.g., Perrig & Kintsch, 1985). Situation models, under certain circumstances, may thus be like Johnson-Laird's (1983) mental models.[9]

The theory of knowledge use in discourse comprehension has been presented here at two levels: first, it is presented in terms of a general computational mechanism, at the level of what Pylyshyn (1985) called the "cognitive virtual machine"; and second, as a particular model that specifies how this mechanism is used in word identification in discourse and in understanding and solving word problems. The function of the model is primarily explanatory. Certain phenomena can now be interpreted within the framework of the model; for example, why a particular formulation of a word problem is especially hard or easy. Unlike less complex theories, however, there is no direct link between explanation and prediction in the present case. Unqualified experimental predictions are hard to come by in a model as complex as the present one. At

---

[8]Longer problems in which the macrostructure does play a role have been investigated by Dellarosa et al. (1988). Depending on whether a word problem establishes a theme of competition or cooperation between two protagonists, *compare or combine* problems will be solved most easily.

[9]Unlike the representation of the text itself--the text base, which is always propositional--situation models may have a different representation format, although this possibility was not considered in the present article. Both text bases and situation models are mental models of one kind or another in the sense of Gentner and Stevens (1983), though not necessarily in the more restrictive sense of Johnson-Laird (1983).

best, one might predict that a particular problem should be a difficult one, but that might mean several different things at the empirical level: that the solution fails, that a particular error occurs, that extra memory resources are required, that a reinstatement search will occur, that the problem must be read twice, and so forth. Even if we knew precisely what the "knowledge-use virtual machine" was like, our ability to make precise experimental predictions that are testable in conventional ways would still be severely limited. That, however, is not to say that such theories are without empirical consequences. Although we cannot predict particular events, predictions concerning classes of events are quite feasible (e.g., the different ways people might have trouble with word problems). Furthermore, our new-found understanding of why and how certain things happen can have important consequences for how certain texts are created in the first place or for instructional practices designed to help people with particular comprehension tasks.

Note. This paper, originally published in Psychological Review, was reproduced with the permission of the author.

## REFERENCES

Aaronson, D., & Scarborough, H.S. (1977). Performance theories for sentence coding:Some quantitative models. *Journal of Verbal Learning and Verbal Behavior, 16,* 277-303.

Anderson, J.R. (1980). Concepts, propositions, and achemata: What are the cognitive units? *Nebraska Symposium on Motivation, 28,* 121-162.

Anderson, J.R. (1983). *The architecture of cognition.* Cambridge: Harvard University Press.

Becker, C.A. (1976). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2* , 556-566.

Briars, D., & Larkin, J.H. (1984). An integrated model of skill in solving elementary word problems. *Cognition and Instruction, 1,* 245-296.

Broadbent, D.A. (1958). *Perception and communication.* New York: Pergamon Press.

Bruner, J.S., & Postman, L. (1949). Perception, cognition, and behavior, *Journal of Personality. 18,* 15-31.

Dellarosa, O. (1986). A computer simulation of children's arithmetic word problem solving. *Behavior Research Methods, Instruments, & Computers. 18,* 147-154.

Dellarosa, D., Kintsch, W., Reusser, K., & Weimer, R. (in press). The role of understanding in solving word problems. *Cognitive Psychology.*

Dik, S.C. (1980). *Studies in functional grammar.* London: Academic Press.

Fischler, I., & Goodman, G.O. (1978). Latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 455-470.

Fletcher, C.R. (1985). Understanding and solving arithmetic word problems: A computer simulation. *Behavior Research Methods. Instruments, and Computers, 17,* 565-571.

Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales & E.C.T. Walker (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North Holland.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension. Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178-210.

Frederiksen, J.R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes, 4,* 323-347.

Gentner, D., & Stevens, A.L. (1983). *Mental models.* Hillsdale, NJ: Erlbaum.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512-521.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development, 54,* 84-90.

Johnson-Laird, P.N. (1983). *Mental models.* Cambridge: Harvard University Press.

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

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

Kintsch, W. (1984). *Strategies for solving word arithmetic problems: Extensions of the Kintsch & Greeno Model* (Technical Report No. 133). Boulder: Institute of Cognitive Science, University of Colorado.

Kintsch, W. (1985). Text processing: A psychological model. In T.A. van Dijk (Ed.), *Handbook of discourse analysis* (Vol. 2, pp. 231-243). London: Academic Press.

Kintsch, W. (1987). Understanding word problems: Linguistic factors in problem solving. In M. Nagao (Ed.), *Language and artificial intelligence* (pp. 197-208). Amsterdam: North Holland.

Kintsch, W. (in press). The representation of knowledge and the use of knowledge in discourse comprehension. In C. Graumann & R. Dietrich (Ed.), *Language in the social context*. Amsterdam: Elsevier.

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

Kintsch, W., & Greeno, J.G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92,* 109-129.

Kintsch, W., & Mross, E.F. (1985). Context effects in word identification. *Journal of Memory and Language, 24,* 336-349.

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognition, 10,* 29-63.

McClelland, J.L. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science, 9,* 113-146.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 82-91.

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Miller, J.R., & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 335-354.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165-178.

Morton, J. (1979). Word recognition. In J. Moron & J.C. Marshall (Eds.), *Psycholinguistics, Series 2: Structures and processes* pp. 107-156). London: Paul Elek.

Mudersbach, K. (1982). Hol-Atomismus als Vereinheitlichung von Atomismus und Holismus. In *Proceedings of the 7th International Wittgenstein Symposium* (pp. 347-349). Wien: Hoelder-Pichler-Tempsky.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition, 22,* 93-136.

Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24*, 503-518.

Pylyshyn, Z.W. (1985). *Computation and cognition.* Cambridge: MIT Press.

Raaijmakers, J.G., & Shiffrin, R.M. (1981). A theory of probabilistic search of associative memory. *Psychological Review, 88*, 93-134.

Riley, M.S., Greeno, J.G., & Heller, J.I. (1983). Development of children's problem solving ability in arithmetic. In H.P. Ginsburg (Ed.), *The Development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Rumelhart, D.E., & McClelland, J.L. (1986). *Parallel distributed processing. Explorations in the microstructure of cognition. Vol. I: Foundations.* Cambridge: MIT Press.

Sanford, A.J., & Garrod, S.C. (1981). *Understanding written language.* New York: Wiley.

Schank, R.C. (1978). Predictive understanding. In R.N. Campbell & P.T. Smith (Eds.), *Recent advances in the psychology of language--Formal and experimental approaches* (pp. 91-101). New York: Plenum Press.

Schank, R.C. (1982). *Dynamic memory.* New York: Cambridge University Press.

Schank, R.C., & Abelson, R.P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., & Bienkowsky, M. (1982). Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489-537.

Sharkey, N.E. (1986). A model of knowledge-based expectations in text comprehension. In J.A. Galambos, R.P. Abelson, & J.B. Black (Eds.), *Knowledge structures* (pp. 49-70). Hillsdale, NJ: Erlbaum.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of Harmony Theory. In D.E. Rumelhart & J.L. McClelland, *Parallel distributed processing. Explorations in the microstructure of cognition. Vol. I: Foundations* (pp. 194-281). Cambridge: MIT Press.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (Re)Consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 523-534.

Till, R., Mross, E.F., & Kintsch, W. (in press). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition.*

Turner, A.A., McCutchen, D., & Kintsch, W. (1986). *The generation of macropropositions during comprehension.* Paper presented at the meeting of the Psychonomic Society, New Orleans.

van Dijk, T.A. (1980). *Macrostructures.* Hillsdale, NJ: Erlbaum.

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

Waltz, D.L., & Pollack, J.B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science, 9,* 51-74.

Wilensky, R. (1986). *Some problems and proposals for knowledge representation* (Tech. Rep. No. UCB/CSD 86/294). Berkeley: Computer Science Division, University of California.

Willoughby, S.S., Bereiter, C., Hilton, P., & Rubinstein, J.H. (1981). *Heal math: Thinking stories book. Level 2.* La Salle, IL: Open Court.

Winograd, T. (1983). *Language as a cognitive process. Volume I: Syntax.* Reading, MA: Addison-Wesley.

Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151-175.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# MENTAL MODELS AND ACQUISITION OF KNOWLEDGE FROM TEXT: REPRESENTATION AND ACQUISITION OF FUNCTIONAL SYSTEMS

Serge BAUDET and Guy DENHIERE

U.R.A. au C.N.R.S. D 1297 :
"Psychologie cognitive du traitement
de l'information symbolique"
Equipe TEXTIMA
Université de Paris VIII
2, rue de la Liberté
93526 St. Denis, France

Taken as a whole, research on the cognitive processing of linguistic information (particularly textual information) leads to the observation that textual structures and specificially linguistic representations are insufficient in accounting for the comprehension and memorization of texts, and, more importantly, in the acquisition of knowledge from texts (Denhière and Baudet, 1988). Results from this research require reconsideration of the hypothesis that cognitive structures put into play by the individual when processing texts are a projection --the psychological equivalent of --textual structures. In processing a text, the individual invokes not only linguistic knowledge --representations which are specific to the structure of the text (from predicative structure to textual superstructure) -- but also knowledge of the domain that the text represents and communicative context.

Taking into account the interaction between the characteristics of the system of communication --conceived as a system of transformation of information put into play by individuals, their linguistic and general knowledge, and their relations, as well as contextual constraints -- leads us to consider that at the end of the processing of a text, the individual has constructed several types of representations of information born by the text, which are more or less distant from the surface of the text and which are more or less privileged as a function of the demands of the task.

In describing these representations, an important step has been realized with the introduction of the notion of mental models for conceptualizing the representation of the world that the individual constructs for himself from what he has experienced and learned, and that they are activated during the reading of texts. The notion of model, however, is of interest only if it is rigorously defined and if the descriptions of the world and the representations produced by the individual are precise.

The goal of this article is to present, for a large class of domains, a conceptual framework allowing such descriptions: a formalization in terms of systems or relationals and/or transformationals, or teleologics (functionals or intentionals) proposed for describing the representations which we produce of a complex domain -- involving several individuals or objects, their characteristics and their relations -- and our representation of the representation that is constructed by the learner. This analysis in terms of systems is applied to the description of a technically complex object: an internal combustion engine. Finally, experimental results are presented which are compatible with the hypothesis of a gradual construction by the learner of a representation organized in terms of systems.

## 1. KNOWLEDGE

### 1. 1. Main characteristics.

By knowledge we mean a stable mental representation of a possible world. This representation can be modeled as a semantic network of propositions or by a coherent set of propositions. By proposition we mean a representation of a state, an event, or an action in a possible world, as well as of an interpropositional relation. The concept of coherence refers here to the property of a representation to create an entity which exists due to characteristics of its structure: relations link propositions which are organized hierarchically. Drawing on the research relevant to this topic in the various disciplines within the cognitive science, we attribute certain characteristics to this representation which satisfy our conceptual requirements:

1. Conceptually, the statement "x knows that p" can be analyzed (Itkonen, 1983) as "x thinks that p", "x has good reasons to believe that p" and "p is true". The first characteristic of knowledge is that it possesses a truth value which distinguishes it from beliefs. Representation has a truth value if it satisfies certain conditions of truth such as can be guaranteed by the experimental methodology. It can be assigned a truth value : true or false. Three criteria are generally put forward in epistemology to characterize the truth of a belief (Puntel, 1978):

- correspondence: p is true if it corresponds to an extralinguistic entity in the world,
- coherence: p is identified by its location in a pattern of beliefs
- consensus: p is accepted by a community of individuals.

The two latter criteria lead us to retain as knowledge those beliefs that are empirically non falsifiable, that represent norms. For example, a grammatical rule which attempts to describe a norm is not falsified because it is not applied in a corpus. If a speaker/writer does not respect this rule, it is the person who is mistaken, and not the description of the norm. In this case, which is concerned with knowledge of actions, it is preferable to speak of the rational

aspect rather than the truth value (Itkonen,1983). Knowledge2 has a rational character if it represents a rationally justified belief, that is, if it conforms to a norm inferred from a body of accepted beliefs. It can also be assigned a value of conformity to a rational norm : right or wrong.
Most frequently, knowledge is the representation of the real world and refers, therefore, to individuals and objects in this world as well as relationships between individuals and objects. It can also be the representation of a hypothetical world which conforms to a rational norm.

2. "Knowledge is explicit in an individual's head" (Le Ny, 1985a) and is associated with a high confidence.

3. Its availability and accessibility in memory as a permanent semantic structure (Ehrlich, 1976) and especially its durable nature, permit us to characterize it as a type representation (see Le Ny, 1985b). This distinguishes it from a token representation which is defined as an "event, of a limited length of time, which is supposed to occur briefly and as transitory in the cognitive activity of the subjects" (Le Ny, 1985b). This accessibility is a fundamental functional characteristic of knowledge. Thus, knowledge must intervene in all phases of processing of new information for the following purposes:

- providing a database which, along with new information, will serve as input to the processing,
- permitting a categorization of new information, its structuration and integration into an individual's prior knowledge.

4. Finally, knowledge is characterized as a situation model, that we conceive as a representation which is a configuration with a structure corresponding either to a part of the universe, or to the structure of a rational theory.

The characteristics of knowledge which we put forth have been determined primarily on the basis of what we know of the human cognitive information processing, and the conceptual requirements needed to study this cognitive processing. This excludes from our research a set of fundamental questions and leads us to define these concepts sometimes differently than would be done in other disciplines of cognitive science. Thus, we propose that knowledge is associated with a high confidence. We have chosen to use the term "confidence" instead of "certitude" because an epistemological knowledge is fundamentally uncertain. For us, confidence is information associated with a cognitive representation and is revealed through observable behavior such as explicit judgements, resistance to change, etc.

It is possible that for some psychologists certain characteristics that are retained are not pertinent. This is probably the case for the truth or rational character of knowledge. However, both practical and theoretical reasons lead us to believe that it is important for us to have the means to maintain the distinction between knowledge and belief. First, we are interested in problems of learning : that is, the acquisition of information that we might require to be true or at least rational. Then, we presume that knowledge, as a result of the social construction, has the capacity to integrate into the highly structured sets of theories. In terms of memory representation, this could be translated as the

---

[2] "practical knowledge" of Putman (1978).

ability to integrate knowledge into more coherent structures —thus making it more available and easier to access— than those which integrate beliefs. This distinction appears even more important when we consider cognitive processing. For an individual, the difference in meaning an individual establishes between knowledge —which is a representation that is considered true— and belief —which is a representation that cannot be estimated in terms of its truth value— is pertinent in the sense that it affects the behavior that involves such representations. The same can be said about the certainty or doubt that is associated with beliefs and knowledge once the individual has been informed that it is a question of beliefs or knowledge. Moreover, we hypothesize that the three criteria epistemologists have proposed for characterizing the truth of a belief —correspondence, coherence and consensus— are cognitively relevant in the sense that they represent meaningful elements added to the meaning of the belief when it is considered true. Thus, knowledge of the world will be considered as referring to a certain entity or relation of this world; the structure that the learner will assign to it will thus correspond to that of the real world, or at least to the experience of that reality. It is this hypothesis which principally legitimates for us the use of concepts such as mental models.

Knowledge is assumed to be organized at various levels in the semantic memory of individuals : from the semantic features ("sèmes" in French) to the general schemata of problem solving.
Two broad types of knowledge organization : schematic and taxonomic are generally considered in the literature (see Mandler, 1979). The former can be described by a predicative type of structure with empty slots that are likely to be filled by appropriate arguments. The latter corresponds to a categorical organization involving superordinate relations.

### 1. 2. Knowledge acquisition from texts

By knowledge acquisition we mean, in the most general sense, the construction of a stable mental representation of a possible world. This representation is stored and maintained in memory and can be retrieved, with a certain likelihood, and used later. Cognitive research on text processing has demonstrated that a text-based approach or an approach based solely on linguistic knowledge is insufficient (Denhière and Baudet, 1988). In view of this research, it is necessary to re-evaluate the cognitive invariants. Experimental results show that when people process texts, they draw on type-representations of the text structure, as well as on their knowledge of both the domain of the text and the communicative situation. This redefinition of the cognitive invariants accompanies a theoretical re-evaluation of the text invariants. These exist only as relations that can be understood and interpreted by the receiver in a communication context. Their invariance is a result of invariances of the communication system set in motion to communicate a representation of the world. Taking into account the interaction among the characteristics of the communication system which is conceived as a system of information transformation involving individuals, their linguistic and world

knowledge, and their relations, as well as the constraints on the situation, it leads us to believe that after processing a text, the comprehender has constructed cognitive representations that are not thought of as text representations but rather of what is being expressed in the text. In other words, when a learner acquires knowledge from a text, the cognitive representations generated are not those of the text but rather of the objects, states, events, actions and relations represented in the text. These representations are built with information from two sources: the text and the knowledge and beliefs of the learner about the possible world represented in the text. Knowledge acquisition from a text occurs basically by means of gradual attachment during text processing of what is being said in the text to the representation already built in memory.
An important step has been realized in the description of such representations by the introduction of the concept of mental models or situation models for representing the representation of the world that individuals build through their experiences, and that they use during the reading of a text.

### 1. 3. Mental models and situation models

A number of authors have used the notion of mental model (Garnham, 1981; Gentner, 1983; Johnson-Laird, 1980, 1983; Norman, 1983) or situation model (van Dijk and Kintsch, 1983; Kintsch, 1986; Perrig and Kintsch, 1985; Mannes, 1988; Mannes and Kintsch, 1987) to conceptualize the representation of the world that learners build up from experience and learning, and which they activate during the reading of text. Van Dijk and Kintsch (1983) for example define a situation model as:

> "a cognitive representation of the events, actions, persons, and in general the situation, a text is about.... A situation model may incorporate previous experiences,.and hence also previous textbases, regarding the same or similar situtations. At the same time, the model may incorporate instantiations from semantic memory about such situations" (pp. 11-12).

To the extent that these models of the world are broader and richer than the information contained in the text and that they are used to process the text information, the representations of the world will obviously have a pronounced effect on learners' performance.

The psychological validation of these models consists of either keeping the formal and semantic features of the sentences constant while varying the distance from the underlying mental models (Garnham, 1981) or by modifying the information provided prior to reading the text so as to trigger preferential recovery of situation models of varying compatibility with the text, and then testing their effect on text processing (Mannes and Kintsch, 1987). The results of these experiments are compatible with the assumption that mental models or situation models affect comprehension and memorization of text.

These validations were carried out before the concept of a mental or situation model had been rigorously defined. In addition, neither the domains referred to by these models (states, events, relations, etc. ) nor the learner's representations of the domains, have been precisely described. Clearly, formal definitions of both the domain and the representations of such domains by people are necessary if we are to study comprehension. More importantly, if we are to examine an individual's learning from the standpoint of assessing initial knowledge, establishing goals for learning, developing suitable written and pictorial materials or constructing individualized computer learning aid systems, then a precise and more detailed description is essential.

### 1. 4. Systems

Our objective of providing help during learning and our theoretical choices constrain us in formulating, in addition to a representation of the cognitive functioning of the learner, a detailed representation of the domain to be learned and its cognitive representation for the learner. In order to describe the representation of a complex domain —involving many individuals or objects, their characteristics and their relationships— as well as our representation of the representation the learner constructs, we have adopted a formalization in terms of systems. The focus on learning prompted us to formalize three types of systems:

- A relational system represents complex stative situations defined as states of individuals or objects of a possible world.
- A transformational system represents complex events defined as transformation sequences of stative situations.
- A teleological system represents structures and behaviors of organisms, defined as functional units. When those units —human beings and by extension, animals— are credited with intentionality, they will be considered to be intentional systems; when they are components of a technical or biological device they will be referred to as functional systems. Complex technical objects (e.g., machines, engines, information processing systems) and production systems (e.g., blast furnaces, rolling mills) are exemplars of the functional system.

The model so far described has four main features :

- it draws on concepts from the philosophy of action related to causality (von Wright, 1967; Mackie, 1974);
- it integrates knowledge of human cognitive functioning, in particular knowledge related to the construction of representations of states, events, actions, stative relations, naïve causality, as well as knowledge on constructing the local and global coherence of a representation (Denhière and Baudet, 1988; see Trabasso's chapter in this book);
- it uses formal constructs developed within artificial intelligence such as hierarchical network analysis, procedural rules, etc. to guarantee a level of rigor in description and application;
- it has numerous potential applications and can serve as an example for constructing individualized learning support systems and can aid in detecting, understanding and correcting malfunctions in the functional systems.

As an example, we will present in the following section the theoretical description of a functional system.

## 2. THEORETICAL DESCRIPTION OF A FUNCTIONAL SYSTEM

2. 1. Defined as intensional, a functional system is:

- relational, that is, a collection of objects affected by the definition of one or several relations between these objects. Formally, a relational system is a sequence <I, R1, ..., Rn> where I is a non-null set and R1, ..., Rn are relations that apply to elements of I (see Coombs, Dawes and Tversky, 1975).
- transformational, that is, in an interval of time t : i, j, it is characterized by modifications in the "normal course of events" of successive states of the system either to remain as is (conservation of states) or to change (events) (von Wright, 1967; François, 1985).
- teleological, that is to say, it is a functional unit. In an interval of time t : i, j, the initial state of the system, defined by the objects present, their relationships and the initial values of their attributes, is modified so as to attain a configuration which constitutes the target goal of the system (final state). Each change occurring during the interval t : i, j, functions as a means of achieving this goal, that is, it creates the necessary conditions, in the circumstances (see Mackie, 1974), to reach this goal.

Complex technical objects (e.g., machines, engines, information processing systems) and production systems (e.g., blast furnaces, rolling mills) are exemplars of the functional system.

2. 2. Description of three aspects of a functional system : relational, transformational, and teleological.

The behavior of a functional system, that is, the sequence of transformations that characterize it, is described by a teleological sequence of events and states that follow one another in chronological order, are causally related and goal-oriented

2. 2. 1. States and events.

The description of a system state (ST(s)) can be represented by:

$$\boxed{ST(s) \ =\hat{}= \ <Wi,ti>}$$

where, =^= refers to the semantic relation "is represented by"

**Wi** refers to a set of propositions that represent a configuration of objects, their attributes and values, as well as stative relations among them (spatial locations, component, membership, inclusion, comparison, etc.)

ti refers to a temporal mark (instant or interval).

An event is defined as a change in the state of the system. A change is a process that causes the system to go from an initial state (STi) to a final state (STf), where both STi and STf are different states of a possible world:
---->, arrow indicates the directionality

| STi -------- change -------> STf |
|---|

An event can be described :
a) either in terms of the difference between two stative descriptions of the system:

< Wi, ti >, < Wj, tj >

These differences can be due to:

- the appearance or disappearance of an object,
- the modification of the value of an attribute belonging to an object,
- the modification of a relation between objects.

These modifications take place between time **ti** and **tj**.

b) or by a list of propositions representing :
(i) the event proper,
(ii) and that which is not modified but pertinent for the representation of the sequence of events.
(i) The event proper can be :
- either a modification of the value of an attribute belonging to an object of the system (Le Ny, 1979) :

| EV(O) ------> CHANGE (O, on A, of ( ti, am ) at ( tj, an )) |
|---|

where,
**EV** : event predicate
**O** : its argument
**A** : the attribute being modified
**ti,tj** : two successive moments in time
**am, an** : two values of the attribute **A**

- or a modification of the relation between objects,
- or the appearance or disappearance of an object in the system.

(ii) This is not modified in the system but is pertinent for the representation of the sequence of events (see the concept of causal field, 2.2.2).

These propositions represent:
- either a non-modified state of an object in the system (Le Ny, 1979),

| ST (O) ----> on A (ti, ai) (O) ) |
|---|

the object **O** is in a state **ST**, if at the time **ti**, this object takes the value **ai** for the attribute **A**.
- or a non-modified relation among objects in the system,
- or the existence of an object in the system.

At a certain level of description, many events of the system (successive or simultaneous) are conceived of and described as a single event (compound or complex). The level of description adopted is constrained by :

- the level of analysis which is determined by the information available on the states of the system as well as by the purpose of the analysis: the goal of the description and the assessment of knowledge available for whom the description is carried out; and the language and the conventional means used to describe these events.

Consequently, the identification of an event in the description will be based on the system modifications as well as on the cognitive and conventional units.

2. 2. 2. Causal relations

Causal connections are prominent in transformational and teleological systems. Therefore, a precise description and conceptualization of these relations are essential. The present model draws on the philosophy of action (von Wright, 1967; Mackie, 1974) while taking into account knowledge about the construction of the representation of naïve causality.

*Interventionist concept of causality* .
The occurrence of a causal relation between events or states instantiates the concept of causation. This concept presupposes the concept of intervention into "the normal course of events" (von Wright, 1971) which brings about a change in the natural state of affairs by allowing a state either to remain stable, or to change. Two important consequences stem from this interventionist concept of causality :

- the notion of condition is an essential component of the concept of causality;
- causation is dependant upon a certain context or circumstances; it is a modification in the normal course of events, or in other words, the creation of an abnormal condition (Hart and Honoré, 1959).

*General concept of causality* .
We borrow the interpretation of the general concept of causality from Mackie (1974), as below :

| |
|---|
| X Causes Y : **the occurrence of X is a sufficient and necessary condition, in the circumstances, for the occurrence of Y .** |

"X is a necessary condition for Y" means: "occurrence of X and occurrence of Y, and, in the circumstances, non occurrence of Y if non occurrence of X"

"X is a sufficient condition for Y" means: "non occurrence of X and non occurrence of Y, and, in the circumstances, occurrence of Y if occurrence of X".

*Naïve concept of causality*

The aim of the model is to construct a representation (a psychologist's model) of the cognitive representation that the learner has of causality (in the real world or presented through language or any other medium). The focus is on the conception of "naïve" causality which is the product of cognitive processing used by a non-expert in analytical philosophy of information available on the sequence of events in the world. This information comes from two sources : memory and the individual's environment, such as the world or a material representation of a possible world, e. g, a text.

Models put forward in modern analytical philosophy need to be reinterpreted in the light of what is known about cognitive processing of causality. Recent experimental work (Hilton and Slugoski, 1986) is in line with the hypothesis that the cognitive representation of causal relations is built up via a causal explanatory operation (a search through a set of events for the cause of a given event or state) which uses :

- counterfactual reasoning based on the criterion of necessary condition (given X and Y, can there be Y if there is not X?); and
- judgment of normality of events in the situation.

Thus, the naïve concept of causality which will be activated by the occurrence of a causal connection in language or in the world, can be interpreted as follows (Mackie, 1974) :

| |
|---|
| **X Causes Y : the occurrence of X is a necessary condition, in the circumstances, for the occurrence of Y.** |

*Circumstances and the causal field*

Causal relations are built with respect to a specific context: the circumstances. Thus, propositions which represent them are constructed in relation to a representation of this context (the "cognitive background", Mackie, 1974): knowledge of the normal course of events. A certain amount of this knowledge covers the causal field, that is, the set of events and states which do not appear as modifications of the normal course of events but which counterfactual reasoning sets up as conditions for the manifestation of this effect.

The cause and the effect represent modifications introduced into the normal course of events: Cause is an INUS condition (Insufficient but Necessary part of an Unnecessary but Sufficient condition), a modification introduced into the causal field which, in the absence of any other modification, is a necessary condition for the effect to occur (Mackie, 1974). Events or states which are part of the causal field are not interpreted as being the cause or the effect. They are considered to be a "supplementary condition" to causation.

*The description, within a system, of the occurrence of a causal relation comprises three elements :*

- the description of a causal field,
- the INUS condition: the cause, and
- the effect.

The choice of the causal field is determined by the individual's knowledge of the "normal course of events". The description of the causal field is determined by the goal assigned to this description. Thus, if our goal is to train thermal internal combustion engine experts, the selection of the causal field elements and their level of description will be made as a function of the type of expertise that is sought. If the goal is to train repair men, the elements to be described will be selected in terms of their probability of accidental modification (e.g., mechanical breakdown) and their level of description will be in terms of the intervention possibilities of the repair man (parts that can be repaired vs. parts that can only be replaced). If the goal is to train design engineers, then other selection criteria for the elements and other description levels of these elements will be required. Furthermore, it will not be possible to consider some of these elements as legitimate descriptions of the causal field. They will have to be causally explained in order to be modified eventually during the design. It is therefore essential that the chosen conceptualization allows adapted descriptions. In a system, each occurrence of a causal relation modifies the circumstances for the occurrence of the next modification. After attempting to causally interpret each modification pair, an individual must build in memory, in order to understand the system's functioning, a representation of the new normal course of events, and therefore of a new causal field against which the new event might be causally interpreted. Thus, when an event or a state appears as a modification of the normal course of events at a certain point in the causal path, it will be possible to consider it as a legitimate element of the causal field in a subsequent step.

### 2.2.3. Teleological description of a system

Causal relations linking events of a functional system are essentially part of the causal domain of the physical world. We distinguish this type of causality which characterizes relations among agentless events, from intentional causality in transformations performed by agents. Nevertheless, both descriptions have a teleological nature since they explain the cause by the effect. Consequently, in a functional system description, events are considered as means to attain the assigned goal.

A teleological description of a system can be expressed as follows (see Mackie, 1974) :

**I has C in E because C in E leads to B**

in which

- **I** = individual or set of individuals,
- **C** = behavior of a system described by the sequence of states and events,
- **E** = environment or circumstances,
- **B** = goal of system (final state aimed at by the system designer).

This type of description is feasible since, by construction, the behavior of the system is driven by the final goal for which the system was designed. In addition, this type of description is necessary to account for the selection, by the designer, within a possible world, of events considered as causal conditions.

### 2. 2 .4. Decomposition of a system into sub-systems

The description of a functional system can be simple or complex. The degree of complexity is defined by the number of objects that the system is composed of as well as the number of branches and the length of the causal path. Obviously, the complexity of the description is a function of the complexity of the system. But, more fundamentally, it is determined by the level of analysis of reality taken by the person carrying out the description (for a discussion of the level of analysis, see Le Ny,1985c). Specifically it is determined by :

- the unit retained, that is the material object resulting from the analysis of reality that was retained as an object of the system: a workshop, a machine, a part of the machine, ..., an atom, etc., and
- the number of events causally related that are identified in its description.

At the same level of analysis, in the case of a complex description, the system must be decomposed into sub-systems, whose description is more elementary. The teleological nature of the functional system, that is, its hierarchically structured organization into goals/sub-goals makes this decomposition possible. This decomposition reflects the structure of reality : the functional system is tangible evidence of problem solving activity implemented by the designer.

Each sub-system, like the system itself, makes up a functional unit. It is characterized by its role in the general functioning of the system: the final state of the sub-system serves as a sub-goal of the system. The final event leading to this final state is the cause, the necessary condition in the circumstances, of a modification of the adjacent (sub-)system which is immediately superordinate or adjacent to the same hierarchical level.

The system is thus analyzed as a set of hierarchically organized functional units: a tree structure. These units are related causally, temporally and topologically. The hierarchy is defined by the organization of goals and sub-goals.

A tree structure representation was adopted for the teleological description of a complex functional system. The superordinate node represents the main goal of the system. The subordinate nodes represent sub-goals of the system, which must be attained for the main goal to be accomplished. These subordinate nodes are the macro-events of the system. The occurrence of these macro-events is itself conditioned by the occurrence of events represented by the nodes immediately below. Links among nodes of the same level represent conditional relations.

The construction of a structure of this type thus consists of a categorization and a hierarchical organization of events as a function of the goals assigned to the system, which may or may not violate the temporal sequence. This makes it possible to render functioning systems in text form. The goal structure ensures the global coherence of the representation of the system.

### 2.2.5. Internal and external causality relations

In the analysis of systems into sub-systems, internal causality must be distinguished from external causality. Causal relations are internal when they pertain to states and events of the functional unit under consideration, and are external when they link these states and events to those of the adjacent system. The state or event of the (sub-)system which is a condition of the modification of the state of the adjacent (sub-)system can be considered as an output of its own (sub-)system and as an input (an external cause) to the adjacent (sub-)system.

### 2.2.6. Degree of decomposition of the system

The degree to which a system can be decomposed into sub-systems is constrained by the necessity that the sub-system has to be a functional system, and thus conform to the definition of this type of system (see 2.1.). But the depth of the goal/sub-goal structure is chosen by the individual carrying out the description, and who therefore, has some control over the degree of decomposition. However, the necessity of describing a causal path will dictate a certain type of decomposition. Given that the modifications of the system have been identified, and that the causal path from the initial to the final state of the system has been established, for a certain level of analysis we will have the following :

a) the transformation sequence which forms a direct branch to the path that goes from the initial to the final state will be considered as a characteristic of the behavior of a sub-system directly subordinate (level 2 sub-system), and

b) the sequence that branches from this other branches will be considered as a characteristic of the behavior of a sub-system subordinated to the previous one (level 3 sub-system), etc.

Each modification sequence that will branch off another will therefore dictate the description of a sub-system. Within this framework, it is interesting to distinguish a particular type of functional system: the production system. Its behavior is finalized by a goal : to assign a set of characteristics to one of its objects: "la matière d'oeuvre". Priority in determining the degree of decomposition of the complex system into sub-systems, will be given to the description of the transformational path of the "la matière d'oeuvre". Each of the transformations will be described as an event caused by the output of a sub-system.

## 3. ANALYSIS OF THE STARTER SYSTEM (WITH A POSITIVE ELECTROMAGNETIC SWITCH) OF THE THERMAL INTERNAL COMBUSTION ENGINE.

This analysis was developed in the context of student learning in the subject of automotive mechanics at the "Lycée Professionnel Jean-Pierre Timbaud d'Aubervilliers" (Baudet, Blaizet, Denhière, 1987).

### 3. 1. Description of the actions, events and states of the system

Events are described in terms of differences between two stative descriptions (see 2.2.1.). In those descriptions we take into account only the attribute-value couples and the relations which either undergo modification or are pertinent to the description of the causal field. Table 1 below presents the descriptions of two events in the starter system. The underlined elements signify those which are modified.

**TABLE 1 : DESCRIPTION OF THE EVENTS 1 & 2 OF THE GEARING SYSTEM**

**EVENT 1 : CIRCULATION OF ELECTRIC CURRENT i IN THE CABLE BETWEEN THE ANTI-THEFT SWITCH AND THE ELECTROMAGNETIC CONTACTOR'S COIL.**

| INITIAL STATE | ATTRIBUTE | VALUE | FINAL STATE | ATTRIBUTE | VALUE |
|---|---|---|---|---|---|
| **INDIVIDUALS** | | | | | |
| XXXXXXXXXXXXXXXXXXXXXXXX | | | Current | Intensity | i |
| Cable | Section | Small | Cable | Section | Small |
| Anti-theft switch | Position | "Start" | Anti-theft switch | Position | "Start" |
| **RELATIONS :** | | | | | |
| **LOCALISATION :** | | | | | |
| XXXXXXXXXXXXXXXXXXXXXXXX | | | Inside (Current, Cable) | | |
| Between (Cable, Anti-theft switch, electromagnetic contactor's coil) | | | Between (Cable, Anti-theft switch, electromagnetic contactor's coil) | | |

EVENT 1 CAUSE EVENT 2

**EVENT 2 : CIRCULATION OF CURRENT i IN THE ELECTROMAGNETIC CONTACTOR'S COIL**

| INITIAL STATE | ATTRIBUTE | VALUE | FINAL STATE | ATTRIBUTE | VALUE |
|---|---|---|---|---|---|
| **INDIVIDUALS** | | | | | |
| Attraction coil | Number of rings | x | Attraction coil | Number of rings | x |
| Attraction | Diameter of wire | y + 1 | Attraction coil | Diameter | y+1 |
| Maintenance coil | Number of rings | x + 1 | Maintenance coil | Number of rings | x +1 |
| Maintenance coil | Diameter of wire | y | Maintenance coil | Diameter of wire | y |
| Electromagnetic contactor.......................... | | | .Electromagnetic contactor | | |
| ........XXXXXXXXXXXXXXXXXXXXXXX | | | Current | Intensity | i |

RELATIONS:

| | |
|---|---|
| LOCALISATION : | |
| XXXXXXXXXXXXXXXXXXXXXXX. | Inside (Current, coil's wire) |
| Inside (Attraction coil, Maintenance coil) | Inside (Attraction coil, Maintenance coil) |
| PART OF : | |
| (Attraction coil,Electromagnetic contactor) | (Attraction coil,Electromagnetic contactor) |
| (Maintenance coil, Electromagnetic contactor) | (Maintenance coil, Electromagnetic contactor) |
| COMPARISON | |
| x+1 > x | x+1 > x |
| y+1 > y | y+1 > y |

EVENT 2 **CAUSE** EVENT 3.

3. 2. Graphic description of the functioning of the starter system (with a positive electromagnetic switch) in terms of a causal path

Figure 1 below presents the sequence of actions, events and states which characterize the functioning of the first sub-system of the stater system, that is, the gearing sub-system. Each action, event, and state which is identified by a numbered label in figure 1 refers to an entry in table 2.

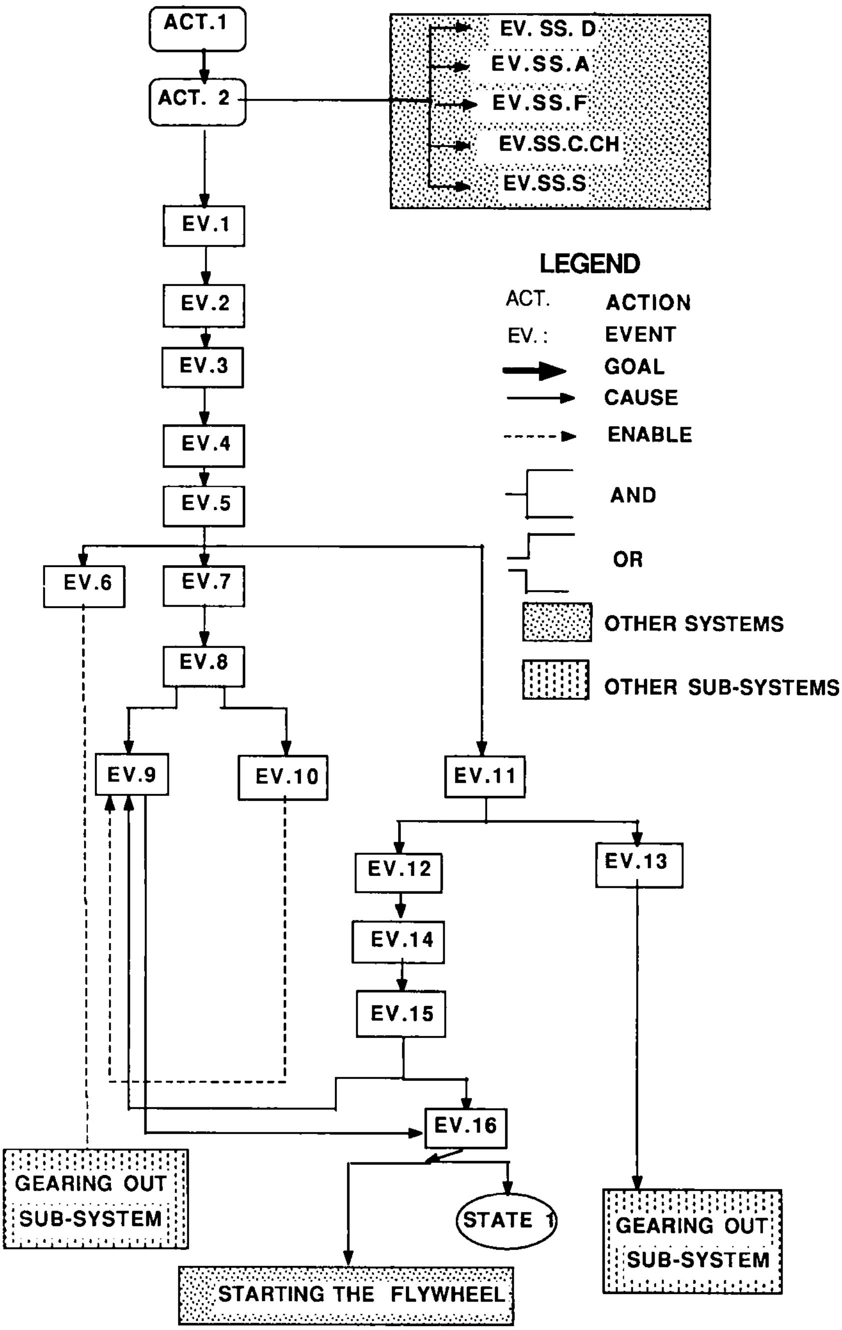



FIGURE 1 : GRAPHIC DESCRIPTION, IN TERMS OF CAUSAL PATH, OF THE FUNCTIONING OF THE GEARING SUB-SYSTEM OF THE STARTER SYSTEM.

In the description of the relation among actions, events and states, only the CAUSAL attribute has been retained. Other attributes exist and may be added as a function of the goal of learning. The legend for Figure 1 is as follows :

- the lines separated by hyphens mark the boundaries of the system.
- the double arrows indicate the relation GOAL (intentional causality).
- the single arrows indicate the relation CAUSE (INUS. condition).
- the dotted lines indicate the relation ENABLING (relation between an element of the causal field and the event or resulting state).
- the double dotted arrow indicates the relation TO CAUSE A MENTAL EVENT (intention) WHICH CAUSES AN ACT (see van Dijk, 1977; Searle, 1983).

This sequence of actions, events and states provides a chronicle of how the system works. The explanation of the relations among the states, events and actions ensures the local coherence of the representation of either the verbal or the graphic-text or of the cognitive representation.

**Table 2 : List of actions, events, and states of the gearing sub-system.**

| | |
|---|---|
| ACTION1 ......... | The ignition key is inserted in the anti-theft switch. |
| ACTION2 ......... | Turn the ignition key to the "start" position. |
| EVENT 1 ......... | Circulation of electric current in the cable which goes from the anti- theft switch to the electromagnetic contactor's coil. |
| EVENT 2 ......... | Circulation of electric current in the electromagnetic contactor's coils. |
| EVENT 3 ......... | Generation of a electromagnetic field around the electromagnetic contactor's coil. |
| EVENT 4 ......... | Magnetization of the magnetic core. |
| EVENT 5 ......... | The rectilinear movement of the plunger core. |
| EVENT 6 ......... | Compression of the return spring of the plunger core. |
| EVENT 7 ......... | The tilting of the lever of the positive switch. |
| EVENT 8 ......... | The moving of the gearing system of the starter's pinion. |
| EVENT 9 ......... | The coming into gear of the starter's pinion. |
| EVENT 10 ........ | The thrust starter's pinion. |
| EVENT 11 ........ | The switching off of the electromagnetic contactor's contacts. |
| EVENT 12 ........ | Circulation of electric current from the battery to the cable of the starter's electric motor. |
| EVENT 13 ........ | Deactivation of the attraction coil of the electromagnetic contactor. |
| EVENT 14 ........ | The starting of the starter's electric motor. |
| EVENT 15 ........ | Rotation of the starter's pinion. |
| EVENT 16 ........ | Generation of a force which puts in contact the starter's pinion and the dented crown of the flywheel. |
| STATE 1 ......... | The pinion's dents are facing in between the dents of the dented crown of the flywheel. |
| STATE 2 ......... | The pinion's dents are facing the dents of the dented crown of the flywheel. |
| STATE 3 ......... | Maintenance in the gearing position of the starter's pinion. |

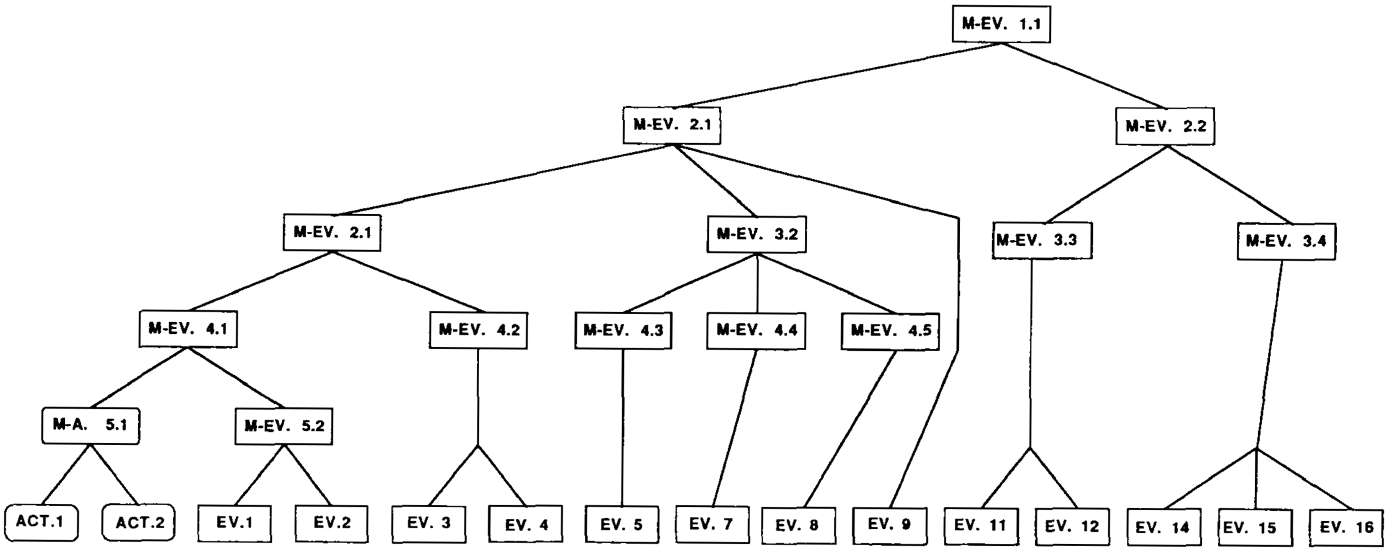



FIGURE 2 : GRAPHIC REPRESENTATION OF THE GOAL-STRUCTURE OF THE GEARING SUB-SYSTEM OF THE STARTER SYSTEM

LEGEND : M-EV = MACRO-EVENT; M-A = MACF; ACT = ACTION; EV = EVENT.

### 3. 3. Teleological description of the functioning of the system

Figure 2 describes graphically the resulting structure in terms of goals/sub-goals of the sequence of events presented in 3.1.2.
This structure corresponds to the achievement of the target goal of the gearing sub-system of the starter system: the creation of an event. This event is an external cause of a modification of the superordinate system (the thermal internal combustion engine) which consists of starting the flywheel.
The nodes subordinate to the original node represent the sub-goals of the starter system whose attainment is conditional for the realization of the main goal of the system. They correspond to the macro-events of the system. The occurrence of these macro-events is itself conditional on the events represented by the nodes which are immediately subordinate. The degree of decomposition of macro-events into elementary component events is determined by the level of description adopted. The construction of such a structure therefore consists of a categorization and a hierarchical organization of the events with respect to the goals of the system which might not respect the chronological order of these events. This ensures the global semantic coherence of the cognitive representation and of the text.

### 3. 4. Construction of explanatory texts

It follows that we can construct two types of texts explaining how a functional system works. The two types differ mainly with respect to their features of semantic coherence. Therefore from the perspective of previous research on text comprehension and memorization, one can say that the forms of semantic coherence will determine the characteristics of the mental representations which the learners elaborate while reading these texts. The two types of texts will induce two types of representation which adapt more or less readily to other uses. If in reality both are required by the expert, the learner must build these representations progressively and know how to access and apply them differentially depending on the nature of the task.

## 4. CONSTRUCTION OF AN INDIVIDUALIZED SYSTEM TO ASSIST IN LEARNING

Poitrenaud (1987) has designed a system of questioning and information presentation called **QASH : Questioning and the Acquisition of Hierarchical Structures**, implemented on the Macintosh microcomputer. This program allows us to model our representation of domains in terms of systems and to test our hypotheses about the representations and operations involved in learning. Using this educational software, QASH allows us to evaluate the initial state of knowledge of the learner in order to present the information necessary to attain the desired final state (see Tapiero, Poitrenaud and Denhière, 1988).

This program consists of four files which contain : the presentation of the questions (verbal or graphic) and the possible responses; the linearisation of the question tree in the form of attribute-value couples; the scoring grid for the answers; and the texts --both verbal or graphic--associated with the questions and answers of the learner. The use of operators (Identity, Equality of sets, Inclusion) permits us to identify four types of responses : true, true by default, true by excess, and false. The automatic management of the questions and the presentation of subsequent information depend on the results of this analysis.

### 4. 1. Principles of question construction

The construction of the questionnaire and its operation in QASH are regulated as described in section 3.1.The goal structure allows one to present questions at different levels of the hierarchy of the representation of the functional system. The precision of the responses to questions at one level determines the presentation of subsequent questions at the next level which is immediately subordinate. The questions correspond to the three aspects —relational, transformational and teleological— of the whole functional system.
For the relational aspect the questions pertain to the existence of objects and the pertinent attribute-value couples (that is to say, those whose value will be modified or which are necessary for the description of the causal field), and the stative relations between objects (location and comparison).
For the transformational aspect the questions pertain to the objects (modified attribute-value couples, appearance and disappearance of objects in the system), and the modified stative relations (location and comparison).
For the teleological aspect only the causal relation between two events is the focus of the questions, the other components being accounted for by the dynamic aspect of the questioning process.
For each question, the correct answer appears within a set of distractors which are developed by the application of rules adapted to the different aspects of the questions.

### 4.2. Principles for construction of micro-texts

The nature of the response to each question given by the learner determines the nature of the explanatory text presented:

- in the case of an exact response, no information is presented,
- in the case of an incorrect response, all information associated with the question is presented,
- in the case of a response which is true by default (the absence of one or n elements of the response), the portion which is correct is positively reinforced and complementary information is presented,
- in the case of a correct answer by excess, a correction is provided and the correct information is repeated.

### 4.3. Example : Questions pertaining to the starter system.

Questioning begins with the interrogation of the superordinate macro-event of the structure which corresponds to the achievement of the target goal in the starter system, that is, starting the flywheel. Thus the first question pertains to this event.

If the answer to this question is correct, then an initial group of questions pertaining to the relational aspect of the system-event is presented. The questions first pertain to the existence of two objects : "dented crown of the flywheel" and "starter pinion (gearwheel)" which are part of this event. The questions then focus on the pertinent attribute-value couples for a description of these objects, the form in the present case, and finally they pertain to the stative relations between the two objects (locative relations or relative position).

A second group of questions related to the transformational aspect of the system event is then presented. Here the questions deal with modification of the locative relation between two objects (movement of the starter pinion from the outside to the inside of the dented crown of the flywheel), and of the attribute "movement" between two objects (going from immobility to rotation).

Because of the superordinate position of this macro-event, the third aspect, the teleological aspect of the system, cannot be the object of questions.

All the subordinate events give rise to questions whose structure follows the principles presented and illustrated above.

## 5. VALIDATION OF THE CONCEPT OF SYSTEM

We carried out four tasks (free interview, causal questioning, completing lacunary event triples, and multiple choice questionnaires on the existence of events and causal relations) with three groups of students having various levels of knowledge of car mechanics (G1 > G2 > G3). Our objective was to identify their representation of the starter system. The level of knowledge of students had been determined using a questionnaire on the existence of events from the starter system; students from the first two groups, unlike those from the third, had been taught about this system.

We put forward the general hypothesis (H1) that the acquisition of knowledge about a functional system, on the basis of a text or by observation of the system, is an activity which results in the construction of a coherent signification which corresponds to the description of a functional system as proposed above. In this context, representations of states, events, and actions are structured in functional units of a higher order: the sub-systems, which are hierarchically organized as a goal/sub-goals structure. This leads us to predict a distance effect (D) for systems and among sub-systems, and a system and sub-system boundary effect (B) on the recovery from memory of objects, states, events, actions and relations.

Our principal differential hypothesis (H2) is that there are stages in the acquisition of knowledge about a functional system. These stages are characterized essentially by the various degrees of structuration of the students'

representations in terms of systems and sub-systems. Thus, subjects in group G1 with a good knowledge of the starter system will have a representation of this system which corresponds to our description of a functional system. Subjects from group G3, who have not yet learned about this system but who do have general knowledge of mechanics, will not have constructed a representation of the starter system as an autonomous functional system. Subjects from group G2, who have learned about the starter system but who have demonstrated, by their performance on a questionnaire, a lower level of acquisition than group G1 should either lack a representation of the starter system organized as an autonomous functional system but they should know more elements of this system than G3; or have a representation of the starter system organized as an autonomous functional system but, in relation to group G1, this should be less elaborated insofar as it should contain fewer elements and be less well organized with respect to sub-systems. Consequently, the distance (D) and boundary (B) effects predicted in (H1) will vary according to the subjects' level of knowledge, which leads us to predict an interaction between these factors.
Tasks executed by the subjects are assumed to vary in terms of the activities involved in the recovery of knowledge from memory (see Baudet, 1988). On the one hand they differ in the extent to which the deliberate processes intervene in the recovery process : it is at maximal during the interview, minimal in the multiple choice questionnaires task, and intermediary in the causal questioning and in completing lacunary event triples. On the other hand, they differ by the effect of the local and global coherence characteristics of the cognitive representation on the recovery performance: local coherence is the determining effect in the task involving the multiple choice questionnaire, and global coherence is the determining effect in the interview task. Therefore, we predict a simple interaction between the type of task (T) and the level of elements in the representation hierarchy (micro-events vs macro-events (S), intra-system relations vs inter-systems relations (N)). We also predict a double interaction between these factors and the knowledge level of the subjects.

### 5.1 Free interview and causal questioning

#### 5.1.1 Procedure

The subject is invited to verbally relate to the experimenter all that he/she knows about the phenomena (and related systems) which occur between "the insertion of the key into the ignition switch and the start of the car motor". After this interview, each state, event or action present in the protocol is questioned twice by the questions "why" and "how". This probing should allow us to infer not only the propositional content but also the structure of the subjects' representation.

#### 5.1.2 Protocol analysis

For each subject we thus have an interview protocol and a questioning protocol. First, an inventory of all objects mentioned in the protocols is made and they are categorized as a function of their system membership: starter system, systems adjacent to the starter system, and other systems (see Baudet, Blaizet, Denhière, 1987; Blaizet, Cheritel, Legros, 1988). We then retain, for each protocol, the propositions describing states, events and actions. A first categorization of these propositions is done in terms of their system membership, as identified above. They are then categorized in terms of their membership level within a sub-system: initial, intermediate and final positions. Then they are categorized in micro- and macro-propositions depending on whether they represent micro- or macro-events. Finally, we make an inventory of the relations made by subjects among the states, events and actions. They are first classified ---with respect to their position in the structure analyzed--- in terms of systems: starter intra-system, inter-systems when the relation is established among two elements belonging to two different systems, and other intra-systems (not the starter system). The relations are then categorized in terms of their nature: CAUSE (C), ENABLING (E), GOAL (G), TEMPORAL (T) : non conditional relation marking sequential or simultaneous occurrence of two elements, and SPECIFICATION (SP) : part-whole relation between two states, events or actions.

5.1.3 Summary of major results

5.1.3.1 Distance effect (D) for systems

The average number of objects ($F_{2,36} = 55.6$ ; $p<.001$) and events ($F_{2,36} = 33.6$; $p<.01$) mentioned in the interview protocols and questionnaires were ordered according to the predicted hierarchy :

STARTER SYSTEM > ADJACENT SYSTEMS > OTHER SYSTEMS.

The difference between each pair of means was significant at the p<.01 level.

| | SYST | > | ADJSYST | > | OTHER SYST |
|---|---|---|---|---|---|
| OBJECTS | 16.2 | > | 6.1 | > | 1.0 |
| EVENTS | 4.9 | > | 2.0 | > | 0.4 |

**Table 3.** Average number of objects and events mentioned in the interview protocols and questionnaires.

This distance effect can also be observed for membership relations ($F_{2,36} = 10.6$; $p<.01$). The average number of relations mentioned in the interview protocols and questionnaires is ordered according to the predicted hierarchy: the internal relations within the starter system (INTRASYST) are significantly more numerous than those established for systems (INTERSYST) the latter being more numerous than internal relations within other systems (INTRASYST≠).

| | | | | |
|---|---|---|---|---|
| **INTRASYST** | **>** | **INTERSYST** | **>** | **INTRASYST≠** |
| **10.6** | **>** | **3.8** | **>** | **3.0** |

**Table 4**. Average number of relations mentioned in the interview protocols and questionnaires

This distance effect (D) varies with the knowledge level of subjects (C). The interaction between those two factors is significant for objects ($F_{4,36} = 9.33$; $p<.01$) and for events ($F_{4,36} = 7.8$; $p<.01$). In both cases, there is a significant difference between the mean number of objects and relations belonging to the starter system and those objects belonging to adjacents systems for groups G1 and G2, but not for G3. For the relations, the interaction with the knowledge level of subjects is significant ($F_{4,36} = 4.124$; $p<.001$). Thus for groups G1 and G2 the number of relations internal to the starter system was significantly greater than the relations belonging to either of the other two system categories. This was not the case for group G3.

5.1.3.2 Boundary effect (B)
The position effect for the events in the sub-system is significant: ($F_{2,36} = 115.0$; $p<.001$). The average number of events mentioned in the interview protocols and questionnaires was ordered according to the predicted hierarchy:

FINAL > INITIAL > INTERMEDIATE.

The difference between each pair of means was significant at the $p<.01$ level.

| | | | | |
|---|---|---|---|---|
| **FINAL** | **>** | **INITIAL** | **>** | **INTERMEDIATE** |
| **13.2** | **>** | **6.4** | **>** | **2.4** |

**Table 5**. Average number of events mentioned in the interview protocols and questionnaires as a function of their position in a sub-system.

There is also a significant interaction between the factors position (P) and knowledge level (K) ($F_{4,36} = 7.8$; $p<.01$). An analysis of the simple effects shows that only in the final position do significant differences occur between G1 (the group with more knowledge) and the other two groups, G2 and G3.

5.1.3.3 Task effect (T)
We observe a significant interaction between the type of task (T) (interview vs. causal questionnaire) and the level in the hierarchy of the representation of items recovered in memory and mentioned in the subjects' protocols. The interaction between the type of task (T) and the level in the structure (S) shows that the questioning facilitates the recovery of micro-events, but not of macro-events ($F_{1,18} = 6.11$; $p = .02$). On the other hand, the interaction between the type of task (T) and the level of relations (N) shows that the questioning procedure was effective in recovering relations internal to the starter system which are subordinate to the inter-system relations ($F_{2,36} = 13.50$ $p<.001$). The

interaction T * S * K approaches statistical significance ($F_{2,18} = 2.7$; $p = .09$). The questioning resulted in a relative improvement of the recovery of micro-events for groups G1 and G2 but not for G3. A significant interaction N * T * K ($F_{4,36} = 3.70$; $p=.01$) shows an improvement in the recovery of subordinate relations for groups G1 and G2 but not for G3.

### 5.1.4 Conclusion on the free interview and causal questioning

The results obtained are compatible with the main hypothesis (H1) according to which the cognitive representation of a complex functional system corresponds to our description in terms of a functional system.
These results are also compatible with the differential hypothesis (H2) which states that there are stages in the acquisition of knowledge, and that those stages are characterized by the various degrees of structuration in terms of systems and sub-systems. Thus, subjects in group G1 possessing a high level of knowledge of the starter system have a representation of this system which corresponds to our description in terms of functional systems while subjects in group G3 who have not had instruction on the starter system but who possess general knowledge about mechanics did not build a representation of the starter system organized as an autonomous functional system. Group G2 which had received training on the starter system but which had acquired a lower level of knowledge than subjects in G1, performed in a manner compatible with the hypothesis which states that while they have a representation of the starter system organized as an autonomous functional system, in relation to G1 it is less elaborated insofar as it contains fewer elements and/or is not as well organized with respect to sub-systems.
The interactions between the factors type of task and level of elements in the representation hierarchy allow us to infer the existence of local and global coherence characteristics of the representation which are compatible with our description in terms of functional systems (H1) and with the hypothesis of acquisition of knowledge by stages (H2).

## 5.2 Incomplete event triples

### 5.2.1 Procedure

At the end of the questioning task subjects received a booklet with eight pages. Two events were presented on each page and they were separated by a gap which the subject was invited to fill with an event from the starter system occurring between the two events presented. The eight event pairs were chosen according to the following principles:

- 4 pairs expressed macro-events and 4 pairs expressed micro-events,
- half of the triples (event 1 - gap - event 3) formed a sequence of adjacent events in our description. The other half formed a sequence of events which were not adjacent,
- for one half of the triples the gap was located at the boundary of two sub-systems, the two events given belonging to two different sub-systems. For the

other half, the gap was located within a sub-system, with both given events belonging to the same sub-system.

### 5.2.2 Analysis of the responses

Subjects responses are categorized in terms of :

a) answers belonging to the same sub-system as the expected answer (S-SUBSYST), answers belonging to the same system but to a different sub-system (S-SYST) and answers which belong to another system (OTHER SYST) as well as omissions (OMIS).

b) correct and incorrect responses. An answer is considered correct if it expresses an identical or similar proposition to one or many of the expected propositions based on our a priori analysis of the starter system.

### 5.2.3 Main results

#### 5.2.3.1 Distance effect (D) for systems

The frequency of response types follow the predicted hierarchy ($F_{3,54} = 11.9$; $p<.01$.). There are more responses belonging to the same sub-system as the expected response (S-SUBSYST) than there are responses belonging to the same system but to a different sub-system (S-SYST), which are more frequent than those belonging to other systems (OTHER SYST).

| S-SUBSYST | > | S-SYST | > | OTHER SYST | (OMISSIONS) |
|---|---|---|---|---|---|
| .363 | > | .179 | > | .095 | (.363) |

The differences between the adjacent frequencies are significant at the $p<.05$ level.

As it is shown in table 6 below, the different patterns of response frequencies differ according to the knowledge level ($F_{6,54} = 12.0$: $p<.01$).

G1: S-SUBSYST > S-SYST > OTHER SYST
G2: S-SUBSYST = S-SYST > OTHER SYST
G3: no significant difference

| | G1 | G2 | G3 |
|---|---|---|---|
| S-SUBSYST | .625 | .375 | .090 |
| S-SYST | .250 | .232 | .050 |
| OTHER SYST | .017 | .054 | .021 |
| OMISSIONS | .100 | .330 | .640 |

**Table 6.** Frequencies of different response types for groups of subjects.

5.2.3.2 Boundary effect (B)
The correct responses are more numerous when the gap occurs at the border of two sub-systems rather than within a system ($F3,54 = 5.9$: $p<.01$). The boundary effect was observed for that group with the highest level of knowledge, G1 (.714 vs .535), but not for G2 (.321 vs .429), or G3 ( .40 vs .144). This double interaction approaches the level of significance ($F6,54 = 2.0$ : $p<.10$).

5.2.4. Conclusion on the incomplete event triples

Two main conclusions can be made on the basis of the results presented.
The majority of events expressed in the subjects' responses belong to the same sub-system as one or other of the events in the event pair; the number of intrusions of events belonging to other sub-systems of the starter system is greater than the number of intrusions of events belonging to other systems. This distance effect to the sub-system can be observed in the performance of subjects possessing a good understanding of the starter system, (G1). For subjects in group G2 it was not possible to ascertain significant differences between the frequency of responses belonging to the same sub-system and the frequency of responses belonging to another sub-system within the same system. Nonetheless within the same sub-system, both are clearly superior to the frequency of responses which belong to a system different from the events in the event pair. This observation is compatible with the hypothesis that while subjects in group G2 have a representation organized in terms of a functional system, it is less well organized than for subjects in group G1 with respect to sub-systems.
The results are also compatible with the hypothesis that a boundary effect of the sub-systems exists which is determined by the greater accessibility of events in initial and final positions. The probability that the response will be correct is greater when the gap is situated at the boundary of two sub-systems rather than within a sub-system. This boundary effect is observed only for those subjects whose representation is organized in terms of functional systems and sub-systems, G1.

5.3. Multiple choice questionnaires on the existence of events and causal relations

5.3.1 Procedure

One week after participating in the previous experiments subjects were invited to complete two questionnaires.
The first questionnaire consists of fourteen questions on the existence of events related to the workings of the starter system. The questions are presented in the following format: a sentence describing a macro-event is presented in writing; the subject is then required to choose from an array of presented events, those events which form part of that macro-event. The presented events include the correct response along with three distractors constructed as a function of their distance (with respect to systems and sub-systems) from the correct answer:

- distance 1 (D1) : the event belongs to the same sub-system as the correct response,
- distance 2 (D2) : the event belongs to the same system but to a different sub-system,
- distance 3 (D3) : the event belongs to another system.

The second questionnaire includes 19 questions on the existence of causal relations. A sentence describing a micro-event is presented in writing. The subject is then required to choose from an array of presented events, those events which are directly caused by the micro-event. The events presented to the subject include the correct response (the result) as well as four distractors which respectively describe:

- either the GOAL of a sub-system to which the event in question belongs,
- or the GOAL of a sub-system adjacent to the one where the event belongs,
- or an event belonging to this adjacent sub-system,
- or the GOAL of a sub-system which is not adjacent or of another system entirely.

5.3.2 Hypotheses and predictions

5.3.2.1. Predictions about events

Following is a restatement of the predictions concerning the distance effect (D) for systems (D0 > D1 > D2 > D3), and of an interaction between this distance effect (D) and the knowledge level of subjects (K).

5.3.2.2. Predictions about causal relations.

The hypothesis of a cognitive representation organized as a functional system and sub-system allows us to predict that events which function as goals and sub-goals have a greater probability of being available and accessed in memory. This will reduce the influence of the distance effect when trying to establish causal relations. This prediction of a dominant teleological effect should produce the following order of response frequencies:

<u>Prediction A</u>

RESULT = GOAL S-SYST > GOAL S-SADJ > GOAL OTHER S-SYST > EV S-SADJ

We predict a dominant effect of distance for subjects with a low level of knowledge which should produce the following order of response frequencies:

<u>Prediction B</u>

RESULT > GOAL S-SYST > EV S-SADJ > GOAL S-SADJ > GOAL OTHER S-SYST

5.3.3 Main results

The weak performance of subjects in group G3 prevented us from including their results in the analysis.

5.3.3.1 Questionnaire on the existence of events

Distance effect (D) among systems. Response frequencies are ordered as predicted ($F_{3,72} = 43.5$; $p<.001$) and, as shown in table 7 below, the distance effect for sub-systems is observed only for those subjects having a high level of knowledge, group G1 (K * D : $F_{9,72} = 5.40$; $p<.01$).

| | G1 | G2 |
|---|---|---|
| **- Distance 0 (correct response)** | .58 | .39 |
| **- Distance 1 (same sub-system)** | .37 | .10 |
| **- Distance 2 (same system)** | .24 | .28 |
| **- Distance 3 (other system)** | .12 | .17 |

Table 7. Frequency of response as a function of the distance to the sub-system for groups G1 and G2.

5.3.3.2. Questionnaire on causal relations

An analysis of variance shows that prediction A of a dominant teleological effect is compatible with the results from group G1 and that prediction B is compatible with the results of group G2.

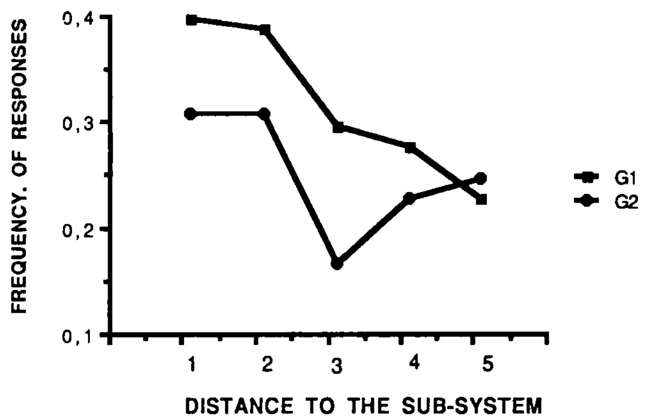



Figure 3 : Frequency of responses as a function of the distance to the sub-system (D1= Result, D2= Goal Sub-System, D3= Goal Adjacent System, D4= Goal Other System, D5= Event Adjacent system).

### 5.3.4. Conclusions on multiple choice questionnaires

There are two major conclusions :

1. For the group having the highest level of knowledge (G1), the number of event intrusions belonging to other sub-systems is inversely proportional to the distance (measured in sub-systems) to the sub-system of the event in question. This distance effect in terms of sub-systems is not evident with subjects having less knowledge of the topic (G2) even though the previous experiments have demonstrated that for the same subjects a distance effect with respect to systems is observed. This observation is compatible with the hypothesis that subjects in group G2 have a representation organized as a functional system which is less well organized with respect to sub-systems than that of the subjects in group G1.
2. The results are compatible with the hypothesis that organization in terms of goals/sub-goals reduces the influence of the distance effect for subjects whose cognitive representation of the starter system is organized as a functional system and sub-systems.

### 5. 4. Conclusion on the results of the three experiments

The three experiments provide a body of results compatible with the hypothesis that acquiring knowledge about a functional system is an activity which culminates in the construction of a coherent network, which corresponds to the description of the proposed functional system. They also provide information as to what might be the appropriate steps that are necessary to achieve mastery of a complex technical system: from the incoherent representation of some events, states or actions to a cohesive organization. The cohesive unit is a functional system capable of differentiating this organization into sub-units of the same type (sub-systems units at a high level of the functional system).

## 6. GENERAL CONCLUSIONS

The emergence of cognitive research on the acquisition of knowledge from texts in a domain which up until now has been confined to the study of comprehension and memorization of texts, especially the narrative text, has led to a rejection of an approach which centered exclusively on texts and the use of linguistic knowledge. The study of the cognitive representations built at the end of text processing remains a favored approach to research. However, taking into account the specific characteristics of those type-representations that are knowledge and beliefs and considering interaction among characteristics of the communication system --conceived as a transformation system of information involving individuals, their linguistic and world knowledge, and their relations, as well as any constraints on the situation-- has led us to a different perspective. We have observed that after processing a text, an individual constructs many types of representations of the information in the text. These representations may or may not be reflected in the surface information of the text and may or may not be considered important depending on the task demands. In the description of those representations, an important step has been the introduction of the concept of mental model or situation model to conceptualize the world representation that individuals build through their experiences and their learning, and which they activate during the reading of a text. The concept of a model, however, only has value if the model is rigorously defined and if the descriptions of the world and other representations that an individual has of it are clearly specified. Our analysis in terms of systems is an attempt to produce a conceptualization that allows for a precise description of the representation that we have of the world and of the representation that the learner builds for himself. It allows us to formulate in new terms, questions regarding comprehension and text production. It is a necessary condition for the detailed study of the processing of complex verbal information such as the interplay of the linguistic and logico-linguistic elements of the text, the cognitive characteristics of individuals and the characteristics of the world represented by a text. Our model allows us to seriously consider the construction of computerized systems to assist in learning in complex domains. A first series of experimental work has produced a body of results which support the "psychological validity" of our proposed model.

## REFERENCES

BAUDET, S. (1986). Compréhension et mémorisation de récit chez l'enfant. Effet de l'origine sociale. Thèse de Doctorat de 3ème cycle, Université de Paris VIII.

BAUDET, S. (1990). Relative importance of information and retrieval from memory. In H. Mandl, E. De Corte, N. Bennett & H. F. Friedrich (eds.), Learning and Instruction. Oxford : Pergamon Press ltd, 199-214.

BAUDET, S., BLAIZET, F. et DENHIERE, G. (1987). Acquisition individualisée de connaissances et représentation cognitive d'un système fonctionnel complexe. Communication au cycle de formation "Approche systémique, Méthode d'Analyse Descendante", Laboratoire d'Automatique et d'Analyse des Systèmes, C.N.R.S., Toulouse, 30 Novembre-3 Décembre 1987. Document de l'Equipe " Psychologie cognitive du Traitement de l'information symbolique", Université de Paris VIII.

BLAIZET, F., CHERITEL, J., & LEGROS, D. (1987). Représentation cognitive et compréhension du fonctionnement d'un système fonctionnel complexe : le moteur thermique à combustion interne, Communication au Colloque : Culture technique et formation, Cité des Sciences et de l'Industrie de la Villette, Paris, 17-18 Décembre 1987.

COOMBS, C.H., DAWES, R.M. & TVERSKY, A (1975). Psychologie mathématique. Paris: P.U.F.

DENHIERE ,G . (1984). Il était une fois .... Compréhension et souvenir de récits, Lille: Presses Universitaires de Lille.

DENHIERE, G. (1988). Story comprehension and memorization by children : the role of input-, conservation- and output processes. In: F. WEINERT & M. PERLMUTTER (eds), Memory development : Universal changes and individual differences, Hillsdale, N.J.: Erlbaum.

DENHIERE, G, & BAUDET, S. (1989). Cognitive Psychology and Text Processing : From text representation to text-world. In P. Ouellet (Ed.) Cognition and Artificial Intelligence, Special Issue, Semiotica, 77-1/3, 271-293.

EHRLICH, S. (1976). Structures sémantiques circonstancielles et permanentes. Variables de situation et de connaissance. Bulletin de Psychologie, Numéro spécial : La mémoire sémantique, 34-45.

FRANCOIS, J. (1989). Changement, Causation, Action : Trois catégories fondamentales de la description sémantique du lexique verbal à l'exemple de l'allemand et du français. Genève, Droz.

Garnham, A. (1981). Mental models as representation of text. Memory and Cognition, 9, 560-565.

Gentner, D. & STEVENS, A.L. (1983). Mental Models. Hillsdale, N.J.: Lawrence Erlbaum

GRAESSER, A.C. (1981). Prose comprehension beyond the word. New York, Springer-Verlag.

HART, H.L.M. & HONORE, A.M. (1959). Causation and the Law. Oxford: Clarendon Press.

HILTON, D.J. & SLUGOSKI, B.R. (1986). Knowledge - Based Causal Attribution : The Abnormal Conditions Focus Model. Psychological Review, vol. 93, 1, 75-88.

ITKONEN, E. (1983). Causality in Linguistic Theory. London, Croom Helm.

JOHNSON-LAIRD, P.N. (1980). Mental models in cognitive science. Cognitive Science, 4, 71-115.

JOHNSON-LAIRD, P.N. (1983). Mental models. Cambridge, MA: Harvard University Press.

KINTSCH, W. (1986). Learning from text. Cognition and Instruction, 3(2), 87-108.

LEHNERT, W. (1977). Human and computational question-answering. Cognitive Science, 1, 47-73.

LE NY, J.F. (1979). La sémantique psychologique. Paris: Presses Universitaires de France.

LE NY, J. F. (1985a). Comment (se) représenter les représentations. Psychologie Française, 30, 231-238.

LE NY, J. F. (1985b). Depliements. Doc n° 19, CEPCO, Université de Paris-Sud.

LE NY, J. F. (1985c).Le réductionnisme, sans la turpitude. Doc. n°26, CEPCO, Université de Paris-Sud.

MACKIE , J. L. (1974). The cement of universe. A study of causation. Oxford: Clarendon Press.

MANDLER, J. M. (1979). Categorial and schematic organization in memory. In C. R. PUFF (ed.), Memory Organization and Structure, N.Y., Academic Press.

MANNES, S. M. (1988). A theoretical interpretation of learning vs. memorizing text. European Journal of Psychology of Education, Special Issue, Acquisition of Knowledge from Text, G. DENHIERE & H. MANDL (Eds).

MANNES, S. M. & KINTSCH, W. (1987). Knowledge organization and text organization. Cognition and Instruction, 4(2), 91-115.
NORMAN, D.A. (1983). Some observations on mental models. In D. GENTNER & A.L. STEVENS, Mental Models. Hillsdale, N.J.: Lawrence Erlbaum
PERRIG, W. & KINTSCH, W. (1985). Propositional and situational representations of text. Journal of Memory and Language, 24, 503-518.
POITRENAUD, S. (1987). QASH : manuel de référence. Université de Paris VIII
PUNTEL, L. B. (1978). Wahrheitstheorien in der neueren Philosophie. Darmstadt, Wissenschaftliche Buchgesellschaft.
PUTNAM, H. (1978). Meaning and the Moral Sciences. London, Routledge.
SEARLE, J. R. (1983). Intentionality. Cambridge University Press, 1983.
TAPIERO, I (1987). L'acquisition inidividualisée de connaissances à l'aide de textes et d'illustrations : Contribution à l'étude de la mesure automatisée des connaissances initiales. Mémoire de D.E.A., Université de Paris VIII.
TAPIERO, I., POITRENAUD, S., & DENHIERE , G. (1988). Individualized acquisition of knowledge with the computer: interrogation and learning guided by the structure of knowledge. European Journal of Psychology of Education, Special Issue: Acquisition of knowledge from text.
van DIJK,T.A. (1977). Text and Context. London, Longman, 1977.
van DIJK, T.A. & KINTSCH, W. (1983). Strategies of discourse comprehension. N.Y.:Academic Press.
von WRIGHT, G.H. (1967) : The Logic of Action : a Sketch. In N. RESCHER (ed.), The Logic of decision and action. Pittsburgh: University of Pittsburgh Press.
von WRIGHT, G.H. (1971). Explanation and Understanding. London: Routledge.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)
Elsevier Science Publishers B.V. (North-Holland), 1991


# KNOWLEDGE MODIFICATION DURING READING

Steffen-Peter BALLSTAEDT and Heinz MANDL

Deutsches Institut für Fernstudien
an der Universität Tübingen
Tübingen, Federal Republic of Germany

To date the psychology of knowledge lacks adequate methods to assess modifications of concept structures during text comprehension. A study is presented, where the Heidelberger-Struktur-Lege-Technik (SLT) was used successfully to assess complex processes of concept-building in comprehension of instructional texts. The study examines the concept alterations occurring with a segmented factual text containing contextual dissonance.

## 1. ASSESSMENT OF KNOWLEDGE STRUCTURES

Knowledge has been considered a structure composed of interconnected cognitive elements ever since the development of Aristotele's association theory and this view is still accepted even by artificial intelligence research which presents knowledge as neo-associationistic networks. The basic building blocks of knowledge are concepts linked together by relations. The simplest unit is a proposition which in turn is connected to other propositions through relations, and thus become integrated into a complex knowledge structure (overview in Hoffmann, 1986; Tergan, 1986). When a person reads a text, and if he/she understands the contents, and thus learns from it, then a part of his/her knowledge structure becomes activated and will be changed continuously while reading. Thus new concepts and relations are actually introduced under the guidance of the author, and the reader's existing structure becomes more differentiated, or even completely restructured (cf. accretion, tuning, restructuring by Rumelhart & Norman, 1978). If one assumes that the reader has learned something, then after reading a more elaborated concept has been placed behind the same old word. The knowledge changes are not only dependent on the externalized knowledge provided by the author, but also on the reader's prior knowledge. To date little attention has been paid to idiosyncratic aspects of reader-text-interaction. Research concerned with text comprehension usually infers from group performances to cognitive processing by testing with recall and recognition after a short text has been read. Only some studies have attempted to assess knowledge changes during reading in an experimental or quasi-experimental fashion. Revealing would be time series on quantitative and qualitative knowledge changes when dealing with extensive study material (books, pictures in texts, audiovisual units, etc.). Relating to this topic two approaches seem to be quite interesting. Shavelson (1974) and Shavelson & Stanton (1975) assessed knowledge structures on physical concepts of novices before and after taking a physics course. They used three different assessment methods: word association, card sorting, and graph building. Both structures were compared with the structure of an expert as well as with each other. These three structures can be understood to represent three different stages of knowledge acquisition. The knowledge structures of the novices became similar to that of the expert once they had completed the course: they contained more relations and more adequate concept formations. The findings in the Shavelson studies were based on strictly limited, well-defined concepts. In most cases the acquisition of knowledge is much more diffuse, but also more creative. One opportunity to register such processes is provided by the thinking aloud method. It was developed by the German psychology of thinking for problem solving. Yvonne Waern (1980) and Graesser (1981) have used them for on-line text comprehension. Here the reader reveals after reading a short text paragraph his/her thoughts, ideas, and associations. The analysis of thinking-aloud-protocols is quite tedious, but offers rich, qualitative material. Its interpretation dependends on how one estimates the possibility of verbalizing cognitive processes (cf. Nisbett & Wilson, 1977; Ericsson & Simon, 1980).

In this section we are going to introduce a method with which knowledge structures and their changes are externalized via a structure-formation-technique consisting of concept cards and relation cards. Scheele & Groeben (1984) constructed this game and called it the Heidelberger-Struktur-Lege-Technik (SLT), which is used for assessing subjective theories. We modified this procedure in order to use it to register the individual knowledge structures and changes as precisely as possible. Therefore, this procedure can be used to diagnose knowledge states, evaluate study material and realize ethnographic studies. We shall first describe the Structure-Formation-Technique and then present a study where this technique was employed to analyze individual knowledge changes. The final section of this paper will describe the advantages and disadvantages of this procedure.

## 2. A STRUCTURE-FORMATION-TECHNIQUE (SFT)

The SFT consists of empty concept cards (4x4 cm) as well as smaller relation cards with printed symbols. Eighteen possible relations are provided and can be used to connect the different concepts. Table 1 shows the various arrangements. Two kinds of relations are to be distinguished: defining relations which correspond to the usual relations used in semantic networks for the representation of concepts. According to Klix's (1976) terminology these relations are classified as intraconceptual relations: identity, classinclusion, property, example.

The second group of relations are the empirical relations according to Klix's interconceptual relations. Here the structure-formation-technique deviates radically from the usual existing network structures and offers a much simpler list of relations. This list is based on relations that may exist between empirical variables: positive and negative dependence, reciprocal and curvilinear dependence, mediated dependence (suppressor or moderator variables). How usable and adequate these descriptions are for knowledge structures will be discussed when evaluating this procedure.

Table 1: Graphical representation and verbal examples of relations used between concepts in the Structure-Formation-Technique

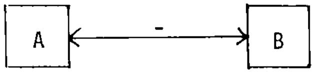
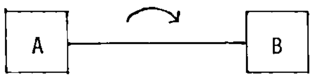
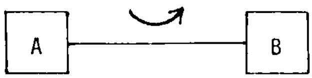
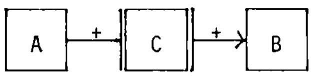
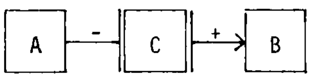
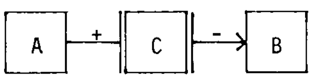
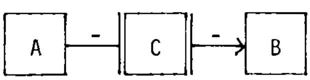
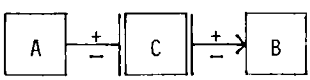
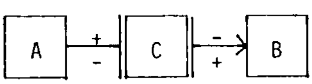
**Intraconceptual relations**

| | | |
|---|---|---|
| identity | $A \overset{=}{\text{———}} B$ | A text (a) is a sentence structure (B) |
| inclusion | $A \xrightarrow{i} B$ | Novels (A) are a kind of text (b) |
| example | $A \xrightarrow{e} B$ | "Germinal" (A) is a novel (B) |
| property | $A \xrightarrow{p} B$ | Smoke (B) belongs to fire (A) |
| conjunction of properties | $A \xrightarrow{p} B \wedge C$ | Tomatoes (A) are red (B) and healthy (C) |
| disjunction of properties | $A \xrightarrow{p} B \vee C$ | Tomatoes (A) are red (B) or green (C) |

**Simple interconceptual relations**

| | | |
|---|---|---|
| positive dependence | $A \xrightarrow{+} B$ | The greater the heat (A) the more sweat (B) is secreted |
| negative dependence | $A \xleftrightarrow{-} B$ | The greater the fear (A) the poorer the performance (B) |
| mutual positive dependence | $A \xleftrightarrow{+} B$ | Success (A) increases motivation (B), this leads to more success (A) |

Table 1 (continued)

| | | |
|---|---|---|
| mutual negative dependence |  | Low self-assurance (A) leads to failure (B), failure leads to decreased self-assurance (A) |
| curvilinear dependence with summit (maximum) |  | With increasing age (A) nearsightedness (B) increases at first and later on decreases again |
| curvilinear dependence with low point (minimum) |  | With increasing age (A) leisure time (B) decreases at first and later increases again |
| Mediated interconceptual relations | | |
| positively mediated positive dependence |  | Increasing alcohol consumption (A) leads to addiction (B) when conflicts (C) increase |
| negatively mediated positive dependence |  | Frustrations (A) lead to greater agressions (B), if no solutions (C) are available |
| positively mediated negative dependence |  | Car accidents (A) will incur fewer deaths (B) if safety belts (C) are used |
| negatively mediated negative dependence |  | The more incomprehensible a text (A) is the less it will be read (B) if the interest (C) is missing |
| equi-directional alternating interaction |  | Self-determination in learning (A) with high intelligence (C) results in good performance (B) while low intelligence results in poor performance |
| contra-directional alternating interaction |  | Group work (A) with great social fear (C) results in poor performance (B) while low fear results in high performance |

The subjects were given written instructions to get acquainted with the relations which they had to practise at home. This was followed by a training session at the institute where the subjects with the help of the experimenter learned to lay out a structure with a simple concept. After this initial practise the actual experiment began. The experiment was always done in the following manner: The respective critical concept was presented on a concept card and the subject was asked how he/she would explain this concept to another person. All the important concepts which came to mind were supposed to be written down on available concept cards. Following this the subject had to connect these concepts with relation cards, i.e. he/she formed a structure which integrated all these concepts. Every relation that was laid out was again verbalized by the experimenter in order to either stimulate its confirmation or correction. Otherwise the experimenter behaved in a non-directive manner, i.e. he/she did not provide any suggestions or evaluate the activities of the subject. The subject was allowed to add concepts, remove concepts, or rearrange concepts until he/she considered his/her knowledge structure complete. Now the experimenter verbalized the completed structure once more so that the subject was given a final opportunity to make corrections. Only then was the final structure preserved in form of a drawing or a photograph. This concept structure is considered prior knowledge and now becomes confronted with new information, i.e. a text which deals with the critical concept. After having read the text the subject was asked whether he/she wanted to change his/her prior knowledge structure, i.e. wanted to insert new or remove old concepts. Again the experimenter verbalized any newly arranged relation. After the subject has changed the structure a new text segment was offered, etc.: Structure-formation and reading phases alternated until the subject had read all the information that was to be related.

Now we would like to discuss the analysis of the externalized knowledge structures. Scheele & Groeben (1984) developed their Heidelberger Struktur-Lege-Technik (SLT) in order to reconstruct subjective theories. Therefore they were more interested in an individually centered, idiographic evaluation and less interested in nomothetical group comparisons. They do, however, consider such comparisons feasible. In order to evaluate idiographically as well as nomothetically a systematic comparison of intraindividual and interindividual structures becomes necessary. Quantitative and qualitative measures can be used. First of all the number of concepts and their addition during knowledge acquisition should be determined. Then these concepts can be grouped together with a content analysis and counted in the categories. Correspondingly, the number of relations and kind of relations can also be determined. They can be ranked according to increasing complexity both theoretically as well as empirically.

These complexities range from simple dependencies to reciprocal interactions as well as to mediated interactions. On the basis of the just mentioned listings different measures of complexity concerning content and formal structure, differentiatedness, and coherence can be calculated. The analysis of a laid out structure can be especially useful when the ideal knowledge structure of the concept is known, i.e. as for physical and mathematical concepts. In this case the deviation of the laid out structure from the target structure can be assessed. If such a target structure is unknown or cannot be constructed the structure can still be investigated with respect to internal coherence, degree of contradiction and empty spaces. Depending on the hypotheses one could also imagine the use of other evaluation methods, since methodological creativity knows no limits.

## 3. EXPERIMENT: KNOWLEDGE ACQUISITION WITH TEXTS

### 3.1. Research questions

The following study belongs to a group of experiments in which the utilization of the structure-formation-technique in knowledge acquisition with texts is examined (Ballstaedt & Mandl, 1985; Mandl & Ballstaedt, 1986). In order to not only get processes of accretion a text was employed which stimulated the restructuring of knowledge. To accomplish this, two theories contradicting each other as to one of their major statements were presented to the subjects. According to the cognitive dissonance theory (Festinger, 1974) it was expected that the contradictory statements would act as provocation causing the activation of restructuring processes.

### 3.2. Experimental text

The selected topic in the text dealt with dreams because we believed that for the concept "dream" every subject would possess some prior knowledge which is retrievable. Secondly, two contradictory dream theories exist. A text composed of three segments was written (T1, T2, T3). The contents were constructed in such a manner that each segment could be read individually, but when read together resulted in one continuous text. Every segment consisted of two type-written pages:

> T1: Empirical findings in experimental dream research: The dream as the brain activity ; sleep cycles registered with the electroencephalogram and the electroocculogram; eye-movements during sleep (REM-phases).
>
> T2: The psychoanalytic dream theory of Freud (1972): Latent dream thoughts and manifest dream contents; dream work, especially its symbolization; the dream as a meaningful, interpretable content.
>
> T3: The neurophysiological dream theory by Crick & Mitchison (1983): Neural activity of the brainstem during REM-phases; cortical interpretations of random activation patterns: the dream as a meaningless combination of various contents.

The text segments could be easily understood, but did not spare the reader the scientific terminology. T1 talked about the findings which were in agreement with the psychoanalytical theory as well as with the neurophysiological theory. T2 and T3 contradicted each other in one central point: one text segment presented the dream as a meaningful combination of various contents while the other said that dreams were meaningless combinations. Since a solution was not offered it was assumed that the two contradictory text segments would lead to cognitive dissonance.

### 3.3. Experimental subjects

The subjects were twelve college students who did not study psychology. They were paid for their participation.

### 3.4. Procedure

The structure-formation-technique was employed in the manner previously described. After a training session with the concept DESERT the major experiment was done. The subjects designed the prior knowledge structure for the concept DREAM (= S0). Following this they read T1 and were allowed to restructure their prior knowledge design (= S1). The second text segment given to them was either T2 or T3. This was followed by a restructuring of S1 (= S2). The same procedure occurred for the third text segment, either T2 or T3. Afterwards the final knowledge structure lay on the table (= S3). Now this was followed by a short interview where the attitude for the two dream theories was ascertained from the subjects and their evaluation of the SFT was registered. The main experiment lasted approximately 4 - 6 hours. The described procedure can be placed between an experiment and casuistry. Though knowledge modifications are described as individual cases, they are registered and assessed under partially controlled conditions. For this kind of procedure Dörner (1983) has created the term "experimental casuistry".

### 3.5. Results: Idiographic-casuistic evaluation

Each subject laid out four different knowledge structures at successive points of time between which a text segment was read. We first examined the individual learning procedures and then looked for the tendencies exhibited by the group as a whole.

From the 12 learning sequences the subject No. 7 (female) was selected. Figure 1 a-d shows the successive knowledge structures.

The boldly edged concept cards here represent the new concepts the subjects constructed after having read a text segment, the hachured concept cards represent old, but restructured concepts. The prior knowledge of this subject can be expressed verbally in this manner:

> S0: The dream belongs to the subconscious which represents a kind of experience processing. Sleep is a prerequisite to dreaming. One kind of dream is the nightmare. Dreams consist of pictures which under specific circumstances can evoke memories that can be interpreted.

The subject externalized a simple structure consisting of common knowledge concepts (DREAM, PICTURES, NIGHTMARE) interspersed with psychoanalytic terminology (SUBCONSCIOUS, INTERPRETATION). The expression "under specific circumstances" is the verbalization of an empty space located in the structure. Here an empty concept card was added. The subject knows that dream pictures can evoke memories, but according to her view situations have to exist which she does not know at the moment. Altogether her knowledge seems to be somewhat fuzzy (Does the dream belong to the subconscious? Do memories need to be interpreted? etc.). Supposedly an expert on dream theory would call such knowledge "superficial knowledge".

After having read T1 the subject changed her structure by adding and rearranging concept cards. The new structure can be described in the following manner:

> S1: The dream belongs to the subconscious which represents a kind of experience processing. *All three are examples of brain activity. The dream is a special kind of sleep.* One kind of dream is the nightmare. *During the REM-phase the dream evokes pictures and during the deep-sleep phase abstract contents.* Under specific circumstances both evoke memories that can be interpreted. *The REM-phase is accompanied by eye movements which are recorded and measured by the EOG; and by brain waves which are recorded with the EEG.*

Some new neurophysiological concepts (REM-PHASE, DEEP-SLEEP PHASE, EYE MOVEMENTS, BRAIN WAVES, EEG, EOG) were incorporated while the basic structure and the concepts of prior knowledge were maintained. Only three old concepts were restructured, i.e. they appeared in new relations. For subject No. 7 the second text segment was the psychoanalytic theory. The subsequent knowledge structure reads as follows:

> S2: *The dream develops out of wishes which provide the latent dream thoughts.* The dream belongs to the subconscious, and represents a kind of brain activity. The dream is a special kind of sleep. One kind of dream is the nightmare. During the REM-phase the dream evokes pictures and during the deep-sleep phase abstract contents. *Both represent the manifest dream content and can be interpreted as a coded text. When waking up and with the aid of free association* under specific circumstances memories can be retrieved which can be interpreted. *Contents from longterm memory and actual experiences lead to the manifest dream content by using symbolization, condensation, reversal into the opposite, and displacement.* The REM-phase is accompanied by eye movements measured by the EOG and brain waves which are recorded by the EEG.

This text segment obviously suited this subject since she had already incorporated into her prior knowledge structure two psychoanalytical concepts: She accepted thirteen new concepts and continued to use four old concepts in new relations. That the subject favours the psychoanalytic theory became obvious after she read the text segment T3 which contained the neurophysiological theory. She only added two new concepts:

> S3: The dream develops out of wishes which provide the latent dream thoughts. Wishes belong to the subconscious which represents a kind of brain activity. The dream is a special kind of sleep. One kind of dream is the nightmare. During the REM-phase the dream evokes pictures and during the deep sleep phase abstract contents. Both represent the manifest dream content which can be interpreted as a coded text. When waking up and with the aid of free association under specific conditions memories can be retrieved which can be interpreted. Contents in longterm memory and actual experiences lead to the manifest dream content through symbolization, condensation,

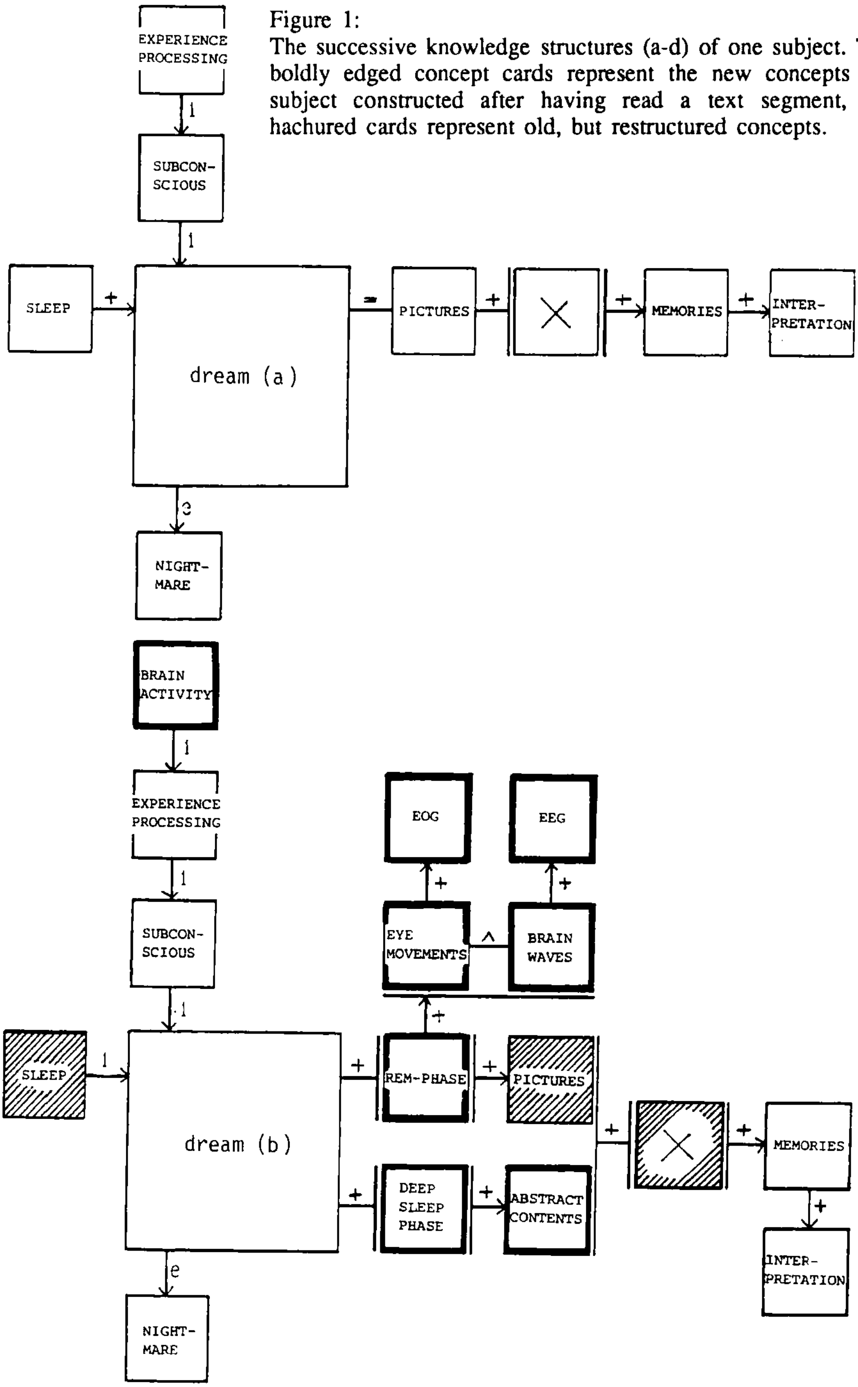
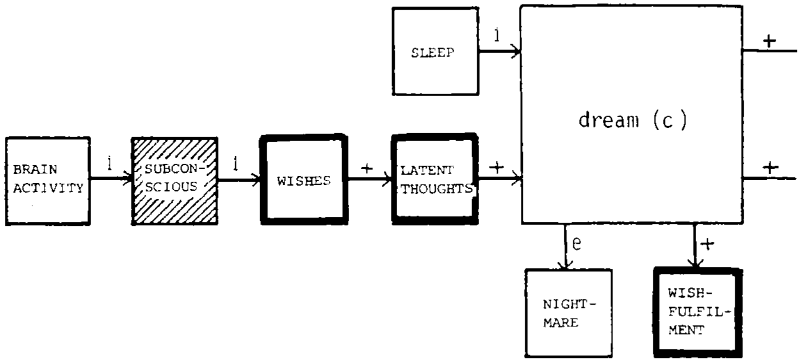



Figure 1:
The successive knowledge structures (a-d) of one subject. The boldly edged concept cards represent the new concepts the subject constructed after having read a text segment, the hachured cards represent old, but restructured concepts.

Figure 1 (continued)



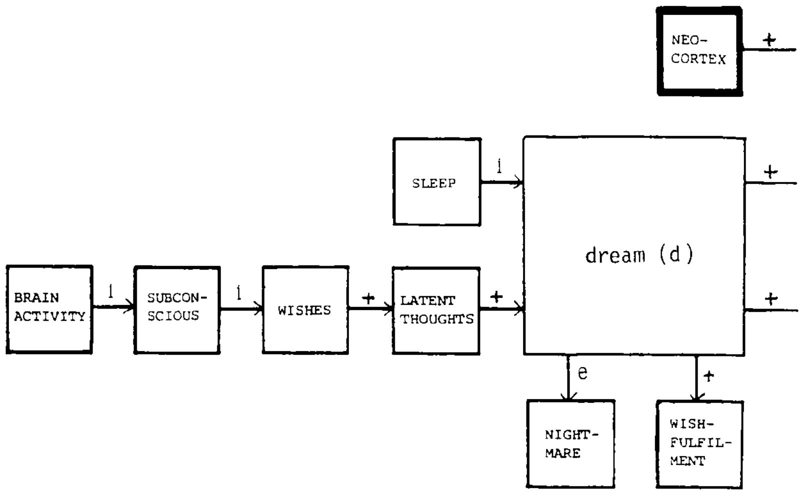
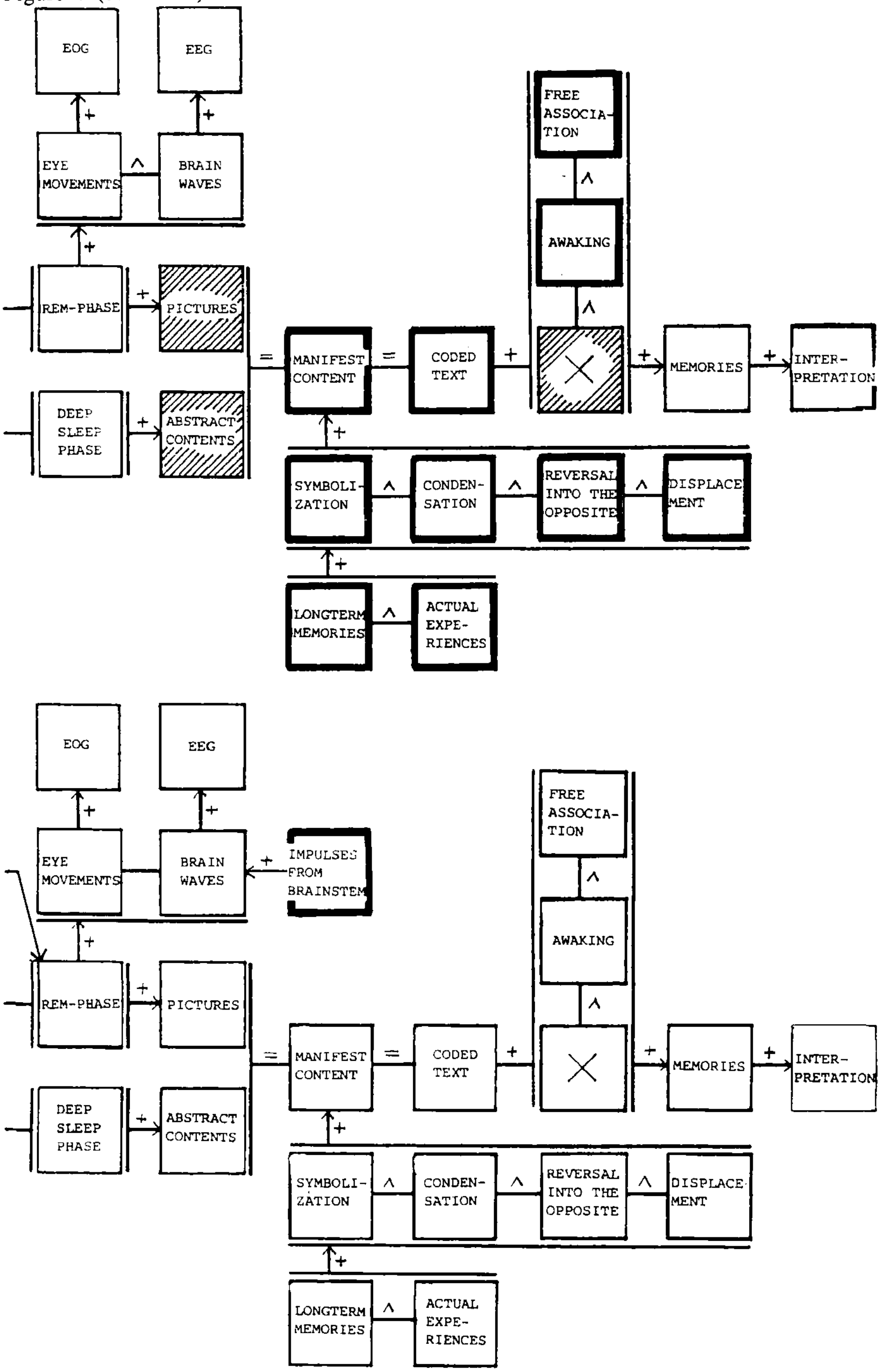



Figure 1 (continued)



> reversal into the opposite, and displacement. The REM-phase is accompanied by eye movements measured by the EOG, and brain waves measured by the EEG. *The REM-phase is a function of the neocortex. Impulses originating from the brainstem lead to brain waves.*

When questioned, the subject admitted a preference for the psychoanalytic dream theory, even though she was only able to retrieve superficial information according to her prior knowledge structure. *Accordingly she only integrated those concepts into her knowledge* structure which agreed with the psychoanalytical theory. The contradiction between meaningful and meaningless dream content was ignored.

If we take a look at the knowledge changes which occurred for all twelve subjects we can to ascertain four different strategies the subjects applied to the two contradictory theories, although the procedures employed differed greatly:

CUMULATION: Seven more externally oriented subjects simply adopted concepts located throughout the text segments without paying any attention to the existing contradiction. They accumulated knowledge without attempting to integrate or consider the contradiction.

IGNORING: Two subjects, one being our example, mostly ignored those concepts of the theory which did not interest them. In eleven cases those were the neurophysiological concepts. They were simply dropped because the subjects favored the psychoanalytical theory.

SEGREGATION: Two subjects recognized the partial incompatibility of both dream theories and designed two loosely connected structures, one neurophysiological structure and one psychoanalytical. The contradiction of the two theories was consciously maintained.

*INTEGRATION: Only one subject attempted to integrate both approaches,* but gave up very soon since the necessary restructuring was much too complicated.

These four strategies can be classified or ranked according to the amount of effort invested from least to most: ignoring, cumulation, segregation, integration. It is quite noticeable that the more difficult strategies appear less often and that the subjects tried to avoid restructuring. This effect can probably be traced back to the experimental situation which did not require further activities. Furthermore, the subjects thought that this was a learning experiment and expected a test and probably assumed that they would be best prepared with a cumulative strategy.

### 3.6. Statistical evaluation

A statistical evaluation of the SFT is not meaningful for small samples since the individual structures are very different in content and format. The amount and kind of concepts and the relations vary so much that it is almost impossible to calculate the mean. To describe the quantitative growth of structures we nevertheless attempted some classifications. The contents were divided into three categories with a content analysis: everyday concepts, psychoanalytic concepts, neurophysiological concepts. Table 2 shows the average growth of concepts for two text sequences.

Table 2: Average growth of concepts for two text sequences in three categories (n = 12)

| | $\overline{x}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_1$ | $T_2$ | $T_3$ | | $T_1$ | $T_3$ | $T_2$ | |
| | $S_0$ | $S_1$ | $S_2$ | $S_3$ | $S_0$ | $S_1$ | $S_2$ | $S_3$ |
| everyday concepts | 6.1 | 7.1 | 7.1 | 7.1 | 10.5 | 11.0 | 11.3 | 11.1 |
| psycho-analytic concepts | 5.1 | 4.8 | 14.7 | 14.7 | 3.3 | 3.3 | 2.7 | 7.8 |
| neurophysio-logical concepts | 1.0 | 7.7 | 7.7 | 12.1 | 1.7 | 6.1 | 9.8 | 9.5 |

It is not surprising that the structures became more complex since more concepts were added or that the preceding text segment influenced the categories of the adopted concepts. It is, however, very interesting that the amount of everyday concepts remained the same for both groups. This corresponds with the observation that prior knowledge structures rarely change.

The relations were also divided into three categories: defining (intra-conceptual) relations, simple interconceptual relations, and mediated interconceptual relations. The average increase for each category is shown in Table 3. All relations show an increase (a). However, if one calculates the relative share in the structure then the defining relations decrease slightly, the simple interconceptual relations increase, and the mediating relations remain constant (b). Complexity and differentiatedness was not calculated since the size of the laid out knowledge structures was relatively small.

Table 3: Average number (a) and relative proportions (b) of three relation categories (n = 12)

(a)

| | $\bar{x}$ | | | |
|---|---|---|---|---|
| | $S_0$ | $S_1$ | $S_2$ | $S_3$ |
| intraconceptual relations | 9.0 | 13.4 | 16.5 | 19.3 |
| simple interconceptual relation | 2.4 | 3.9 | 5.6 | 7.5 |
| mediated interconceptual relations | 1.5 | 2.2 | 2.8 | 3.3 |

(b)

| | % | | | |
|---|---|---|---|---|
| | $S_0$ | $S_1$ | $S_2$ | $S_3$ |
| interconceptual relations | 69 | 68 | 66 | 63 |
| simple interconceptual relation | 18 | 20 | 22 | 25 |
| mediated interconceptual relations | 12 | 11 | 11 | 11 |

## 4. DISCUSSION

Compared to the sophisticated network presentation of knowledge the SFT offers simpler connections for interconceptual relations that are based on a variable model of empirical psychology. Two objections can be made. First of all the danger could exist that the methodological thinking of the empirical social scientist does not correspond to the thinking of a novice. In this case the scientist's offer of relations would present a Procrustean bed to his/her knowledge. Secondly, interconceptual relations are too simple to represent complex knowledge. Here the utilization of relations would lead to a simplification of knowledge. Both objections should be taken seriously since in this study none of the subjects complained about lack of opportunity to externalize knowledge. Conversely, in a study done with experts (Mandl & Ballstaedt, 1985) the subjects did complain about missing relations. One gets the impression that the SFT seems to be suited for novices, i.e. the externalization of everyday theories. Nevertheless, one should contemplate the development of a technique based on the semantic networks utilized in cognitive science. Such a technique would be extremely useful in comparative studies concerned with text comprehension and artificial intelligence research.

All subjects found the SFT stimulating and confirmed that they became aware of their knowledge when they saw the structure spread out in front of them. However, the face validity of this procedure should not be accepted too quickly. An unpublished study reports that the SFT showed great intraindividual differences for the same concept with respect to structure, even though time intervals were short. With such findings the validity of the procedure may be questionable: Does the structure really represent knowledge? To answer this question one has to understand that the idea of solidly connected brain structures is on offspring of artificial intelligence research used in computer simulations and not applicable to the human.

At the moment the brain still contains many secrets. We still don't know much about cortical spreading processes and resonances which, with the activation of a concept, can lead to a different awareness each time this concept becomes activated. This is why a subject depending on mood, situation, and task will never construct the same knowledge structure twice, except if he/she learned it by heart. It should be noted that the frequently exercised practice in cognitive psychology to infer from performance to representation or to competence seems to be a rather precarious undertaking. One should not expect test theoretical reliability or validity, but rather utilize the SFT for the investigation and study of idiosyncratic and situational variations for every single knowledge activation. Concerning the quality of relations and the handling of the procedure several improvements will be necessary to optimize the method for individual knowledge diagnosis.

## REFERENCES

Ballstaedt, St.-P. & Mandl, H. (1985). Veränderungen von Wissensstrukturen beim Lesen. In D. Albert (Hg.), Bericht über den 34. Kongreß der Deutschen Gesellschaft für Psychologie in Wien 1984, Bd. I: Grundlagenforschung (S. 252-254). Göttingen: Hogrefe

Berlyne, D.E. (1974). Konflikt, Erregung, Neugier. Zur Psychologie der kognitiven Motivation. Stuttgart: Klett

Crick, F. & Mitchison, G. (1983). The function of dream sleep. Nature, 304, 111-114

Dörner, D. et al. (1983). Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität. Bern: Huber

Ericsson, K.A. & Simon, H.A. (1980). Verbal reports as data. Psychological Review, 87, 215-251

Festinger, L. (1978). Theorie der kognitiven Dissonanz. Bern: Huber

Freud, S. (1972). Die Traumdeutung. Frankfurt/M.: S. Fischer

Graesser, A.C. (1981). Prose comprehension beyond the word. New York/Heidelberg/Berlin: Springer

Hoffmann, J. (1986). Die Welt der Begriffe. Psychologische Untersuchungen zur Organisation des menschlichen Wissens. Weinheim: Beltz, 1986

Klix, F. (1984). Über Wissensrepräsentation im menschlichen Gedächtnis. In F. Klix (Hg.), Gedächtnis, Wissen, Wissensnutzung (S. 9-73). Berlin: VEB Deutscher Verlag der Wissenschaften

Mandl, H. & Ballstaedt, St.-P. (1985). Assessment of concept building in text comprehension. In F. Klix (Ed.), Symposium in Memoriam Hermann Ebbinghaus. Amsterdam: Elsevier Science Publishers (North-Holland)

Nisbett, R.E. & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. Psychological Review, 84, 231-259

Rumelhart, D.E. & Norman, D.A. (1978). Accretion, tuning and restructuring: three modes of learning. In J.W. Cotton & R.L. Klatzky (Eds.), Semantic factors in cognition (pp. 37-53). Hillsdale, N.J.: Erlbaum

Scheele, B. & Groeben, N. (1984). Die Heidelberger Struktur-Lege-Technik (SLT). Weinheim: Beltz

Shavelson, R.J. (1974). Methods for examing representations of a subject-matter structure in a student's memory. Journal of Research in Science Teaching, 11, 231-249

Shavelson, R.J. & Stanton, G.C. (1975). Construct validation: methodology and application to three measures of cognitive structure. Journal of Educational Measurement, 12, 67-85

Tergan, S.-O. (1986). Modelle der Wissensrepräsentation als Grundlage qualitativer Wissensdiagnostik. Opladen: Westdeutscher Verlag

Waern, Y. (1980). Thinking aloud during reading. A descriptive model and its application. Scandinavian Journal of Psychology, 21, 123-132

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# COHERENCE IN SEMANTIC REPRESENTATIONS : TEXT COMPREHENSION AND ACQUISITION OF CONCEPTS

Jean-François LE NY

Institut des Sciences Cognitives et de la Communication
Centre d'Etudes de Psychologie Cognitive
Centre Scientifique d'Orsay
Bâtiment 335
91405 ORSAY CEDEX - FRANCE

Text comprehension may be considered to involve a series of specific processing phases whose final result is a complete semantic, mental representation, which is not a representation of the text, but rather, of what the text is about. As a rule, this representation can cover either a particular or a general state of affairs. From a psychological, cognitive point of view, the word "coherent" must be applied to this mental representation, once it is constructed. If a text or any other form of discourse may be described as "coherent", it is by virtue of its ability to produce such a coherent semantic mental representation in one or several reader(s).

This approach differs from that of many linguists, who are mainly interested in identifying the marks in discourse that render coherence manifest. In our perspective these marks provide instructions which allow the reader to construct an appropriate, coherent representation of the state of affairs described by the text : it is easier for the reader to construct a coherent representation when he or she is guided by such marks. However, because the coherence of representations is mainly a function of internal content relationships, it may be achieved even in the absence of such surface marks.

This view is certainly compatible with, and may be seen as equivalent to that of J. Petöfi (this volume) : our "coherence of the representation" in a cognitive, psychological model, corresponds to the concept "constringency" in his linguistic model. The differences between our approaches concern the status of "worlds" and "states of affairs" : in formal semantics they are abstract constructs whereas, in our model, there is one and only one real world, and nothing else but representations.

These representations are about individuals, situations and events that can be real or fictitious ; that is, having a content - or an "intention" in Brentano's sense. An important question is the extent to which differences in contents of these mental representations are relevant for text comprehension and memorization. The processes used in comprehension, and the general structure of the representations constructed from texts are usually considered to be content independent. However, certain characteristics of content are relevant, especially the internal relationships that make the particular structure of a content more or less plausible (see below). In the elaboration and transmission of scientific knowledge it obviously is important to know whether texts are true, false, or fictitious with a reasonable partial plausibility (as in science fiction).

These questions will be examined by considering the processes readers use to comprehend, i.e., in constructing mental, semantic representations from texts, the forms of representations readers possess in memory while reading texts, and the acquisition of scientific representations, i.e., conceptual schemata.

Construction of semantic representations from texts is a natural activity, i.e., ultimately, a set of events or states in the brain. Cognitive psychology studies these representations within its own framework, i.e., as natural events or states in the mind, irrespective of their neurophysiological or neurochemical nature. Only their functional properties are relevant at this level.

Such processes and representations can therefore be modelled in a way that either uses concepts relevant to neurobiology - as it will be argued in part one of this paper - or that can be simulated in a computer - as is presented in the final section.

## The role of memory in establishment of coherence in semantic representations

Psychologists generally agree that texts are processed in large units, sentences or chunks, from which partial representations are constructed : these partial representations are then inter-connected in such a way that a whole representation is progressively constructed (Le Ny, 1989a).

The entire mental representation constructed from a text can be conveniently modelled in the form of a semantic network composed of propositions. In this type of description, the semantic network is an external, physical, figurative representation (in the first sense of the word), which is assumed to correspond to the internal, mental, semantic representation (in the second sense) : the network consists of nodes, which correspond to bundles of information, as a rule conveyed by words in the text, and links, which correspond to semantic relations. Several types of networks can be generated which have relations represented by these links (for example, predicates, cognitive relations viewed as "instance of", semantic cases or roles, etc.). However in all cases they are semantic relations. They are outputs of operations in the comprehension process that either decode the syntactic relations present in the text sentences (for example word order, grammatical voice, inflexions, etc.), or infer these semantic relations from the word content. For example, semantic relations such as agent, patient, or instrument may have been obtained from identification of an active, passive or reflexive voice in the text, word order (in some languages), surface cases (in other languages), and word meaning, i.e., pre-existent knowledge of causal relations in the subject's long term memory.

In such models, text comprehension thus results from two sets of operations : the first set consists of syntactico-semantic parsing of successive sentences or chunks. The second set can be modelled as the addition of new branches to the extant semantic network. Parsing of each particular sentence or chunk is a sequence of operations yielding a partial semantic representation, in an appropriate format, which is added to the current incomplete general representation, modelled as a partial network.

One advantage of this view is that it is highly compatible with recent advances in automatic text interpretation in Artifical Intelligence. Machine parsing, for example, as carried out by an Augmented Transition Network automaton, provides a partial semantic representation of sentence meaning, which can be added to and integrated with the representation constructed from previous text sentences.

Important problems, such as the interpretation of indexicals or anaphors (see Garrod, this volume) are involved in the second set of operations. However, the main problem is the attachment of the newly constructed partial representations of a sentence or chunk to the current whole representation. This is where the issue of coherence emerges : in attaching the new representation constructed from sentence parsing to the current whole semantic representation.

From a psychological rather than an A.I. point of view, attachment is further complicated by the process of forgetting, or mnemonic decay, which is inherent to human memory and comprehension, and is obviously inapplicable to automatic interpretation.

In the Kintsch and van Dijk model (Kintsch and van Dijk, 1978 ; van Dijk and Kintsch, 1983), sharing arguments among propositions is considered to be a fundamental condition for the construction of coherence, in this case referential coherence. As this construction is assumed to take place in a limited capacity working memory, a special mechanism is postulated for maintaining relevant units in working memory. Conservation of the propositional representation after processing is assumed to take place in episodic memory.

Data contradicting this view have been collected recently. In a series of experiments, (Achour and Le Ny, 1983 ; Hyodo, Le Ny and Achour, 1986) ; Le Ny, Achour, Carfantan and Verstiggel, 1983 ; Le Ny, Carfantan and Verstiggel, 1982) it has been shown that : 1) mnemonic decay of newly processed information takes place during comprehension ; 2) this decay begins very early on (presumably just after presentation of a word) and continues down to an asymptote ; 3) this decay is semantic in nature, and concerns not only discrete words, but also parts of word meaning - as evidenced by testing recognition of synonyms or related words, titles, semantic relations of predication, propositions, etc. ; 4) the slope of this decay (expressed in averaged as well as individual data) is gradual, monotonous, and asymptotic, and shows no evidence of any transfer from a memory store to another memory store, as postulated in the Kintsch and van Dijk model ; 5) this decay is elective, and is steeper for less important than for more important bundles of information ; 6) one may compensate for this decay by reprocessing.

The picture that emerges from these results is that the whole representation constructed from the text changes constantly in memory. This representation can be viewed as involving activation of memory units and structures (Anderson, 1976 ; 1983), progressive de-activation of these units and structures as a function of time (i.e., mnemonic decay),

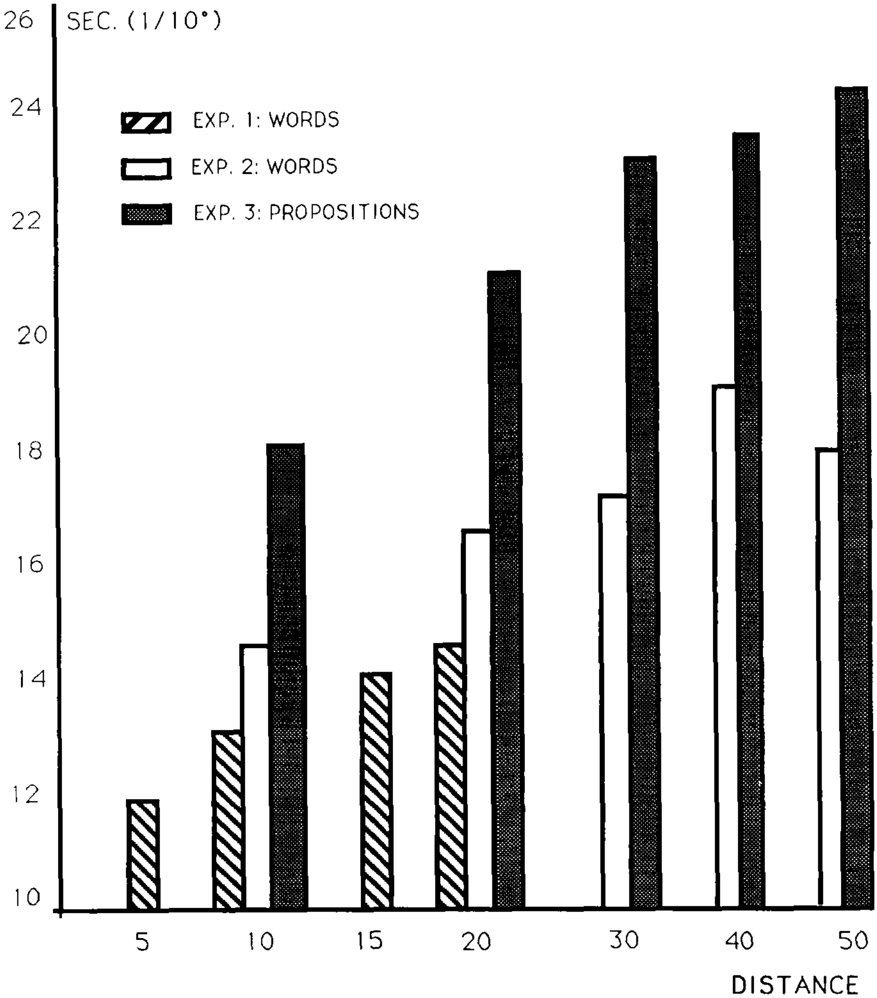



Figure 1: Mean recognition times of words, or atomic propositions, after comprehension of paragraphs as a function of words appearing between the target and the test word. The results of the three different experiments are summarized in this graph.
Individual data show the same trends as these averaged data (after Achour and Le Ny, 1983; Hyodo, Le Ny and Achour, 1986).

and reactivation of some of these when old information is reiterated. The resulting representation should be viewed as a dynamic one : its elements (represented in a semantic network as nodes and links) are first activated, then decay, then reactivated, then decay again, reactivated again, etc. Thus, the role attributed to an episodic memory, and even the concept itself of episodic memory is brought into serious question.

Coherence of the current whole representation can only be maintained if decay of its components is incomplete, so that attachment of new branches can be established. This process depends on reactivation of the main components, which tend to reappear in most types of text. When successive propositions in a text share arguments, there is increased opportunity for a re-processing of information.

This description can account for thematic coherence. Stories, for example, are usually about one or a small number of characters : the representation of these characters - indexed in the text by names or anaphors - is thus re-processed again and again. Moreover, stories are often about a given category of events concerning characters : this type of text is very similar to experimental situations such as Reder's (Reder, 1979 ; Reder and Anderson, 1980; Reder and Ross, 1983), in which a set of facts are memorized, and the facts are related to given themes (such as running). Expository texts are typically organized as a presentation of interrelated themes, which, as a rule, correspond to a set of facts or concepts. Thus, re-processing plays an important role in maintaining coherence in the representation constructed from these texts.

In brief, "coherence of texts" can be regarded as the set of conditions that maintain coherence in the mental representation constructed from them : conditions of re-processing are basic in this respect.

## Connectedness between representations constructed from words or sentences

The centrality of conditions of re-processing can be incorporated into models that use features as an analytical unit in the description of meaning and knowledge. In this case it is assumed that features can also be processed and activated, undergo decay, be re-processed, and be reactivated.

This kind of model describes connectedness between two related word meanings or concepts as the presence of one or several common feature(s) in the two units. The consistence of a mental representation is subordinated to the condition that no one given feature carries an affirmative or a negative value simultaneously. I will provide some illustrations of this below when dealing with schemata.

Models which include features can also account for phenomena of semantic transfer and interference : thematic coherence, as mentioned above in Reder's sense, occurs in the presence of common features in the meaning of many words used in various sentences. In a previous experiment (Le Ny, Carfantan and Verstiggel, 1982) involving yes-no judgments of words with respect to sentences, it was shown that acceptance of words as titles of sentences depended on the presence of common semantic elements in the representation evoked by the title word and the representation constructed from the sentence. For example, an experimental sentence referred to "snow that rushed down" in a determined setting (in short : mountain). The word "avalanche" was accepted as a title for such a sentence with a short response time. Response times increased if the distance between the presentation of the relevant information in the sentence (in this case : snow, rush down) and presentation of the title word was increased.

Titles work like micro-summaries of texts or sentences. They contain the most important, focal information in the whole representation, and thus are an expression of the way the representation is organized and made cohesive in the subject's memory.

In an experiment, P. Duclos (1986) studied how thematic coherence controlled activation of representations in paragraph comprehension. This study used the same method as mentioned above. The general procedure consisted of presenting subjects with paragraphs composed of three sentences on a screen, and asking them to judge two words - called probe words - presented immediately after the paragraph for their semantic relationship to the content of the paragraph. Half of these words were chosen in such a way that they could be used as titles for one sentence of the previous paragraph, or for the whole paragraph. The other 50 percent were words unrelated to the paragraph or any part of it. Each paragraph was constructed from three sentences i, j and k, which concerned the same theme. The second and third sentences could be inverted, thus producing an i, k, j, order, but no observable change in the whole meaning of the paragraph. Here is an example of a paragraph :

Sentence i : This morning in Lyon, two armed men attacked an armored car belonging to the Duval bank.
Sentence j : A security guard was wounded in the shoulder by a bullet as he was at the wheel of the car.
Sentence k : The robbers got away with an estimated seven million new francs.
Probe words : on j : victim
on k : loot
overall : hold-up

For the sake of simplicity we will hereafter disregard overall probe words, which did not produce any result of interest to this question.

Only one probe word corresponding to a sentence was presented to each subject. Thus inversion of the i and k sentence order resulted in a shorter or longer distance from the presentation of the sentence to the presentation of the probe word. Groups were counter-balanced to test for the i or the k sentence.

It was hypothesized that this procedure would show the state of the semantic representation constructed from the sentence and its evolution as a function of time. This state of the representation for a given sentence is assumed to depend on the information presented before or after the given sentence.

Let us now directly examine the problem of thematic coherence. The material exemplified above only concerned one type of paragraph. Sentence i was assumed to be an introduction to the whole paragraph : this type of construction with an introduction and a content body is common in press releases.

A second type of paragraph was used in the experiment. In these, the first sentence was replaced by a series of dots of the same length. Thus, sentences j and k (or k and j) were presented without an introduction.

A third type of paragraph was also used : in these, the first sentence was replaced by a sentence that was semantically unrelated to the two following sentences. For the quoted example, this sentence was :

Sentence u : The French soil produces tons of grain for consumption in the country and for exportation.

Observe that no surface mark is used for expressing between-sentence relationships : the subjects must establish semantic coherence between the sub-representations corresponding to these successive sentences solely on a semantic thematic basis, i.e., as a result of the process of comprehension itself. The problem is whether such activity, or its results, can be shown experimentally.

The reasoning that allows us to derive predictions on this issue is as follows :

In the second type of paragraph, thematic uniqueness was maintained in the text. But if coherence is basically a property of the mental semantic representation, and if it is viewed as an overall product of the many semantic relationships existing within it, (as they are modelled by the links in a semantic network), then it must be hypothesized that the composed activation of the semantic sub-network corresponding to each sentence j and k will be weaker in the absence of sentence i than when it is present. Thus response times should be longer for these paragraphs than for the first type of paragraph.

In the third type of paragraph, thematic uniqueness was not maintained. The first sentence of the paragraph, u, introduced the readers to a theme at variance with the two subsequent sentences. Similar reasoning leads us to the hypothesis that the composed activation in the representations constructed from sentences j and k should be weaker in this case, i.e., when these sentences follow a different-theme introductory sentence than no introductory sentence at all, and, a fortiori, a same-theme introductory sentence.

The main prediction was thus that the response times to the title probe words would be the shortest for the first type of paragraph, longer for the second type, and the longest for the third type. Additional assumptions concerned evolution over time of the semantic representations. It was assumed that they would undergo decay, thus producing longer response times for longer sentence/probe word distances. In addition, a possible statistical interaction was hypothesized between this decay and the type of paragraph.

Figure 2 presents the results of this experiment. Response times to the title probe words were clearly a function of the type of paragraph. Further, response times increased significantly with distance, supporting the hypothesis of a progressive decay of the representation. This is not a new result : what is notable is that no significant interaction between

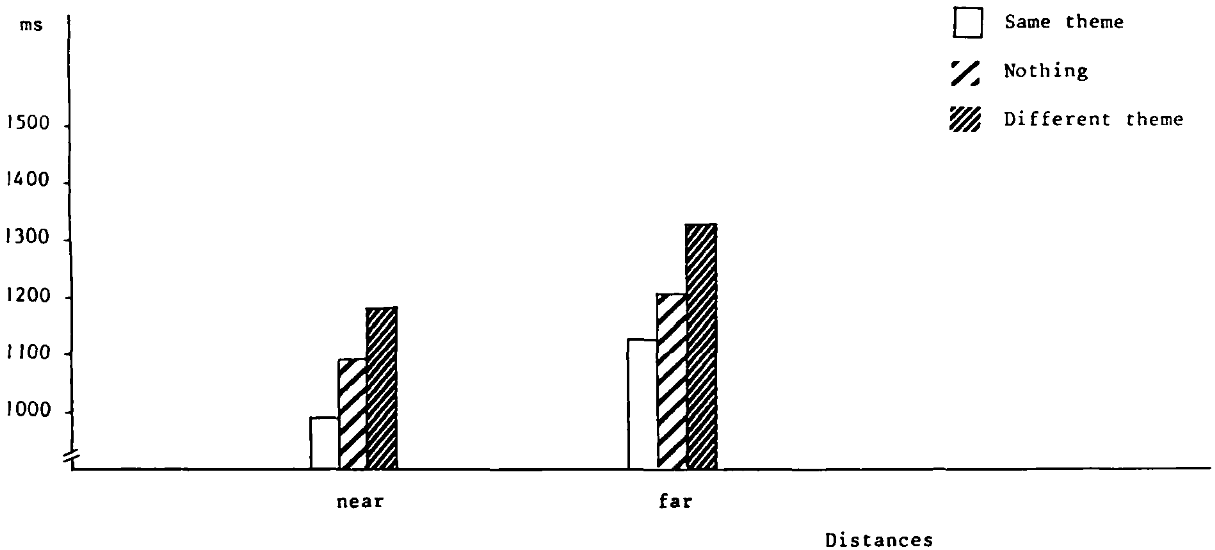



Figure 2. Mean recognition times of titles for sentences after comprehension of paragraphs as a function of the number of words appearing between the target sentence and the test word (two intervals).
The initial sentence in the paragraph had the same theme as the target sentence, was absent, or had a different theme (after Duclos, 1986).

this decay and the type of paragraph was observed.

I would like to draw the following conclusions from these results : thematic coherence was automatically computed during comprehension as a result of complex processing, which included sentence parsing and establishment of relationships between partial representations constructed from successive sentences. Such thematic coherence had a quantitative counterpart in the representation constructed from sentence processing -- in the different level of activation in elementary parts of this representation. In addition, an opposite effect was observed when thematic plurality was present in the sequence of sentences composing the paragraph : the third type of paragraph produced a poorer representation than the first type and even the second type, i.e., the neutral one. In other words, thematic plurality produced effects of interference, or inhibition, on the construction of the mental semantic representation.

A detailed model of these effects can be generated by extending the notion of spreading activation (Anderson, 1976, 1983 ; Collins and Loftus, 1975 ; Le Ny, 1989a). However, space does not allow for presentation of such a detailed model here.

## Acquisition of new concepts through text processing

Two categories of instructional texts can be defined : those which are intended to present readers with specific or general facts, and those which are intended to teach concepts. The former describe events or properties concerning various entities, and produce knowledge in the readers that has a propositional form : for example
"The Chinese recorded the appearance of comets as early as 613 B.C.", or
"The shorter a laser beam's wavelength, the deeper it goes".

Such text processing involves concepts (i.e., uses previously acquired concepts), but it does not itself lead to acquisition of new concepts. To a large extent instructional texts are similar to story-telling texts as described above : the main difference is that they relate true facts, whereas stories are fictitious. But as stated above, truth and falsehood are clearly irrelevant as concerns psychological processing of texts.

The second type of text deals with concepts acquired through language. These concepts differ in this respect from natural-kind concepts, which are acquired through practice -- mainly from the perception of things, and the effects of actions upon them.

Concepts acquired through language are often more abstract and more complex than natural-kind concepts ; their acquisition is based on relationships constructed from previously acquired natural concepts, which are conveyed by instructional discourses or texts (Le Ny, 1986, 1989b).

A main property of these texts is that they involve universality. A concept constructed from a text, as any concept, is a representation that can be used to subsume various objects or facts. Thus their content - their intension - can be described by clauses preceded by one or several universal quantifiers.

A classical schema for the description of generic concepts is :
( x) A(x) -- P(x), in which A is the concept under study, and P is another concept ; for example, a concept denoting a superordinate category or a property. A concrete example of this is, in non-formal language : for any x, 'x is a dog' entails 'x is a mammal'. Or : for any x, 'x is a dog' entails 'x has a heart'.

The same general schema could be used with two or more variables. This schema can be applied to natural concepts irrespective of their mode of acquisition, be it by repeated natural use of language, or by one-time systematic definition.

This schema brings out the relationship between the content of one concept, in this case A, and the content of another concept, in this case P. This relationship is expressed by the universal quantifier, "all", and by the connector (in the above example, "entails").

Such relationships between concepts, or word meanings, must be abided by in texts in order to allow readers to construct coherent representations. By insisting that universality of the quantifier must be perfect (i.e., with no exceptions) and that the internal relationship (symbolized above by an arrow, --, or by "entails") must be a logical conditional, we have formulated conditions of consistence which are more severe than those for coherence.

But the schema presented above may be extended to looser semantic relationships, as, for example, in analysis of typicality. In this case a non-standard quantifier such as "most" can be used, together with a less strict use of "entails". For example : for most x, 'x is a bird' entails 'x can fly'. Or in another version : for any x, 'x is a bird' entails 'x presumably can fly'.

This relationship has been called plausibility, and is viewed as an important condition of text cohesion and of construction of coherent semantic representations (Black, Freeman and Johnson-Laird, 1986 ; van Dijk, 1977 ; Haliday and Hasan, 1976).

Especially, a text which introduces new concepts for acquisition by a reader as a rule emphasizes the relationships that constitute the core of the concept. This can be done by explicitly mentioning the quantifier(s) and the main relationships in the concept : definitions, or at least complete characterizations are examples of such texts. In other cases, the text only presents a description of a typical object or case exemplifying the concept, leaving the reader to generalize from the description.

This is presumably the main difference between acquisition of concepts from educational texts and natural acquisition : the latter involves abstraction, or gradual separation of relevant from irrelevant features. To the contrary, acquisition of new concepts through language is based upon a presentation of the relevant features, and the explicit or implicit instruction to form from these a sub-schema possessing universality, that is, a capacity to subsume many instances.

Thus, the coherence inherent to a concept consists of the relationships existing between the semantic components at its core. For complex concepts, these relationships are between other concepts, which themselves are involved in lower-order components and relationships. The types of relationships which concepts contain determine the various kinds of conceptual schemata. Existence of such schemata is a condition of coherence in the readers' representations. These issues must be considered in connection with those of knowledge representation (Bobrow and Winograd, 1977 ; Brachman and Schmolze, 1985 ; Le Ny, 1989c ; Minsky, 1981 ; Rabinowitz and Mandler, 1983 ; Rumelhart and Ortony, 1977).

## Analysis of a concept and use of schemata

Concept analysis may be exemplified by an illustration borrowing from animal psychology, more precisely from the field of animal behavior and learning.

Let us assume that the concepts of "classical" and "instrumental (or operant) conditioning", and "extinction", have been formed in a group of students. From these a new concept can be acquired : "spontaneous recovery".

A cognitive description of the concept "spontaneous recovery" involves three main components :

1. in the presence of the conditioned stimulus or the experimental situation, the animal produces the conditioned response ;
2. the conditioned response was not produced by the animal in the presence of the same stimulus or situation previously ;
3. this change in the animal's behavior re-occurs after simple rest that is, with no particular modification in the animal's environment or in the relationships (contingency) between the animal's behavior and its environment.

This third component is critical if we want students to clearly distinguish the concept of "spontaneous recovery" from the related concept of "relearning" : in the latter the same two first components are also present, but the change in the animal's behavior has been caused by a modification in the environment. "Relearning" occurs when the initial conditioned stimulus - or response in instrumental conditioning - is again followed by the reinforcing stimulus.

Coherence of the concept "spontaneous recovery" - as of "relearning - is, therefore, dependent on the relationships among the schema's components. The concept of "spontaneous recovery" is itself a schema, under which many various cases or instances of particular spontaneous recoveries in various animals, various circumstances, with various stimuli, can be subsumed. This is the manifold universality of the concept : as stated above, the cognitive description we have just presented should be preceded by several quantifiers expressing that this description holds "for any animal, any conditioned stimulus, any conditioned response, any relevant situation".

But in addition, the concept of "spontaneous recovery" is also a particularization -- a sub-schema of a more general schema, that can be termed : "schema of scientific phenomena". This is, in turn, a sub-schema of the very general super-schema of process.

A cognitive characterization of such a super-schema can be summarized with the following three points :

1. Let us specify an anterior state concerning an aspect of a given object or entity : in the case of "spontaneous recovery" this is a state concerning an aspect of an animal ; the state is "cannot produce a given conditioned response" ;

2. Let us specify a different posterior state concerning the same aspect of the same object or entity ; in this case : "can produce the given conditioned response".

Note that the posterior state is different from the anterior in two respects : a) it is associated with a time that is after the other state; b) it has a negative value if the anterior state had a positive value, or a positive value if the anterior state had a negative value. Here, the conditioned response can be produced whereas it could not be produced before.

Such a contrast between two successive states is very common. It appears in all processes. This is why we can say that "spontaneous recovery" is a sub-schema of the super-schema of process. This super-schema also has many other natural sub-concepts (for example in verbs as "to whiten", "to grow", "to go", etc.). At an even higher, more general, level, the contrast existing in "process" is encompassed by the concept expressed in the verb "to change".

3. Let us specify the conditions or cause that produce the change from the anterior to the posterior state. This causal relationship is the component that makes the concept of "spontaneous recovery" a sub-schema of the schema of "scientific phenomenon".

In brief, so far we have four levels of conceptualization, all containing schemata, but with an increasing degree of abstraction (figure 3) : the schema of "spontaneous recovery", at the lowest level, is a sub-schema of the schema of "scientific phenomenon", at an intermediate level, which is a sub-schema of the schema of "process", at a higher level, which is a sub-schema of the schema of "change" at the highest level. All schemata can be characterized by their respective variable and constant components.

Another observation can be made on the concept of "spontaneous recovery", and many other concepts having the same structure, regarding their internal coherence. "Spontaneous recovery" involves "conditioning" and "extinction", with specific relationships between them. For the concept to apply, a given animal must first be conditioned to produce a given response, and then the response submitted to extinction. In other words, the concept of "spontaneous recovery" is partially composed of a script, that is, a type of schemata that encompass a series of constant events or sub-processes involving constant roles that can be filled by entities having constant characteristics ; the fillers have the role and characteristics of "the animal", "the stimuli", "the environmental conditions", and "the response". The presence of such constants (i.e., common entities in all sub-processes of the process), is also a form of coherence, in this case higher-level referential coherence, internal to the concept. An instanciation of "spontaneous recovery" would be

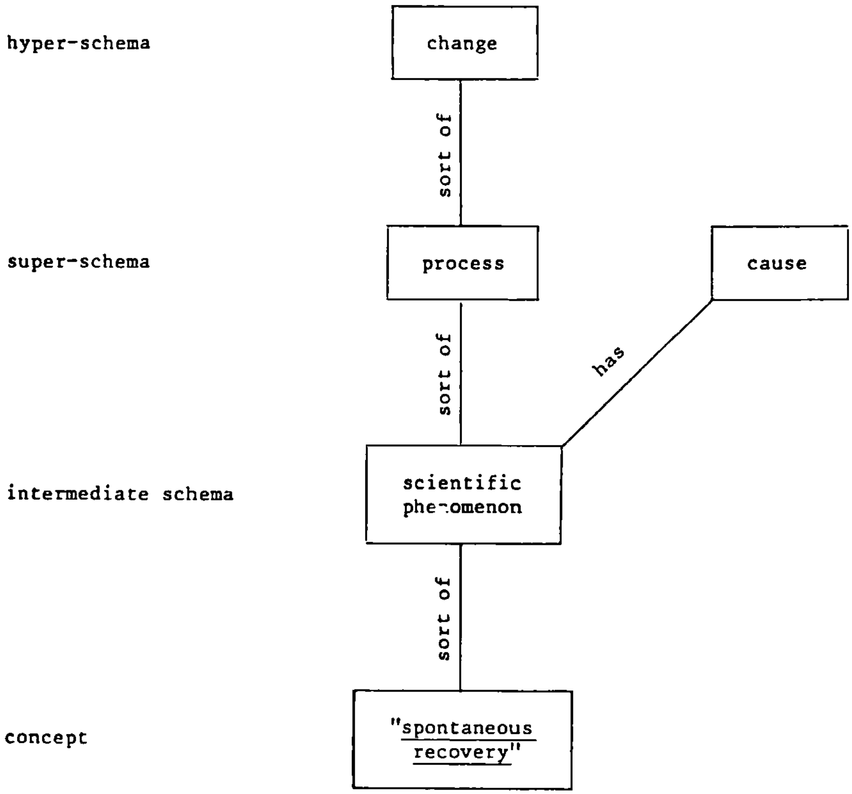



Fig. 3

A hierarchy of schemata subsuming the concept of "spontaneous recovery".

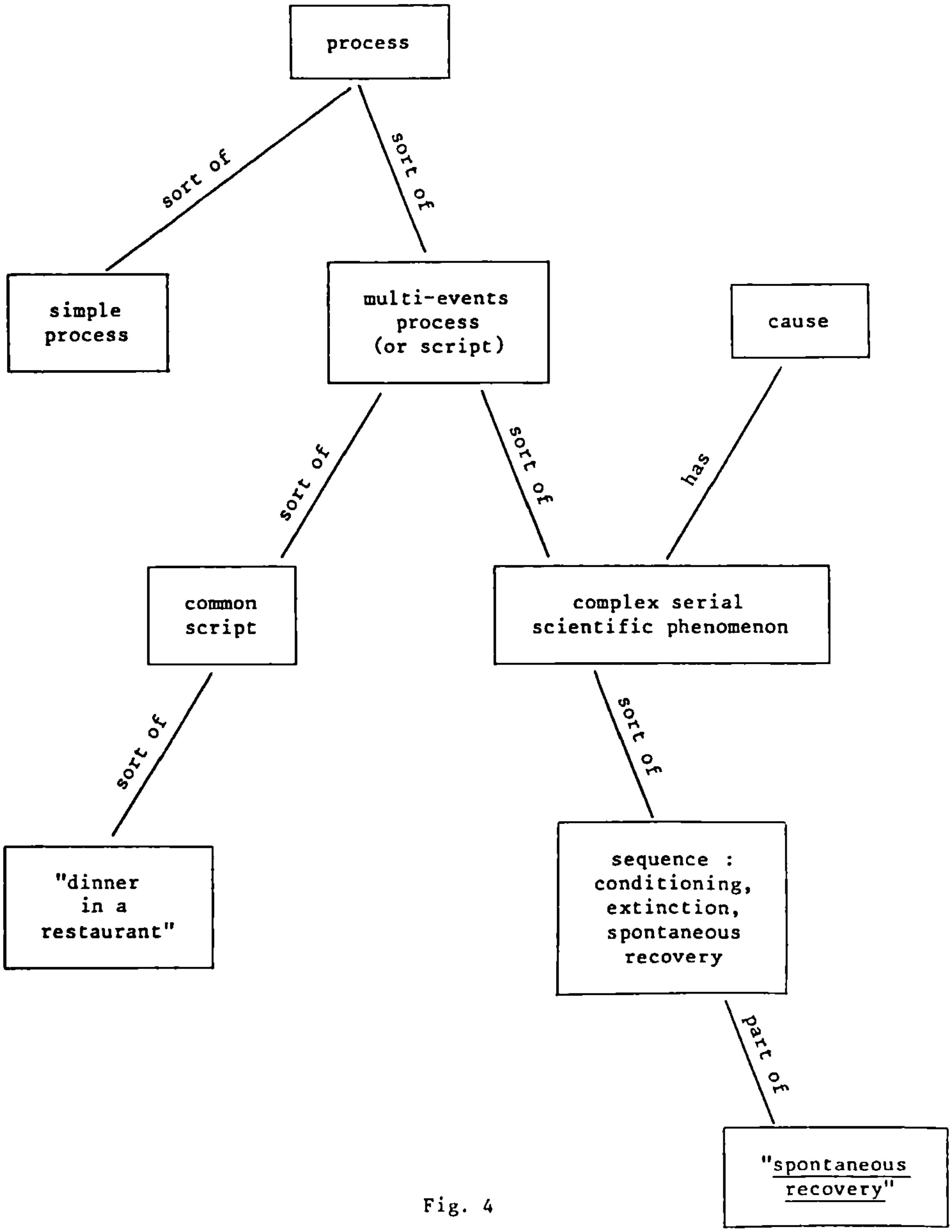



Fig. 4

A hierarchy of schemata and scripts including the concept of "spontaneous recovery".

incoherent if, for example, the animal that showed spontaneous recovery was different from the animal that was previously conditioned.

All the fillers of the schema must have the above-mentioned characteristics, and the relationships among them or their parts or aspects be maintained, for the concept of "spontaneous recovery" to be coherently applied. Therefore, because this concept involves a series of successive events and states (i.e., as a multi-event process) it is an instance of a script. It belongs, mutatis mutandis, to the same general and abstract type as the script of "dinner in a restaurant" (figure 4).

Activation of such scripts, once they have been acquired, can be studied by methods similar to those described above (see Bower, Black and Turner, 1979 ; Walker and Yekovich, 1984, 1987 ; Yekovich and Walker, 1986). However, the script described here does not encompass the entire concept. The complete description of "spontaneous recovery" also involves the relationship of "cause", or "condition", which is a typical component of the schema "scientific phenomenon".

We will not discuss the complex issue of causality here. However, this analysis demonstrates how its mention as in our specification 3 above, can be assessed via the concept, irrespective of the various theoretical ways causality may be viewed.

## An Artificial Intelligence approach

It is clear from the previous analysis, which was couched in non-formal terms, that this detailed model of the conceptual representation to be acquired by the student can be expressed in a very precise language of representation, and used for forms of Artificial Intelligence processing.

An automaton for artificial interpretation of short conceptual texts has recently been developed (see Le Ny, 1986, 1989b ; Le Ny, Carité and Poitrenaud, 1986). It can furnish a machine representation of response texts produced by students for the characterization of concepts (for example, "spontaneous recovery"). These texts are produced in response to conceptual questions such as : "What is an X ?", in which X is the name of a concept. Machine interpretation of these student responses allows correction of their individual concepts if the components and links do not fit the scientific model of the concept.

In this context the problem of coherence in representations appears at two distincts levels : 1. it concerns automatic interpretation, that is, the construction of a machine representation from the text produced by the student ; 2. it concerns mental representation in the student ; in this case, this mental representation is a durable one, as acquisition of concepts requires mnemonic fixation in long-term memory of the transient semantic representation constructed by comprehension of the text. At these two levels coherence depends on the presence of fixed relationships.

In summary, coherence has been viewed in this paper as a property of mental, semantic representations that presupposes establishment or presence of relationships between bundles of information. In comprehension these relationships allow for the attachment of successive partial representations, constructed in working memory by parsing of the successive sentences, to an overall cumulative representation, conserved as a trace of previous sentence processing. Activation of relevant bundles of semantic information, and decay of previously activated bundles play a major role in this process. One effect of thematic uniqueness is to heighten the activation of a given representation in the subject's mind if related information has been processed previously ; previous presentation and processing of unrelated information produces inhibition.

When a subject reads a text in order to acquire new concepts, the information that is conserved in his or her long term memory is basically relational : variable information can be disregarded whereas constant information must be conserved. Schemata serve to describe these two kinds of information in a hierarchical structure ; schemata are themselves acquired as belonging to higher order schemata, i.e. as sub-schemata. The existence of constant relationships, which constitute the concepts viewed as schemata, is also the pre-condition for coherence in representations. Extraction of such constant relationships and the identification of their roles in text processing is one means for cognitive psychology and related sciences to explore comprehension.

## References

(1) Achour, L., and Le Ny, J.F., 1983. L'évolution de la représentation évoquée par un mot au cours de la compréhension de phrases : étude par la technique du sondage. Année Psychologique, 83, 409-422.

(2) Anderson, J.R., 1976. Language, memory and thought. Hillsdale, N.J.: Erlbaum.

(3) Anderson, J.R., 1983. The architecture of cognition, Cambridge, Ma., London : Harvard University Press.

(4) Black, A., Freeman, P., and Johnson-Laird, P.N., 1986. Plausibility and the comprehension of text, British Journal of Psychology, 77, 51-62.

(5) Bower, G.H., Black, J.B., and Turner, T.J., 1979. Scripts in memory for text. Cognitive Psychology, 11, 177-220.

(6) Bobrow, D.G., and Winograd, T., 1977. An overview of KRL, a Knowledge Representation Language, Cognitive Science, 1, 3-46.

(7) Brachman, R.J., and Schmolze, J.G., 1985. An overview of the KL-ONE knowledge representation system, Cognitive Science, 9, 171-216.

(8) Collins, A., and Loftus, E., 1975. A spreading activation theory of semantic processing, Psychological Review, 82, 407-428.

(9) van Dijk, T., 1977. Text and context : explorations in the semantics and pragmatics of discourse. London, Longman.

(10) van Dijk, T. and Kintsch, W., 1983. Strategies of discourse comprehension, New-York : Academic Press.

(11) Duclos, P., 1986. Niveau d'activation et compréhension de phrases : étude par la technique du sondage. Paris, Orsay : Dissertation.

(12) Haliday, M.A.K. and Hasan, R., 1976. Cohesion in English, London, Longman.

(13) Hyodo, M., Le Ny, J.F., Achour, L., 1986. Medium term decay of semantic representations during the comprehension of paragraphs,

Orsay, CEPCO, document n° 55.
(14) Kintsch, W. and van Dijk, T., 1978. Toward a model of discourse comprehension and production, Psychological Review, 85, 363-394.
(15) Le Ny, J.F., 1986. Discourse comprehension and memory for concepts. In : F. Klix and H. Hagendorf (eds.) : Human memory and cognitive capabilities, Amsterdam : Elsevier-North-Holland.
(16) Le Ny, J.F., 1989a. Science cognitive et compréhension du langage, Paris : Presses Universitaires de France.
(17) Le Ny, J.F., 1989b. Automatic induction of concept contents from verbal characterizations. In : G. Tiberghien (ed.) : Advances in Cognitive Science, vol. 2, Chichester : Ellis Horwood.
(18) Le Ny, J.F., 1989c. The role of schemata in concept acquisition and diagnosis. In : H. Maurer (ed.) : Computer Assisted Learning, 2nd International Conference on Computer-Assisted Learning, Dallas Berlin, Springer-Verlag.
(19) Le Ny, J.F., Achour, L., Carfantan, M., and Verstiggel, J.C., 1983. Discourse comprehension and selective decay of meaning's accessibility in recent memory. In : G. Rickheit and M. Bock (eds.): Psycholinguistic studies in language processing. Berlin, New-York : de Gruyter.
(20) Le Ny, J.F., Carité, L., Poiternaud, S., 1986. Construction of individualized texts for the transmission of knowledge through discourse. In : J.H. Danks, I. Kurcz and G. Shugar (eds.) : Knowledge and language. Amsterdam : Elsevier-North-Holland.
(21) Le Ny, J.F., Carfantan, M., and Verstiggel, J.C., 1982. Accessibility in working memory and role of reprocessing in sentence comprehension. In : J.F. Le Ny and W. Kintsch : Language and comprehension. Amsterdam : North-Holland.
(22) Minsky, M., 1981. A framework for representing knowledge. In : J. Haugeland, (ed.) : Mind design, Cambridge, Ma : MIT Press.
(23) Reder, L.M., 1979. The role of elaboration in memory for prose. Cognitive Psychology, 11, 221-234.
(24) Reder, L.M., 1982. Plausibility judgments versus fact retrieval : alternative strategies for sentence verification. Psychological Review, 89, 250-280.
(25) Reder, L.M., and Anderson, J.R., 1980. A partial resolution of the paradox of interference : the role of integrating knowledge. Cognitive Psychology, 12, 447-472.
(26) Reder, L.M., and Ross, B.H., 1983. Integrated knowledge in different tasks : Positive and negative fan effects. Journal of Experimental Psychology : Learning, Memory and Cognition, 9, 55-72.
(27) Rabinowitz, M., and Mandler, J.M., 1983. Organization and information retrieval. Journal of Experimental Psychology : Learning, Memory and Cognition, 9, 430-439.
(28) Rumelhart, D.E., and Ortony, A., 1977. The representation of knowledge in memory. In : R.C. Anderson, R. Spiro and W.E. Montague (eds.) : Schooling and the acquisition of knowledge, Hillsdale, N.H. : Erlbaum.
(29) Walker, C.H., and Yekovich, F.R., 1984. Script-based inferences : effects of text and knowledge variables on recognition memory. Journal of Verbal Learning and Verbal Behavior, 23, 357-370.
(30) Walker, C.H., and Yekovich, F.R., 1987. Activation and use of script-based antecedents in anaphoric reference, Journal of Memory and Language, 26, 673-691.
(31) Yekovich, F.R., and Walker, C.H., 1986. Retrieval of scripted concepts, Journal of Memory and Language, 25, 627-644.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)
Elsevier Science Publishers B.V. (North-Holland), 1991


# ACQUISITION, UPDATING AND ORDERING OF KNOWLEDGE IN THE READING OF PRESS AGENCY WIRES

Jacques FRANCOIS

**Université de Paris VIII**
U.R.A. D1297 au C.N.R.S.
Equipe Textima
2. Rue de la Liberté
93526 SAINT-DENIS. CEDEX 02
(France)

When a news editor receives a series of agency wires on a current topic within his/her area of competence and must write a story that summarizes the facts in them, routine professional experience triggers a series of cognitive operations wich enable the editor to **acquire** pieces of information (that can be integrated into the conceptual frames instantiated by the general structure of these texts), **update** and **order** them (temporally, causally, teleologically) in such a way that the outcome of the reading process is a mental (Johnson-Laird, 1983) or situational model (van Dijk and Kintsch, 1983) of the information in the wires.

This paper presents a stepwise description of a sequence of cognitive processes enabling an experienced news editor to acquire, update, and organize pieces of (generally randomly ordered) information about a macro-event from a series of wires.

## 1. KNOWLEDGE UPDATING THROUGH TEXT PROCESSING

In the early 80s, the Schank and Abelson group (see De Jong, 1982), as well as Larsen (1982) took an interest in analysis and simulation of knowledge updating from informational text processing, in particular news articles .

The Fluent Reading, Updating and Memorization Program (FRUMP) developed by Schank and Abelson at al. is equipped to summarize stories via what are termes "sketchy scripts" which contain a given number of open slots. For example the "vehicle accident" script contains six open slots for type of vehicle, object of collision, geographical location, number of deaths, number injured, and allocation of fault. When FRUMP encounters a newpaper article identified as a vehicle accident, it instantiates this script and searches and retains the information that fits in the slots. It disregards all other information.

When FRUMP encounters a new article on the same type of vehicle accident, it can update its knowledge by **insertion** of data in an unfilled slot from the first article, or by **substituting** information from the new article if the new information differs from information in slots filled in the previous article. For example, if a vehicle accident article on day D reports 1 death and 1 injured, and the day D+1 article states 2 dead and no injured, FRUMP instantiates a comparison procedure for data in these slots and replaces day D information by day D+1 information on the basis of the rule that the information in the second article is more accurate than the information in article 1 unless proved otherwise.

Larsen (1982) discusses three weaknesses of the FRUMP program :
1. Since the open script variables must be defined in advance, FRUMP (in contrast to human readers) cannot retain unusual information.
2. FRUMP can only handle short descriptions of news stories where events are instantaneous. It cannot incorporate chains of subsequent events.
3. FRUMP automatically deletes obsolete information after the substitution process. Thus there is no way of storing previous slot values.

Larsen suggests that FRUMP's shortcomings can be overcome by creating four types of **updating corrections** :
- no change of old slot value
- insertion of value in empty slot
- substitution of new value for old value
- addition of new variable to accept unexpected values.

He also stresses the importance of storage of outdated slot information in memory. Larsen also discusses the role of bottom-up and top-down processes in knowledge updating. He defines two types of recognition modes used in the updating of the existing knowledge base. Preferential use is made of the **bottom-up mode** (comparison of pieces of information) when the reader processes a second article and does not intend to update his/her knowledge. The **top-down mode** (script-driven) supplants the bottom-up mode when existing knowledge needs to be extended or updated.

## 2. A SKETCHY MODEL OF COGNITIVE OPERATIONS IMPLEMENTED IN THE UNDERSTANDING OF A SERIES OF AGENCY WIRES

The comprehension process implemented by an experienced reader (a news editor for exemple) involves sorting the information in the wires as a function of topic and newsworthiness. This process can be broken down into the series of cognitive operations summarized below.

Agency wires are teletyped continuously to news desks and are immediately channelled to the news editor who handles this type of macro-event (in the example below, the editor is a Middle East specialist). The frame (see § 3 below) provides him with classificatory and evaluative information before he begins to read.

The four new wires presented below serve as examples of the cognitive operations involved in understanding a series of wires of the hijacking of a Middle East Airlines jet in June 1985 (henceforth MEA1, ..., MEA 4). The editor skims the frame of wire MEA 1 and reads the key ***hijacking-Lebanon***, the instructional field ***lead***, the prefix 0 and the beeps : "", and the title ***A MEA jet hijacked at Larnaca***. This instantiates the conceptual frame "agency wire" (comparable to the sketchy scripts in the FRUMP program described in § 1) which allows him to extract classificatory and evaluative information from the articles.

The frame can be represented as follows:

| | |
|---|---|
| KEY : | information on topic : .................. |
| INSTRUCTIONAL FIELD | one wire / first wire / summary : .......... |
| PREFIX : | priority 1/2/3/4 |
| BEEPS : | newsworthiness 1/2/3/4/5/6 |
| TITLE : | nature of macro-event : ......................... |

The news editor extracts the following information from MEA 1 (i.e. he assigns the following values to the variables of the conceptual frame) :

- MEA 1 is about the hijacking of a plane in Lebanon
- MEA 1 is the lead wire in a series covering the sequence of events surrounding the hijacking
- the information has been assigned low priority (prefix 0 by the French News Agency "Agence France Press"),
- the information is average in newsworthiness (4 beeps out of a maximum of 6),

- the hijacked airplane is a MEA jet and the event took place in Larnaca (Cyprus).

Reading the key is probably enough to activate the news editor's "hijacking" frame (or another more restricted one based on the specific features of terrorist hijackings in the Middle East, see §3). Thus even before reading the wire the editor instantiates a mental model of an "agency wire" frame wich contains press wire specifications, and a "Middle East hijacking" frame based on his/her knowledge of the world and the Middle East.

While reading MEA 1 the editor activates a propositional representation of the facts in the title wich **feeds** mental model 1 directly, or indirectly via bottom-up inferences. In other words certain empty or provisional default slots are filled. The end result of the reading of MEA 1 is mental model 2 composed of slots filled through propositional analysis. Bottom-up inferences and default information assign information to the remaining blanks in his "Middle East hijacking" frame.

MEA 2 (immediately followed by MEA 3) arrives 1 and 1/2 hours later on the news desk. The lag means that the propositional representation of MEA 1 is no longer in working memory. In all likelihood the news editor reads MEA 2 while instantiating mental model 2. It is highly improbable that the editor will first re-read MEA 1 when he receives MEA 2. In contrast he may read MEA 1 again after reading MEA 2 if MEA 2 does not allow him to construct mental model 3 integrating mental model 2 and the information from the second wire coherently (for example if there is no way of resolving conflicting information between values assigned to variables in the Middle East hijacking frame on MEA 1 and MEA 2).

To sum up, the sequence of cognitive operations is as follows :

1) activation of the "agency wire" frame for wire specifications,

2) activation of the "Middle East hijacking" frame deriving from the editor's knowledge of the Middle East,

3) construction of mental model 1 of the macro-event reported in wires 1 and 2,

4) propositional analysis of the first wire and generation of bottom-up inferences based on the editor's semantic knowledge base,

5) (in all probability simultaneously) construction of mental model 2 from wires 3 and 4,

6) storing of mental model 2,

7) reactivation of mental model 2 upon arrival of wire 2,

8) propositional analysis of wire 2,
9) construction of mental model 3 of the macro-event from wires 7 and 8,
10) loop 6 to 9 each time a new wire in the series is received.

This sketchy model assumes that propositional analyses and bottom-up inferences feed the mental model immediately. Clearly the interaction between propositional analysis, bottom-up inferencing and the continuous updating of the mental model of the macro-event is far more complex (see the Kintsch "construction-integration model" (1988) for the most recent attempt at formalizing these interactions).

Three separate operating systems are involved in the integration of the propositional analysis and the current mental model :
- **acquisition** of new knowledge which fill blank slots or default values,
- **updating** of outdated knowledge, by specification (for example when the editor learns from MEA 4 that the hijacked plane was a Boeing 727), or by substituting presumably accurate information for information now known to be erroneous (when for instance the editor learns in MEA 2 that only three members of the crew rather than the whole crew as indicated in MEA 1 were taken hostage by the hijacker),
- **ordering** (chronological and causal) of old and new knowledge.

Section 3 describes how mental model 1 is constructed. Sections 4-6 present an exemplary propositional analysis of the first wire and the stepwise building up of the final mental model from the events reported in the four wires.

## 3. CONSTRUCTION OF MENTAL MODEL 1 AFTER READING THE FIXED STRUCTURE OF MEA WIRE 1

The first cognitive operation performed by the editor when he receives the lead wire is the activation of the "agency wire" frame. This frame contains two sub routines (a) the "fixed" structure; (b) the "open" structure.

The fixed structure is an architecture of **instructional constraints** found in the Press Agency Handbook. The open structure handles **stylistic guidelines** (primarily the inverted pyramid principle : a wire starts with the most important information and then gives less important information in descending order in the following paragraphs ; the latter can be deleted starting from the end of the text without distorting the facts ; for a discussion, see Nef & François, 1987, pp. 206-208).

The instructional constraints can be expressed as substitution rules :

| | |
|---|---|
| (1) WIRE | -> COMMAND AREA + TEXT AREA + DATE |
| (2) TEXT AREA | -> INSTRUCTION (+PREFIX) + TEXT |
| (3) INSTRUCTION | -> KEY (+INSTRUCTIONAL FIELD) (+REFERENCE FIELD) |
| (4) PREFIX | -> FLASH/BULLETIN/URGENT |
| (5) TEXT | -> (BEEPS) + LEAD + DATELINE + $\Sigma$ |
| (6) DATELINE | -> ORIGIN + DATE (+AGENCY) |
| (7) $\Sigma$ | -> S* (= sequence of sentences) |
| (8) S | -> $P_{info}$ (+P/ADJUNCT$_{source}$) |
| (9) DATE | ->AGENCY + DAY + HOUR + MINUTE + MONTH + YEAR |

The command area is a code indicating type of wire and press organization receiving them. The text area contains a key for sorting wires, the instructional field, ("lead" in MEA 1), a reference field, and optionally a prefix ("flash", "bulletin", or "urgent" in decreasing order of newsworthiness ; MEA 2 is classified "urgent"), and then the text which is obligatory. The text proper may be preceded by beeps indicating the newsworthiness of the wire (up to 6 beeps), the title dateline (where the wire originated) date and source (for example Nicosia, June 12, AFP). $\Sigma$ indicates a text composed of several sentences (in MEA wires 1-4 each sentence was a new paragraph).

The fixed structure also acts as a parser of syntactic structure of each sentence in the wire, since each sentence introduces a piece of information from a new source (thus at least the first sentence should indicate this source) in the form of a matrix proposition $P_{source}$ (ex: *reliable sources indicate that P...*), an interpolated clause (ex: *P, as indicated by airport authorities*), or a source-adjunct (*P, according to reporters on the scene*). The end of the wire gives exact date and time the report was filed.

Readers familiar with agency wires have probably integrated these instructional constraints and possess a frame which allows them to locate classificatory and evaluative data immediately from the text architecture and identify sources from the body of the text.
The four agency wires below illustrate how a news editor can breakdown information using a "agency wire" frame.

**MEA WIRE 1**

| | |
|---|---|
| KEY | Lebanon-Hijacking |
| INSTRUCTIONAL FIELD | Lead |
| REFERENCE FIELD | Ø |
| PREFIX | Ø |
| BEEPS | ' ' ' |
| TITLE | **A MEA jet hijacked at Larnaca** |
| DATELINE | NICOSIA, June 12 (AFP) -- |
| $S1/P_{info}$ | A Lebanese Middle East Airlines jet on a daily flight from Beirut to Larnaca (Cyprus) was hijacked by a Palestinian and forced to land at 5:40 p.m., local time in Larnaca, (...) |
| $S1/P_{source}$ | (....) as indicated by airport authorities. |
| $S2/P_{info}$ | The Palestinian hijacker who released all the passengers and is holding the crew hostage, has demanded to meet with Cypriot reporters and foreign newsmen on the island to explain the motives for the hijacking. |
| $S2/P_{source}$ | Ø |
| $S3/P_{info}$ | At about 4:00 P.M. (G.M.T.), the plane is still on the runway of the airport which has been closed and where the police had reinforced its strength (...) |
| $S3/P_{source}$ | (...) as indicated by airport authorities. |
| DATE | AFP 12 (d) 18 (hr) 10 (mn) JUN 85 |

**MEA WIRE 2**

| | |
|---|---|
| DATE | AFP 12 (d) 19 (hr) 39 (mn) JUN 85 |
| KEY | Lebanon-Hijacking |
| INSTRUCTIONAL FIELD | Ø |
| REFERENCE FIELD | Ø |
| PREFIX | urgent |
| BEEPS | ' ' ' |
| TITLE | **Palestinian hijacker surrenders to Cypriot authorities** |
| DATELINE | NICOSIA, June 12 (AFP) -- |
| S1/$P_{info}$ | The Palestinian who was holding three crew members hostage on a jet belonging to the Lebanese Airline Company Middle East Airlines at Larnaca airport surrendered to Cypriot authorities at 20:15 local time (17:15 GMT). |
| S1/$P_{source}$ | (....) according to reporters on the scene. |

**MEA WIRE 3**

| | |
|---|---|
| DATE | AFP 12 (d) 19 (hr) 43 (mn) JUN 85 |
| KEY | Lebanon-Hijacking |
| INSTRUCTIONAL FIELD | add |
| REFERENCE FIELD | Ø |
| PREFIX | Ø |
| BEEPS | Ø |
| TITLE | **The Palestinian hijacker surrendered to Cypriot Authorities** |
| DATELINE | NICOSIE, June 12 (AFP) -- |
| S1/$P_{info}$ | The Palestinian hijacker relinquished the grenade used to threaten the crew and was placed on a Jordanian ALIA airline heading for Amman (....) |
| S1/$CIRC_{source}$ | (....) according to reporters |
| S2/$P_{info}$ | The plane on a scheduled daily flight from Beirut to Larnaca, had 151 passengers on board. |
| S2/$P_{source}$ | Ø |
| S3/$P_{info}$ | The hijacker had threatened passengers with a grenade during the flight and had released all the passengers when the plane landed in Larnaca at 14:40 GMT. |
| S3/$P_{source}$ | Ø |

**MEA WIRE 4**

| | |
|---|---|
| DATE | AFP 12 (d) 21(hr) 37 (mn) JUN 85 |
| KEY | Lebanon-Hijacking |
| INSTRUCTIONAL FIELD | Ø |
| REFERENCE FIELD | Ø |
| PREFIX | Ø |
| BEEPS | Ø |
| TITLE | **The Palestinian hijacker has left Cyprus** |
| DATELINE | NICOSIE, June 12 (AFP) -- |
| $S1/P_{info}$ | The Palestinian hijacker who took over a Lebanese Middle East Airlines Boeing 727 jet Wednesday for |
| three | hours left Larnaca (Cyprus) airport slightly after |
| 19:00 | GMT on board a Jordanian ALIA plane (....) |
| $S1/P_{source}$ | (....) airport authorities indicate. |
| $S2/P_{source}$ | Well informed sources indicate that .... |
| $S2/P_{info}$ | (....) the Jordanian plane was heading towards the Jordanian capital |
| $S3/P_{info}$ | The MEA Boeing cancelled its scheduled flight Wednesday evening to Beirut and will resume flights Thursday morning (....) |
| $S3/P_{source}$ | (....) indicated by a Middle East Airline spokesman |

Reading the lexical items **lebanon hijacking** and the key **larnaca** in the title should activate the news editor's "Middle East plane hijacking" frame presented above. The conceptual structure of this frame can be categorized into three systems derived from G. van Wright's (1967) time-action semantics (see Denhière and Baudet , 1988) as follows:

(a) **relational systems** - representing complex static situations
(b) **transformational systems** - representing complex events ("natural" changes from one state to another)
(c) **teleological systems** - (intentional for humans, functional for machines) representing individual's structure and behavior.

The crucial phase in hijacking negotiations involves at least two conflicting intentional systems: the hijacker's system and the intentional system of the authorities (and the hostages). This has been termed the "intentional intersystem" in the frame below.

**CONCEPTUAL ACTION FRAME (= MENTAL MODEL 1)**
**"Plane hijacking in the Middle East "**

(Default values are in brackets )

**A. Classification and Evaluation of Event**
**(see fixed structure for each wire)**

| | |
|---|---|
| * type of event | plane hijacking |
| * part of world | Middle East |
| * importance | ............................ |

**B. Sources of Information**
**(see open structure for each wire)**

**C. Sequence of Events**
**(see open structure for each wire)**

**Phase 1 : previous situation**
**(static relational system)**

| | |
|---|---|
| type of plane | ............................ |
| airline company | ............................ |
| destination | ............................ |
| number of passengers | (>Ø)..................... |
| additional information | ............................ |

**Phase 2 : hijacking :** (interaction of 2 intentional systems : hijacker, authorities, crew, passengers)

**2.1. Initial state (transformational system)**

| | |
|---|---|
| actor | (terrorist)............ |
| place | ............................ |
| time | ............................ |
| orders to crew | (land in a safe area for hijacker) |
| tactics | (taking hostages) |
| additional information | ............................ |

**2.2. Negotiations (intentional intersystem)**

| | |
|---|---|
| sequence of events | (contact with control tower ; request for fuel and/or negotiator) |
| means used by hijacker | (taking hostages) |
| goals of hijacker (intentional system) | (freeing prisoner/attracting attention of media) |
| tactics and goals of crew, passengers and authorities (conflicting intentional systems) | (arrest of hijacker, freeing any hostages) |
| outcome | ............................ |

**2.3 Final State (transformational system to initial variable system)**

| | |
|---|---|
| if hijacker surrenders: | |
| conditions for surrender | .............................. |
| if hijacker escapes: | |
| how escaped | .............................. |
| if hijacker is apprehended: | |
| how regained control of plane | .............................. |

---

**Phase 3: outcome situation (static relational system)**

| | |
|---|---|
| hijacker | .............................. |
| plane | .............................. |
| passengers | .............................. |
| crew | .............................. |
| moral implications | .............................. |
| political implications | .............................. |

## 4. CONSTRUCTION OF MENTAL MODEL 2 FROM READING OF MEA WIRE 1

Mental Model 2 is based on a propositional analysis of MEA 1 (enhanced if necessary by bottom-up inferences) and the frames activated by the fixed structure of the wire. The propositional analysis presented below is designed to yield **referential, case-valence-semantic and classificatory representations.**

**-reference:** a referential symbol is assigned to each description of a process or individual.

**-case-valence:** a) process actors are differentiated from complementary actors (based on semantic valence theory ). b) a case is assigned to each argument associated to a predicate (Agent of, Causer of, object of, experiencer of, etc .).

**-classification:** each process predicate is assigned to a conceptual class (action, process, state, etc., see François , forthcoming).

A propositional analysis of the title and the first sentence of MEA Wire 1 are presented below. The right hand column lists processes involved in recognition of syntactic and semantic properties in the left hand column. Properties are classified by level :
a) level 1 : syntactic and semantic analysis (possibly yielding two alternative descriptions)
b) level 2 : each string defining an individual or a process i.e. is assigned a reference symbol used to modelize semantic relationships between individuals ($X_i$, $X_j$) anaphora; between individuals and processes (r: $X_i$, $e_k$ ): case-valence; between processes (rr: $e_i$, $e_j$ ); interpropositional connections.

Propositional analysis of the three wires shows that numerous descriptions apply to the same individual.

*x1 : a Lebanese MEA company plane/ the plane/ the plane on its daily route from Beirut to Larnaca/ the airplane/ A Lebanese MEA Boeing 727/ The MEA Boeing....*

*x5 : A Palestinian hijacker/ the Palestinian hijacker/ the hijacker holding three crew members of a Lebanese MEA plane hostage at Larnaca airport/ the Palestinian/ the hijacker/ the Palestinian hijacker who took over a Lebanese airlines Boeing 727 for three hours Wesnesday.....*

*x26 : the grenade he used to threaten the crew/ a grenade*

*x27 : a Jordanian ALIA airlines plane headed for Amman/ A Jordanian ALIA airplane/ the Jordanian plane*

*x29 : Amman/ the Jordanian capital*

c) level 3 : The model also lists case-valence, interpropositional connections, and the logical and discourse properties of the processes derived from the processes in the right-hand column.

### PROPOSITIONAL ANALYSIS OF MEA WIRE 1
### (title and first sentence)

*Un avion de la MEA détourné à Larnaca*
*A MEA plane hijacked at Larnaca*

| | **identifier** |
|---|---|
| <u>syntactic/semantic analysis</u> | |
| PAST PART /"être détourné" ("un avion ...") | morphosyntactic of past part. |
| & part-of ("un avion", "la MEA") | interp. of preposition "de" |
| & loc ("être détourné", "à Larnaca") | interp. of preposition "à" |
| <u>Referential analysis</u> | |
| [un avion de la MEA] : $x_1$ | introduces new individual |
| [MEA] : x 2 | idem |
| [Larnaca] : x 3 | idem |
| [être détourné] : $e_1$ | introduces new process : action |
| <u>cases</u> | |
| OBJECT-OF (x1, e1) | subject of passive construction |
| CIRC-LOC (x3, e1) | locative object of non-loc. action |

Interpropositional connections

| | |
|---|---|
| e1 < e Statement | semantics of past participle in title |

SENTENCE 1

*Un avion de la compagnie libanaise MEA qui assurait la liaison quotidienne régulière B-L (Chypre)........*
*A Lebanese MEA airplane (making its) daily scheduled flight from Beirut to Larnaca (Cyprus)........*

syntactic/semantic analysis

| | |
|---|---|
| IMPERFECT TENSE / "assurer" ("qui", "la liaison quotidienne...") | syntax of "assurer" |

referential analysis

| | |
|---|---|
| [qui ] : $x_1$ | relative pronoun x 1 antecedent |
| [la liaison .... ] : $x_4$ | introduces new individual |
| [assurer ] : $e_2$ | introduces new process : activity |

cases

| | |
|---|---|
| CAUSE-OF ($x_1$, $e_2$) | inanimate gramm. subject of activity |
| OBJECT-OF ($x_4$, $e_2$) | direct object of activity |

*...... a été détourné par un pirate de l'air palestinien.......*
*...... was hijacked by a Palestinian*

syntactic/semantic analysis

| | |
|---|---|
| PAST TENSE / "être détourné par"("un avion ... Chypre","un.... palestinien") | syntax of "détourner" + passive |

referential analysis

| | |
|---|---|
| [un avion ... qui ...] : $x_1$ | coherence with MEA 1 title |
| [un palestinien] : $x_5$ | introduces new individual |
| [être détourné par] : $e_1$ | coherence with MEA 1 title |

cases

| | |
|---|---|
| OBJECT-OF ($x_1$, $e_1$) | subject of passive verb |
| AGENT-OF ($x_5$, $e_1$) | complement of agent |

*........et a atterri à 17:40 locales à Larnaca*
*........and landed at 17:40 local time at Larnaca*

<u>syntactic/semantic analysis</u>

| | |
|---|---|
| PAST TENSE /"atterrir" (".....") | tense + syntax of land |
| & Loc ("atterrir", "à Larnaca") | interp. preposition "at" |
| & Chrono ("atterrir" "à 17:40 locales ") | interp. preposition "at" |

<u>referential analysis</u>

| | |
|---|---|
| [Ø] : $x_1$ | ellipsis, antecedent x 1 |
| [Larnaca] : $x_3$ | see Title, MEA 1 |
| [17:40 locales] : $x_6$ | introduces new temp. individ. |
| [atterrir] : $e_3$ | introduces new loc. process |

<u>cases</u>

| | |
|---|---|
| OBJECT-OF ($x_1$,$e_3$) | gram. subject of loc process |
| SPATIAL LOC ($x_3$, $e_3$) | direct object. of loc. process |
| TEMPORAL LOC ($x_6$, $e_3$) | temporal complem. of loc process |

<u>interpropositional relations</u>

| | |
|---|---|
| e1 pendant $e_2$ | associating imperfect and past |
| e1 et ensuite $e_3$ | coordination of 2 past tenses |

*(......indique-t-on de sources aéronautiques.....)*
*(...as indicated by airport authorities....)* see analysis of "fixed" structure of wire

Mental Model 2 (identical architecture as the Middle East Hijacking frame) is built up from the knowledge acquired through propositional analysis of MEA wire 1. New information appears in the central column and missing information is indicated by "?" in brackets. The source of each piece of information appears in the right hand column. Only one inference is ambiguous. The wire states that a Palestinian hijacked the MEA airplane in Larnaca. On the basis of information in the first wire, the hijacking could have taken place during the flight, thus forcing the pilot to land in Larnaca, or alternatively while the plane was landing. The first hypothesis is more probable and the reader infers that there is a power struggle between the pilot and the hijacker, who forces him to land at Larnaca.

The second wire disambiguates this point since it states that the hijacker indeed forced the pilot to land at Larnaca but also that Larnaca was the original destination. This explains why the agency reporter preferred to use the term "took control" of the plane (in Wire 4) rather than the term "hijacked plane" used in the previous wires.

**MENTAL MODEL 2 FROM MEA WIRE 1**

| Conceptual Frame | Knowledge | Source |
|---|---|---|
| A. | | |
| type of event | plane hijacking | frame |
| part of world | Lebanon/Middle East | key |
| importance | average | Ø prefix<br>+ beeps : "" |
| B. news sources | airport authorities | S1, S3 |
| C. | | |
| phase 1 previous situation | | |
| type of aircraft | ? | |
| company | MEA | title |
| route | daily Beirut-Larnaca | S1 |
| number of passengers | ? | |
| other data | MEA Lebanese co. | S1 |
| phase 2 hijacking | | |
| 2.1 start | | |
| perpetrator | a Palestinian | S1 |
| where | ? | |
| when | ? | |
| demands on crew | land at Larnaca | frame inference<br>(default value) |
| tactics | ? | |
| other data | Ø | |
| 2.2 negotiations | | |
| sequence of events | 1. landing at Larnaca<br>17: 40 local time | S1 |
| | 2. release of passengers | S2 |
| | 3. crew held hostage | S3 |
| hijacker's tactics and motives: | | |
| tactics | taking of hostages | S2 |
| short term goal | meet with reporters | S2 |
| long term goals | explain motives | S2 |

| | | |
|---|---|---|
| tactics and goals of crew, passengers and airport authorities | 1. keep airline on ground | S3 |
| | 2. close airport to traffic | S3 |
| | 3. strengthen police force | S3 |
| result | (premature) | |
| other data | coverage up to 16: 00 GMT | S3 |
| 2.3 end of hijacking | (premature) | |
| phase 3 | (premature) | |

## 5. CONSTRUCTION OF MENTAL MODEL 3 AFTER READING WIRES 1-3

As stated above (§ 2), generation of a propositional analysis for each wire and its resulting mental model may be less cumbersome, but grossly misrepresents the probable degree of complexity involved in acquisition, updating and ordering of knowledge as wires reach the news desk. However since only fuzzy hypotheses can be made on the interrelations between these cognitive operations, a simplified presentation serves present purposes.

Note however than mental model 3 can be built up either from a) the integration of the propositional analysis of wires 2 and 3 (combined since wire 3 arrived 4 mins after wire 2 suggesting that they were read one after the other) and mental model 2 constructed from the reading of MEA 1 ; b) or alternatively, if ambiguities are spotted, by integration of a propositional analysis of MEA 1 (after re-reading).

However, the only ambiguity which would account for re-reading MEA 1 after processing MEA 2 is the problem mentioned above of the adequacy of the terms *détourner/détournement* ("hijack/hijacking") to describe the macro-event since the plane landed where it was supposed to land. It is likely that the reader in this case simply substituted "the plane landed where the hijacker decided to land" in his/her conceptual frame for the original "the plane landed in a different airport than its original destination".

Space prevents me from reproducing propositional analyses for wires 2 and 3 and I will move directly to mental model 3. New or updated knowledge has been underlined. The effect of updating and ordering of knowledge is striking : as expected, new information deals with the ongoing events in the hijacking during the half hour period separating MEA 1 from MEA 2 ; the information is about phases 2.3 (the outcome of the hijacking) and 3 (final state).

The circumstances surrounding surrender raise the problem of the outcome of negotiations. Since the hijacker was transferred to a Jordanian plane heading for Amman, the reporter probably used a default procedure to infer that transfer one of hijacker's ultimatims for surrender and release of the 3 remaining hostages.

Several pieces of information contained in MEA 3 refer to previous phases :

- the number of passengers on board (phase 1 previous situation)
- the information that the "hijacking" began **in flight** and not while the plane was landing at Larnaca as the reader may have correctly inferred from MEA 1 (phase 2.1 : start of hijacking)
- information about hijacker tactics (threatening the crew with a grenade).

The other piece of ambiguous information in MEA 1 is that the **crew** was taken hostage. This conflicts with later information indicating that the hijacker only took **three crew members** hostage. Here the experienced reader may adopt an informativeness maxim that current information replaces previous information. However (as the AFP reporter's Handbook points out) this maxim only applies when AFP prints a summary wire that officially cancels all earlier wires on the same macro event. An (unpublished) study on a longer plane hijacking indicates that the number of passengers presumed to be on board and the number of hijackers can vary considerably from one wire to another without AFP making explicit mention of the contradiction in figures between wires.

| MENTAL MODEL 3 FROM WIRES 1, 2 and 3 | | |
|---|---|---|
| **Conceptual Frame** | **Knowledge** | **Source** |
| A. | | |
| type of event | plane hijacking | MEA 1 : frame |
| part of world | Lebanon/Middle East | MEA 1 : key |
| importance | average | MEA 1 : prefix Ø beeps "" |
| | MEA 2 fairly important | MEA 2 : prefix "urgent" |
| B. news sources | MEA 1 airport auth. | MEA 1 S1 + S3 |
| | MEA 2 and 3 : reporters on the scene | MEA 2 : S1, MEA 3, S1 |
| C. | | |
| phase 1 previous situation | | |
| type of aircraft | ? | |
| company | MEA | MEA 1 : title |
| route | daily Beirut-Larnaca | MEA 1 : S1 |
| number of passengers | 151 | MEA 3, S3 |
| other | MEA Lebanese co. | MEA 1, S1 |
| Phase 2 hijacking | | |
| 2.1. : start | | |
| perpetrator | 1 Palestinian | MEA 1 : S1 |
| where | during flight | MEA 3 : S3 |
| when | idem | idem |
| orders to crew | land at Larnaca | MEA 1 : frame inference |
| tactics | threaten crew with grenade | MEA 3 : S3 |
| 2.2. : negotiations | | |
| sequence of events | 1. landing at Larnaca | |
| | 17:40 local time | MEA 1 : S1 |
| | 2. release of passengers | MEA 1 : S2 |
| | 3. taking 3 crew members hostage | MEA 3 : S1 |
| hijacker's tactics and goals | (intentional system) | |
| tactics | taking of hostages | MEA 1 : S1 |
| short term goals | meet with reporters | MEA 1 : S1 |
| long term goals | explain motives | MEA 1 : S1 |

| | | |
|---|---|---|
| tactics/goals<br>of crew, passengers and<br>airport authorities | | |
| airport authority tactics | 1. keep plane on ground | MEA 1 : S1 |
| | 2. close airport to traffic | MEA 1 : S3 |
| | 3. increase police force | MEA 1 : S3 |
| outcome | agreement of both parties to place hijacker on plane for Amman | backward-inference |
| 2.3. : end | | |
| how resolved | negotiated surrender | MEA 3 : title |
| means | hijacker relinquishes grenade | MEA 3 : S1 |
| phase 3 : outcome situation | | |
| hijacker | in plane heading for Amman (ALIA Jordanian Airlines) | MEA 3 : S1 |
| passengers | released | MEA 1 : S2 |
| crew held hostage | released | frame inference |
| moral implications | ? | |
| political implications | ? | |

## 6. MENTAL MODEL 4 CONSTRUCTED FROM THE FOUR MEA WIRES

As mentioned above (§ 4), the description of individuals can vary from one wire to another, in particular x1 (the airplane) and x5 (the hijacker). Propositional analysis of MEA wire 4 encounters a similar problem regarding event reference.

The title of MEA 4 indicates that the hijacker left **Cyprus** whereas sentence 1 states that **he left Larnaca airport** and specifies means of transportation and time of departure. From a purely logical standpoint, leaving Cyprus does not imply leaving the Larnaca airport or the reverse. Should the two action descriptions be classified separately ? Clearly in terms of information value it is more important to know that the hijacker left Cyprus than to know that he left Larnaca airport. This is why the reporter chose to put Cyprus in the title. Indicating **Larnaca airport** in the first sentence adheres to the inverted pyramid rule, i.e. the first sentence should state the when/how and why of the departure (exact place, time, means of transportation) which readers are assumed to want to know from reading the title.

As in § 5 I will omit the propositional analysis for MEA 4 and turn to mental model 4. Like mental model 3, mental model 4 contains new information in the title and reports information on the sequence of events after the hijacker surrendered during the 2 hour interval between MEA 2-3 and the filing of MEA 4:

- the take off of the ALIA plane to Amman with the hijacker on board
- the change in scheduled time for the flight back to Beirut of the hijacked MEA plane

However this wire also contains information on the situation prior to the hijacking : the number of passengers. Thus in integrating propositional analysis into the ongoing construction of the mental model, the datum "a Lebanese MEA Boeing 727" fills the empty slot for the heading "phase 1 : type of plane".

**MENTAL MODEL 4 GENERATED FROM READING OF THE FOUR MEA WIRES**

| Conceptual Frame | Knowledge | Source |
|---|---|---|
| A.<br>type of event<br>part of world<br>importance<br>no beeps | <br>plane hijacking<br>Lebanon/Middle East<br>average<br><br>MEA 2 average news value<br>MEA 4 low news value | <br>MEA 1 : frame<br>MEA 1 : key<br>MEA 1 : prefix Ø<br><br>"urgent" |
| B. news sources | MEA 1 airport auth.<br>MEA 2 and 3 : reporters on the scene | MEA 1 S1 + S3<br><br>MEA 2 : S1,<br>MEA 3, S1 |

| | | |
|---|---|---|
| C. | | |
| phase 1 : previous situation | | |
| type of aircraft | Boeing 727 | MEA 4 : S1 |
| company | MEA | MEA 1 : title |
| route | daily Beirut-Larnaca | MEA 1 : S1 |
| number of passengers | 151 | MEA 3, S3 |
| other | MEA Lebanese co. | MEA 1, S1 |
| Phase 2 : hijacking | | |
| 2.1. : start | | |
| perpetrator | 1 Palestinian | MEA 1 : S1 |
| where | during flight | MEA 3 : S3 |
| when | idem | idem |
| orders to crew | land at Larnaca | MEA 1 : frame inference |
| tactics | threaten crew with grenade | MEA 3 : S3 |
| 2.2. : negotiations | | |
| sequence of events | 1. landing at Larnaca 17:40 local time | MEA 1 : S1 |
| | 2. release of passengers | MEA 1 : S2 |
| | 3. release of remaining crew members | MEA 3 : S1 |
| hijacker's tactics and goals | | |
| tactics | taking of hostages | MEA 1 : S1 |
| short term goals | meet with reporters | MEA 1 : S1 |
| long term goals | explain motives | MEA 1 : S1 |
| tactics/goals of crew, passengers and airport authorities | | |
| airport authority tactics | 1. keep plane on ground | MEA 1 : S1 |
| | 2. close airport to traffic | MEA 1 : S3 |
| | 3. increase police force | MEA 1 : S3 |
| outcome | agreement of both parties to place hijacker on plane for Amman | backward-inference |
| 2.3. : end | | |
| how resolved | negotiated surrender | MEA 3 : title |
| means | hijacker relinquishes grenade | MEA 3 : S1 |

| | | |
|---|---|---|
| phase 3 : outcome situation | | |
| hijacker | in plane heading for Amman | MEA 3 : S1+ MEA 4 : S2 |
| | ALIA Jordanian Airlines | |
| | plane took off shortly after 9:00 GMT | MEA 4 : title+ S1 |
| passengers | released | MEA 1 : S2 |
| crew held hostage | released | frame inference |
| plane | return flight to Beirut postponed to Thursday | MEA 4 : S3 |
| moral implications | ? | |
| political implications | ? | |

Attentive reading of the four wires reveals that at least three fairly important pieces of information were not stated but rather were inferred by default. The three inferences are the following :

(a) None of the wires explicitly state that the hijacker met with journalists to explain why he hijacked the plane. Thus we have no information about the hijacker's motives. However it is likely that the hijacker only agreed to surrender because this condition had been fulfilled. Second, it is obvious that this condition is less problematical (reporters simply agree to being present) than the release of political prisoners or payment of a ransom, both of which require government approval. Further it is apparent from the fact the hijacker released all the passengers and some of the crew immediately upon landing in Larnaca that he was attempting to convince the Cypriot authorities to let him stage a media operation. To arrive at this conclusion, the attentive reader must make a number of inferences on the basis of his knowledge of the world, Middle East conflicts, and the tactics used by all sides to achieve their goals.

(b) We do not know why the hijacker was transferred to a Jordanian airplane headed for Amman. The transfer is obviously favorable to the hijacker --- since he is unlikely to be apprehended on Jordanian soil --- and suggests that going to Jordan was one of the preconditions for surrender.

(c) AFP makes no statements concerning the moral or political implications of the hijacking. The informativeness maxim implies that the implications are practically nil : if there were important ramifications, the AFP would have filed a supplementary wire devoted to comment and reactions to the hijacking or a sumary wire.

## 7. CONCLUSION

The pivotal feature of this sketchy model of the processes involved in the understanding of the four AFP wires on the MEA hijacking is the construction of a mental model of the macro event presented in discontinuous form in the four wires (see Johnson-Laird, 1983 and Schnotz, 1988). This however does not imply that propositional analysis is less important. Mental model 1, derived from the "Middle East Hijacking" frame which is activated by the key and the title of the first wire, is practically empty and the few slots that are filled contain default information (for example the assumption that a hijacking in the Middle East is a terrorist hijacking designed to obtain the release of political prisoners or create a media event). Slots are filled primarily through propositional analysis. However propositional analysis itself is not equipped to recognize the referential, chronological, causal or teleological coherence of the macro event reported in the wires. This presentation of the functions of propositional analysis (forming the text base) in conjunction with a mental (or situation) model thus fits the description of the cognitive model described by van Dijk and Kintsch (1983, chapter 10).

## NOTE

The study reported in this paper was initially aimed at analysis of temporal and aspectual relations in short French and German narrative texts (in collaboration with F. Nef, see François 1984; François and Nef, 1984; Nef and François, 1987; François, 1987) and adheres closely to cognitive text semantic model developped in Constance (West Germany) by the Christoph Schwarze (1987, 1989) and Eberhard pause (1988a, 1988b) groups. The author wishes to thank Guy Denhière for introducing him to the methodology used by cognitive psychologists in analysis of informational text comprehension and Serge Baudet, Frédéric Nef, Eberhard Pause, Marie-Thérèse Schepping, Wolfgang Schnotz and Christoph Schwarze for their careful reading and helpful comments on earlier versions of this paper.

## REFERENCES

AGENCE FRANCE-PRESSE. *Manuel de l'agencier.* Paris, manuscrit non daté.

DE JONG, G (1982). An overview of the FRUMP system. In W. LEHNERT & D. RINGLE, *Strategies for natural language processing.* Hillsdale, N.J. : Erlbaum, 149-176.

DENHIERE, G. & BAUDET, S. (1989). Cognitive psychology and text processing : from text representation to text-world. SEMIOTICA, Special issue : *Cognition and Artificial Intelligence.* P. Ouellet (ed.), 77, 1/3, 271-293.

FRANCOIS, J. (1984). Le cheminement du temps narratif. Essai d'interprétation temporelle de mini-séquences narratatives du français et de l'allemand. *DRLAV*, 31, 81-96.

FRANCOIS, J. (1987). Résolution d'ambiguités aspectuelles dans la traduction (le problème du transfert morphosémantique du prétérit allemand en français). In C. FUCHS (ed.) *L'ambiguité et la paraphrase.* Centre de publications de l'Université de Caen, 117-122.

FRANCOIS, J. & NEF, F. (1984). Deux méthodes convergentes d'analyse et de synthèse de l'enchaînement événementiel dans les micro-récits. *Cahiers de Grammaire*, 8, Toulouse : Université Paul Sabatier, 77-115.

JOHNSON-LAIRD, P. (1983). *Mental models. Towards a cognitive science of language, inference, and consciousness.* Cambridge : Cambridge University Press.

KINTSCH, W. (1988). The role of knowledge in discourse comprehension : a construction-integration model. *Psychological Review*, 95, 2, 163-182.

LARSEN, S.F. (1982). Knowledge updating in text processing. In A. FLAMMER & W. KINTSCH (eds.), *Discourse processing.* Amsterdam : North-Holland, 205-218.

MACKIE, J.L. (1974). *The cement of universe. A study of causation.* Oxford : Clarendon Press.

NEF, F. & FRANCOIS, J. (1987). Qu'apportent les modèles cognitifs à la sémantique du discours ? Reconnaissance des relations temporelles entre événements et constitution d'un univers de savoir à la lecture d'une dépèche de presse. *DRLAV*, 36-37, 199-241.

PAUSE, E. (1988a). Anaphern im Text. Fachgruppe *Sprachwissenschaft Universität Konstanz*, Arbeitspapier Nr 4.

PAUSE, E. et al. (1988b). "Plötzlich startet eine ducati" : das Textverstehens-modell von Con-3-Tra. Fachgruppe *Sprachwissenschaft Universität Konstanz*, Arbeitspapier Nr 2.

SCHNOTZ, W. (1988). Textverstehen als Aufbau mentaler Modelle. In H. MANDL & L. SPADA (eds.), *Wissenspsychologie.* München : Psychologie Verlags Union, 299-330.

SCHWARZE, C. (1987). Text understanding and lexical knowledge. In H. VERSCHUEREN & L. BERTUCCELLI-PAPI, (eds.), *The pragmatic perspective*. Amsterdam : Benjamins, 587-612.

SCHWARZE, C. (1989). Polysemie als Prozedur, am Beispiel von frz. A TRAVERS und CHEZ. In C. HABEL, D. HERWEG, & L. REHKAEMPER, (eds.) *Raumkonzepte in Verstehensprozessen*. Tübingen : Niemayer.

VAN DIJK, T. A. & KINTSCH, W. (1983). *Strategies in discourse comprehension*. New-York : Academic Press.

VON WRIGHT, G. H. (1967). *The logic of action : a sketch*. In : N. RESCHER (ed.) The logic of decision and action. Pittsburgh : University of Pittsburgh Press.

This Page Intentionally Left Blank

# COHERENCE: PERSPECTIVES IN LINGUISTIC AND COGNITIVE RESEARCH

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# ASPECTS OF TEXTUAL CONTINUITY LINGUISTIC APPROACHES

Michel CHAROLLES and Marie-France EHRLICH

UFR des Sciences du Langage
Univ. de Nancy II
23 Bld Albert 1er
54015 NANCY

Lab. de Psycho. Expérimentale
Univ. Paris V, EPHE, EHESS,
CNRS UA 316 - 28 rue Serpente
75006 PARIS

Among the operations involved in discourse comprehension, the most fundamental are those which allow the subject to recognize or construct how text contents are ordered and form a whole, because it is these that are the essence of textuality itself. This is clear in that what differentiates a text from a random sequence of sentences (a non-text) is precisely its continuity and homogeneity.

The issues of text coherence are thus central both for the linguist, whose interest lies in the description of discourse structure and its governing principles, and for the psycholinguist, whose goal is to gain a better understanding of how subjects produce and understand texts.

Beyond the simple observation that both linguists and psychologists have studied text coherence, it is extremely difficult to single out what constitutes a shared empirical base, or how one field has borrowed from the other. This is not surprizing in that objects and methodologies in linguistics and psychology are not identical. Furthermore, theoretical frameworks, often divergent within a given field, are even more so when two scientific fields are involved. In order to define common ground and to pave the way for future collaboration, it seems useful as a first step to chart the changes that have taken place as each of these fields has explored issues in textually. Therefore, we have decided to present a two-part survey devoted to research in linguistics (in this section) and research in psychology (in the next section). We do not claim that either of these presentations is exhaustive, since the literature in each of these fields is enormous. Our aim is more modestly to describe the main orientations in current research. We hope in this way to accomplish two goals:

- to provide a groundwork which should help situate new studies to be presented here;

- to set the stage for a general discussion regarding the potential interest of an intersection of approaches in linguistics and psycholinguistics.

This overview of developments in linguistics begins with a discussion of how text grammarians studied issues of coherence and text continuity until the mid 1970s, and how their methods differed from procedures in discourse analysis inspired by the works of Benveniste and Jakobson. We will then show how a number of recalcitrant examples led text linguists in the late 1970s to view coherence in a new light, and to adopt a radically different model, termed "procedural". Interest in linguistics in the 80's has focused on identifying the roles and nature of different markers of text continuity or connectedness used for text composition and construction in different languages. We will close with a discussion regarding current issues in the growing body of literature devoted to markers of text cohesion and connexity.

## 1. COHERENCE IN TEXT GRAMMARS

Starting from the Chomskian principle that the native speakers of a given language L are able to produce and understand an infinite number of well-formed discourse sequences in language L, text grammarians proposed (1) that speakers share the ability to distinguish a sequence of grammatically acceptable sentences making up a text from a sequence of grammatically acceptable sentences failing to do so, (2) can recognize structural similarities between superficially different texts, and (3) can summarize a given text while maintaining a given text structure. The proponents of this model (Harweg, Petöfi, Gülich, Ballmer, Rieser, Dressler, Van Dijk), argue that all sequences of grammatically acceptable sentences are not equally acceptable as texts, since if this were the case, it would be unnecessary to define the set of rules governing text construction. According to this model, example 2 is a text but example 1 is not "grammatical", to use Van Dijk's term (1972):

1. Peter is a botanist. The sum of the angles of a triangle is 180°.

2. Peter is a botanist. He loves nature.

Examples and counter examples of this nature quickly led researchers to recognize the importance for text coherence of both linguistic markers of continuity (pronominal anaphora, definite descriptions, conjunctions, etc.) and of the knowledge readers bring to the text. Authors posited that the presence of continuity markers do not guarantee, and may not even be necessary for a sequence to be perceived as coherent (cf. Harweg, 1968; Bellert, 1970), they also investigated the role of intersentential connectors...(see in particular Gülich, 1970; Gülich and Raible, 1977). All these "surface" (as they were termed) or syntactic markers of text continuity, in addition to markers of temporality (Koch, 1973), presuppositional implication, and topicalization, were formalized by Van Dijk in 1972. Aside from elements of coherence thought to be accessible to all native speakers because they are conventional, researchers were quick to point out the importance of the role of external knowledge, especially its effect on interpretation of quasi-implications (as they were called by Bellert, 1970) that serve to link sentences. Authors such as Van Dijk and Petöfi proposed theories incorporating world and related belief components in their text models. For instance, Petöfi's TestWest Theory (1973) associates a set of worlds and sub-worlds (related to personal, temporal, local instances or to propositional attitudes) with type of text. Van Dijk (1978a, b) extended the notion of coherence to relevance, defined as those intentions vehiculed by the illocutionary force in the sentences making up the text.

This generally accepted framework allowed researchers to concentrate on more specific aspects of text, such as the role of topicalization or the lexicon in a text grammar (Petöfi, 1975, 1977), and the construction of macrostructures (Van Dijk, 1977).

## 2. COHERENCE IN RESEARCH ON DISCOURSE AND ENONCIATION

During this same period (1965-1975), work in France in particular was deeply influenced by the theories of Benveniste and Jakobson, based on postulates diametrically opposed to those of the text grammarians.

In "Problèmes de linguistique générale" (1966), Benveniste takes the position that languages are structured on different levels: the merismatic, phonematic, morphemic and sentence levels. The constituent units of these levels combine distributionally (i.e., contrast for a given level) and hierarchically (from one level to the next). Thus, merisms combine to form phonemes which combine to form morphemes and so on. The two boundaries are a lower limit formed by the merisms, which have no constituent smaller units, and an upper limit, the sentence, which although resulting from the combining of morphemes does not itself integrate into a higher level. According to Benveniste, a sentence cannot be considered a constituent unit of a larger whole because there are no rules governing formation beyond the sentence level. Once on a sentence level, we are outside the domain of language defined as a system of signs, and we enter another universe, where language is defined as an instrument of communication (1966, pp.129-130). A similar argument appears in Jakobson (1963): linguistic units are arranged in an ascending order, when we reach the top of the ladder, the combination of phrases into sentences, the syntactic restriction rules cease to operate, and all combinations are acceptable <u>a priori</u>.

However, Benveniste and Jakobson do not conclude that there is no linguistics beyond the sentence level. According to Benveniste, the sentence level constitutes the boundary of the semiotic universe (i.e., a system order), whereas beyond the sentence level is a semantic universe, the universe of "la parole", which is the manifestation of language in communication. The only object of study in this universe is how the subject appropriates the formal apparatus of language to signify his position as a speaker through the use of specific markers (1970, p.14). Thus, the goal of discourse analysis is to assess how speakers use the possibilities available to them in a language to mark their subjectivity in different communicative situations. In the numerous studies on enonciation inspired by Benveniste and Jakobson, (Dubois, 1969; Provost-Chauveau, 1971; Guespin, 1971, 1976; and for an overview Maingeneau, 1976; Kerbrat-Orecchioni, 1980; Charolles, 1986) and also in studies generated by Harris (1952) (cf. Bierwisch, 1965), problems of coherence receive practically no attention; in fact, the issue does not even arise. These studies are descriptive and contrastive; they were not aimed at identifying empirically testable text composition rules-rather, they were designed to shed light upon formal regularities in the use of markers of subjectivity (person, modalizations, etc.) and the social relations between speakers (thus explaining the interest in political corpora).

Thus the principles, aims and methods guiding research on enonciation in general, and the analysis of political discourse in particular which flourished in France from 1970-1980 were at the antipodes of text grammar. The coexistence of these two approaches underscores the fundamental question of whether there are criteria for acceptability on the text level. Without going so far as to adhere to Benveniste's and Jakobson's positions, text grammarians were gradually forced to admit that it is far from easy to state when a text (i.e., a sequence of two sentences) is coherent and when it is not.

## 3. CHANGES IN THE EMPIRICAL BASIS OF TEXT GRAMMAR

By the mid 1970's, researchers had gradually realized that certain examples of incoherent sequences were in fact not as incoherent as had been thought. In his thesis, Van Dijk (1972) gives the following example of an "agrammatical" sequence:

> 3. We will have guests for lunch. Calderon was a great writer.

This text indeed appears to be incoherent since the two sentences composing it seem to have been chosen at random. Nevertheless, Dahl and Dahl (1974) rightly point out in their critical review of Van Dijk (1972) that the sequence is not absolutely devoid of coherency. It is only necessary, these authors say, to imagine a family who celebrates Calderon's death every year with a meal, in order for the sequence to cease to appear incoherent. (Similar remarks have been made by Lenhert, 1979, and Randquist, 1985). This type of discussion is highly characteristic of debates on coherence during this period. Each time an author proposed a set of text coherence rules (see for example recent works by Reinhart, 1980, Giora, 1985), another author was able to contest the generality of these rules. Once the imagination is put to work to create ad-hoc communicative situations, the connection between the most obscure sequence of sentences can become feasable. This clearly demonstrates that there are no rules governing correct text-formation which are generalizable to all situations and will receive unanimous approval, as was the case for syntax rules on the sentence level.

Thus, the position originally adopted by text grammarians that certain sequences of sentences or macropropositions could be inherently coherent or incoherent no longer appeared valid. This reassessment of the empirical bases of text grammars, and -more broadly speaking- the new perspective on coherence phenomena that gradually took hold at the end of the 70's had a number of important consequences.

## 4. NEW ORIENTATIONS

### 4.1. Conceptual and terminological distinctions

Although the terminology is not entirely consistent, there is general agreement as to the distinction between coherence, which defines text interpretability, and markers of relationships between sequences or constituents of sequences. As a result of Halliday and Hasan's work (1976) these markers are normally termed cohesion (see de Beaugrande, 1980; Dressler and de Beaugrande, 1981; Martin, 1983; Enkvist, 1989). Under the heading of cohesion authors tend to place phenomena such as anaphora (pronominal, nominal, deictic, possessive), repetition, ellipsis, connectors, presuppositional relations, and thematization markers. Some, such as Hatakeyama, Petöfi and Sozer (1984), who argue for a more sophisticated terminology, reserve the term cohesion for features related to thematic continuity, and contrast this term with connexity, which implies grouping together the other markers. However, since morphemes such as "but", "for", and "thus" are generally called connectors, the use of the term connexity to refer exclusively to them also has a certain following. In this case, an opposition is set up between anaphoric, thematic, etc. cohesion markers, which serve to signal identity, inclusion or association among constituents affecting text constitution, and connexity which signals relationships between propositional contents and/or speech acts and indicates text composition (for the difference between text constitution/text composition see Motsch and Viehweger, 1981; Gülich and Kotschi, 1983).

### 4.2. Coherence and Interpretability

Once coherence has been circumscribed, it remains to define what this notion covers. Most authors (Enkvist, Reiser, Danes, van de Velde, 1989) associate coherence and intelligibility within a communicative context (discourse context). Authors who have made notable contributions in this area include Petöfi, 1979; and 1985; Heydrich and Petöfi, 1981; Hatakeyama, Petöfi and Sozer, 1984; Petöfi and Olivy, 1986.

In Petöfi's model of theoretical meaning interpretation, the text content, once processed syntaxically and semantically (intension) using a canonical descriptive vocabulary, is associated with text worlds (T WR). This theoretical interpretation phase is independent of knowledge or beliefs the interpreter may have concerning the states of affairs expressed in the text, since the objective is purely to use formal cues (such as markers of time, place, persons, de dicto modalization, "world creative predicates") to represent the different complexes of states of affairs in the text (for one application, see Eco, 1979). At this point the interpreter has available a "text producer" model (TCo R = "text correlate") which can be compared against his/her own representational model -<mR>- of states of affairs. If the two models coincide, (compatibility or accessibility) the interpreter will judge the text to be expressing acceptable or plausible relationships, and therefore, as being coherent. If not, a reinterpretation will take place and the interpreter will "complement" the "text producer model" until he is satisfied. The system includes the possibility of correcting prior models in order to obtain an intelligible representation of what the speaker wishes to convey in the text. Because of the extensional dimension of the Petöfi system, it can be placed in the "theory of models" framework, illustrated among others by Kamp's (1979) "structures of discourse representation". His formalism, highly divergent from Petöfi's, is oriented towards a description of temporal reference and makes use of categories (events, states) which are not found in TestWest.

### 4.3. Coherence as a principle of interpretation

Although coherence is not a property of texts, and although it is impossible to make absolute judgments as to whether a sequence is coherent or not, it can be shown (Charolles, 1983) in contrast that coherence is a type of a-priori requirement for discourse processing. Coherence, in this light, is a principle of interpretation and reinterpretation comparable, and in fact highly similar, to the maxim of relevance in Grice (1975) and to the principle of coherence mentioned in Ducrot (1972). This principle can be justified by the fact that when subjects search for coherence, they do not stop at surface markers of sentence boundaries (e.g., they are not satisfied with a simple identity of arguments across two successive propositions).

### 4.4. Role of inference

Because the processes of interpretation and reinterpretation are governed by coherence rules, they lead the subject to construct relations which are not explicit in the text (see Van de Velde, 1981, 1985, 1989; Langleben 1981, 1983, 1985). Not all the inferences subjects make are linguistically based. Some are extrapolations from real knowledge supposedly common to the reader and the text. One of the aims of linguists working in this area is to define which inferences are structurally based and which are not. This question is difficult, and goes beyond the scope of pure linguistics. The answer depends essentially on what role is assigned to the lexicon. Linguists (see for example Petöfi, 1973; and Van de Velde, 1981) tend to weight the lexico-encyclopedic component of their systems a way so as to control a maximum of inferences while remaining structurally within the realm of linguistics.

### 4.5. Changes in forms and goals of models

Observations similar to the ones above led a certain number of researchers (Dressler and de Beaugrande, 1981; de Beaugrande, 1980; Eikmeyer, 1983; Rieser, 1989) to propose new models, or to adopt positions radically different from those defended by text grammarians in the 1970's. The aim of these "procedural" models or theories is not to produce rules for assessing the well-formedness of texts, but rather <u>to represent how subjects process and reprocess text when constructing their interpretation of meaning</u>. These models can be situated within the general framework as Petöfi's system described above, but they also integrate certain constraints which take into account psychological factors. For example, in Eikmeyer (1983) and Rieser (1989), the model is governed in part by an "immediacy assumption" stipulating that the system should be able to resolve any interpretational difficulty immediately (for example a pronominal ambiguity), and a "reanalysis assumption", which allows the subject to backtrack in order to rectify inferences. This system is constructed to process information on line, which explains its recourse to a parser (cf. Lundquist, in press). As research along these lines develops, the boundary becomes fuzzier between artificial intelligence and psycholinguistics. Authors are now forced to borrow from the problem-solving literature issues such as processing cost, strategies, and heuristics (cf. Van Dijk and Kintsch, 1983) which until very recently were not part of the linguist's conceptual repertoire.

## 5. LINGUISTIC ANALYSIS OF THE MARKERS OF TEXT CONTINUITY

A wealth of studies have examined markers of text continuity, but the watershed publication was Halliday and Hasan's <u>Cohesion in English</u> which appeared in 1976 (cf. Lundquist, 1980, for an example of the impact of this work). Work in linguistics (on a variety of languages) dealing with text cohesion and connexity are so numerous and are appearing so rapidly that no overview can be up to date. Even if we narrow our focus to anaphoric devices, the number of references is enormous. Nevertheless, several essential conclusions supported by the current literature may be made related to the issues developed here. The main task for the linguist is to define the specificity of each marker (or sets of markers) in a given language at a given moment. For example, imagine a speaker who realizes he has said something which may be confusing, and wishes to avoid having his listeners mistake one person for another. In French the speaker has a range of devices at his disposal: he can baptize his characters; he can use definite descriptions, identificatory clauses, or specific pronouns. However, these options are not equivalent, and are not equally powerful. Take for instance such very highly specialized anaphoric devices such as "this one", "that one", "the former", "the latter", "the one", or "the other one". Linguists reason that if all these possibilities are present in the language, they must all serve a specific function. The next logical step for the linguist is to test for the specific functions of these forms by constructing examples, and finding instances where one form is more effective than another, or where only one is acceptable. The same principle can be applied to paraphrastic reformulation markers such as in French "c'est-à-dire", "autrement dit", etc., which although semantically similar, in fact have clearly distinct functions (see Gülich and Kotschi, 1983). When these analyses are extended to all types of markers, it is evident that languages provide the means for extremely subtle coding of relational distinctions, which actually go beyond the categories traditionally reserved for them.

The interpretational cues associated with cohesion and connexity are difficult to describe in detail. We inevitably must use very fine descriptive categories even to explain the difference between "but" and "however" (cf. the difference between implication and what Ducrot calls argumentative orientation), which are hard to integrate into broad theoretical bases (cf. Ducrot, 1983; Roulet et al., 1983; Moeschler, 1985). Relatedly, it should be noted that the analysis of authentic conversations -aside from the fact that this analysis reveals the process of discourse construction and coherence- is the result of mutual and fundamentally interactive efforts on the part of the participants (cf. Schegloff, 1982). The use of authentic discourse draws attention to the types of markers and their range of use which often escape notice when linguists work on artificial corpora (on analysis of discourse more or less inspired by ethnomethodology see Sinclair and Couthard, 1975; Stubb, 1983; M. Couthard and Montgomery, 1981; Brown and Yule, 1983; Werth, 1981).

Although capable of providing complex interpretational information, linguistic markers of cohesion and coherence also usually appear to be underdetermined. In the case of pronominal anaphora, most devices are not sufficient in themselves to identify the antecedent. This is true for "missing antecedents", as in :

4. John is a guitarist because he thinks it is a nice instrument.

or opaque referents such as :

5. Oedipus wants to marry his mother. Jocaste avoids him.

or pronouns referring to a previous mention as in:

6. I saw a rhinoceros.
A what ? Spell it !

or to individuals located outside the boundaries of the "mental space":

7. In 1960 I loved my boss but now I hate him.

On these phenomena, see among others Karttunen, 1969; Bolinger, 1979; Hall, 1972; Yule, 1982; Bosch, 1983 and 1985; Conte, 1981; Fauconnier, 1983; Webber, 1980.

A number of markers are thus underdetermined, in that they contain insufficient information to automatically detect the antecedent, or the relationship between sentences. This implies that the amount of effort the interpreter needs to resolve ambiguity remains decisive, and that the cohesion/connexity -coherence distinction is, in the final analysis, far from resolved (cf. Martin, 1983).

Markers may be underdetermined, but they are at the same time sufficiently specific to make certain sequences unacceptable. For example, in French, a subsequent mention of a nominal expression by a second nominal expression can take the form of either a demonstrative or a definite description. However, depending on the co-text, the forms are not interchangeable. For example, in the following example, only the second sequence is possible:

8. J'ai vu une voiture. Cette voiture roulait vite.

*9. J'ai vu une voiture. La voiture roulait vite.

(I saw a car. This car was going fast/I saw a car. The car was going fast).

In the next pair of examples, only the definite form is acceptable:

10. J'ai vu une voiture et un camion. La voiture roulait vite.

*11. J'ai vu une voiture et un camion. Cette voiture roulait vite.

(I saw a car and a truck. The car was going fast/I saw a car and a truck. This car was going fast).

The explanation for this "paradox", termed "immediate nominal anaphore", presupposes a careful analysis of the way demonstratives and definite articles are used to refer to a preceding expression, ways which evidently differ markedly (cf. Corblin, 1983; Kleiber, 1986).

This last comment brings us to a discussion of the points raised in parts one and two, when we mentioned the antinomy between the positions of text grammarians and the followers of discourse analysis regarding whether there are criteria for acceptability beyond the sentence level. We can now conclude:

- Sequences such as (1, 2, 3), where there are no relational markers between the constituent sequences, are theoretically never unacceptable, since it will always be possible to invent a situation that suggests a plausible relationship between states of affairs.

- On the other hand, for sequences such as (4, 5...10, 11) which contain relational markers between sentences or parts of sentences, some are acceptable and some are not, since they contain markers having a certain degree of specificity in the language system and consequently cannot be used randomly.

The following pairs of examples illustrate in another way this point.

12. Max a presque la moyenne en maths, il a des chances d'avoir son examen.

*13. Max a presque la moyenne en maths, il a des chances de rater son examen.

*14. Max a à peine la moyenne en maths, il a des chances d'avoir son examen.

15. Max a à peine la moyenne en maths, il a des chances de rater son examen.

(Max has almost passing grades in math, he may pass the exam/fail the exam; Max has just passing grades in math, he may pass the exam/fail the exam);

The variable acceptability of the set 12-15 is due to the fact (Ducrot, 1983) that "presque" (almost) is positively oriented (towards a favorable outcome) which would be better-captured by a sentence like "Max has passing or better than passing grades in math", whereas "à peine" (just) suggests the reverse (i.e., "Max does not have passing grades"). These argumentative polarities are generated by the conventional meaning assigned to the quantifiers "presque" and "à peine".

The use of either of these in an utterance S1 thus places constraints on the set of utterances (S2...Sn). If we compare:

16. The King of France is bald because he had typhoid.

*17. The King of France is bald because France is a monarchy.

(examples drawn from Raccah, 1982) it can be seen that a continuation of a presupposition is unacceptable. 17 is poorly constructed because the author fails to introduce into S2 an utterance containing the cause of a fact denoted by a presupposed proposition in S1 (i.e., there is a king of France). Here again, the way information is presented is a constraint on sentence continuation. A speaker cannot introduce presupposed content (and thus treat it as though it were already tacitly recognized; or operate on the basis that it serves as a mutual framework for what follows) and then, immediately afterwards, explain it (as though the information could be placed in doubt) without introducing inconsequentiality.

The final examples deal with acceptability phenomena related to thematization:

18. Qui Pierre voudrait-il épouser ?
- Pierre, il voudrait se marier avec Sophie.

(Who does Pierre want to marry ? Pierre wants to marry Sophie).

*19. Qui Pierre voudrait-il épouser ?
- Sophie, Pierre voudrait se marier avec elle.

(Who does Pierre want to marry ? Sophie, Pierre wants to marry her).

In 18, "Pierre" works as a local presupposition (Martin, 1983), and can be thematized perfectly well in the response. In contrast, "Sophie" (which saturates the indefinite of the local presupposition "Pierre wants to marry someone") necessarily has a rhematic status and thus cannot be the object of a thematization in the answer, as shown in 19.

## 6. CONCLUSION

Markers of cohesion and connexity, which are designed to facilitate the interpretation and the construction of coherence, may also produce incoherence. Since these terms (or operations) have a specific function, set by convention, they cannot be used outside of the rules governing their use. When they are used in ways which go against their function, the result is incoherency or <u>lack of consequentiality</u>, to the extent that the listener/reader is unable to recover the intelligibility of the process which led to the production of the utterance.

## REFERENCES

de BEAUGRANDE, R., (1980), Text, discourse and process, Hillsdale, LEA.

de BEAUGRANDE, R. & DRESSLER, W., (1981), Introduction to text linguistics, London, Longman.

BELLERT, I., (1970), On a condition of the coherence of texts, Semiotica, 4, 335-363.

BENVENISTE, E., (1966), Problèmes de linguistique générale, T1, Paris, Gallimard, T2, 1974.

BIERWISH, M., (1965) Compte-rendu de Z.S. Harris, Discourse analysis reprints, Linguistics, 61-73.

BOLINGER, D., (1979), Pronouns in discourse, in E. Givon (ed), Syntax and Semantics, Vol.12, Discourse and Syntax, New-York, Academic Press.

BOSCH, P., (1983) Anaphora and agreement, New-York, Academic Press.

BOSCH, P., (1985), Constraints, Coherence, Comprehension, in E. Sözer (ed), Text connexity, Text coherence, Aspects methods results, Hamburg, Buske, pp.299-520.

BROWN, G. & YULE, G., (1983), Discourse analysis, Cambridge, University Press.

CHAROLLES, M., (1978), Introduction aux problèmes de la cohérence des textes, Langue Française, 38, 7-41.

CHAROLLES, M., (1983a), Coherence as a principal in the interpretation of discourse, Text, 3, 71-99.

CHAROLLES, M., (1986), Le problème de la cohérence dans les études françaises sur le discours durant la période 1965-1975, in M. Charolles, J.-S. Petöfi & E. Sözer (eds.) Research in text connexity and text coherence, Hamburg, Buske.

CHAROLLES, M., (ed), (in press), The resolution of discourse, to appear, Hamburg, Buske.

CHAROLLES, M., PETOFI, J.-S. & SOZER, E., (eds), (1986), Research in text connexity and text coherence, Hamburg, Buske.

CONTE, M.-E., (ed), (1981), Deissi testuale, Anaphora, in

Sull'Anaphora, Firinze, Presso l'Academia della Crusca, 37-54.

CORBLIN, F., (1983), Défini et démonstratif dans la reprise immédiate, Le Français Moderne, 2, 118-134.

COULTHARD, M. & MONTGOMERY, M., (eds), (1981), Studies in discourse analysis, London, Routledge Kegan, 1981.

DAHL, O. & DAHL, J., (1974), Review of T.-A. van DIJK's, Some aspects of text grammars, in Probleme und Perspectiven der neueren textgrammatishen Forshung, Papiere zur textlinguistik Hamburg, Buske, pp.121-134.

DANES, F., (1989), On van de VELDE'S paper, Man, verbal text, inferencing and coherence, in W. Heydrich & al., Connexity and coherence : analysis of text and discourse, Berlin, de Gruyter, pp.228-240.

DRESSLER, W., (1977), Current trends in textlinguistics, Berlin, New-York, De Gruyter.

DUBOIS, J., (1969), Enoncé-Enonciation, Langages, 13, 100-110.

DUCROT, O., (1983), Les mots du discours, Paris, Minuit.

ECO, U., (1979), The role of the reader, Boomington, Indiana UP.

EIKMEYER, H.-J., (1983), Procedural analysis of discourse, Text, 3, 11-39.

ENKVIST, N., (1989), From text to interpretability : a contribution to the discussion of basic terms in text linguistics, in W. Heydrich & al. (eds), Connexity and coherence : analysis of text and discourse, Berlin, de Gruyter, pp.369-383.

FAUCONNIER, G., (1983), Espaces mentaux, Paris, Minuit.

GIORA, R., (1985), What is a coherent text, in Sözer (ed), Text connexity, Text coherence, Aspects methods results, Hamburg, Buske, pp.16-35.

GRICE, H.P. (1975), Logic and conversation, in P. Cole & J.L. Morgan (eds), Syntax and Semantics : Speech Acts, (Vol.3), New York, Academic Press, pp.41-58.

GROSZ, B., (1977), The representation of use of focus in dialogue understanding, Technical note 15, SRI International Artificial Intelligence Center.

GULICH, E., (1970), Makrosyntaxder Gliederungssignale in gesprochenen, FranzÖsich, Munchen, Fink.

GULICH, E. & RAIBLE, W., (1979), Linguistiche textmodelle, Munchen, Fink.

GULICH, E. & KOTSCHI, T., (1983), Les marqueurs de la reformulation paraphrastique, Cahiers de Linguistique Française, Université de Genève, 5, 305-351.

GUESPIN, L., (1971), Problématique des travaux sur le discours politique, Langages, 23, 3-24.

GUESPIN, L., (1976), Types de discours ou fonctionnements discursifs, Langages, 41, 3-13.

HALLIDAY, M.A.K. & HASAN, R., (1976), Cohesion in English, London, Longman.

HALL-PARTEE, B., (1972), Opacity, coreference and pronouns, in D. Davidson et G. Harman (eds), Semantics of Natural Language, Dordrecht, D. Reidel.

HARRIS, Z.-S., (1952), Discourse analysis, reprints, The Hague, Mouton.

HARWEG, R., (1968), Pronomina und textkonstitution, Munchen, Fink.

HARWEG, R., (1973), Text grammar and literary texts, Poetics, 9.

HATAKEYAMA, K., PETOFI, J.-S. & SOZER, E., (1984), Text, connexity, cohesion, coherence, in E. Sözer (ed), 1984, Trad. Fr. Centro International di Semiotica e di Linguistica. Documents de travail.

HEYDRICH, W. & PETOFI, J.-S., (1981), Pragmatic considerations within a text theoretical framework, in H. Parret, M. Sbisia & J. Verchuren (eds), Possibilities and limitation of pragmatics, Amsterdam, J. Benjamin.

HEYDRICH, W., NEUBAUER, F., PETOFI, J.S. & SOZER, E., (eds), (1989), Connexity and coherence : analysis of text and discourse, Berlin, de Gruyter.

HIRST, G., (1981), Anaphora in Natural Language understanding, a survey, lecture notes in computer science, New-York, Springer Verlag.

HOBBS, J.-R., (1983), Why is discourse coherent, SRI international 1978, also in F. Neubauer (ed), Coherence in natural language texts, pp.29-71.

JAKOBSON, R., (1963), Essais de linguistique générale, Paris, Minuit.

KAMP, H., (1979), Events, instants and temporal reference, in A. von Setchow (ed) Semantics from a multiple point of view, Berlin, W. de Gruyter, pp.376-417.

KARTTUNEN, L., (1976), Discourse referents, 1969, reprints in J.-D. Mac Cawley (ed), Syntax and semantics, notes from the linguistic underground, New-York, Academic Press.

KERBRAT-ORECCHIONI, C., (1980), L'énonciation de la subjectivité dans la langue, Paris, A. Colin.

KLEIBER, G., (1986), Adjectif démonstratif et article défini en anaphore fidèle, in J. David & G. Kleiber (eds), Déterminants, syntaxe et sémantique, Paris, Klincsieck.

KOCH, W.-K., (1973), Time and Text, in J.-S. Petöfi & H. Rieser (ed), Studies in text grammar, Dordrecht, D. Reidel, pp.173-201.

KUMMER, W., (1972), Outlines of a model for a grammar of discourse, Poetics, 3.

LANGLEBEN, M., (1981), Latent coherence, contextual meanings, interpretation of a text, Text, 1, 279-313.

LANGLEBEN, M., (1983), An approach to the microcoherence of a text, in F. Neubauer (ed), Coherence in natural language texts, Hamburg, Buske, pp.71-99.

LANGLEBEN, M., (1985), A long way to the full interpretation, in E. Sözer (ed), Text connexity, text coherence, aspects methods results, Hamburg, Buske, pp.106-138.

LUNDQUIST, L., (1980), La cohérence textuelle, Kobenbaum, Nyt Nordisk Forlag, Ar. Busck.

LUNDQUIST, L., (in press), Argumentative text structure, a Procedural Model, in M. Charolles (ed). The resolution of discourse, Hamburg, Buske.

MAINGUENEAU, D., (1976), Initiation aux méthodes de l'analyse du discours, Paris, Hachette.

MARTIN, R., (1983), La logique du sens, Paris, PUF.

MOESCHLER, J., (1985), Argumentation et conversation, éléments pour une analyse pragmatique du discours, Paris, Hatier.

MOTSCH, W. & VIEHWEGER, O., (1980), Sprachhandlung, Satz und Text, in J. Rosenberg (ed), Sprache und Pragmatik, Lunder Symposium, Lund, pp.125-153.

NEUBAUER, F., (ed), (1983), Coherence in natural language texts, Hamburg, Buske.

PETOFI, J.-S., (1973a), Towards an empirically motivated grammatical theory of verbal texts, in J.S.Petöfi & H. Rieser (eds), Studies in text-grammar (ed), Dordrecht, Reidel, pp. 205-273.

PETOFI, J.-S., (1975), Modalité et topic-comment, dans une GT à base logique, Semiotica, 15.

PETOFI, J.-S., (1978), A formal semiotic text theory as an integrated theory of natural language, in Dressler (ed), Current trends in text linguistics, Berlin, de Gruyter.

PETOFI, J.-S., (1979), Some philosophical aspects of the investigation of natural language, in Language, Logic and Phylosophy, Wieu, HPT.

PETOFI, J.-S., (ed), (1979), Text vs sentence, Hamburg, Buske, Vol.1 et 2.

PETOFI, J.-S., (ed), (1981), Text vs sentence continued, Hamburg, Buske.

PETOFI, J.-S., (1985), Procedural aspects of text interpretation, interaction between text and interpreter in the meaning constitutive interpretation process, in E. Sözer (ed), Text connexity, text coherence, aspects methodes results, pp.176-189.

PETOFI, J.-S. & OLIVI, T., (1986), Texture, composition, signification, Degrès 46.

PETOFI, J.-S., (ed), (1989), Connexity and Coherence, Berlin, de Gruyter.

PROVOST-CHAUVEAU, G., (1971), Problèmes théoriques et méthodologiques en analyse du discours, Langages, 9, 6-22.

RACCAH, P.-V., (1982), Présupposition, signification et implication, Semantikos, 1, 63-93.

RANDQUIST, M., (1985), The barely visible glue. Some aspects of textual connectedness, in E.Sözer (ed), Text connexity, text coherence , aspects, methods, results, Hamburg, Buske.

REINHART, T., (1980), Conditions for text coherence, Poetics to day, 1, 161-180.

RIESER, H., (1978), On the development of text-grammar, in W. Dressler (ed), Current trends in textlinguistics, Berlin, de Gruyter, pp.6-21.

ROULET, E., & al., (1985), L'articulation du discours en Français contemporain, Bern, Peter Lang.

SCHEGLOFF, E.A., (1982), Discourse as an interactional achievement, some used of uh, husch and other things that come between sentences, in D. Tanen (ed) Analyzing discourse, text and talk, Washington, George-town University Press, pp.71-94.

SINCLAIR, J.-M. & COULTHARD, M., (1975), Towards an analysis of discourse, Oxford, University Press.

SOZER, E., (ed), (1985), Text connexity, Text coherence, Aspects methods results, Hamburg, Buske.

Van de VELDE, R.-G., (1981), Textuality and human reasoning, Text, 1, 385-407.

Van de VELDE, R.-G., (1985), Inferences and coherene in text interpretation, in Sözer (ed), Text connexity, text coherence, aspects methodes results, Hamburg, Buske.

Van de VELDE, R.-G., (1989), Man, verbal text, inference and coherence, in Heydrich, W., Neubauer, F., Petöfi, J.S. & Sozer, E., (eds), Connexity and coherence : analysis of text and discourse, Berlin, de Gruyter.

Van DIJK, T.-A., (1972), Some aspects of text-grammars, The Hague, Mouton.

Van DIJK, T.-A., (1978), Text and context, London, Longman.

Van DIJK, T.-A. & KINTSCH, W., (1983), Strategies of discourse comprehension, New-York, Academic Press.

WEBBER, B.-L., (1980), Syntax beyond the sentence, anaphora, in R. Spiro, B.-C. Bruce & W.-F. Brewer, (eds), Theoretical issues in reading comprehension, Hillsdale, LEA, pp.141-164.

WERTH, P., (ed), (1981), Conversation and discourse, London, Croom Helm.

YULE, G., (1982), Interpreting anaphora without identifying reference, Journal of semantics, 1.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# ASPECTS OF TEXTUAL CONTINUITY PSYCHOLINGUISTIC APPROACHES

Marie-France EHRLICH and Michel CHAROLLES

Lab. de Psycho. Expérimentale
Univ. Paris V, EPHE, EHESS,
CNRS UA 316 - 28 rue Serpente,
75006 PARIS.

UFR des Sciences du Langage
Univ. de Nancy II
23 Bld Albert 1er
54015 NANCY

Psychologists whose goal is to describe the cognitive processes used in text processing have a strong interest in phenomena of text continuity. Summarizing the literature on this topic is difficult and is perhaps premature for two main reasons: first of all, part of the research in this field has focused on the use of linguistic markers of text continuity during language acquisition, whereas other researchers have concentrated on adult subjects; second, text continuity has been studied both from the angle of production and comprehension. Not only is it clear that production and comprehension involve different mental operations, but these operations themselves are currently the object of theoretical debate.

Only studies on text continuity related to comprehension processes in adult subjects will be presented in this paper. Even when the field is narrowed in this way, the fuzziness of the constructs is still considerable. For instance, the terms coherence, cohesion, cohesiveness, connexity, connectivity, connectedness, and referential continuity are assigned different meanings by different authors, are not always defined, and are at times used without reference to a particular linguistic theory.

Three major features characterize the diversity of studies in this area:

a) Investigations of text continuity have used widely varying materials, sentence pairs, or texts of varying length. For the most part, materials have been restricted to narratives, with rare examples of descriptive or expository texts. The theoretical issues and the experimental designs behind such experiments cannot be identical. Especially in the case of long texts, text continuity must be treated both on the local and global levels, and in terms of interrelations.

b) Methods have changed over the last ten years. This evolution is due to replacement of classic off-line procedures such as sentence completion, coherence judgment, and recall, by on-line paradigms which are thought to provide a more direct approach to mental operations (e.g., priming, eye movement recording, error detection time).

c) The third feature concerns theory. Whereas early studies were aimed at identifying the procedures subjects use to search for intersentential referents, researchers quickly became aware of the necessity of recasting the issue on the level of mental representation. The main question thus became: how does a subject construct a coherent and organized mental representation of discourse ? Framing the issue in this way implies a consideration of the functional characteristics of the subjects' processing system and in particular, the limited capacity of working memory.

Although the notion of mental representation is currently widely accepted, researchers are much less in agreement regarding the characteristics of these representations. These divergencies are reflected in the variety of approaches used to study text continuity: propositional representations in Kintsch (1974), Kintsch and Van Dijk (1978); causal connections between narrative sequences in Trabasso, Secco and Van Den Broek (1984); mental models in Johnson-Laird (1981), Sanford and Garrod (1980), Garrod and Sanford (1981); or Rickheit and Strohner's (1986) functional approach.

We propose to examine here the mental operations involved in processing of cohesion markers (the early works (1), and the new perspectives (2)) and the problem of the immediacy of referential resolution (3).

## 1. MENTAL OPERATIONS INVOLVED IN THE PROCESSING OF COHESION MARKERS AND THE CONSTRUCTION OF TEXT COHERENCE. EARLY WORKS

### 1.1. Clark's given-new strategy

Given that Clark's works are well known, the main principles will be only briefly summarized. Based on Halliday's (1967) distinction, Haviland and Clark (1974) posit that the subject uses a given-new strategy which is composed of three steps:

Step 1: The subject uses intonational and syntactic cues to identify given and new information.

Step 2: The subject searches his memory for a unique entity (object, state or event) to which the given information refers. This search for an antecedent results in fairly rapid matching if it can be located in the immediate linguistic context. If the search fails, the subject uses real world knowledge to construct a plausible bridging inference to integrate the referent in the text. This search is time consuming.

Step 3: The subject adds the new information to memory by linking it to the referent that has been identified or created.

A number of experiments measuring comprehension time for the reading of pairs of sentences supported the step 2 hypothesis. Later, Clark and Sengul (1979) demonstrated that the distance separating an anaphore (a definite noun phrase or pronoun) from the referent affects comprehension time, confirming the discontinuity model of working memory. Identical results were obtained for both anaphoric devices, suggesting that regarding the distance factor, coreference is equally direct for a noun and for a pronoun. However, the authors point out that the availability of referents does not depend solely upon their presence in working memory, but also upon the status of the referent in relation to the topic (Lesgold, Roth, Curtis, 1979).

Three points should be stressed regarding the Clark model:

a) Search for the sentence's referent begins only after the entire sentence has been mapped as given-new information.

b) This search operates on sentences prior to (n). In other words, starting from sentence n, the reader back tracks to sentence n-1, n-2, n-3,... etc to reinstate text continuity.

c) The construction of bridging inferences consumes a larger amount of cognitive resources than does direct matching.

1.2. Kintsch and Van Dijk's Model: local and global interpropositional relationships

Whereas Clark's works derive from a theoretical perspective that emphasizes the tacit communicational agreement between speaker and receiver, Kintsch's studies (Kintsch, 1974; Kintsch and Van Dijk, 1978) are aimed at defining the semantic representation of texts and how this representation is elaborated by a subject during reading. Kintsch's conception of text continuity is nevertheless similar to that of Clark.

For Kintsch, a text is coherent if on a local level the propositions (basic semantic units composed of a predicat and one or several arguments) making up the microstructure of the text are related, and if, on the global level, the propositions are organized within a macrostructure. Understanding a text thus involves constructing a graph of coherent and hierarchized propositions. This construction is constrained by the limited capacity of the working memory, and thus proceeds by processing cycles during each of which a chunk of n propositions is analyzed. Local cohesion is defined by a simple rule: two propositions are cohesive if they share a common argument. This rule provides a means of constructing a cohesive graph of the propositions found in one cycle.

Cohesion between two successive cycles is established in the following way:
- It is assumed that part of the working memory is a short-term memory buffer which can only accept a limited number (s) of propositions. When a chunk of n propositions is processed,

s of them are chosen and stored in the buffer. This selection procedure is based on a twofold criteria which takes into account the importance and the recency of the propositions (leading-edge strategy). The unselected propositions are transferred into long term memory.

- The propositions maintained in the buffer serve to construct cohesion between two successive cycles in a direct and automatic way if one of the propositions being processed shares an argument with one of the propositions in the buffer.

- If there is no shared argument, the subject initiates a search procedure through previously processed propositions stored in long term memory, to locate a proposition that shares an argument with one of the propositions currently being analyzed. This proposition is then reinstated in working memory.

- If the search is inconclusive, the subject makes an inference based on his knowledge of the word, and adds this inferred proposition to those present in the text in order to constitute a coherent graph of propositions.

Kintsch and Van Dijk hypothesize that the reinstatement of a proposition and the formation of an inferred proposition require more time and cognitive resources than direct matching, and consequently texts which require these processes are more difficult to understand.

The findings reported by Kintsch and Vipond (1979) and Miller and Kintsch (1980) which redefine readability of texts in terms of comprehension difficulty tend to confirm this hypothesis. In this latter study, reading time (RT), and the number of propositions recalled (R) formed the ratio RT/R used as an a readability index for 20 texts averaging 77 words adapted from Reader's Digest. The number of reinstatements and inferences was correlated with reading time (.44 and .48) and with readability (.61 and .57) whereas the number of inferences was inversely correlated (-.52) with recall.

These data are interesting, and elicit several comments:

a) Kintsch's framework situates the problem of text continuity directly at the cognitive level of semantic representation. Two propositions are cohesive or not cohesive on the basis of a simple rule, that of the shared argument. This method is, however, only superficially simple: although sentences can be transcribed into propositional form in a rule-governed fashion, (Kintsch, 1974; Turner and Greene, 1978) a certain number of options escape control. Examination of the small number of examples provided in these authors' articles shows that anaphoric devices in the texts (such as pronouns, definite noun phrases) are recoded as their referents in the propositional forms, which thus makes the proposition cohesive. On the other hand, other forms, such as proper nouns, are not handled in the same way, since the proposition is assumed to require a bridging inference. A more systematic use of propositional analysis should help clarify this point. In any case, this approach does not provide a means of studying how the subject uses the variety of linguistic markers of cohesion, since these markers are cancelled out by the propositional analysis itself.

b) Kintsch indicates that referential identity among several propositions is not a sufficient condition for coherence. Nor is it a necessary condition. The propositions of a text base are related if the entities to which they refer are related in a possible world. This view is developed in Van Dijk and Kintsch (1983).

c) These studies mainly focus on local coherence as found in the microstructure. However, global coherence is also present since the propositions selected by the leading-edge strategy are mostly macropropositions, as Miller and Kintsch (1980) specify. The selection of macropropositions is guided by the superstructure of the text and by specific information in the title and introduction (Lorch and Lorch, 1985; Lorch, Lorch and Matthews, 1985).

Local coherence and global coherence are interrelated. However, the small number of studies (Cirilo, 1981; Kieras, 1981) designed to differentiate between them in fact demonstrate the complexity of the system. Ways of investigating global coherence obviously depend upon the representation of mental operations involved in discourse processing. This feature can be seen clearly in Trabasso's works (Warren, Nicholas and Trabasso, 1979; Trabasso, Secco and Van Den Broek, 1984; Trabasso and Sperry, 1985), where causal cohesion and narrative coherence are assumed to be related. A narrative representation is defined as a causal network of events, linked by a causal chain relating the major events in the narrative. Causal cohesion in a narrative can be quantified as the percentage of events linked to the causal chain. Their reanalysis of Stein and Glenn's data (1979) shows that immediate and delayed recall of a story depends on its causal cohesion. In a similar perspective, Kemper (1983) suggested replacing the classic readability formulae by one that would consider inferential load, defining it as the number of inferences necessary to establish a causal chain. However, Trabasso's position is difficult to classify. Aside from the fact that the identification of a causal chain does not correspond to a set of defined rules, his definitions of cohesiveness, cohesion, coherence and inference are different from those used by the majority of other authors.

## 2. NEW ORIENTATIONS: MENTAL MODEL-BASED COHERENCE

### 2.1. Garrod and Sanford's approach

Whereas Kintsch's theoretical position was close to that of Clark's, Garrod and Sanford's model challenged previous views. On the basis of their findings showing a facilitating effect of the specificity of the antecedent, they proposed a radically different representational model of sentence and discourse comprehension which stimulated new hypotheses regarding the resolution of reference (Sanford and Garrod, 1980; Garrod and Sanford, 1981).

They argue that proposition-based representations are not rich or specific enough to account for discourse comprehension, and suggest an alternative hypothesis based on the activation of complex structures, termed "scenarios". Their definition of scenarios is close to the Johnson-Laird definition of mental models (1977; 1983); scenarios are representations of situations including features which are not explicity stated in the text, but rather are derived from the subject's prior knowledge of these situations. The limits of this extended representation are not easy to define, but it is assumed that the discourse topic plays such a role (i.e., limited extended representation). Sentence representation is located in an activated part of the memory termed the focus, a term employed by Grosz (1977), a specialist in artificial intelligence. The focus is composed of 1) an explicit focus, containing the representation of information given in the sentence and a representation of information derived from the text structure, for example, the topic; and 2) an implicit focus, containing the representation of knowledge required to give meaning to the sentence.

In the Garrod and Sanford model, in contrast to the positions of Clark and Kintsch, referential continuity is not achieved by backtracking from sentence n to sentence n-1, n-2, etc. Instead, they posit that the mental representation of sentences appearing before n constrains the interpretation of sentence n. In their model, the subject makes spontaneous inferences to predict what will follow. Thus, referential continuity is not primarily dependent upon bridging inferences: direct matching with the contents of the focus can be sufficient whether or not the antecedent has been stated.

The data reported by Sanford and Garrod (1981) clearly challenge the idea that bridging inferences use up time and cognitive resources. In their experiment, pairs of sentences were preceded by a context sentence and were followed by a question designed to ensure that the subject read for meaning (e.g., related the two sentences). When the antecedent was unstated, reading time of the target sentence was only extended a non-significant 20 ms. No difference in response time for the question was observed between conditions, suggesting that both for stated and unstated antecedents referential continuity had been achieved.

However, this direct matching is restricted to cases where the antecedent, although unstated, was nevertheless part of the sentence representation.

The distinction between explicit and implicit focus provides a means of understanding the difference between the functions of definite noun phrases and pronouns: whereas noun phrases are stored in both components of the focus and can provide additional information, pronoun reference can only be resolved by elements contained in the explicit focus, and serve to maintain a previous referent in memory. Garrod and Sanford's 1982 data are consistent with Marslen-Wilson, Levy and Tyler's (1982) analysis of the frequency of pronouns, zero anaphora, and definite noun phrases in production of narratives.

Other notable experiments by Garrod and Sanford investigated the influence of topic on referential resolution (Garrod and Sanford, 1983), and the accessibility of a pronoun antecedent, which was shown to be dependent not only upon distance, but also on membership in an episode and salience of entities in a given representation (Anderson, Garrod and Sanford, 1983). Note that Garrod and Sanford's results and their theoretical perspective are concordant with results based on sentence completion procedures concerning the role of the semantic features of the verb and topicalization in the assigning of coreference to ambiguous pronouns (Garvey and Caramazza, 1974; Garvey, Caramazza and Yates, 1976; Caramazza, Grober and Garvey, 1977; Kail, 1979; K. Ehrlich, 1980). They also are concordant with recent studies showing an immediacy effect for coreference (cf.3).

Garrod and Sanford's experiments have been conducted with short -several lines long- narrative texts. Our findings for longer narratives (250 words) were consistent with their results regarding reading time, the measured variable indicating construction of a representation, but differed for recall (M.-F. Ehrlich, sous presse). Two versions of a text were constructed: a high cohesion control text, and a low cohesion text which violated certain rules put forward by Charolles in 1978. For the low cohesion text, reading time only increased a non-significant 8%. However, recall dropped by 25%. We concluded from these results that the interpreter may adopt coherence as a principle of interpretability and perform the minimum operations to reinstate coherence, therefore using little time. However, the mental representation obtained is poorly structured and fragile, which is reflected in the recall performance.

### 2.2. Johnson-Laird's approach

The studies described so far were aimed at investigating the mental operations the subject uses to build a coherent representation from cohesion markers (in the text) while considering features of the events mentioned in the sentences, in particular their plausibility vis à vis the prior knowledge. For Johnson-Laird, coherence and plausibility must be distinguished from one another. "A necessary and sufficient condition for discourse to be coherent as opposed to a random assemblage of sentences is: is it possible to construct a single mental model from it" (Johnson-Laird, 1981, p.367). A discourse can be perfectly coherent although it describes a bizarre sequence of events; thus coherence must be evaluated independently of plausibility. Coherence is essentially dependent upon coreference. Plausibility is dependent upon the interpretation the subject is able to assign to a discourse in an appropriate temporal, spatial, causal and intentional framework.

In this case, coherence refers to what linguists have termed cohesion, whereas plausibility refers to what they calls coherence. It is obvious, however, that the views suggested by Johnson-Laird are congruent with recent views in linguistics.

Garnham, Oakhill and Johnson-Laird (1982) compared three versions of a 200-250 word narrative: the original version (A), the randomized version (B) with sentences presented in a random order and the restored version (C) derived from (B) in which the referential continuity was restored by replacing the pronouns with fuller noun phrases. The restored version was easier to understand and recall than the randomized version, although more difficult than the original version. However, the restored version described a sequence of somewhat unplausible events. In contrast, no difference was observed for the versions of a descriptive text (1). In a second experiment with 7-and 8-year-old good and poor readers, only the good readers were able to take advantage of the referential continuity of the stories. More recent studies by Black, Freeman and Johnson-Laird (1986) used the same procedure but varied the plausibility of the text while keeping the referential continuity constant.

The guiding principle in Johnson-Laird's theory is that referential continuity and plausibility, or in other words linguistic markers of cohesion and coherence, all enter into the construction of a unified mental text model. Ackerman's (1986) work with young children and adults clearly demonstrates the importance of distinguishing between cohesion and coherence.

## 3. THE IMMEDIACY EFFECT IN COHERENCE: THE CONTRIBUTION OF ON-LINE PARADIGMS

The theoretical arguments developed by a number of authors led them to the conclusion that reinstatement of coherence is immediate or quasi immediate. Data tends to support this view. For example, Segui and Kail (1984) found this result for sentences involving the assignment of a single referent to an ambiguous pronoun; for example, "le tigre effraie le chasseur parce qu'il est énorme (ou désarmé)" (the tiger scares the hunter because it is enormous (or unarmed) (2). Aware of the potential methodological pitfalls involved in experimental measurement of mental events, these authors designed three separate techniques: acceptability judgment intervening after the sentence has been read, acceptability of the final adjective presented after the beginning of the sentence has disappeared (off the screen), and lexical decision time for the final adjective as a function of variation in the context-producing sentence. Although these techniques, when applied judiciously, provide a more direct approach than assessment of reading time or sentence completion, they do not have all the advantages of on-line techniques.

The aim of on-line techniques is to provide a means of studying mental events as on-going phenomena rather than on the basis of their outcomes. Three of these techniques have been used to study the coherence process: priming, detection of prononciation or spelling errors, and eye movement recordings.

Using a sophisticated priming technique, Dell, Mc Koon and Ratcliff (1983) showed that the presentation of an anaphoric noun (the criminal) leads to rapid activation (250 ms) of its antecedent (a burglar) and other concepts in the same proposition as the referent. McKoon and Ratcliff (1984) refer to the three criteria defined by Posner (1978) for automaticity to conclude that the activation of the antecedent by the anaphore is an automatic process.

Garrod and Sanford (1985) adopted the error detection technique devised by Tyler (1983) in a developmental study to investigate how an exhaustive contextual interpretation of anaphoric devices affects subjects' abilities to interpret the end of the sentence. A first experiment used five-to-seven-sentence-long narratives describing a main character named by a proper noun, and a secondary character. The narrative was followed by a test sentence in which the subject of the verb was the anaphoric proper noun (main character) or the anaphoric reiteration of the secondary character. The verb was either consistent or inconsistent with the contextual meaning of the subject of the verb in the text. The distance separating the verb from the subject was also varied by insertion or non-insertion of an adverb between the two. Finally, the verb was mispelled by one or two letters which transformed it into a non-word. The subjects' task was to detect this error as quickly as possible. The test sentence was followed by a question. In a second experiment, the subject of the test sentence was an unambiguous anaphoric pronoun.

The results of the first experiment clearly support the hypothesis of immediate utilization of the contextual significance of the anaphore: detection time of the spelling error was longer with inconsistent verbs (3589 ms vs 2932 ms), and no interaction was observed between the consistency factor and type of anaphore or distance.

However, the increase in error detection time for pronouns was only observed when the antecedent was the main character. Thus the immediacy of exhaustive interpretation of an anaphoric device is not the general rule. Garrod and Sanford integrate these results with their model and provide a lengthy discussion of the specific function of the pronouns (cf Garrod, this volume).

These recent findings coincide with data obtained in the field of eye movement reading research. Early studies supported the "immediacy assumption", associated with the "eye-mind assumption" developed by Just and Carpenter (1980), who argued "Readers interpret a word while they are fixating it and they continue to fixate it until they have processed it as far as they can" (1980, p.350). However, when Carpenter and Just's (1977) data on the assignment of coreference to ambiguous and non-ambiguous pronouns are examined closely, a more subtle interpretation appears necessary. Pronouns produced a high percentage (50%) of regressions towards the antecedent, which was always located on the line preceding the target sentence. Furthermore, the origins of these regressions fell into a bimodal distribution: some occurred after a single fixation on the pronoun, suggesting that search for the referent begins as soon as the pronoun is encountered, whereas others occurred after as many as six fixations, suggesting that the subject continued to explore the line before backtracking to find the antecedent.

Studies by K. Ehrlich and Rayner (1983), and K. Ehrlich (1983) were designed to re-examine this issue. They criticized the artificiality of Carpenter and Just's task -stating whether each line of text is consistent or not with the preceding lines- and modified this procedure by asking subjects to read for meaning. The texts consisted of narratives five or six lines long. The distance separating the pronoun and its antecedent varied: Near, Intermediate, and Far. For each subject, the fixation duration was computed for the following sequence: (1) the fixation prior to encoding the pronoun; (2) the fixation during which the pronoun was encoded; (3) the fixation following encoding of the pronoun and (4) the next subsequent fixation. The major finding was an interaction: the mean fixation durations (msec) were: 224, 248, 224, and 207 for fixations (1), (2), (3), and (4) respectively in the Near condition, and very similar in the Intermediate condition; they were: 220, 242, 269, and 296 (msec) for fixations (1), (2), (3), and (4) respectively, in the Far condition. Thus, when the antecedent was Near or Intermediate to the pronoun, the subject encountered no difficulty, coreference could be assigned during pronoun fixation duration. In the Far condition, more complex operations were apparently instigated.

It would appear that the search for coreference starts during pronoun fixation but is not completed.

The authors concluded that these data are incompatible with Just and Carpenter's immediacy and eye-mind assumptions, but fit with "the process monitoring hypothesis" suggested by Rayner in 1978. According to Rayner, certain processes such as perceptual encoding of words, lexical access, and some syntactic parsing take place during fixations as predicted by Just and Carpenter, but integration processes such as assignment of coreference are not immediate.

These results must be interpreted cautiously, since the relationship between eye movements and mental operations is not a direct one (Lévy-Schöen, 1983 and 1988; Jacobs and Lévy-Schoen, 1987; M.-F. Ehrlich and Rossi, 1986; M.-F. Ehrlich, 1988).

The essential point is that different experimental paradigms have produced congruent results. It is only in certain "easy" cases that coherence is reinstated immediately.

## 4. CONCLUSION

Even though this review was restricted to the operations governing textual continuity during comprehension in adult subjects, it remains incomplete. We scarcely mentioned studies on connectives since this area has been largely neglected in studies with adults, in contrast to children (see Caron, 1983, Haberlandt, 1982). A special section was not devoted to inference processes since the issues of that area are extremely varied and are not directly connected with the purpose of this paper. For the same reason, we did not present the metacognitive features of text continuity processing, which have been receiving increasing attention in recent literature (see Baker, 1985a and b) .

Although psychologists have been highly active both on the theoretical and on the experimental levels, much remains to be accomplished: investigation of the specificity and the efficiency of cohesion markers, processing of intersentential organizers in the case of long texts, systematic examination of markers characteristic of different types of texts, and so forth.

New methods are required to address these questions, in particular, methods should be better designed to assess the coherence of the mental representations that subjects construct. A theoretical framework that acknowledges the complexity of these issues must also be elaborated. Richkeit and Strohner (1985) suggest a functional approach which takes into consideration: (1) the text and text variables, (2) the author of the text and his/her purpose; (3) the reader/ listener: his/her processing system, his/her knowledge and beliefs, and his/her hypotheses about the author's purpose.

To embark upon a task of this complexity, it is obvious that psychologists must narrowly cooperate with linguists.

We have attempted to describe the major empirical works and the main theoretical orientations in the literature devoted to discourse coherence in linguistics and in psychology. Any presentation of this type can only be partial since a systematic overview is pratically impossible in so vast a field. Time has prevented us from citing numerous works, and we have deliberately restricted ourselves to the watershed publications in each area.

The purpose of this paper and this in the previous section, has been to provide an introductory framework for the papers to be presented in this symposium and to prepare for the general discussion, rather than to draw our own conclusions. If, however, an essential point should be retained from this survey, it would be in the form of a question. We have seen that there are determinations in linguistics which constrain the ways coherence is expressed (in each natural language, the markers of continuity cannot be used indiscriminately). We have also seen that there are cognitive determinations involved in discourse processing (certain comprehension "strategies" appear more frequently than others). Given these observations, to what extent would it be worthwhile to investigate how these two types of determination are related ? The question is clearly open-ended enough to inaugurate debate.

## NOTES

(1) K. Ehrlich and Johnson-Laird (1982) have shown, however, that referential continuity does affect a three-sentence spatial description.
(2) The ambiguity is only present in French, since there is no neutral pronoun.

## REFERENCES

ACKERMAN, B.P., (1986) Referential and causal coherence in the story comprehension of children and adults, Journal of Experimental Child Psychology, 42, 336-366.

ANDERSON, A., GARROD, S.C. & SANFORD, A.J., (1983) The accessibility of pronominal antecedents as a function of episode shifts in narrative text, Quarterly Journal of Experimental Psychology, 35, 427-440.

BAKER, L., (1985a) How do we know when we don't understand ? Standards for evaluating text comprehension, in D.L Forrest-Pressley, G.E. Mac Kinnon and T.G. Waller (eds.), Metacognition, Cognition and Human Performance, Vol.1, New-York, Academic Press, pp.155-206.

BAKER, L., (1985b) Difference in the standards used by college students to evaluate their comprehension of expository prose, Reading Research Quarterly, 20, 297-313.

BLACK, A., FREEMAN, P. & JOHNSON-LAIRD, P.N., (1986) Plausibility and the comprehension of text, The British Journal of Psychology, 77, 51-62.

CARAMAZZA, A., GROBER, E. & GARVEY, C., (1977) Comprehension of anaphoric pronoums, Journal of Verbal Learning and Verbal Behavior, 16, 601-609.

CARON, J., (1983), Les régulations du discours, Paris, PUF.

CARPENTER, P.A. & JUST, M.A., (1977) Reading comprehension as eyes see it, in M.A. Just and P.A Carpenter (eds.), Cognitive processes in comprehension, Hillsdale, N.J., Erlbaum Associates, pp.109-139.

CHAROLLES, M., (1978) Introduction aux problèmes de la cohérence des textes, Langue Française, 38, 7-41.

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

CLARK, H. & SENGUL, C., (1979) In search of referents for nouns and pronouns, Memory and Cognition, 7, 35-41.

DELL, G., Mc KOON, G. & RATCLIFF, R., (1983) The activation of antecedent information during the processing of anaphoric reference in reading, Journal of Verbal Learning and Verbal Behavior, 22, 121-132.

EHRLICH, K., (1980) Comprehension of pronouns, Quarterly Journal of Experimental Psychology, 32, 247-255.

EHRLICH, K., (1983) Eye movements in pronoun assignment : a study of sentence integration in K. Rayner (ed.), Eye movements in reading, N.Y., Academic Press, pp.253-268.

EHRLICH, K. & JOHNSON-LAIRD, P.N., (1982) Spatial descriptions and referential continuity, Journal of Verbal Learning and Verbal Behavior, 21, 296-306.

EHRLICH, K. & RAYNER, K., (1983) Pronoun assignment and semantic integration during reading : eye movements and immediacy of processing, Journal of Verbal Learning and Verbal Behavior, 22, 75-87.

EHRLICH, M.-F., (1988) La lecture : processus fondamentaux. Apports et limites de l'analyse des mouvements des yeux, in La

Psychologie du Langage : Recherche fondamentale et perspectives psychopédagogiques, Dijon, CRDP, pp.51-75.

EHRLICH, M.-F., (in press) The processing of cohesion devices in text comprehension. Psychological Research.

EHRLICH, M.-F. & ROSSI, J.-P., (1986) Exploration oculaire et opérations cognitives de compréhension au cours de la lecture d'un texte, L'Année Psychologique, 86, 63-82.

GARNHAM, A., OAKHILL, J. & JOHNSON-LAIRD, P.N., (1982), Referential continuity and the coherence of discourse, Cognition, 11, 29-46.

GARROD, S.C. & SANFORD, A.J., (1981) Bridging inferences and the extended domain of reference, in J. Long & A. Baddeley (eds.), Attention and Performance Vol.IX, Hillsdale, N.J., L. Erlbaum Associates, pp.331-346.

GARROD, S.C. & SANFORD, A.J., (1982) The mental representation of discourse in a focussed memory system : implications for the interpretation of anaphoric noun phrases, Journal of Semantics, 1, 21-41.

GARROD, S.C. & SANFORD, A.J., (1983) Topic dependent effects in language processing, in G.B. Flores d'Arcais & R.J. Jarvella (eds.), The process of language understanding, N.Y., J. Wiley and Sons, pp.271-296.

GARROD, S.C. & SANFORD, A.J., (1985) On the real-time character of interpretation during reading, Language and cognitive processes, 1, 43-59.

GARVEY, C. & CARAMAZZA, A., (1974) Implicit causality in verbs, Linguistic Inquiry, 5, 439-464.

GARVEY, C., CARAMAZZA, A. & YATES, J., (1974) Factors influencing assignment of pronoun antecedent, Cognition, 3, 227-243.

GROSZ, B., (1977) The representation of use of focus in dialogue understanding, Technical note 15, SRI International Artificial intelligence Center.

HABERLANDT, K., (1981-1982) Les expectations du lecteur dans la compréhension du texte, in J.-F. Le Ny & W. Kintsch (eds) Langage et Compréhension, Bulletin de Psychologie, , 35, 733-740.

HALLIDAY, M.A.K., (1967) Notes on transitivity and theme in English II, Journal of Linguistics, 3, 199-244.

HAVILAND, S.E. & CLARK, H.M., (1974) What's new ? Acquiring new information as a process in comprehension, Journal of Verbal Learning and Verbal Behavior, 13, 512-521.

JACOBS, A.-M. & LEVY-SCHOEN, A., (1987) Le contrôle des mouvements des yeux dans la lecture : questions actuelles. L'Année Psychologique, 87, 55-72.

JOHNSON-LAIRD, P.N., (1977) Procedural semantics, Cognition, 5, 189-214.

JOHNSON-LAIRD, P.N., (1981) Comprehension as the construction of mental models, in H.C. Longuet-Higgins, J. Lyons & D.E. Broadbent, The psychological mechanisms of language, The Royal society, London, pp.353-374.

JOHNSON-LAIRD, P.N., (1983) Mental models, Cambridge, Cambridge University Press.

JUST, M.A. & CARPENTER, P.A., (1980) A theory of reading : from eye fixations to comprehension, Psychological Review, 87, 329-354.

KAIL, M., (1979) Coréférence et thématisation, L'Année Psychologique, 79, 411-427.

KEMPER, S., (1983) Measuring the inference load of a text, Journal of Educational Psychology, 75, 391-401.

KIERAS, D.E., (1981) The role of major referents and sentence topics in the construction of passage macrostructure, Discourse processes, 4, 1-15.

KINTSCH, W., (1974) The representation of meaning in memory, Hillsdale, N.J., L. Erlbaum Associates.

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

KINTSCH, W. & VIPOND, D., (1979) Reading comprehension and readability in educational practice and psychological theory, in L.G. Nilsson (ed.), Perspectives on memory research, Hillsdale, N.J., L. Erlbaum Associates, pp.329-365.

LESGOLD, A.M., ROTH, S.F., & CURTIS, M.E., (1979) Foregrounding effects in discourse comprehension, Journal of Verbal Learning and Verbal Behavior, 18, 291-308.

LEVY-SCHOEN, A., (1983) Mesurer les mouvements des yeux : pour quoi faire ? Le travail humain, 46, 3-9.

LEVY-SCHOEN, A., (1988) Les mouvements des yeux comme indicateurs des processus cognitifs, in J.-P. Caverni, C. Bastien, P. Mendelsohn & G. Tiberghien (eds), Psychologie Cognitive: modèles et méthodes, Grenoble, Presses

Universitaires de Grenoble, pp.329-347.

LORCH, Jr. & R.F., LORCH, E.P., (1985) Topic structure representation and text recall, Journal of Educational Psychology, 77, 137-148.

LORCH, Jr., R.F., LORCH, E.P. & MATTEWS, P.D., (1985) On-line processing of the topic structure of a text, Journal of Memory and Language, 24, 350-362.

MARSLEN-WILSON, W.D., LEVY, E. & TYLER, L.K., (1982) Producing interpretable discourse : the establishment and maintenance of reference, in R.J. Jarvella & W. Klein (eds.), Speech place and action, Chichester, J. Wiley, pp.339-378.

Mc KOON, G. & RATCLIFF, R., (1984) Priming and on-line text comprehension, in D. Kieras & M. Just (eds.), New methods in reading comprehension research, Hillsdale, N.J., L. Erlbaum Associates, 1984, pp.119-128.

MILLER, J.R. & KINTSCH, W., (1980) Readibility and recall of short prose passages : a theoretical analysis, Journal of Experimental Psychology and Human Learning Memory, 6, 335-354.

POSNER, M., (1978) Chronometric explorations of mind, Hillsdale, N.J., L. Erlbaum Associates.

RAYNER, K., (1978) Eye movements in reading and information processing, Psychological Bulletin, 85, 618-660.

RICKHEIT, G. & STROHNER, H., (1986) Towards a functional approach to text connectedness, in J. Petöfi (ed.), Text connectedness from psychological point of view, Helmut, Hamburg, Buske Verlag.

SANFORD, A.J. & GARROD, S.C., (1980) Memory and attention in text comprehension : the problem of reference, in R.S. Nickerson (ed.), Attention and Performance Vol. VIII, Hillsdale, N.J., L. Erlbaum Associates, pp.459-474.

SANFORD, A.J. & GARROD, S.C, (1981-1982) Vers la construction d'un modèle psychologique de la compréhension du langage écrit, in J.-F. Le Ny and W. Kintsch (eds) Langage et Compréhension, Bulletin de Psychologie, 35, 643-648.

SEGUI, J. & KAIL, M., (1984) Le traitement de phrases localement ambiguës : l'attribution de la coréférence des pronoms, in M. Moscato & G. Pieraut-le Bonniec (eds.), Le langage, Publication de l'Université de Rouen, n°98, 137-149.

STEIN, N.L. & GLENN, C.G., (1979) An analysis of story comprehension in elementary school children, in R.O. Freedle (ed.), New directions in discourse processing, Vol.II Hillsdale, N.J., L. Erlbaum Associates, pp.53-120.

TRABASSO, T., SECCO, T. & van Den BROEK, P., (1984) Causal cohesion and story coherence, in H. Mandl, N.L. Stein & T. Trabasso (eds.), Learning and Comprehension of Text, Hillsdale, N.J., L. Erlbaum Associates, pp.83-112.

TRABASSO T. & SPERRY L.L., (1985), Causal relatedness and importance of story events, Journal of Memory and Language, 24, 595-611.

TURNER, A. & GREENE, E., (1978) Construction and use of a propositional text base, JSAS, Catalog of selected documents in Psychology, 3, 58.

TYLER, L.T., (1983) The development of discourse mapping processes : the on-line interpretation of anaphoric expressions, Cognition, 1983, 13, 309-341.

VAN DIJK, T.A. & KINTSCH, W., (1983) Strategies of discourse comprehension, New-York, Academic Press.

WARREN, W.H., NICHOLAS, D.W. & TRABASSO, T., (1979) Event chains and inferences in understanding narratives, in R.O. Freedle (ed.), New directions in discourse processing, Vol.II, Nordwood Ablex, pp.23-52.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# PRONOUNS AND COGNITIVE CONNEXITY

Simon GARROD

Department of Psychology
University of Glasgow
Adam Smith Building
Glasgow. G12 9YR, Scotland

## 1. INTRODUCTION

This paper presents a cognitive view of pronouns. The view reflects an approach to comprehension which not only concerns the nature of the various processes which might occur when we interpet or generate pronouns but also tries to take into account how these processes could be implemented given the severe constraints on human attention and immediate memory which we know to obtain.

Such a cognitive account compliments other types of account as found, for instance, in philosophy or linguistics by offering a different perspective on how natural languages might be designed to take advantage of the well established cognitive constraints from which we all suffer. It seems plausible that many inherent features of natural as opposed to artificial languages (such as logical and mathematical calculi or high level computer languages) reflect basic properties of our cognitive systems. For instance, why should ellipsis and the other forms of anaphora found in every piece of natural text cause such problems in designing effective machine interpreters (i.e. devices for converting natural languages into artificial ones), when the widespread use of such expressions usually enhances the intelligibility of a text. According to the account which I will sketch here the efficient use of all forms of elliptical expression relies upon the fact that we can only operate within a very restricted processing domain and this limitation effectively constrains the number of ways of filling an elliptical gap or assigning an appropriate anaphoric interpretation. In any processing system that does not suffer in this way such expressions seem to be open to a multitude of alternative interpretations not intended by the speaker or writer.

As a background to the discussion I will contrast too broadly different approaches to the problem of how we deal with pronouns and other forms of anaphora, one taking its lead from computational linguistics and the other reflecting the sort of cognitive constraints discussed above. I will then briefly mention two experiments which seem to support the cognitive view and discuss some of their implications.

### 1.1. Anaphora and processing theories

An anaphor can be defined as any linguistic expression whose interpretation depends in some well-defined way on the interpretation of some other, usually previously mentioned, expression, which is commonly called the antecedent of the anaphor. Anaphors, so defined, represent a wide range of expressions including certain definite noun-phrases, elliptical verb-phrases, and of course referential pronouns like 'he', 'she' and 'it'. From a processing point of view they all present two major problems for any theory of comprehension. First, there is the rhetorical problem of determining which antecedent the anaphor is intended to relate to and then there is the subsequent semantic problem of determining how it takes its meaning from that of the antecedent once this has been discovered.

One approach to these problems which is associated with the computational linguistic literature is to concentrate on the relationship between the anaphor and the prior text, the aim being to define a search algorithm which could operate on some representation of the prior text in order to discover a matching antecedent for the pronoun or definite noun-phrase. For such a procedure to succeed one has to formulate a simple and well defined set of principles which reflect our actual use of these expressions. For pronouns these might include such rules as that the antecedent agree in number and gender with the pronoun and that it preferably have been mentioned most recently in the text and so on.

At first, such accounts looked very promising. For instance, Norman et al. [1] are able to report that a program which used a relatively simple search algorithm of this sort yielded about 90% correct interpretations of the pronouns used in a sample of narrative texts. At the same time, there were a number of psychological studies which seemed to indicate that as the distance in the text between antecedent and anaphor increased so pronouns and other forms of anaphor became increasingly more difficult to resolve (e.g. Clark & Sengul, [2]; Erhlich & Rayner, [3]), suggesting that such expressions instigate a search for the antecedent starting with a record of the most recently encountered text and working its way back, as Norman et al's program does for pronouns.

However the early promise of such models has never really been fulfilled, particularly when applied to the resolution of pronouns. Even though they can handle a high proportion of instances they always seem to leave a residue of cases, which cause us no great problem, yet stubbornly defy assimilation into any coherent system of heuristic rules. This is apparent, for instance, in the example shown below where one pronoun 'them' does not match its intended antecedent "a prize" in gender or number and the other 'she' does not pick out its nearest potential antecedent "the best pupil".

(1) Every week the teacher gives the best pupil in the class a prize.

(2) <u>She</u> buys <u>them</u> in the local supermarket.

Two factors seem to complicate the issue, one is the nature of the representation over which the search is to operate and the other is what might be called the dynamics of the search domain, that is the principles which govern how the domain changes as the interpretation proceeds. The importance of the first of these factors becomes apparent in accounting for the semantic interpretation of pronouns and other forms of noun-phrase anaphor. Both psychologists and logicians have argued that this must depend upon some model of the text domain, as opposed to a purely linguistic representation (e.g. Kamp, [4] Stenning, [5] Sanford & Garrod,

[6]). Thus in terms of sentence (2) above it is only in relation to an abstract model that one can readily account for the interpretation of the pronoun 'them' as the set of prizes suggested in sentence (1).

The second complicating factor arises in part from the idea of model-based as opposed to text based interpretation of pronouns. Unlike texts, models are not linear representations where it is possible to define distance between antecedents and anaphor. Furthermore, there are a number of cases when a pronoun is used to pick out a textually distant anaphor in preference to a closer one (Grosz, [7]). The problem therefore is to account for the relative accessibility of antecedents for anaphoric reference with respect to an abstract model of the domain. Such complications have led to the formulation of Focus accounts of anaphoric processing.

Focus accounts of processing stem from a conviction held by most cognitive scientists that whenever we interpret anything we do so with reference to a limited representation of the domain. This limited representation is then the changing focus of our attention. In A.I. such focus systems are represented by restricted knowledge partitions (see Grosz, [7]) and have been used to account for the interpretation of pronouns and other anaphors (see Sidner, [8]). Comparable accounts can also be found in the psychological literature (Sanford & Garrod, [6]; Garrod & Sanford, [9]) where focus is equated with the representation in a limited capacity working memory system.

Whereas search accounts of anaphoric processing concentrate on the specification of a search procedure, focus accounts therefore concentrate on specifying the status of various potential antecedents within the processing domain. Before considering the details of such an account, the basic idea can be illustrated in relation to some experiments design to establish both the accessibiity of different antecedents as a function of how they were introduced and how this relates to the ease with which readers interpret pronouns and definite noun-phrases.

### 1.2 Antecedent accessibility and pronominal interpretation

The first of these studies which is reported more thoroughly in Sanford, Moar & Garrod [10] consisted in two parts. Initially we took a set of simple two line narrative texts, which could be read as introductions to a short story. We then constructed versions of each text which varied the way each of two characters were introduced into the story with a view to establishing what factors might determine the perceived prominence of the character. An example of the contrasting versions of one set of passages is shown in Table 1., where it can be seen that each of the characters can be introduced either by proper name or description of their role or both.

Table 1

Example of materials used in part one of the Sanford, Moar and Garrod experiment (here the customer is judged to be less scenario-dependent than the waiter). Below are shown the results of the continuation experiment as a function of primacy, scenario dependence and naming of antecedent.

[Mary / the customer] entered the restaurant amd sat down.

[Alphonso / the waiter ] wearily limped over to take her order.

| 1st | versus | 2nd |
|---|---|---|
| 43 % | | 46 % |
| dependent | versus | independent |
| 44 % | | 48 % |
| named | versus | not named |
| 61 % | | 30 % ** |

** = $p < 0.001$. (sign test).

In order to determine how these factors might affect the perceived prominence of each character we employed a simple procedure of having subjects write a continuation sentence to follow each snippett of text. These continuations could then be analysed according to which of the two characters were mentioned and in what way (e.g. by pronoun or a fuller description). The relative likelihood of mentioning a character can then be taken as a measure of its status within the reader's current focus of attention. The frequency of mentions is shown in Table 1 as a function of each of the factors, and it will be apparent that the only factor which produces any effeact is that of initial form of introduction, where the use of the proper name has a substantial and reliable effect on the likelihood of continuing the story with a reference to that character.

The second part of the experiment used the same materials but with the addition of two extra sentences. A filler sentence in which no reference is made to either character and a fourth target sentence in which one of the characters could be referred to either by pronoun or a fuller description matching that in the prior text (see table 2). However, in this case we recorded the time subjects spent reading this critical target sentence using a self-paced reading time procedure (see Garrod & Sanford, [11]). The idea here was to estabish the relation between the perceived prominence of the character as estimated in the first part with that of processing difficulty for the sentence containing the anaphor. The results of this experiment are shown in Table 2 as a function of distance and form of initial mention. The striking result here is again that the only factor which seems to affect reading time is that of form of initial introduction and it is only for the sentences containing pronoun anaphors that this effect emerges.

Table 2

(a) Example of the materials used in the Sanford, Moar & Garrod reading time experiment.

Order of exposure

(1) Mr. Bloggs/The manager was dictating a letter.
(2) Claire/The secretary was taking shorthand.
(3) It was getting to be late in the afternoon.
(4) He/She/Mr. Bloggs/Claire was feeling hungry.

(5) ** QUESTION **
(6) Was Claire/The secretary doing the filing?

(b) Reading time in Msecs. for the sentence containing the anaphor (4 above) as a function of anaphor and antecedent.

| | Anaphor | | | |
|---|---|---|---|---|
| | Pronoun | | Full Description | |
| Antecedent | named | not named | named | not named |
| 1st | 2170 | 2456* | 2631 | 2626 |
| 2nd | 2174 | 2466* | 2633 | 2626 |

* = Min $F'_{(1,61)} = 9.47$ ($P < 0.01$)

The results from this experiment are broadly in line with the predictions of a focusing account for pronominal anaphors. Any difficulty in interpreting the pronoun does not seem to reflect the fact that there is only one antecedent which matches in gender and number or that in one case the antecedent is closer in the text, it simply reflects antecedent prominence at the time of encountering the pronoun. The general explanation for these results, argued in more detail elsewhere (see Sanford, Moar & Garrod, [10], Garrod & Sanford [12]) is that introducing a character into a narrative with a proper name marks that character as the thematic subject of the story and hence puts it in the focus of the reader's attention. The pronoun simply serves to indicate that the story is maintaining reference to this thematic subject and in this sense its interpretation comes as much from the cognitive connection between the antecedent material and the present sentence as it does from the linguistic connection.

However, such an explanation raises questions about how the pronoun which does not pick out the focussed thematic subject receives its interpretation. All that the reading time result indicates is that the sentence as a whole takes longer to interpret. In fact if one considers the general problem of anaphora, some writers (e.g. Marslen-Wilson, Levy & Tyler [13]) have suggested that it should be recaste as a problem of resolving a sentence rather than just the anaphors it might contain. In other words, it is really a problem of locating one's interpretation of the sentence within some representation of the text as a whole. According to such a view, any particular anaphoric expression, say a pronoun, might on occasion be instrumental in resolving the sentence whereas on other occasions it might be the resolution of the sentence as a whole which determines the interpretation on the pronoun (e.g. when it only makes sense if the pronoun refers to one antecedent as opposed to another, as in sentence (1) and (2) above). Maybe this is what happens when a pronoun does not pick out a focussed antecedent.

The second experiment (reported in Garrod & Sanford [14]) was designed in part to test such a possibility. This experiment relied upon a rather different procedure, which gives an indication of whether the pronoun is interpreted as soon as it is encountered in the sentence or left unresolved until more information has been sampled, as would be expected if the sentence was being used to help resolve the pronoun. Again the experiment was carried out in two parts and can best be explained by looking first at an example of the materials used in part one.

Table 3

Materials used in spelling error detection study (Garrod & Sanford, [14]). Below are shown the magnitude of the consistency effects as a function of antecedent and anaphor.

Title: A dangerous incident at the pool.

Context: Elizabeth was a very inexperienced swimmer and wouldn't have gone into the pool if the lifeguard hadn't been standing nearby. But as soon as she was out of her depth she started to panic and wave her hands about in a frenzy.

Target sentences:

(1) Within seconds Elizabeth jumped* into the pool.
(2) Within seconds the lifeguard jumped into the pool.
(3) Within seconds Elizabeth sank* beneath the surface.
(4) Within seconds the lifeguard sank beneath the surface.

* Misspellings: jumped - jimped, sank - senk.

| Consistency effects (msecs.) | |
|---|---|
| Proper name | 813 |
| Definite description | 500 |
| Main character pronoun | 551 |
| Secondary character pronoun | -159 |

In these materials two characters are introduced into the story, "Elizabeth" and "the lifeguard" and a critical target sentence occurs at the end of each story in one of four conditions. It either contains a reference to Elizabeth or the lifeguard which may be followed immediately by a verb which indicates an action either consistent or inconsistent with the full contextual interpretation of the its subject noun-phrase. So, for instance, while (3) and (4) below yield consistent interpretations, (5) and (6) yield anomalous ones.

(3) Within seconds, Elizabeth sank beneath the surface.
(4) Within seconds, the lifeguard jumped into the pool.
(5) Within seconds, Elizabeth jumped into the pool.
(6) Within seconds, the lifeguard sank beneath the surface.

Furthermore, the contextual anomaly (where it occurs) depends upon *appreciating the full contextual significance of the anaphor.* For instance, in (5) the fact that it is inconsistent for Elizabeth to jump at this point in the story follows from our knowledge of her current physical state: suspended in water out of her depth. This knowledge can only become available after we have identified the antecedent of the anaphor and then recovered the relevant contextual knowledge about the referent in question.

In the first experiment the time to detect misspellings on these critical verbs was recorded on the assumption that such spelling error judgements would be a function of the consistency of the verb in its context (see Cole & Jakimik, [15]), for evidence of such effects in the detection of mispronounciations in speech). In fact, just such an effect was observed; spelling errors on contextually inconsistent verbs took about 600 msecs. longer to detect during reading than those on consistent verbs. This result indicates that for both proper name and definite NP anaphors the readers have been able to resolve the reference by the time they encounter the next word in the text, that is the critical verb.

This same technique could then be used with pronouns replacing the fuller descriptions in the critical sentences, and one can thereby establish whether pronouns used to pick out characters that are not thematic subjects (i.e. those introduced according to their role in the story) receive immediate interpretation like the fuller anaphors. This second experiment produced a very striking result since the consistency effects were only observed when the subject of the inconsistent verb was a pronoun referring back to the thematic subject of the story, that is the character introduced by proper name. The results are shown in Table 3, which compares the magnitude of the consistency effect across the various conditions of the two experiments.

This study therefore suggests that there are two types of process underlying the resolution of these pronouns. It can either occur at the time of encountering the pronoun, but only when the antecedent is promimently in focus (as in the case of the current thematic subject) or it is left, at least until the main verb of the sentence has been interpreted and then occurs as a consequence of attempting to resolve the sentence as a whole against the prior representation of the text and this occurs even in the case where the pronoun uniquely identifies its antecedent in gender and number.

## 2. CONCLUSIONS

In this paper a contrast has been sketched between two rather different approaches to the problem of how we understand pronouns, one deriving from a computational linguistic viewpoint and the other from a more cognitive psychological viewpoint. According to the former the main problem is cast as one of defining the search algorithm used in discovering antecedents for the pronoun, according to the latter it is seen as a problem of defining the nature of the limited mental representation which a reader or listener has available of a text which might serve as a basis for anaphoric interpretation.

In effect the two approaches differ with respect to what determines the connections between our interpretation of the different expressions in a piece of text, is it a purely linguistic connection or does it have a cognitive component? Although many writers have supposed that for full NP anaphors it is just such cognitive linkages between a representation of what has been said and what is currently under interpretation that determines how they are resolved, the studies cited here seem to implicate a special type of cognitive linkage namely thematic subjecthood which also dominates pronominal interpretation (see Garrod & Sanford [12]). Such an account suggests that the paradigm usage of pronouns is one in which they maintain reference to something already in the forefront of the reader's mind. Any shift in reference should be signalled by a fuller description such as a proper name, which in turn may serve to introduce a new thematic subject which will become the default reference for the subsequent segment of text and so on.

## REFERENCES


[1] Norman, D., Rummelhart, D. and LNR., Explorations in Cognition (San Fransisco, Freeman 1975)

[2] Clark, H. H. and Sengul, C. J., In Search of Referents for Nouns and Pronouns, Memory and Cognition (1979) 7 35-41

[3] Ehrlich, K. and Rayner, K., Pronoun Assignment and Semantic Integration During Reading; Eye Movements and the Immediacy of Processing, Journal of Verbal Learning and Verbal Behavior (1983) 22 45-87

[4] Kamp, H., A Theory of Truth and Semantic Representation, in: Groenendijk, J., Janssen, T. and Stockhof, M., (eds.), Formal Methods in the Study of Language (Mathematisch Centrum, Amsterdam, 1981)

[5] Stenning, K., Quantifiers and their Encoding in Textual Comprehension, in: Freedle, R., (ed.), Discourse Processes, Advances in Research and Theory (Norwood, N.J., 1971)

[6] Sanford, A. J. and Garrod, S., Understanding Written Language (Chichester, J. Wiley, 1981)

[7] Grosz, B., The Representation and Use of Focus in Dialogue Understanding, Technical Note 15. (SRI International Artificial Intelligence Center, 1977)

[8] Sidner, C. L., Focusing for Interpretation of Pronouns, American Journal of Computational Linguistics (1981) 7 217-231

[9] Garrod, S. and Sanford A. J., The Mental Representation of Discourse in a Focused Memory System, Journal of Semantics (1982) 1 21-41

[10] Sanford, A. J., Moar, K. and Garrod, S., Proper Names as Controllers of Discourse Focus, Language and Speech (1988) 31 1 43-56

[11] Garrod, S and Sanford, A. J., Integration of Semantic Information During Reading, Journal of Verbal Learning and Verbal Behavior (1977) 17 77-90

[12] Garrod, S. and Sanford, T., Thematic Subjecthood and Cognitive Constraints on Discourse Structure, Journal of Pragmatics (1988) 12 519-534

[13] Marslen-Wilson, W. D., Levy, E. and Tyler, L. K., Producing Interpretable Discourse: the Establishment and Maintenance of Reference, in: Jarvella, R. J. and Klein, W., (eds.), Speech, Place and Action (Chichester, J. Wiley & Sons, 1982)
[14] Garrod, S. and Sanford A. J., On the Real-time Character of Interpretation During Reading, Language and Cognitive Processes (1985) 1 1 43-59
[15] Cole, R. A. and Jakimik, J., A Model of Speech Perception, in: Cole, R. A. (ed.), Perception and Production of Fluent Speech (Hillsdale, N.J. LEA, (1980)

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# The Development of Coherence in Narratives by Understanding Intentional Action

Tom Trabasso
Department of Psychology
University of Chicago
5848 S. University Avenue
Chicago, IL 60637

On analytic and developmental grounds, global coherence in story comprehension may develop out of local understanding of states and actions organized around goals. The construction of a coherent representation of a story follows from inferring relations between states and actions, integrating these events into episodes, integrating episodes by relating their actions or states, and organizing episodes into higher-order structures.

To demonstrate how a coherent structure arises out of local causal inferences based on knowledge of intentional action, stories of contrasting are analyzed by a causal network model. Evidence for the parallel between the construction of coherent representations and the development of story understanding is presented. Data from judgments of the importance of statements ( van den Broek [1,2] ) and from answering questions by children (Liu [3] ) show that children understand the role of goals in motivating actions within an episode by age 5. The integration within-episodes is well established by age 8. However, from 8 through 11, changes understanding how states and actions are related across episodes become more marked. This kind of integration requires understanding how failed outcomes cause new goals to be generated and how successful actions and outcomes enable actions in different episodes. By age 14, between episode integration and the importance of the role of topical statements in organizing episodes into higher-order, story structures both become more apparent.

The data indicate that development of story understanding follows a sequence in which integration via inferences within episodes precedes that between episodes, and that both forms of integration change with development. The causal network model builds story structure in a parallel fashion. Together, they suggest that global coherence in story representation may follows from local understandings of states and actions organized around goals.

## 1. Introduction

This paper discusses how one might construct a coherent interpretation by application of local inferences from one's knowledge of intentional action. To do this, a causal network model for discourse analysis (Trabasso, van den Broek and Suh [4]) is outlined and is used to describe how one constructs representations of stories that differ greatly in structure and meaning of content. The constructive process, based upon the analytical model, is then shown to find support in developmental data of children who judge the importance of or who answer questions on story statements that they read or hear. The analysis and data indicate that story understanding and its development proceed in a parallel manner, from simple to complex integration of the content by causal inferences that result from the application of common but developing knowledge about intentional action.

### 1.1. The Development of Coherence in Storytelling.

When children tell stories, they do so in a systematic, developmental manner (Stein [5], Stein and Policastro [6] ). Children's early stories are composed mostly of unrelated, descriptive states. Later, they describe actions that are only temporally related. Still later, they begin to relate actions by causal and temporal inferences. Later still, they explicitly include goals and purposes that direct the actions in more complete episodes. Finally, they produce story themes in which the protagonist's goals are blocked or come into conflict with those of others. In these stories, subgoals are often generated and embedded. These processes result in the formation of complex, hierarchical or sequential episodic structures.

The development of the ability to generate and to understand narrative discourse depends upon growth in knowledge of intentional action, and growth in the ability to apply this knowledge to everyday events as well as to narratives. This application entails knowledge about (1) the kinds of problems that animate beings frequently encounter and attempt to solve, (2) the goals and plans that result from such problems, (3) the actions that may be carried out as attempts to attain the goals, (4) the consequences or outcomes of pursuing a given course of action, and (5) the internal states other than goals such as thoughts, beliefs, and feelings as well as involuntary actions that accompany problems, failures and solutions.

In the corpus of stories studied by Stein [5], as narratives develop from simple to complex episodes, they also become increasingly coherent. The stories form more than a list or a series of descriptive states or temporally successive actions. They form causally connected, integrated wholes that correspond to adult definitions of well-formed stories (Stein and Policastro [6]). The question in this chapter is: How is this development of coherence achieved, not only by children but also in the interpretation of a story itself?

In answer, global coherence is achieved by the making local, causal inferences between states and actions. This application of knowledge of intentional action a story's content is assumed to arise from a common knowledge base that develops in a manner parallel to the construction of the representation of a given story. Local coherence results from understanding how goals and other internal states arise from undesirable changes of state produced by events, how goals direct actions toward changing the undesired state of affairs, and how goal-directed actions result in changes of state that either satisfy or fail to achieve the goals. Thus individual events occur and activate general knowledge of intentional action. This knowledge is then used to interpret the events and their implications. As a result, expectations are established that guide interpretation of subsequent events and the making of inferences that interrelate them. The content of the events constrain what is activated, what is expected and what types of relations are inferred between states and actions. Together, these activities form a continuously changing set of circumstances that constrain the generation and interpretation of subsequent states and actions and contribute to the subjective feeling of coherence. Since knowledge of intentional action develops, and is limited, what develops and how it is applied determine and reflect the constructive process. We now turn to how one might describe this process analytically. We then present data from developmental studies that parallel the analysis of processing.

## 2. The Construction of Coherent Understanding.

How coherence is built from local constraints is now illustrated with an analytic model for discourse representation by Trabasso, van den Broek and Suh [4]. Suppose that you were reading the content of the story in Table 1 about a boy named Jimmy (taken from van den Broek [7]). When you read the setting statement (S1), that there once was a boy named Jimmy, you begin to infer a set of circumstances concerning a boy by the name of Jimmy in some place and time. These circumstances constrain and enable the states and actions that follow and can serve as a topic for the organization of a series of episodes. The initiating event (IE1) of Jimmy's seeing Tom's new 10-speed bike sets up expectations that he is going to react to it in some way so that when you experience the goal statement (G1) that he wants such a bike, you are not surprised given the circumstances and knowledge about boys and bikes. The goal is very important since it directs actions that are carried out in attempts to achieve it. As such, this analysis of understanding amounts to application of naive theories of intentional action. Jimmy's action (A1) of counting money is best understood as an attempt to find out if he could buy a bike. An expectation is set up but is uncertain as to the outcome of the goal with respect to success or failure. The story informs you that Jimmy does not, in fact, have enough money--a failed outcome (FO1). At this point, you have integrated or organized via inferences five states and actions into a structure called an episode.

My colleagues and I (van den Broek, [1,2,7]; van den Broek & Trabasso [8], Liu [3], Trabasso & van den Broek [9]; Trabasso, van den Broek & Liu [10]; Trabasso, Secco & van den Broek [11], Trabasso & Sperry [12]; Trabasso, van den Broek & Suh [4]) have evolved a general transition network model that is shown in the top part of Figure 1 for narrative analysis and representation in which the basic unit is an episode composed of causally connected categories. In the model, Settings (S) provide the circumstances or conditions that enable the states and actions to occur. Initiating Events (IE) begin the story with a change in state that psychologically causes the protagonist to react (R) emotionally or cognitively and to set up a goal (G) to alter the state change. The goal (G) motivates other goals or actions (A). Actions (A) enable other actions; they also enable or physically cause outcomes (O) that either attain or fail to attain goals. The outcomes (O), as goal failure or success, psychologically cause reactions (R) or goals (G). They may also enable or physically cause other outcomes (O). Each of the events are recursive and may lead to categories of the same kind, i. e. cognitions such as thoughts or beliefs can determine other thoughts or beliefs, etc.

Table 1: Hierarchical Version of Jimmy Story

| Episode | Category, Notation | Content |
|---|---|---|
| 1 | Setting, S1 | There once was a boy named Jimmy. |
| 1 | Initiating Event, IE1 | Jimmy saw Tom's new 10-speed bike. |
| 1 | Goal, G1 | Jimmy wanted a 10-speed bike. |
| 1 | Action, A1 | Jimmy counted the money he had with him. |
| 1 | Failed Outcome, FO1 | Jimmy did not have enough money. |
| 1 | Goal, G2 | Jimmy wanted to save $100. |
| 2 | Action, A2 | Jimmy asked his mother for some money. |
| 2 | Failed Outcome, FO2 | Jimmy's mother said that he should earn his own money. |
| 3 | Goal, G3 | Jimmy wanted to get a paper route. |
| 3 | Action, A3 | Jimmy went to a newspaper office. |
| 3 | Success Outcome, SO3 | Jimmy accepted a paper route. |
| 2 | Action, A2* | Jimmy delivered the newspapers before sunrise without missing a house |
| 2 | Success Outcome, SO2 | Jimmy saved $100. |
| 2 | Action, A1* | Jimmy took the money he had saved. |
| 1 | Success Outcome, SO1 | Jimmy bought a beautiful bike. |

The preceding discussion of the Jimmy story was done in terms of this model and indicates how one can integrate a series of states and actions into an episodic unit. Thus, in the Jimmy story, the five states and actions are not seen as a list or even as a simple chain of events. Rather, they are seen as an integrated series having coherence. Integration between episodes is made by inferences between states and actions but the constituents now belong to different goal structures. When this kind of inference occurs, it leads to a coherence that is more global in nature. Underlying the inferences that causally relate the states and actions is an assumption that each causal state or action is necessary in the circumstances of the story for its consequent state or action (Mackie, [13]). That is, if the cause A had not occurred, then the consequent B would not have occurred in the circumstances of the narrative. In the discourse analysis of a wide range of stories, this counterfactual criterion has served to aid in the identification of causal inferences that reliably structure the story and provide a basis for predicting memory, judgment of importance, summarizing, and answering questions on the narrative events (van den Broek [1],[2],[7]; van den Broek & Trabasso [8]; Trabasso, Secco and van den Broek [11]; Trabasso & van den Broek [9]; Trabasso & Sperry [12]; Trabasso, van den Broek & Liu [10]; Trabasso, van den Broek & Suh [4]).

In the model, the categorical nature of the content in the episode and the causal chaining between the categorized content can be generalized by allowing the categories to be recursive. By this process, one can constrain relations between states, actions and state changes and generate or describe a large variety of story structures.

## 2.1. Representation of Stories of Different Structures

Figure 1 (after van den Broek [7]) shows the model and two specific realizations that contrast markedly in their structure and meaning. In the first, the episodes follow one another sequentially; in the second representation, the episodes are embedded by failed goals and outcomes or are enabled by successful attainment of goals. The upper figure depicts a hierarchically structured story in which the episodes are imbedded; the lower figure depicts a sequentially structured story in which the episodes follow one another in succession.

Figure 1. Causal network model and representations of hierarchical and sequential versions of the Jimmy Story

Causal transition network model: S--> IE--> G-->A-->O

Representation of Hierarchical Version of Jimmy Story:

S1-->IE1-->G1-->A1-->FO1 A1*-->SO1

G2<-->A2-->FO2 A2*-->SO2

G3-->A3-->SO3

Representation of Sequential Version of Jimmy Story

S1-->IE3-->G3-->A3-->SO3

IE2-->G2-->A2-->FO2-->A2*-->S02

IE1-->G1-->A1-->FO1-->A1*-->SO1

The hierarchical story structure in Figure 1 is that derived by applying the general transition network model of Trabasso, van den Broek and Suh [4] in the top of Figure 1 to the stories in Tables 1 and 2. The hierarchical story of Table 1 and its representation in Figure 1 consist of three episodes that describe or represent (1) how Jimmy buys a bike, (2) how he earns money, and (3) how he obtains a job. The numbers in the left column of the hierarchical story in Table 1 identify the three episodes. The episodes of the Jimmy story in Table 1 form a hierarchical story because the second episode of earning money is subordinate to the first episode of obtaining a bike, and the third episode of getting a job is subordinate to the second episode. The statement categories are named in the second and notated in the third columns of Table 1. An asterisk marks the second occurrence of a statement category within a given episode.

As can be seen in Figure 1, causal relations can occur *within and between* episodes. That is, within an episode a goal motivates an attempt, an attempt enables an outcome, and an outcome contains an attempt that is motivated by a goal. That is, in episode three, Jimmy lands a job. The episode begins with a goal: Jimmy wants a paper route (G3). The goal motivates an attempt: Jimmy goes to the newspaper office (A3). The attempt enables an attempt in a successful outcome: Jimmy accepts a paper route (SO3). This attempt in the outcome is also motivated by the goal of wanting a job (G3).

Causal relations between episodes result when a statement in one episode psychologically causes, motivates or enables statements in another episode. Between episode relations are those between failed outcomes and goals, or between goals, successful outcomes and actions. For example, in the first episode, Jimmy wants to buy a bike (G1). This motivates him to count his money (A1). Counting his money enables him to find out that he does not have enough money, a failed outcome (F01). Finally, the goal of wanting a bike (G1) motivates and the failed outcome of not having enough money (FO1) conjoin to cause psychologically the subgoal of wanting to earn money (G2). The subgoal of wanting to earn money starts the second episode. One can also infer that the successful outcome of the third episode (SO3), Jimmy's obtainment of a job as a paperboy, enables an attempt in the second episode (A2*), the delivering of newspapers. This action is also motivated by Jimmy's goal of earning $100 (G2). The goal (G2), the attempt (A2*) and the subordinate successful outcome (SO3) jointly are necessary and sufficient for the successful outcome (SO2) of earning the $100. Thus, successful outcomes may be causally antecedent to superordinate actions and outcomes.

Table 2: Sequential Version of the Jimmy Story

| Episode | Category, Notation | Content |
|---|---|---|
| 3 | Setting S1 | There once was a boy named Jimmy. |
| 3 | Initiating Event IE3 | Jimmy had a lot of spare time. |
| 3 | Goal G3 | Jimmy wanted to get a paper route. |
| 3 | Action, A3 | Jimmy went to a newspaper office. |
| 3 | Success Outcome, SO3 | Jimmy accepted a paper route. |
| 2 | Initiating Event, IE2 | Jimmy did not have enough money. |
| 2 | Goal, G2 | Jimmy wanted to save $100. |
| 2 | Action, A2 | Jimmy asked his mother for some money. |
| 2 | Failed Outcome, FO2 | Jimmy's mother said that he should earn his own money. |
| 2 | Action, A2* | Jimmy delivered the newspapers before sunrise without missing a house. |
| 2 | Success Outcome, SO2 | Jimmy saved $100. |
| 1 | Initiating Event, IE1 | Jimmy saw Tom's new 10-speed bike. |
| 1 | Goal, G1 | Jimmy wanted a 10-speed bike. |
| 1 | Action, A1 | Jimmy counted the money he had with him. |
| 1 | Failed Outcome, FO1 | The money was not enough to buy the bike. |
| 1 | Action, A1* | Jimmy took the money he had saved. |
| 1 | Success Outcome, SO1 | Jimmy rode his new bike home. |

A sequential version of the Jimmy story is given in Table 2. Its structure is shown in the lower part of Figure 1. The contrast between the two stories is structural and in meaning but not in content per se: In the sequential story, Jimmy wants a job and gets one, he wants to earn $100 and uses his job to do so, and he wants a bike and buys one with his $100. Although the content of the episodes is the same in the two versions, the reasons why Jimmy wants a job, why he earns $100 and why he buys a bike are radically different. In the hierarchical version, the reasons or causes are based upon a hierarchy of goals; in the sequential version, the reasons for the goals are not as clearly specified. As a result, the hierarchical version permits a greater number of interepisodic inferences than does the sequential version. In the hierarchical version, subordinate goals are caused by superordinate goals and failures to obtain them. Superordinate successful outcomes are enabled by the attainment of subordinate goals. In the sequential version, there are no relations between the goals and between the goals and failed outcomes. There are, however, similar enabling relations between successful outcomes and actions. Thus, the sequential version serves as a control for the hierarchical, interepisodic relations on goals and failed outcomes.

### 2.2. Importance Judgments and Integration of States and Actions.

Trabasso & Sperry [12] ( see also Trabasso & van den Broek [9]) examined the basis for the making of judgments of importance of individual states and actions in narratives. The stories were taken from collections of Japanese and Chinese folktales (Sakade [14]) and the importance judgments were obtained by Brown and Smiley [15] on adult judges. In the studies by Trabasso and Sperry analyzed each of these stories as a causal network and found the number of direct causal inferences (including both antecedents and consequents) for each state and action. They then found a strong linear relationship between the degree of importance of a state or action and its number of direct, causal relations. This finding indicates that when people judge the importance of a state or action in relation to the text as a whole, they apprehend the degree to which that state or action is directly related causally to other states or actions. If one then assumes that the making of causal relations integrates states and actions into a representation of the narrative, then one can use the importance judgment as an index of the amount and kind of integration that occurs. That is, if the goal state varies in the number of connections that it has to actions *within* an episode, then judgments in importance of goals would index the ability to integrate information into episodes. On the other hand, if goals and outcomes cause goals between episodes, then judgments of importance should index the amount of integration *between* episodes.

### 2.3. Within Episodic Integration.

In a doctoral thesis carried out under my direction, Paul van den Broek ([1],[2],[3]) had 756 students from four age groups rate the importance of each state and action in two hierarchical or two sequential stories. The average age (and number of students) in the respective groups was 8 years 9 months (n = 193), 11 years 7 months (n = 205), 14 years 9 months (n = 179), and 18 years 2 months (n =180).

van den Broek found that all the students were sensitive to the number of actions that were directly motivated by goals. The goals had either 4, 8 or 12 direct connections to actions within an episode. The corresponding average judgments of importance were: 3.35, 3.74 and 4.2, showing a strong, linear relationship. On the basis of this finding, one would conclude that at least by age eight, children can use intentional knowledge about goals to infer the reasons for actions. Since goal-action inferences integrate goals, actions and outcomes into episodes, we would conclude that episodic integration is well along in understanding of stories by the age of eight years.

The value of eight years probably underestimates the ability and overestimates the age at which children can make such inferences and integrate information *within* episodes. Goldman [16], Goldman and Varnhagen [17], Stein [5], and Stein and Glenn [18] have evidence from causal questioning after children hear stories that they can make both within- and between-episode causal inferences on actions and on goals. Data from our laboratory (Liu, [3]) on questioning during the hearing of a story by 5- and eight-year old children are presented below.

The present point is not one of competence but the relative degree and order in which within- and between-episode inferences occur as indexed by importance judgments over the age groups that are studied. Thus, for our purposes, within-episode integration is well established and occurs to the same degree across all age groups in the van den Broek [1] study. However, it is the between-episode relations that show the most striking developmental differences. Developmentally, therefore, within-episode integration appears to develop more fully before between-episode integration.

### 2.4. Between-Episode Integration.

#### 2.4.1. Goals, Actions and Outcomes.

Goals and outcomes interrelate episodes in the two story structures whereas actions have only intraepisode relations. Let us suppose that as children become more knowledgeable, they are able to make more inferences that not only serve to integrate goals and actions within episodes but begin to relate episodes themselves. If so, since goals and outcomes allow more inferences to be made than do the actions in the stories, children should, as they become older, make more of these inferences and differentiate goals and outcomes from actions by rating them as more important.

This was, in fact, the case in van den Broek's ([1]) study. The average differences in importance ratings for states that interrelate episodes (goals and successful outcomes) and actions that relate actions and states within episodes for the four respective age groups were .22, .32, .36 and .65. Note that although there is evidence for between-episodic integration in all four age groups, the initial difference appears at age eleven and the largest difference occurs in the eighteen-year-old group.

2.4.2. Failed and Successful Outcomes.

Comparing the importance judgments across the hierarchical and sequential stories for failed-outcomes tells us whether or not the children made interepisodic inferences and used these as a basis for their judgments. The hierarchical stories, it will be recalled, allowed such inferences for failed outcomes whereas the sequential versions did not. The data on failed outcomes parallels exactly those in the previous section on goals and successful outcomes versus actions. The average difference in importance judgment ratings between the hierarchical and the sequential versions increased with age. The mean differences in importance for the respective four age groups were: .11, .52, .50 and .83. Again, although the eight-year olds show evidence for between-episode integration in their judgments, the increase is more noticeable by ages eleven and fourteen years, and was markedly increased by age eighteen.

While failed outcomes varied in their numbers of causal connections in the two structures, successful outcomes had identical numbers and kinds of connections in the two story versions. Hence, no differences should occur with age in importance judgments on successful outcomes as a function of the structure of the story. This was, in fact, the case. The mean differences in importance judgments were negligible across the respective four age groups: .01, .10, -.05 and .03.

2.4.3. Settings.

Setting statements such as those that introduce the character provide important background information or conditions for the action of a narrative. In addition, a setting that introduces the protagonist may take on additional importance in that it serves as a topic of organization of the narrative. In van den Broek's study, the sequential story consists of three episodes that are related by temporal and enabling relations. These relations are weak in comparison to the causal ones that embed the episodes in the hierarchical story versions. In the hierarchical stories, the initiating event of Jimmy seeing Tom's bike begins a long chain of events that finally leads to a purchase of the bike. In the sequential stories, one could organize the episodes around the topic of three events in the life of Jimmy. Thus, the setting statements in the sequential story provide a better basis for a higher order, between-episode integration than they do in the hierarchical story where they merely introduce the main protagonist.

van den Broek's data on the differences in the importance rating for the setting statements between the two versions show that this kind of higher order integration increased with age and that the most noticeable difference occurs at age fourteen. The average differences in importance ratings for the hierarchical and sequential versions were .14, .09, .97 and .96 for the four respective age groups. Thus, topical integration of episodes, a more global form of coherence, is a later achievement than the more local but between-episode integration due to goals and outcomes. If one assumes that integration of the information within episodes is necessary to the formation of episodes, and that topical organization occurs after relations between episodes, then topical organization of episodes is a later development, not only in age but also in processing.

These data on goal connectivity (within-episode integration) and on goals, actions and outcomes or failed outcomes or settings (between-episode and higher order integration) indicate that global coherence follows a developmental time course. By the age of eight years, the children show strong evidence of integration of actions and outcomes by inferences based upon goals. By this age, they also show some evidence of integration of goals and outcomes between episodes but the increase is more noticeable by ages eleven and fourteen, and is most marked by age eighteen. Integration of episodes topically by relations to the setting information appears most clearly by ages fourteen and eighteen.

These data, then, show a parallel between levels of integration within- and between-episodes and development. The integration, however, is made possible by developing knowledge of intentional action and the application of this knowledge in story understanding. Local inferences that relate goals, actions and outcomes are the building blocks of higher-order structure. Global coherence is achieved out of such local understanding. The development of knowledge of intentional action allows the child to construct representations of stories of increasing complexity and global coherence.

## 2.5. Assessing Integration Within and Between Episodes by Questions.

The above evidence on integration of information via inferences was indirect. Importance judgments are apparently determined by the number of direct causal connections that exist between states and actions. However, they require both the making of these inferences and the taking of them into account in relation to other states and actions. A more direct way is to ask specific causal questions about why and how states and actions occur when they occur. As was indicated above, several investigators, notably Goldman and Varnhagen [17], and Stein and Glenn [18] have used questioning after children heard and recalled stories in order to assess comprehension. Further, Graesser and his colleagues (Graesser, [19]; Graesser, Robertson and Anderson, [20]; Graesser and Clark, [21]) have used questioning of narratives in order to decide which inferences should be included in the story representation. The present analytic model is based upon an analysis of the inferences between the surface states and actions of the story and indexes the inferences by arcs and not by content. The model points to the inferences that could be made, not those that are actually made. Which inferences are made are to be determined empirically.

Trabasso, van den Broek, and Liu ([4]) provide a systematic procedure for assessing and promoting the making of causal inferences by questioning either during or following the listening to or the reading of a narrative. Essentially, their method is to first apply the causal network model to the narrative. Once one has a causal network representation of the text, one can be systematic about which nodes to query and which questions to use. Suppose, for example, one wishes to ask causal questions about the hierarchical version of the Jimmy story in Table 1. Following the temporal-causal sequence of the story, one can ask: Why did Jimmy want to buy a bike (G1)? Why did Jimmy count the money (A1)? and What did Jimmy do so that he found out that there was not enough money to buy a bike (FO1)? Note that these questions ask for causal and enabling antecedents for goals, actions and outcomes <u>within</u> an episode. A <u>between-</u>episode question follows: Why did Jimmy want to save $100 (G2)? Since this question focuses on the reasons found in the first episode for a subordinate goal in the second episode. The questioning series of Trabasso et al [4] continues by cycling through why questions on goals and actions, and how questions on outcomes until the end of the story is reached.

Lisa Liu ([3]) has carried out a recent dissertation in our laboratory that follows from this analytic approach to narrative representation and causal questioning. Liu asked 5- and 8-year-old children to answer why questions on goals and actions, and how questions on outcomes while they listened to hierarchical and sequential stories similar in structure to those depicted in Figure 1 and in Tables 1 and 2. Here, we shall focus only on the data she found on these young children's abilities to answer questions via inferences that integrate information within and between episodes.

First, with respect to answers that came from within an episode, the older children showed only a slight advantage over their younger counterparts on providing causes for goals and actions. The respective mean numbers of answers for the five- and eight-year groups were .32 and .38. However, the older children were able to provide a much larger number of goals and actions as answers for how outcomes occurred that denoted goal attainment or failure: 1.31 versus 2.92 average answers per question. Thus, although the five-year-olds were able to make causal inferences within an episode, they made fewer than the eight-year olds. These data are consistent with the findings of Goldman and Varnhagen [17] and Stein and Glenn [18].

The data on answers that involve inferences between episodes are of more interest and compare directly with those obtained by van den Broek ([1] [2]) on importance judgments. Recall that van den Broek found that the importance judgments increased with age for goals and outcomes relative to actions. Since goals and outcomes begin, end and interrelate episodes, part of this judgment lies in the making and use of between-episode inferences. Liu's data parallels those of van den Broek. The mean number of between-episode answers for goals, actions and outcomes for the five-year olds were .33, .11 and .24, showing some ability to provide more reasons for goals and outcomes over actions. These inferences increased substantially in the eight-year olds to .96, .34 and .40, respectively.

Comparing how age and development interact with the two kinds of integration, the average number of within-episode answers increased from 1.94 to 3.69, an increase of 90 percent. The average number of between-episode answers increased from .68 to 1.70, an increase of 150 percent. These data, then, clearly show that integration of information within episodes is developmentally advanced over integration between episodes but that greater interepisodic integration occurs with development. Local inferences initially structure episodes and then interrelate them into higher order structures. This developmental sequence is reflected in Stein's [5] analysis of stories that children tell which was discussed in the introduction.

In further support, van den Broek [1,2] reports data on question answering that he obtained from his subjects after they had completed their judgments of importance. For goals, he found that the average proportion of interepisodic answers per age group was correlated with age. The average proportion of interepisodic answers were .43, .62, .74

and .75 for the four respective age groups. For successful outcomes, the respective proportions of interepisodic answers were .42, .51, .53 and .56. These two sets of data show that interepisodic integration increases with age, is clearly present by age eight, but shows marked changes by age eleven. When the inferences required bridging larger units of text, the two older groups do better.

Since Liu [3] found that questioning the children during listening promoted recall and comprehension of the stories, questioning may overestimate the number of inferences that are made spontaneously by children. On the other hand, the importance judgment task may underestimate the younger child's ability to make these inferences. Despite these problems of competence assessment, the developmental trends in both kinds of measures indicate that higher-order structuring is a developmental outcome and is based upon a developing knowledge of intentional action applied locally to states and actions. Global understanding arises out of local understanding, both in the processing of text and in development.

## 3. Footnote

This research was supported by a grant from the National Institute of Child Health and Human Development, HD 17431, and by a grant from the Benton Center for Curriculum and Instruction at The University of Chicago to T. Trabasso. Gratitude is expressed to Paul van den Broek and to Lisa Liu for allowing the use of their data, and for their comments on this chapter. The chapter is based, in part, upon a talk on coherence that was given at the Eleventh Annual Boston University Conference on Story Comprehension and Production.

## 4. References

[1] van den Broek, P. W. Judging the importance of events in stories: the influence of structural variation and grade level. Unpublished doctoral dissertation, Department of Education, University of Chicago (1985).

[2] van den Broek, P. W. . Causal reasoning and inference making in judging the importance of story statements. Child Development (in print).

[3] Liu, L. G. Using causal questions to assess and promote children's understanding and memory for narratives that vary in causal structure. Unpublished doctoral dissertation, Department of Behavioral Sciences, University of Chicago (1988).

[4] Trabasso, T., van den Broek, P. W., & Suh, S. Logical necessity and transitivity of causal relations in stories. Discourse Processes 12 (1989) pp. 1-25.

[5] Stein, N. L. The development of children's storytelling, In Franklin, M.B. and Barten, S. (eds.), Child language: A book of readings. (Cambridge University Press, New York, 1986).

[6] Stein, N.L., & Policastro, M. The concept of a story: A comparison between children's and teachers' perspectives. In Mandl, H. Stein, N.L. and Trabasso, T. (eds.), Learning and comprehension of text. Lawrence Erlbaum Associates, Hillsdale, N.J., 1984).

[7] van den Broek, P. W. The effects of causal relations and hierarchical position on the importance of story statements. Journal of Memory and Language 27 (1988) pp. 1-22.

[8] van den Broek, P. W. and Trabasso, T. Causal networks versus goal hierarchies in summarizing text. Discourse Processes 9 (1986)1-15.

[9] Trabasso, T., & van den Broek, P. W. Causal thinking and the representation of narrative events. Journal of Memory and Language 24 (1985) pp. 612-630.

[10] Trabasso, T., van den Broek, P. W., & Liu, L. A model for generating questions that assess and promote comprehension. Question Exchange 2 (1988) pp. 25-38.

[11] Trabasso, T., Secco, T., & van den Broek, P. W. Causal cohesion and story coherence. In Mandl, H. Stein, N. L. and Trabasso, T. (eds.), Learning and comprehension of text (Lawrence Erlbaum Associates, Hillsdale, NJ, 1984).

[12] Trabasso, T., & Sperry, L. L. The causal basis for deciding the importance of story events. Journal of Memory and Language 24 (1985) pp. 595-611.

[13] Mackie, J. L. The cement of the universe: A study of causation. (Clarendon Press, Oxford, 1980).

[14] Sakade, F. (ed.), Folktales from China. Ser. 2 (Foreign Languages Press, Peking: 1957).

[15] Brown, A. L. and Smiley, S. S. The importance of structural units in prose passages: A problem of metacognitive development. Child Development 48 (1977) pp.1-8.

[16] Goldman, S. R. Inferential reasoning in and about narrative texts. In Graesser, A. & Black, J. (eds.), The psychology of questions (Lawrence Erlbaum Associates, Hillsdale, NJ, 1985).

[17] Goldman, S. R. & Varnhagen, C. K. Memory for embedded and sequential story structures. Journal of Memory and Language 25 (1986) pp. 401-418.

[18] Stein, N. L., & Glenn, C. G. An analysis of story comprehension in elementary school children. In Freedle, R. O. (ed.). Advances in discourse processes: (Vol. 2) New directions in discourse processing (Ablex Press, Norwood, NJ, 1979).

[19] Graesser, A. C. Prose comprehension beyond the word (Springer-Verlag New York, 1981).

[20] Graesser, A. C., & Clark, L. F. The structures and procedures of implicit knowledge (Ablex Press, Norwood, NJ, 1986).

[21] Graesser, A. C., Robertson, S. P., & Anderson, P. A. Incorporating inferences in narrative representations: A study of how and why. Cognitive Psychology 13 (1981) pp. 1-26.

# TEXT PRODUCTION:
# COGNITIVE ASPECTS

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# PRODUCTION: PLANNING AND CONTROL

Eric ESPERET and Annie PIOLAT (1)

Laboratoire de Psychologie du Langage - URA CNRS 666
Université de Poitiers
Poitiers, France

Centre de Recherche en Psychologie Cognitive
Université de Provence
Aix-en-Provence, France

## INTRODUCTION

All psychological research dealing with language production is concerned with analyzing three categories of phenomena and the interactions thereof, namely: 1) the conditions in which texts are produced (context), 2) the product itself and its various characteristics (the text), and 3) the processes that enable the subject to produce a text under pre-specified conditions. Depending on the author, emphasis is placed either upon a binary relationship (context-text, for example) or upon a certain category of elements (the characteristics of the text, for example). It is the heterogeneity of the research methods employed, rather than a scarcity of studies, that makes it difficult to achieve a general consensus in this domain. We have chosen to review those studies which focused on the functional aspects of production (processes), while paying particular attention to the coordination of strategies or control processes. The control processes in question belong to the third category mentioned above. The long-term objective of these studies is to gain an understanding of the step-by-step procedure used by subjects when producing texts. As stressed by Fayol (1985, p.18), one is faced with the task of "coordinating the available strategies, the situational constraints, the capacity limits of information processing, and the subject's competency into procedural models".

## 1. THE STUDY OF PRODUCTION MONITORING

The processes designated by the terms "planning" and "control" vary considerably from one study to another. Defined very generally, the term "control" can be said to designate a set of execution functions that govern (a) the choice or construction of a plan which itself adheres to the situational constraints perceived by the speaker, (b)

(1) The two authors contributed equally to the production of this paper.

the execution and maintenance of this plan, and (c) the interruption thereof if the production does not sufficiently adhere to it, or if the situation changes. The term "planning" encompasses the sequence of choices that are organized into strategies and carried out prior to the enunciation of a verbal production, whether oral or written. Hence, planning refers to the formation of the textual, semantic, syntactic, and lexical aspects of production. Among these anticipatory stages, some will clearly precede verbal enunciation, whereas others will momentarily interrupt it. Finally, the monitoring of a production brings into play the above processes, as well as the various capabilities possessed by the speaker (knowledge, metaknowledge, environmental analysis).

The rather loose definitions adopted here result from the present diversity of the functional characteristics thought to describe these monitoring processes, whose importance, nonetheless, is unanimously recognized. This state of affairs indeed reflects the complexity of the object to be modeled, and is demonstrated by the large number of components in the proposed models, for example: operations (Hayes and Flower, 1980), processing components (Levelt, 1983), levels (Butterworth, 1980; Espéret, 1989; Frederiksen, 1985), strategic plans (van Dijk and Kintsch, 1983). Moreover, these components share complex relationships. Therefore, the most difficult question before us is how to define the hierarchical dependencies and temporal links that exist between these components during step-by-step monitoring. Developmental and/or comparative studies may well provide fruitful paths in answering this question (cf. Burtis, Bereiter, Scardamalia and Tertroe, 1983; or Fredericksen, Fredericksen and Bracewell, 1986).

Given the problems outlined above, we strove to address three main questions in the studies examined.

First, at what level of activity does monitoring play a role, the superstructural, macrostructural, or microstructural? What is the scope of this monitoring?

Second, when exactly does monitoring occur? Is planning concomitant with or prior to enunciation? Is control continuous throughout the entire production process?

Third, does the speaker control the activation of monitoring processes with the help of meta-strategies, or is planning and control activation automatic?

As is the case with all cognitive processes, those implemented in production monitoring are inferred from various types of observations. The text itself and its characteristics (structure, syntactical and lexical choices, etc.), as well as spontaneous actions on the subject's part (pauses, scratch sheets, self-corrections) may reflect the operations that generated them. Resorting to metacognitive tasks (making plans, talking to oneself, judging texts) often proves necessary. Since the type of observation made by the researcher influences the way he conceives of the processes, the mode of observation shall be used here as a framework to review the literature.

Another factor that influences the choice of which aspects of behavior should be observed is whether production is oral or written. From comparative studies, it has become apparent that the governing processes underlying oral and written production differ (Chafe, 1984; Scinto, 1986). This contrast has often been emphasized, with unplanned oral discourse contrasted with planned written discourse (Ochs, 1979). There are at least two reasons why such a strict dichotomy seems inappropriate. First of all, the influence of certain characteristics of monitoring (such as production cadence, or the possibility of correcting oneself) on production have been studied for both of these modes by varying them experimentally (Gould, 1980, or Piolat, 1982). Moreover, the use of new methodology enables the investigation of more specific situations. For example, scientists have investigated the increase in planning and control possibilities brought about by the use of a word processor (Frase, 1986), and how children's performance improves with the combined use of word processors and speech synthesizers (Borgh and Dickson, 1986). Second, one must be careful not to confuse the resulting effect (upon production) of the constraints linked to the situation itself with those linked to the social context of the production (cf. Piolat, 1977; Schneuwly, 1986). Bronckart's analysis (1985), in which "production space" is distinguished from "social interaction space", provides a framework within which this issue can be systematized.

When necessary, we shall organize the studies presented here by production mode, as defined from the standpoint of production constraints rather than that of social context. It should also be noted that the majority of psychological studies of production monitoring use "monogenerated" rather than "ploygenerated" planning, as termed by Bronckart (op. cit.).

## 2. PAUSES AND THEIR INTERPRETATION

Pauses are the most common unit used to measure the time needed for planning oral productions. Recently, this approach has been extended to writing (Matsuhashi, 1981; 1982).

### 2.1. Pauses during speech

According to O'Connel and Kowal (1980, p.8), "Pausology is the behavioral investigation of temporal dimensions of human speech". Pauses that last longer than 250ms are taken to be indicators of the selection, decision, and encoding processes that precede the semantic, syntactic and lexical activities of language. It is now a well-established fact that these pauses are not produced at random during verbal enunciations, and that they enable planning that extends beyond that of the local verbal units to be uttered immediately after having been anticipated. Butterworth

(1980b and c) thus defined two alternating preparatory phases during a 30-second production cycle: a planning phase (up to 18 seconds) during which the speaker constructs a macroplan for specifying content and syntax that governs approximatively eight sentences, and an execution phase during which local, one-time lexical choices are made. Numerous studies have also indicated that, in general, pauses are longer and occur more frequently at complex textual boundaries, that the values of the parameters describing these pauses decrease when the pause occurs near a sentence connector, and that pauses are less frequent and shorter within sentences (Butcher, 1980; Holmes, 1984; Piolat, 1983).

The debate as to the function(s) of pauses is nonetheless alive and well. Ford and Holmes (1978) challenged the role of pauses in planning, having observed that the processing load (click method) is greater at the end than at the beginning of the enunciation of a sentence. Planning for the next sentence would thus be done in parallel with the emission of the preceding sentence. Pauses would allow either for resting, or for retrospective control.

This type of research leads us to believe that the speaker plans his discourse quasi-instantaneously, and above all automatically, without any truly conscious control over his choice of words. This conception may be questioned when the impact of social interaction upon the speaker's behavior is considered (an increase in the number of full pauses: Beattie, 1980; Good and Butterworth, 1980; Piolat, 1983; variation in length as a function of the person addressed: Piolat, 1983). Pauses may thus be subject to conscious monitoring by the speaker, who, within a certain time limit, would adapt his planning and control procedures to the particular social environment in which he is speaking. Furthermore, Greene (1984) demonstrated that metacognitive knowledge of empty structures leads to a reduction in the number of pauses, as though planning were in some way facilitated by this "advance preparation".

### 2.2. Pauses during writing

The studies reviewed here are aimed more at evaluating how various factors influence pauses than at validating a production model (e.g., the effect of the writer's ability, the effect of the type of text). Matsuhashi (1981) showed that the mean time needed for planning is greater when developing an argument than when telling a story. The step-by-step analysis of pauses makes it possible, however, to confirm certain observations made for oral productions: long pauses (15 sec.) at paragraph boundaries would enable more global planning; lexical decisions would be made last. This research presents several characteristics that merit further study: pauses should be studied in relation to text divisions into macropropositional and superstructural units so as not to confine analyses to microproposition planning (cf. comments of Cooper and Matsuhashi, 1983, p.7) . Moreover, the use of new experimental methods can help prevent ad hoc comments (such as the examples analyzed by Martlew, 1983) and reduce the interindividual variations in pause parameters which make interpretation difficult. By using a computer-aided production method, it is possible to lighten the cognitive load of beginning writers; for instance, the statements of a text can be composed by choosing certain keys (for greater detail see Piolat, Mességué and Farioli, 1987). Using this procedure, Piolat and Mességué (1986) studied the temporal characteristics of on-line story production by ten-year-old children as a function of the macropropositional and syntactic structure of statements.

## 3. ERRORS AND SELF-CORRECTIONS

The analysis of errors (slips of the pen, for example) and of error correction is characterized by the emphasis placed upon local phonetic, syntactic, or semantic aspects. Garret (1982) distinguishes two types of errors (permutation and shifting), which pertain either to major grammatical categories (nouns, verbs, adjectives) or to minor categories (pronouns, conjunctions, determiners). On this basis, two autonomous planning levels are defined -- multiphrasal planning and single phrasal planning, each capable of generating errors in phrase production. Hotopf (1983) also proposed an impressive classification of errors, for both writing and speech. He then studied the role of phonological and graphic similarities, as well as that of the facility of phonological lexical access in relation to the frequency of occurrence of various written and spoken errors. These mechanisms do not encompass the textual dimension of productions. Indeed few studies have touched upon this point. This may be due to the difficulty of developing adequate methods for studying the macro- and superstructural aspects of texts.

### 3.1. Oral self-corrections

The crucial question raised here is: how does production control operate? Error analysis and especially error-correction analysis provide us with a few clues. According to Butterworth (1982), a strictly sequential (top-down) model, linking the various types of planning (semantic, syntactic, phonetic) to each other, cannot account for all possible types of errors; each representational level must have its own control structure, and be directly related to the phonological assembler.

However, another type of model may be advanced, based upon Bock's (1982) analysis, whereby the speaker would only have access to the final product of an information plan (inner speech or overt discourse). The intermediate processes would not be accessible, and control would function as an evaluation of the compatibility between the final product, the planned product, and production norms. Such is the view held by Levelt (1983) based on his study of a thousand self-corrections. Levelt thus rejects a model based on a "perceptual theory of monitoring" in favor of those based on a "production theory of monitoring". The essential processes involved in the production of a message are construction, formulation, and articulation. Two additional components, creating a perceptual loop, may be added: analysis and control. The latter compares intention and product, and thus triggers (in less than 200 ms) the instructions to interrupt the enunciation and reformulate the sentence.

### 3.2. Written revisions and self-corrections

Written production situations also allow the study of local self-corrections, permitting the examination of a wider range of corrections than those directly related to the text (due to the existence of possible observations). Starting with the early studies by Hayes and Flower (1980) and by Bereiter and Scardamalia (1982), this approach, which reduces the motor cost of production, has been updated through adolescents' use of word processors at school. At present, two major findings can be noted. First, expectations concerning the use of word processors (cf. Woodruff, Bereiter and Scardamalia, 1981; Bean, 1983) have not been met. Kurth (1986) showed that their use has an essentially motivational value, and that the transformations closely resemble those in writing. Second, only skilled writers benefit from the possibility of structurally modifying texts; half of the subjects in Woodruff, Lindsay, Bryson and Joram's study (1986) only took advantage of local functions (spelling, punctuation, etc.), and handwritten scratch sheets were widely employed.

Due to the multitude of processes engaged in writing, quite a heavy revision load is required of the writer. Unlike skilled writers, however, beginning writers focus their revisions upon the aspects of their production that are the farthest removed from the macrostructure, regardless of the text composition stage being executed. Skilled writers defer surface revisions, choosing to prolong the planning and control cycles, while concentrating on the overall organization of the text (cf. Bryson, Lindsay, Joram and Woodruff, 1986). These studies also lead one to conclude that the processes involved are largely conscious ones, and must indeed become more and more so if the writer wishes to improve his writing ability.

## 4. METACOGNITIVE DATA AND MONITORING

Rather than focusing solely upon production itself, certain studies have investigated the metacognitive behavior manifested by speakers during production. Using a judgment task, one can evalutate whether children can master a narrative superstructure, and what role such mastery plays in story production (Espéret, 1984; Espéret and Gaonac'h, 1981). The method most often used is to analyze the verbal comments of the speaker during production (while not inconceivable for oral production, this method is primarily used for written production). The a posteriori analysis of these verbal protocols (cf. Hayes and Flower, 1983) makes it possible to model the sequence of decisions made by the writer, and hence to infer the corresponding planning and control processes. The first studies conducted by Hayes and Flower (1980) are a typical example of this method. These authors took into account previously written outlines, which can be considered as the trace of metadiscursive activity. From these observations, they defined planning as a complex procedure comprised of generating and organizing ideas in accordance with the goal and the setting, which ultimately leads to the creation of a plan. They also point out that planning occurs at various stages of text composition, and that the relationship between the various processes implemented in writing (planning, translating, reviewing) is recursive. These same authors (1980; 1981) believe that the writer relies on a large variety of specialized plans during text composition, such as "plans to do" or "plans for generating ideas", "plans to say" or "plans for producing a paper" as well as "plans to compose" or "plans controlling the composing act".

The plans "to do" and "to compose" can be most readily identified from comments made out loud, whereas "plans to say" are deduced from written notes. The model articulated by Hayes and Flower differs in one important aspect from Cooper and Matsuhashi's (1983) model, developed by analyzing slips of the pen and the temporal characteristics of writing. Contrary to the former authors, Cooper and Matsuhashi believe that certain plans never become totally automatic (with the exception of grapho-motor plans), and hence that the writer must always devote conscious attention to planning at all levels, especially with respect to the planning of the grammatical aspects of sentences.

Another characteristic of the Hayes and Flower model that tends to be minimized in other studies is the recursiveness of processes. A good example of this is Burtis, Bereiter, Scardamalia and Tetroe's (1983) study, based upon the analysis of preparatory notes requested of subjects before they wrote the actual paper. By comparing final texts, notes, and verbal protocols, these authors devised a slightly different model in which "conceptual planning" is distinguished from "generating contents". Generating contents, used by 10-year-old subjects, is nothing more than listing the contents that can be used directly in the text, whereas conceptual planning, used by 14-year-old subjects before content generation, involves generating notes on ultimate choices to be made (general principles concerning content, linguistic expression). The sequence in which these processes are said to occur appears rather rigid.

In these studies, the role of metacognition in triggering and maintaining production monitoring is greatly emphasized. According to Burtis et al. (1983), planning becomes efficient when the writer is best able to explain, that is, when he can put his intentions down in writing. Planning would therefore be a highly conscious monitoring process. These authors also establish a parallel between children's ability to discern, in thinking aloud protocols, those verbalizations which concern the elaboration of a plan, and the increase in their amount of conceptual planning. This notion is supported by the Bryson, Lyndsay, Joram and Woodruff study cited above (1986).

Several questions arise from this research:
First, how does the writer devise this preliminary plan? Of what observable behavior is it the trace? What exactly is its status?

Second, what processes are affected by "thinking aloud"? According to Ericsson and Simon (1984) for example, verbalized information is not linked to automatic processes (cf. Caverni, 1986). It is conceivable, however, that statement planning is relatively automatic, thereby avoiding an overly heavy mental load for the writer. Nevertheless, verbal protocols provide insight into how statements are generated. How might these data be reconciled?

Third, what exactly does the term "efficient" encompass when one relates cognitive metacontrol to the resulting product? Does this efficiency concern only the structure of the product (such is more or less the approach taken by Gordon and Braun, 1985, with their training technique), or rather, can one evaluate the linguistic production of the text as well?

## 5. MONITORING IN TEXT ANALYSIS

The most widely employed approach to the study of text production has been to analyze the texts themselves. Psychologists have often used this approach in studies of comprehension and memorization. One of the characteristics of these analyses that shoud be noted is that they produce the most general models, i.e., those with the largest scope, describing all of the processes involved in, and the constraints acting upon production. A classical example is provided by Kintsch and van Dijk's model. First articulated in respect to reproduction (Kintsch and van Dijk, 1978), the model was later defined in the form of strategies that act at various levels of planning (interactive, pragmatic, semantic, local coherency, etc.: van Dijk and Kintsch, 1983, chap.8). The notion of strategic planning, whereby importance is attributed to the communicative and pragmatic aspects of discourse monitoring, is also found in the model proposed by Zammuner (1981; 1985; 1986). Control guided by social conditions has been a central point of investigation as well (Quasthoff, 1985; Quasthoff and Nikolaus, 1982).

A comparison of the functional structures and components of the various models presented so far would go beyond the scope of the present paper. We have chosen instead to comment upon them as a whole. First, the existence of planning processes, in general sequential, is widely agreed on, as is the presence of control processes, whether these be central or spread out over the various levels of planning. Another commonly accepted idea is that of long-term planning, leading to the elaboration of a global semantic structure (superstructural or goal-orientated guidance, for example), that precedes the triggering of short-term planning which itself determines the micropropositional or even linguistic levels. Cooper and Matsuhashi (1983) have proposed rather similar components in another approach dealing with discourse analysis.

Two objections may be made concerning these models. First, they do not provide a step-by-step description of the way operations and required processing capabilities are sequentially implicated. Second, it often proves difficult to validate the systems described. Indeed, the post hoc analysis of productions, as seen by the resulting linguistic choices, does not provide a clear picture of the processes implied therein, nor of their interactions. Nor does this type of analysis enable one to reconstruct the temporal sequence of the operations involved. The suggestions made by Frederiksen (1985; Frederiksen, Frederiksen and Bracewell, 1986) are an attempt to remedy this problem with a discourse analysis that distinguishes among various points of view: clause analysis, propositional analysis, frame analysis, topic-comment and cohesion analysis, on-line inference analysis. These autonomous analyses are aimed at providing information on the sequential activation of operations during production, as well as on the interaction of those operations (relationships between discourse structures at the various levels of analysis). As stressed by the author, this approach is primarily centered upon those processes that enable the elaboration of a text's structure. To arrive at an "extended theory", the contextual constraints and their influence upon the described processes must be incorporated into the model.

The psychological validity of these processes can be determined only through a comparative study. Such a study would require differentiating among speakers or writers according to their mastery of a model's components (adults vs. children, beginners vs. experts), or contrasting production situations in which those characteristics that constitute precise constraints upon the implied processes could be determined. In both instances, the model proposed must be sufficiently well-defined to generate hypotheses concerning the effects of the comparisons.

## CONCLUSION

A few conclusions may be drawn from the brief overview given here. Planning processes appear to involve all levels of production, from the construction of pragmatic plans to the preparation of articulatory sequences, for example. Planning is not necessarily a conscious activity. What relationships exist between the various constructed plans is a controversial issue. In several studies these relationships are considered to be embedded, thereby allowing one to speak of pragmatic or superstructural constraints (linked to the plans at these levels) under which macro- or microstructural planning takes place. Control has also been analyzed in various ways; it can be either central and general, or spread out and specific, and is considered to be automatic to a lesser or greater degree, depending upon the model.

We are far from establishing how the processes of planning and control are coordinated during production. The extreme complexity of the production activity is exemplified by the difficulty of coordinating specific models into more general ones. Inversely, when models are too general, they are difficult to validate since they are less procedural. Thus even the type of model that should be used, i.e., mixed algorithmical and production rules (Hayes and Flower, 1980) or based solely upon production rules (Cooper and Matsuhashi, 1983), remains unclear.

These studies raise an important question. The types of behavior observed, which we used as an index to present the various studies, correspond to a difference in the nature and level of the phenomena under observation. The first two categories discussed (pauses; errors and self-corrections) serve as the primary tools for the analysis of planning and local control of verbal enunciations. When a relationship is established at a more global level (planning at the beginning of paragraphs, for example), this relationship is not generally formulated in terms of a precise analysis of the overall text structure. Research on the temporal planning of text structures should thus be emphasized. Conversely, the other two observation categories (metacognitive data and certain production analyses) are characteristic of more global analyses, and of the description of general, text-level processes. It is therefore more difficult to examine through them the exact on-line progress of production operations.

It is both possible and necessary to adopt another approach to this research area, one that allows for the simultaneous and interactive study of the processes at various levels of planning and control. Real-time data, obtained with computer-aided productions, permit step-by-step tracing of the operations which the subject carries out. Such intermediate data ensure a better reconstruction of the processes involved than an a posteriori reconstitution based on the production itself.

In conclusion, the present paper was purposely limited to the processes of step-by-step monitoring that characterize speakers in monologic situations and that do not consider constraints specific to particular languages. It is clear that language production cannot be identified by this class of phenomena alone, and that psychologists also need to study both polygenerated production and questions related to discourse. Indeed, all cognitive approaches should be able to relate the processes specified here to more general processes (the studies by Scinto, 1983, 1984, and 1986 for example, are a step in this direction).

REFERENCES

**Bean, H.C.**, (1983). - Computerized word-processing as an aid to revision **College composition and communication, 34**, 146-148.

**Beattle, G.W.**, 1980). - Encoding units in spontaneous speech: some implications for the dynamics of conversation, in: Dechert, H.W. and Raupach, M. (eds.), **Temporal variables in speech. Studies in honour of F. Goldman-Eisler**. The Hague: Mouton Publishers, pp.131-143.

**Bereiter, C. and Scardamalia, M.**, (1982). - From conversation to composition: the role of instruction in a developmental process, in: Glaser, R. (ed.), **Advances in instructional psychology**, vol.2. Hillsdale: Erlbaum.

**Bock, J.K.**, (1982). - Toward a cognitive psychology of syntax: information processing contributions to sentence formulation, **Psychological Review, 89**, 1-47.

**Borgh, K. and Dickson, W.P.**, (1986). - The effects on children's writing of adding speech synthesis to a word processor. Paper presented at the **Annual Meeting of the American Educational Research Association**. San Francisco, C.A., avril).

**Bronckart, J.P.**, (1985). - **Le fonctionnement des discours. Un modèle psychologique et une méthode d'analyse**. Neuchâtel: Delachaux et Niestlé.

**Bryson, M., Lindsay, P.H., Joram, E. and Woodruff, E.**, (1986). - Augmented word-processing: the influence of task characteristics and mode of production on writers' cognitions. Paper presented at the **Annual Meeting of the American Educational Research Association**, San Francisco, C.A., April.

**Burtis, P.J., Bereiter, C., Scardamalia, M. and Tetroe, J.**, (1983). - The development of planning in writing, in: Kroll, B.M. and Wells, G., (eds.), **Explorations in the development of writing**. Chichester: John Wiley and Sons, pp.153-174.

**Butcher, A.**, (1980). - Pause and syntactic structure, in: Dechert H.W. and Raupach, M. (eds.), **Temporal variables in speech. Studies in honour of F. Goldman-Eisler**. The Hague: Mouton Publishers, pp.87-90.

**Butterworth, B.L.**, (1980a). - Some constraints on models of language production, in: Butterworth, B.L., (ed.), **Language production**. vol.1: speech and talk. London: Academic Press.

**Butterworth, B.L.**, (1980b). - Evidence from pauses in speech, in: Butterworth B.L. (ed.), **Language production**. vol. 1: Speech and talk. London: Academic Press, pp.155-176.

**Butterworth, B.L.**, Preparing to speak, in: Requin, J. (ed.), **Anticipation et comportement**. Paris: CNRS, pp.473-489.

**Butterworth, B.L.**, (1982). - Speech errors: old data in search of new theories, in: Cutler, A. (ed.), **Speech errors and language production**. The Hague: Mouton Publishers, pp.627-662.

**Caverni, J.P.**, (1986). - Les protocoles verbaux comme observables des processus cognitifs. Communication au **Colloque Annuel de la Section de Psychologie Expérimentale (Société Française de Psychologie): "Activités cognitives: modèles de processus et niveaux d'observation"**. Aix-en-Provence, mars.

**Chafe, W.**, (1984). - Speaking, writing and prescriptivism, in: Schiffrin D., (ed.), **Meaning, form and use in context: linguistic applications**. Washington: Georgeton University Press, D.C., pp.95-103.

**Cooper, C.R. and Matsuhashi, A.**, (1983). - A theory of the writing process, in: Martlew, M., **The psychology of written language**. Chichester: John Wiley and Sons, pp.3-40.

**Ericsson, K.A. and Simon, H.A.**, (1984). - **Protocol analysis: verbal reports as data**. Cambridge: Mit Press.

**Espéret, E.**, (1984). - Maîtrise du schéma narratif et conduite de récit. **Colloque CNRS "Neuropsychopathologie et Sciences Cognitives"**. St Riquier.

**Espéret, E.**, (1989). - The development and role of narrative schema in storytelling, in: Piéraut-Le-Bonniec, G. and Dolitsky, M. (eds.), **From basic language to discourse bases**. Amsterdam: Benjamins.

**Espéret, E. and Gaonac'h, D.**, (1981). - What does storytelling mean for children: narrative schema representation and storytelling at different ages. Paper presented at the **International Symposium on Text Processing**. Fribourg.

**Fayol, M.**, (1985). - **Le récit et sa construction. Une approche de psychologie cognitive**. Neuchâtel: Delachaux et Niestlé.

**Flower, L. and Hayes, J.R.**, (1980). - Plans that guide the composing process, in: Frederiksen, C.H. and Dominic, J., **Writing: the nature, development and teaching of written communication**. Hillsdale: Erlbaum, pp.39-58.

**Flower, L. and Hayes, J.R.**, (1980). - The dynamics of composing: making plans and juggling constraints, in: Gregg, L. and Steinberg, E.R. (eds.), **Cognitive processes in writing**. Hillsdale: Erlbaum, pp.31-50.

**Ford, M. and Holmes, V.**, (1978). - Planning units and syntax in sentence production, **Cognition, 6**, 35-53.

**Frase, L.T.**, (1986). - Creating intelligent environments for computer use in writing. Paper presented at the **Annual Meeting of the American Educational Research Association**. San Francisco: C.A., avril.

**Frederiksen, C.H.**, (1985). - Cognitive models and discourse analysis, in: Cooper C.R. and Greenbaum, S., (eds.), **Written communication annual** vol. 1. Studying writing: linguistic approaches. Beverly Hills: Sage.

**Frederiksen, C.H., Frederiksen, J.D. and Bracewell, R.J.**, (1986). - Discourse analysis of children's text production, in: Matsuhashi, A., (ed.), **Writing in real time**. Norwood: Ablex, N.J..

**Garrett, M.F.**, (1982). - A perspective on research in language production, in: Mehler, J., Walker, E.C.T. and Garrett, M.F. (eds.), **Perspective on mental representation. Experimental and theoretical studies of cognitive processes and capacities**. Hillsdale: L.E.A., N.J., pp.185-199.

**Good, D.A. and Butterworth, B.**, (1980). - Hesitancy as a conversational resource: some methological implications, in: Dechert H.W. and Raupach, M. (eds.), **Temporal variables in speech. Studies in Honour of F. Goldman-Eisler**. The Hague: Mouton Pub, pp.145-152.

**Gordon, C.J. and Braun, C.**, (1985). - Metacognitive processes: reading and writing narrative discourse, in: Forrest-Presley, D.L., Mackinnon, G.E. and Waller, T.G. (eds.), **Metacognition, cognition and human performance**, Vol. 12. Orlando: Academic Press.

**Gould, J.D.**, (1980). - Experiments on composing letters: some facts, some myths, and some observations, in: Gregg, L.W. and Steinberg, E.R., (eds.), **Cognitive processes in writing**. Hillsdale: Erlbaum, pp.97-127.

**Greene, J.E.**, (1984). - Speech preparation processes and verbal fluency, **Human Communication Research, 11**, 61-84.

**Hayes, J.R. and Flower, L.**, (1983). - Uncovering cognitive processes in writing: an introduction to protocol analysis, in: Mosenthal, P., Tamor, L. and Walmsley, S. (eds.), **Research in writing**. New-York: Longman.

**Hayes, J.R., and Flower, L.** (1983). - Identifying the organization of written processes, in Gregg, L.W. and Steinberg, E.R. (eds.), **Cognitive processes in writing**. Hillsdale: Erlbaum, pp.3-30.

**Holmes, V.**, (1984). - Sentence planning in a story continuation task, **Language and Speech, 27**, 115-134.

**Hotopf, W.H.N.**, (1983). - Lexical slips of the pen and tongue: what they tell us about language production, in: Butterworth, B. (ed.), **Language production**, Vol.2: Development, writing and other language processes. London: Academic Press, pp.147-200.

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

**Kurth, R.J.**, (1986). - Using word processing to enhance revision strategies during student composing. Paper presented at the **Annual Meeting of the American Educational Research Association** San Francisco: C.A. April.

**Levelt, W.J.M.** (1983). - Monitoring and self-repair in speech. **Cognition, 14**, 41-104.

**Martlew, M.**, (1983). - Problems and difficulties: cognitive and communicative aspects of writing, in: Martlew, M. (ed.), **The psychology of written language**. Chichester: Wiley and Sons, pp.295-333.

**Matsuhashi, A.**, (1981). - Explorations in the real-time production of written discourse, in: Nystrand, M. (ed.), **What writers know: the language and structure of written discourse**. New York: Academic Press.

**Matsuhashi, A.**, (1981). - Pausing and planning: the tempo of written discourse production, **Research in the teaching of english, 15**, 113-134.

**Ochs, E.**, (1979). - Planned and unplanned discourse, in: Givon, T. (ed.), **Discourse and syntax**. New York: Academic Press, pp. 51-80.

**O'Connell, D.C. and Kowal, S.**, (1980). - Prospectus for a science of pausology, in Dechert, H.W. and Raupach, M. (eds.), **Temporal variables in speech. Studies in honour of Frieda Goldman-Eisler**. The Hague: Mouton Pub., pp. 3-12.

**Piolat, A.**, (1977). - Stratégies syntaxiques spécifiques aux modalités d'expression orale et écrite, **Cahiers de Psychologie, 20**, 43-67.

**Piolat, A.**, (1981). - Effets de la représentation du destinataire sur le comportement de pauses du locuteur. Communication aux **Journées d'Etudes de la Section de Psychologie Expérimentale de la S.F.P.: les Représentations**. Poitiers, Mars, Doc. Ronéoté. Crepco, Université de Provence, Aix.

**Piolat, A.**, (1982). - **L'écrit et l'oral comme systèmes de production**, Thèse de doctorat de 3ème cycle de Psychologie, Université de Provence, Aix.

**Piolat, A.**, (1983). - Localisation syntaxique des pauses et planification du discours, **L'Année Psychologique, 83**, 377-394.

**Piolat, A. and Messegue, A.**, (1986). - Caractéristiques temporelles d'un récit produit par des enfants et des adultes à l'aide d'un ordinateur. Communication au troisième **colloque international de didactique du français: apprendre/enseigner à produire des textes écrits**. Namur, Septembre, Doc. ronéoté: Crepco, Université de Provence, Aix.

**Piolat, A., Messegue, A. and Farioli, F.**, (1987). - Mémorisation de récit par des enfants et des adultes: caractéristiques temporelles du rappel assisté par ordinateur, **Cahiers de Psychologie Cognitive, 7**, n°4.

**Quasthoff, U.**, (1985). - Modèles de discours enfantin dans la représentation événementielle: savoir social ou procédure interactive, in: Wittwer, J. (ed.), **La psycholinguistique textuelle, Bulletin de Psychologie, 38**, 759-774.

**Quasthoff, U. and Nikolaus, K.**, (1982). - What makes a good story? Toward the production of conversational narratives, in: Flammer, A. and Kintsch, W. (eds.), **Discourse processing**. Amsterdam: North Holland.

**Schneuwly, B.**, (1986). - **Oral-écrit: deux systèmes?** Doc. ronéoté, Université de Genève.

**Scinto, L.F.M.**, (1983). - The development of text production, in: Fine, J. and Freedle, R.O. (eds.), **Developmental issues in discourse**. Norwood: Ablex.

**Scinto, L.F.M.**, (1984). - The architectonics of texts produced by children and the development of higher cognitive functions, **Discourse processes, 7**, 371-418.

**Scinto, L.F.M.**, (1986). - Written language and psychological development. New-York: Academic Press.

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

**Woodruff, E., Bereiter, C. and Scardamalia, M.**, (1981). - On the road to computer assited compositions, **Journal of Educational Technology Systems, 10**, 133-148.

**Woodruff, E., Lindsay, P., Bryson, M. and Joram, E.**, (1986). - Some cognitive effects of word processors on enriched and average 8th grade writers. Paper presented at the **Annual Meeting of the American Educational Research Association**. San Francisco.: C.A. April.

**Zammuner, V.L.**, (1981). - **Speech production. Strategies in discourse planning**. Hamburg: Helmut Buske Verlag.

**Zammuner, V.L.**, (1985). - Strategies speakers use to acheve a good communication, **1st International Congress of Applied Psycholinguistics**. Barcelone.

**Zammuner, V.L.**, (1986). - Pragmatics factors and strategies in discourse production, in: **Proceedings of the International Pragmatic Conference**. Amsterdam: J. Benjamine.

This Page Intentionally Left Blank

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# The analysis of writers' think-aloud protocols:
## Developing a principled coding scheme for ill-structured tasks*

Alain Breuleux

Laboratory of Applied Cognitive Science
McGill University
3700, McTavish St.
Montreal, Canada
H3A 1Y2

This paper presents a tactic for encoding think-aloud protocols of writers. This tactic involves well-defined semantic grammars at the level of propositions and frames, and produces a reliable encoding without over-constraining the task assigned to subjects. Finally, the tactic leads to the implementation of computer-based tools for the automatic or semi-automatic analysis of protocols.

---

This paper has three objectives: (a) to discuss the problems involved in applying the method of think-aloud protocol analysis to the domain of writing; (b) to present elements of a tactic to overcome these difficulties; and (c) to describe computer-based tools for applying the tactic proposed.

The production of written discourse has been considered as a problem solving activity for more than a decade (e.g. Hayes and Flower, 1977,1980b). This orientation has resulted in a focus on the *process* of writing, with the problem solving paradigm providing a theoretical framework as well as methods for conducting detailed analyses of cognitive processes. Included in these methods is think-aloud protocol analysis. The outcome of this approach are theories of the processes of writing that enable the development of simulation models.

The use of think-aloud protocols in research on writing poses certain problems, and in order to discuss these I will briefly describe the major characteristics of the method as it is used in problem-solving research. The relationship between think-aloud protocol analysis and other research methods in the context of writing will not be discussed here (see Hayes and Flower, 1983). The characteristics of think-aloud protocol analysis will be presented in rela-

---

* This paper is based on the author's doctoral thesis at the University of Montréal. The research reported here has been supported, in part, by the Ministère de l'enseignement supérieur et des sciences du Québec. The author would like to thank Robert Bracewell and Carl Frederiksen for their help and encouragement, and Michael Hoover for his invaluable editorial assistance.

tionship with the nature of the tasks that have traditionally been used in this sort of research. It will then be possible to contrast writing with these tasks, and to consider the differences implied with regards to the method.

## Think-aloud protocols in problem-solving research

Although think-aloud protocol analysis has been justified and generally accepted as a research *strategy* in problem-solving research, the *tactics* involved in actually encoding the protocols are usually not clearly specified, especially in ill-structured and semantically-rich tasks. I will emphasize the nature of think-aloud protocol analysis as a special form of semantic analysis of discourse. The special characteristics of this form of analysis stem from the peculiar properties of think-aloud discourse.

The collection and use of think-aloud protocols is a research method that is closely tied to the theory of human information processing, in particular its application to problem solving. The assumptions behind this method are that search for a solution to a problem consists of the more or less *controlled* application of information processing operators to states of knowledge. This process, called *search,* takes place in a mental work space, the *problem space,* that contains the set of possible operators and the elements that comprise states of knowledge. The problem space is the mental representation used by the problem solver to represent the operations that are under consideration at different points, those that are actually carried out, as well as the various intermediate results of these operations. Generally, the problem space contains, in addition to operators and knowledge state elements, an initial state and a goal state. The process of solving a problem is conceived of as the establishment of a path relating these two states, and such a path constitutes a solution. Search for a solution thus consists of the construction of a path between initial and goal states by successively applying operators to knowledge states. The application of operators is *controlled,* that is, governed by knowledge of a strategic nature, usually conceived of as a set of rules. General purpose control knowledge, which can be applied in a certain variety of problems, but that do not essentially guarantee the determination of a solution path, are termed heuristic rules, or *heuristics.* Such strategies have been studied in research on problem solving (see Groner, Groner, & Bischof, 1983; Newell & Simon, 1972).

It is within this conception of problem solving that the use of think aloud protocols has been developed and can best be understood. The information processing theory of problem solving assumes that, while solving a problem, an individual sequentially processes the states of knowledge and the possible operators that she considers useful intermediate steps toward a solution. The theory further assumes that, under normal circumstances[1], each intermediate step that comes into the focus of attention is available for verbaliza-

---

[1]Situations that are inappropriate for think aloud purposes are essentially those involving highly non-verbal operations (for a thorough discussion, see Ericsson and Simon, 1984).

tion, and that verbalizing does not interfere with the normal solving process. The transcription of these verbalizations, a *think-aloud protocol,* constitutes an accurate trace of the process by which the individual searches for a solution. This trace generally includes all the erroneous sequences of steps, as well as the alternatives that were considered at different points. The protocol, which consists of a sequence of statements in natural language, can then be analyzed in order to extract in a canonical form the knowledge structures and processes that are verbalized. This extraction, or *encoding,* generates two representations that are direct computations from the protocol: a dynamic trace of the information processing behavior of the subject, the *problem behavior graph* (PBG), and a static structure of the information processing operators and knowledge states used by the subject, the *individual problem space.* A PBG is a node and link structure representing the temporal sequence of moves during the solving process. Nodes usually represent states of knowledge, and links represent operators. A triple composed of an operator and two knowledge states form the unit of a PBG. The unit as a whole represents the application of an operator to an existing knowledge state, producing a new resulting knowledge state. Backing up to an earlier knowledge state is represented by drawing a copy of the node (see Newell & Simon, 1972). The analysis of think-aloud protocols aims mainly at describing the problem space of the individual that produced the protocol, essentially to particularize the general problem space that is initially postulated on the basis of a task analysis. The encoding also serves as a basis for developing an explicit model of the individual's processing of information during the search for a solution. The encoding thus generates two additional representations that are indirect computations from the protocol: the information processing model of the individual, and the trace of its execution. Such a model usually takes the form of a *production system* that can be implemented as a computer program to process information in a way similar to that of the individual human subject. The adequacy of the production system model can be verified by comparing the *trace* of its execution with the problem behavior graph of the individual. Thus, the four resulting structures of the analysis are the problem space, the problem behavior graph, the production system, and the trace of the behavior of the production system (Waterman & Newell, 1971).

## The encoding problem

The definition of these representations resulting from think-aloud protocol analysis tells very little about the actual encoding and extraction of information. In order to determine such an encoding scheme, the nature of the initial representation —the natural language statements in protocols— have to be taken into account, and the relations between this representation and the resulting representations must be specified. In addition to using a well-defined representation for entities in the problem-space, we need to take into account as much as possible "how these will be talked about" by subjects (Waterman & Newell, 1971). The discourse produced by subjects while thinking aloud differs from written and, to a certain extent, spoken discourse. One of these differences is a relaxation of the constraints of communicative discourse (e.g., at the level of syntax), and the sometimes elliptical nature of the verbalizations. Some of these characteristics of think-aloud protocols have consequences for the subsequent analysis, and need to be emphasized in order to understand the encoding methods that have been developed in various settings, including in writing research.

Most of the difficulties of analyzing think-aloud protocols have been avoided in research on problem solving by focusing on well-defined, constrained, and semantically simple tasks: cryptarithmetic, missionaries-cannibals puzzles, elementary logic and arithmetics. Cryptarithmetic problems, for example, are simple arithmetic problems in which operandi and their result are represented as nouns or names, all identical letters having the same unique digit value (e.g. SEND + MORE = MONEY; DONALD + GERALD = ROBERT). The task for the subject is to convert each letter back to a distinct digit so that the resulting numbers form an arithmetically correct operation. The subject is often provided with a clue, such as the assignment of a digit to one letter. Within such task domains, knowledge is limited and can be expressed through simple representation formalisms. Furthermore, the discourse used to express this knowledge is more predictable and ellipses can easily be filled in by the coder. This is so, because an analysis of the task reveals almost all of the elementary processes and states that the subject has to consider and use for solving a problem in that domain. There are fewer real ambiguities in the discourse produced by the subject, and any ambiguity that would arise can easily be resolved by examining the context.

A simple problem space for cryptarithmetic can be constructed directly from the explicit problem definition (see Newell and Simon, 1972, p. 145). The knowledge states in this basic problem space are simple expressions of the equivalence relations between letters and digits, and the set of operators contains the single process of adding an expression to the set of knowledge states. As in the case of most problem spaces, a grammar in BNF notation is a simple and effective way of representing this basic problem space (Newell and Simon, 1972). This grammar, however, has no direct relationship with protocol statements, and does not provide an explicit encoding scheme. Defining an explicit encoding scheme would involve establishing the conditions under which verbalizations in protocols or behaviors are to be considered as acceptable instances of the states and operators expressed, in a canonical form, by the BNF grammar. These characteristics define the conditions under which a rule in the grammar is to be activated, much as in a parser capable of choosing among alternative rewrites for a given rule.

Definitions of explicit encoding schemes have been provided in various attempts at developing automatic encoding systems for problem-solving protocols. The design of these systems, based on formal and explicit descriptions of information structures and operations, reveals some of the requirements for well-defined encoding schemes. I will review briefly here two of the most important of these attempts.

## Automatic encoding systems

Waterman and Newell (1971) made an early attempt at developing an automatic computer-based encoding system for think-aloud protocols. Their approach can be characterized along the following lines. The task domain is kept simple and well-defined, cryptarithmetic[1], and an important part of the analysis procedure is based on a key-word grammar, avoiding a general syntactic analysis. The design of Waterman and Newell's (1971) system, PAS-I, is based on a certain number of intermediate representations: linguistic representations, semantic representations, and psychological representations. The linguistic level processes lexical information (including words and prosodic features), topic information (segmentation of lexical representations into units, topic segments), and linguistic rules (including conventional uses of language, idiosyncrasies, and conversational rules). The semantic level concerns task specific information, mainly assertions of basic problem-space elements, including states, operators, and relations between problem-space elements; these elements are formed into groups consisting mostly of an operator, its input, and its output. The psychological level concerns the four resulting information structures discussed earlier: the PBG, the problem-space, the production system, and the trace of the production system's behavior.

In terms of its operation, PAS-I constructs the PBG from the topic representation, given the linguistic rules and the problem-space. A linguistic processor first transforms topic segments into a set of semantic elements. This processing avoids a standard syntactic analysis by using a keyword grammar that is triggered directly by occurrences of certain specific words, such as "then", "because", "not", "is", "be", "equal". For example, a statement like "Each D is 5" is transformed into (EQ D 5) on the basis that D represents a letter and 5 a digit, and that "is" represents an equality (see Waterman & Newell, 1971, p. 298). The semantic elements are defined by a problem-solving theory for cryptarithmetic, as represented in the problem-space. Then, a semantic processing produces tentative groupings of the semantic elements, on the basis of their temporal adjacency and of the equivalence between different verbal forms (e.g. IF, BECAUSE, THEREFORE are mapped onto the single operator BECAUSEOF). These proto-groups are then analyzed further in order to produce groups consisting of one operator and its associated input and output knowledge elements (for a more detailed account of these transformations, see Waterman & Newell, 1971). Finally, the refined groups are incorporated gradually into the PBG, and this operation involves determining unknowns and origins of knowledge states (relating current states with earlier states). The determination of unknowns is roughly equivalent to resolving anaphoric references for variables left unassigned by the linguistic processor. The determination of origins involves finding the operator responsible for a knowledge state, or finding the input knowledge state for an operator. This

---

[1]Provision was made, however, for adapting the system to a large variety of task domains, by separating the task-specific information from general purpose operations (see Waterman & Newell, 1973).

process is based on a table of possible associations between operators and knowledge elements and on a measure of consistency with the current state of the problem. The actual incorporation to the PBG consists of the attachment of a single operator group either to the growing end of the graph (adding a new node to the last currently active node in the tree), or to a prior point in the problem space, abandoning intermediate nodes (in the case of back-ups from contradictory or irrelevant states).

The design of PAS-I underlines the importance of intermediate representations —linguistic and semantic— in the process of building psychological representations. The design, however, suffers from a limited approach to the analysis of natural language. The use of keywords by the linguistic processor provides an awkward treatment of inflections and differences in word meaning due to their position in a sentence, compared to a more conventional morpho-syntactic processing (see Winograd, 1983). The approach is successful to the extent that, in cryptarithmetic, "examples of operators and knowledge elements occur in relatively [...] simple linguistic contexts" (Waterman & Newell, 1971, p. 315). In addition, the semantic processor operates on assumptions that are possible only because of the limited task environment (for example the constraint that a letter correspond to only one digit). That is, much is already known about the problem-space beforehand, and the induction of new information about it is reduced to a minimum. The applicability of this design to ill-structured and semantically-rich tasks, where induction is more difficult, is not obvious.

A more recent attempt was made by Miller and Goldstein (1976) at developing an automatic protocol analysis system. Although their system deals only with behavioral protocols (traces of actions), and not with think-aloud protocols, its design illustrates the application of formalisms and tools developed in computational linguistics to encoding problem-solving protocols.

Miller and Goldstein, in trying to overcome the restriction of problem-solving research to very limited domains, have explored the usefulness of concepts and algorithms developed in computational linguistics. More specifically, they have proposed (a) that *augmented transition networks* (ATN) provide a more structured organization to represent control knowledge than *production systems;* and (b) that *parse trees* can be used instead of *problem behavior graphs* to represent the interpretation of a protocol. Thus protocol analysis can be considered as *parsing,* which involves the definition of a problem solving grammar and its application to a protocol in order to reveal its *constituent structures.*

Miller and Goldstein (1976) have designed a parser, PAZATN, that uses an ATN representation of planning knowledge (PATN, a *grammatical* theory of planning), applied top-down with some bottom-up control, and produces a parse tree delineating the planning and debugging strategies of someone programming in Logo. The protocols used by Miller and Goldstein are "performance" protocols, the keyboard traces of a subject performing a Logo programming task on a computer. The theory of planning is based on a taxonomy that identifies a small number of types of plans: identification, decomposition, and reformulation. Planning is viewed as the process of selecting a plan type and carrying the subgoals defined by its application to a given problem. This work, including the underlying taxonomy of planning strategies, is a generalization of earlier work in AI on planning (Sacerdoti, 1975; Sussman,1975), and is meant to extend and improve information processing models of human problem solving.

An interesting aspect of Miller and Goldstein's work is the use of a standard computational linguistics approach (the ATN formalism and the technique of chart parsing), and a demonstration of its usefulness in the analysis of problem-solving protocols. This approach overcomes some the limitations discussed earlier of Waterman and Newell's (1971) system.

In general, the basic components of any encoding system seem to be: "grammars to deal with surface structures of natural language, representation of knowledge, matching, and inference mechanisms to infer information not directly expressed in the utterances" (Waterman & Newell, 1973, p. 442). These components will be used in the encoding tactic that will be presented in the remainder of this paper. Both Waterman and Newell (1971) and Miller and Goldstein (1976) have defined semantic grammars for expressing problem spaces, and these grammars form the basis of the encoding system. Neither system, however, incorporates a standard computational analysis of *natural language*. Miller and Goldstein have provided clear evidence on the advantages of using a standard computational linguistic approach. The semantic grammar for a problem-space can be expressed as an ATN, and adding the definition of specific information processing sequences with their resulting knowledge states makes it a well-defined encoding scheme that generates from a protocol the parse tree containing all problem-space components that are present. Such an ATN grammar would also become equivalent to a production system (Miller & Goldstein, 1976). Miller and Goldstein, however, have not applied this approach to the analysis of natural-language protocols. Finally, in the encoding systems defined by both Waterman and Newell (1971) and Miller and Goldstein (1976), the role of the task analysis in determining the specific semantic categories that will be used to analyze think-aloud protocols is of primary importance.

The next sections discusses the characteristics of ill-structured and sematically rich tasks, in relationship with the development of rigourous encoding schemes for think-aloud protocols produced by subjects performing in such task domains.

## Ill-structured and semantically-rich tasks

Ill-structured tasks are characterized mainly by the fact that the information needed to solve the problem is not entirely given in the task instructions, and that the evaluation of operations and states, including the determination of whether the goal state is achieved, is more complex and less definite

(Simon, 1978). Relevant information for an ill-structured task is not clearly delimited, and the final consequences of moves are not easily anticipated. We can say that, almost "by definition", ill-structured tasks lead to incomplete theories, or at least to incomplete *a priori* problem-spaces. In addition, semantically-rich tasks impose the need for a more complex and detailed representational formalism, at least for the PBG and the problem space.

Bhakshar and Simon (1977), among others, have acknowledged that problem-solving research has focused mainly on tasks, like cryptarithmetic and the Tower of Hanoi puzzle, that require little specific semantic knowledge. While using the general method of detailed analysis of think-aloud protocols and of computer simulation of problem solving behavior, Bhakshar and Simon have tried to extend the theory of human problem solving to semantically rich domains. This extension entails, according to Bhakshar and Simon, a focus on how information in memory is organized and represented so as to be available to general problem solving processes.

In the course of that research, Bhakshar and Simon (1977) have developed an interactive computer program, SAPA, that serves as a basis for both the encoding and the simulation of individual protocols, in the domain of chemical engineering thermodynamics. The thermodynamics problems that they have used are characterized by underlying algebraic relations and by the fact that some of the information needed to solve them is implicit rather than explicit. In such situations, the generation of a problem representation by the subject becomes an important part of the task.

With regards to protocol analysis, the computer program provides a semi-automatic encoding system that allows a human coder to "perform a sort of semantic parse" (Bhakshar and Simon, 1977, p. 199). A representation of the basic processes involved in solving thermodynamics problems, together with their organization into sequences, underlie SAPA. These representations can be considered as the instantiation of a partial theory of the problem solving process, and the program assumes that protocols can be described in terms of this set of definite sequences of basic processes (although the encoding can deviate from the pre-established structures in order to account for the steps observed in an actual protocol). The completeness and accuracy of the theory is verified by the extent to which protocols can be encoded easily into the normal process sequences.

Bhakshar and Simon's system includes an encoding scheme for think-aloud protocols that is related to parsing natural language: SAPA parses protocols by identifying constituent structures and relating them together according to rules defining their ordering. The design of SAPA also points out the relationships between a well-defined encoding scheme for a task and a model of the problem-solving processes in that task. Finally, this extension to an ill-structured task emphasizes the process of generating a problem representation, since some useful information is not explicit in the problem formulation. In their research on the generation of problem representations, Hayes and Simon (1976) have developed a computer simulation that interprets the natural language instructions for isomorphs of the Tower of Hanoi. The program demonstrates the usefulness of a semantic representation model based on case grammars for representing problem–solving information.

The essential component processes of an encoding scheme, as they arise from attempts to develop automatic protocol analysis systems, have been presented in the preceding section. These processes include the construction of linguistic, semantic and psychological representations. In general, automatic encoding systems design underlines the importance of models for representing and extracting the semantics of natural language. However, by focusing on well-defined and semantically simple task domains, it has been possible to avoid careful and detailed analyses of the verbalization in encoding protocols. Within these task domains, knowledge is limited and can be expressed through simple representational formalisms, and the discourse used to express this knowledge is highly predictable. The semantic encoding entails assumptions that stem from the limited task environment, and much is already known about the problem-space beforehand, by the subject as well as by the coder. As a consequence, the think-aloud discourse is not ambiguous and can be handled through simple analysis procedures based on pattern-matching.

Think-aloud protocols in semantically rich and ill-structured tasks need "deeper" analyses of language, more inferences, and a more complex representation of knowledge. In writing, for example, the analysis of protocols does not benefit from a complete prior specification of elementary processes that can then be easily identified in the protocols. The task analysis does not identify, except for elementary grammatical conventions, clear constraints that writers follow and that would guide the analysis of their thinking-aloud. Consequently, the extraction of knowledge has to be done by postulating only general encoding categories (e.g., operators and goal statements), and by then examining carefully the protocol itself in order to extract specific encoding categories that will further define the problem-space. This kind of approach has been taken, for instance, by Kuipers and Kassirer (1984) in research on causal reasoning in medicine. Thus, there seems to be a trade-off between the ease with which the logically necessary processes involved in a task can be identified, on the one hand, and the grain with which the performance of subjects engaged in this task have to be analyzed, on the other.

## Think-aloud protocols in research on writing

Although some progress was initially made in research on writing by the application of the problem solving theory, methodological improvements are required in order to develop models of writing that are comparable in precision and scope to models of simple problem solving.

Hayes and Flower (1980a), for example, have applied the traditional approach of problem-solving research to the study of writing. They have used the think-aloud method to develop a model of the writing process. The analysis of protocols is based on an analysis of the task in terms of planning, expression and revision, an adapted version of earlier theories of planned behavior, notably Miller, Galanter and Pribram (1960). The definition of these processes into sub-processes eventually provides an intuitively sound basis for the analysis of the writing process. The accuracy of the encoding, however, depends mostly on the shared implicit understanding of how these categories of process are to be distinguished and recognized in think-aloud protocols (i.e., how are statements from a protocol to be matched against these categories, what criteria are used in assigning a given category to a statement). The only index of the quality of this shared understanding is the computation of the inter-rater reliability, which, in the case of Hayes and Flower is quite good, but indicates nothing about the validity of the definitions.

Similarly, the research of Scardamalia and Bereiter and their colleagues (Burtis, Bereiter, Scardamalia, Tetroe, 1983; Scardamalia, Bereiter & Steinbach, 1984) is partly based on various distinctions between types of planning or types of processes that do not seem sufficiently well-defined to ensure a reliable encoding. The encoding tactic is either based on Hayes and Flower's (1980a) analysis of planning, or consists in experimental manipulations that facilitate the desired types of processing and train subjects in verbalizing them.

It seems that a straightforward application of the methods of standard problem-solving research does not take into account the fact that writing is radically more complex and ill-structured than other tasks that have been previously studied through the use of think-aloud protocol analysis.

However, the well-defined models and analytic procedures developped in the area of research on discourse communication and knowledge representation (Frederiksen, 1975, 1986; van Dijk & Kintsch, 1983) complement these methods and make the initial objectives of think-aloud protocol analysis possible in semantically complex tasks. These models, which can be applied to the analysis of cognitive processing in increasingly complex task domains (Frederiksen & Breuleux, in press), provide a *tactic* for addressing the problems of think-aloud protocol analysis in ill-structured tasks. The term tactic refers here to the techniques involved in representing encoding categories, and in identifying instances of these categories in a think-aloud protocol.

The remainder of this paper presents such an encoding tactic for writers' think-aloud protocols, based on a well-defined semantic model of discourse. I will first explain in what respect writing is an ill-structured task, and then emphasize the major characteristics of problem-solving processes which underlie writing.

## Writing as an ill-structured task

Writing tasks vary in the level of structure which they impose on the writer. For example, an extremely constrained writing exercise would be to produce a description of a simple sequence of events following its temporal order. In such cases, writing requires a detailed memory representation for the events, and some skill in expressing that information in linguistic structures; however it involves few problem-solving processes because the structure of the text to be produced is dictated by the structure of the events. Since control over the initial mental representation in an experimental writing task may alter the production process, some researchers believe it should be left relatively unconstrained (e.g., Danks, 1977). Less constrained writing situations, in which the experimenter does not control the input to the task, usually involve more reorganization: the writer has some idea of the goal text, but no immediately available ways to achieve that goal[1], and must also determine important aspects of the task, like the pragmatic context.

Zammuner (this volume) provides an account of some of the differences between writing and the tasks that have been traditionally studied in problem-solving research. Writing is particular in that it usualy involves a large body of possibly relevant semantically complex knowledge states that the writer has in memory. It also involves linguistic and psychological structures and processes associated with the transformation of these knowledge structures into discourse structures. In most writing research contexts these representations and processes cannot be fully specified in advance. The completeness of these specifications, then, cannot be considered as the only starting point for research on writing. In situations in which the semantic content involved in a writing assignment is not precisely determined by the task definition provided to the subject, and where the actual knowledge used by the subject cannot realistically be delimited beforehand, it is possible to infer these representations and processes during the analysis of the think-aloud protocol produced by a subject, provided that the appropriate tools for the analysis and representation of knowledge from discourse are used. For instance, a partial but precise view of protocols is possible by defining semantic grammars for general aspects of the hypothesized processes and structures involved in writing. In the next section, the general components of writing that can be used in developing a coding scheme are presented, and a tactic for deriving an encoding scheme from the definition of these components is described.

## The tactical use of semantic grammars

Aside from its unconstrained nature, there are a number of invariant aspects that a task analysis for writing reveals. The production of discourse consists in the generation (or retrieval) of a conceptual, abstract, network of information that must then be expressed, that is, transformed into a material, grammatical and linear text. A number of levels of representation are implied in this process: at the extremes, a conceptual representation level for semantic information in memory (or content) and a discourse representation related to the language structures used to encode and signal the conceptual representations (Frederiksen, 1986; Kintsch, 1982). Each of these levels can also be

---

[1]Although expertise in writing usually involves also a knowledge organization directed by goals related to writing, and, thus, knowledge is organized adequately for writing purposes.

subdivided according to different emphases. For example, the research of Flower and Hayes (1981) and Scardamalia and Bereiter (1985) focus on the control knowledge used in writing and usually deals with only the two basic levels of representation. In contrast, the work of Beaugrande (1982), Bracewell (1986), and Frederiksen, Donin-Frederiksen and Bracewell (1987), which is oriented toward the definition and mapping among different levels, assume multiple levels of representation for both semantic structures and language structures. These levels correspond to the *proposition* and *frame* structures encoding information in memory, to the *syntactic structures* encoding propositions in text and the *topical patterning* and *cohesive devices* signalling frame components in the text (Frederiksen, 1986).

Processing within and among these various intermediate levels gradually instantiates the meaning intended by a writer into discourse. At the conceptual level, writing implies the retrieval of information from memory, stored as knowledge *frames;* the writer must then select from the frame structure a restricted quantity of information and specify the exact content of the propositions. At the level of surface structure, the task of writing implies operations controlled by the rules of syntax and the elaboration of the objects corresponding to syntactic structures of a language: the clauses. The process of building these linguistic objects encompasses a limited set of operations involving the determination of a propositional content, lexical decisions and topic choice, i.e., transformations of semantic propositions into syntactic structures. This view of writing emphasizes the rules responsible for the coherence of information structures in memory and in discourse, as well as on the mappings between these different structures. Discourse structures are produced according to two major levels of constraints: they must encode information that correspond to propositions, and they must signal the higher level structure of the frame in which the propositions take part.

### *Writing as a transformation problem*

The analysis of writing in terms of representations at different levels being mapped onto each other emphasizes that writers operate on an initial situation —the information in memory— and transform it into another situation —text— that constitutes the goal of the problem. The writer thus has to derive representations from an initial representation according to certain rules. This analysis shows, in terms of Greeno's typology of problems (Greeno, 1978), the strong components of *arrangement* and *transformation* in writing. It is important to note that, in contrast to well-defined problems, the exact rules to attain a solution (mapping one representation onto another) are not given as part of the problem, and the nature of these rules is an empirical matter. The production of discourse representations from knowledge structures implies the ability to generate and evaluate potential solutions. In this sort of arrangement problems, the search space is particularly vast, and the discovery of a solution requires the availability of principles to avoid verifying all possible solutions randomly (Greeno, 1978). Apart from these arrangement processes, writing also implies transformation processes. Once some content has been generated and chosen, a writer must employ a certain number of transformations, at least to build language structures corresponding in specific ways to this content. In this, writing is similar to theorem proving tasks (Newell & Simon, 1972), structure-building in block worlds (Sacerdoti, 1977), and the Tower of Hanoi (Newell & Simon, 1972). I will emphasize here this transformation component in writing, and the planning strategy associated with the solution of transformation problems.

### *Planning as a major strategy in writing.*

In transformation tasks, the ability fundamental to attaining a solution seems to be *planning* (Greeno, 1978). As we will see from the analysis below, planning provides a useful strategy in writing because it enables the exploration at an abstract level of the rich network of possible solutions in situations where the actual undertaking of these solutions would be time consuming and cognitively costly. This exploration is extremely valuable because of its reduced cognitive cost, compared to the actual undertaking of a solution in the actual problem environment. Stepping into the problem without planning would involve producing ultimately worthless text that satisfies a complex set of constraints, including those pertaining to syntactic structures. By planning, a writer can delay much of these lower level decisions, like the specific ordering of goals and the elaboration of surface structures. The importance of planning in writing has been emphasized in most research on language production (Butterworth, 1980; Flower & Hayes, 1980; 1981; Matsuhashi, 1987).

I will give a more detailed account of what is generally meant by planning and by plans. This analysis of planning will provide the basis for an encoding scheme.

### *Analysis of planning*

Planning is "the judicious postponement of commitments". It is generally defined as the hierarchical and sequential organization of goals and subgoals representing a course of action (Miller, Galanter & Pribram, 1960; Newell & Simon, 1972; Volpert, 1982). I will summarize here some of the most important features of planning and of plans.

Newell and Simon (1972) have illustrated how planning consists in the *abstraction* of objects and operators of a problem to create a manageable representation in which a blueprint for the solution in the original problem space is produced (Newell & Simon, 1972). This process permits the exploration of a number of possible solutions. The usefulness of the strategy lies in the cognitive bargain resulting from a search through a small and abstract problem space, as opposed to working out a solution in the original, detailed problem space.

Planning results in a solution outline —a plan— that can then be translated and tried out in the original problem space. A plan coordinates the selection of operators, and its execution involves making behavior conditional on the state of the problem and a hierarchy of goals and sub-goals (Simon, 1975). Complex plans can be represented as networks in which the nodes are *goals*, mainly actions, related one to another by hierarchical relations (specifying levels of detail and decomposition), logical relations (conjunction and disjunction), and temporal relations (specifying an order between actions). The work of Sacerdoti (1977) shows that consistent planning can be achieved in computer systems capable of deriving plans at different levels of abstraction, gradually formulating more detailed plans until executable procedures are organized into a coherent sequence. The execution of plans relies heavily on monitoring, the ability to *evaluate* and compare planned and actual states of a problem.

Finally, since goals represent anticipated situations or actions that will take place in the future with a certain probability, therefore *temporal antecedence* over action and *conditionality* are essential features of goals (Rips & Marcus, 1977; Simon, 1975).

### *Derivation of a semantic grammar for planning*

This analysis of planning provides sufficient information for defining a type of psychological representation (a plan) that should result from the analysis of protocols. By taking into account the major semantic characteristics of the verbalizations and actions during planning, these features of planning and plans can be refined further into an encoding scheme for the analysis of writers' think-aloud protocols. The general assumption behind the encoding scheme is that representations for goal statements, which are the basic entities in plans, can be defined by configurations of proposition arguments.

The semantic representation model developed by Frederiksen (1975, 1986) seems appropriate for representing the content of think-aloud protocols in semantically rich tasks such as writing because it is a well-defined system that enables the representation of complex entities including states and their identifying relations (e.g., attribute, temporal and location relations), events and their case relations (such as agent, source, result and theme), dependency relations (cause, condition, and, or), algebraic relations (order, equivalence), and complex functions. Beaugrande (1981) presents a comparative account of this representation system. Hayes and Simon (1976) have used a similar representation system for simulating the generation of a problem representation. Furthermore, Frederiksen's propositional model has been implemented as an interactive computer-based system, CODA, that guides a user in the construction of propositional representations from natural language.

This representation model can be used as the basis for a precise definition, at the levels of semantic propositions and frames, of the knowledge elements and information processing operators in a problem space. Operators are represented as resultive events, knowledge states as processive events, states or relational propositions. Operators and knowledge states are linked together by an object or theme case relation, except for explicit transformation operators whose associated input and output knowledge states are placed in source and result relations, respectively, with the operator. These definitions of entities revealed by a task analysis can then be mapped onto the discourse level in order to define an encoding scheme for think-aloud protocols of writers. This tactic is similar to the development of a semantic parser, and this issue will be discussed in the next section.

The illustration to follow will focus more specifically on the definition of goal verbalizations and actions, based on the abstraction, temporality, and conditionality features of goals. The features of goals identified from the preceding analysis of planning are summarized in Table 1. Table 1 also gives, for each major characteristic of goals, a set of semantic attributes that characterize the corresponding verbalizations and actions, with examples from a protocol that will serve to illustrate the encoding system later on. In the semantic representation, these semantic attributes involve configurations of case relations and tense, modality, and truth value markers. Goals can consist either of operations that are possible, postponed, intended, or of states that are expected to obtain. Table 1 presents both cases. In verbalizations of goals, states are usually expressed as results of explicitly verbalized operators, although they can be expressed without their effective operator, and operators without their resulting state (see Waterman & Newell, 1971). In addition to the semantic characteristics of verbalizations and actions, the encoding requires taking into account the relationship between verbalizations and actions. Goal states can be identified by the fact that they have no reference to already existing states. The absence of any direct equivalence between what is verbalized and earlier text is one indication that what is verbalized is content that the subject *wants* to write. Similarly, goal operators can be identified as events that are not achieved concurrently, implying a result state (resultive events) that does not already obtain. Similar types of definitions can be constructed for evaluation statements, which are important in the monitoring of planned actions during the execution phase. Statements that consist in the attribution of some quality to an already achieved state, or that consist of a processive event indicating a sort of state of mind, can readily be identified as evaluation statements.

The information in Table 1, along with the definitions given above, represent a first step in deriving a well-defined semantic grammar from the task analysis. The encoding scheme is obviously more complex, involving a representation of the concurrence of actions and verbalizations by both temporal equivalence and content equivalence (see Breuleux, 1988).

**Table 1**

Characteristics of goals and semantic characteristics of goal statements (operators)

| ***Characteristics of goals*** | ***Semantic characteristics of goal statements and actions*** | ***Instances from protocol of SH3*** |
|---|---|---|
| Temporal antecedence over action | temporality marked as future (*FUT*) | 3. it's going to be...<br>16. I will talk about...<br>17. I will bring in...<br>19. I will draw a parallel...<br>28. I will see about that... |
| Conditionality | disjunctive relations between goals; (*ORALT; OREXCL*) | 285. either I drop the show or I drop Reagan |
| | interrogative truth value (*INT*) | 8. what could I talk about...?<br>14. why not the esthetics...?<br>27. why not conclude on...? |
| | marked modality (*QUAL, CAN, COND, ROOT*) | 12. I've got to find...<br>14. it can change...<br>172. I'd talk about...<br>292. I could put... |
| Volition | Volition lexical identifier (lexical) | I want to... |
| Abstraction | Fragmentary, incomplete, illegal actions | 9. demystification space<br>16. (1) intro shuttle<br>21. exactly plane explodes |

Note: Terms in italics refer to categories in Frederiksen's (1975) propositional grammar.

Given that planning consists of the coordination of goals, the identification of a plan in a protocol requires, in addition to a definition of goals, the identification of relationships among goal statements. More specifically, a reliable definition should be provided for constructing the hierarchical, sequential, and logical organization of goals. In the propositional representation used here, hierarchical relations consist mainly of location (LOC:), part-whole (PART:), theme (THM:); sequential organization is represented by the temporal order relation (ORD:TEM), and logical relations are represented by AND and OR relations. The semantic characteristics of the major goal relations are summarized in Table 2, together with examples from actual protocols. Relations between goals can be explicitly verbalized or visible from planning actions (arrows, itemization and ordering marks, etc). The semantics of these, although not fully specified here, can be represented within the same model as the verbalizations (see Frederiksen, 1975). In addition, relations sometimes have to be inferred mainly on the basis of default relations, semantic argument overlap, or the coder's semantic interpretation.

**Table 2**

Semantic characteristics of goal relations

| *Type of relation* | *Semantic characteristics of goal relations* | *Instances from protocol of SH3* |
|---|---|---|
| Sequential Relations | Temporal order (*ORD: TEM*) | I'll keep on with what I'm talking about and **after that** I'll reformulate |
| Hierarchical Relations | Part-whole relations (*PRT*) | I'll develop on that theme **by** putting... |
| | Theme relation (*THM*) | I have to think **about** what I'll write... |
| | Goal relation (*GOAL*) | **In order to** keep on with the idea I'll say... |
| | | **To** conclude I'll put... |
| | Location relation (*LOC*) | I start my paragraph **where** I'll talk about... |
| | Result relation (*RSLT*) | I want to integrate it **so that** it relates to... |
| | indentation, itemization and ordering marks | "**problem 1** fellowships problem **2** then after that **3**..." |
| Logical Relations | Disjunction (*ORALT, OREXCL*) | I'm wondering if I will not qualify more the decision **or** if I'll continue to say what I'm talking about... |

Note: Terms in italics refer to categories in Frederiksen's (1975) propositional grammar.

On the basis of such a definition of plans as goal structures in which nodes and links are explicitly specified a grammar can defined in order to represent the constituent structures that can be extracted from think-aloud protocols. A tentative ATN grammar of this kind is presented in Figure 1, defining a plan as a sequence of goals (two or more) that can be related together. The grammar can be applied to a think-aloud protocol by identifying related goals on the basis of their characteristics defined in Table 1 and Table 2, so that the exact sequences of planning and writing processes in protocols are identified reliably. The grammar can be implemented as a parsing program that takes as input the propositional representation for protocol statements and categorizes the appropriate ones as goals.

The grammar is partial, since the task analysis here deals only with a single, although important, aspect of writing. But it constitutes an explicit encoding scheme for planning processes in writers' think-aloud protocols. The encoding scheme is based on a very limited set of assumptions regarding the

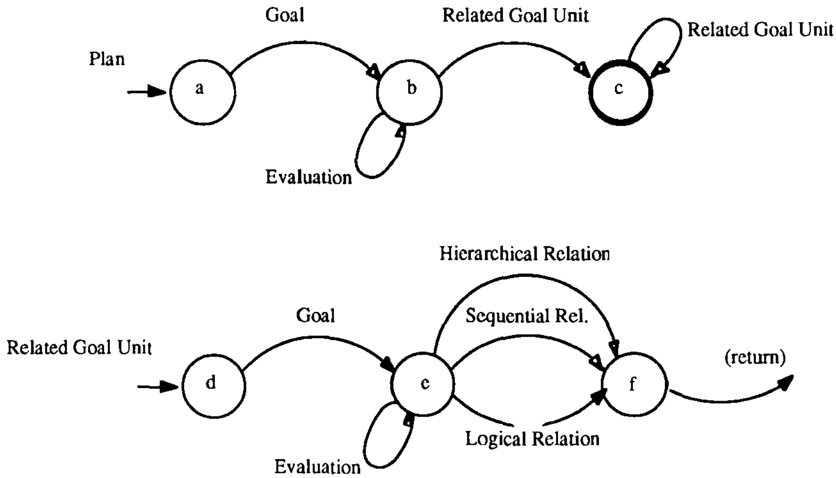



Fig. 1. An ATN grammar for plans

semantics of goal structures (plans) and on the characteristics of their verbalizations in discourse. It is, however, aimed at gradually inducing more specific properties of planning by extracting them from protocols.

### *Illustration of the coding procedure*

The encoding scheme presented above has been used to analyze the protocols of both expert and pre-expert writers in situations that are not very constrained. It has been very useful in determining whether planning occurs, as well as the content and context of planning occurrences (see Breuleux, 1988). I will illustrate it's use with examples from the protocols of two pre-expert writers composing an editorial type of paper about the American space shuttle accident for a student newspaper. The protocols were collected a few days after the accident. The subjects are French-speaking students in a bachelor program in writing and documentation at the University of Sherbrooke. The examples provided here were translated from French by the investigator for the purpose of the illustration. The topic was suggested, rather than imposed to the subjects. It was chosen by the experimenter on the assumptions that the subjects would have an opinion about the event because it had been very important and controversial, while at the same time these opinions would not yet constitute a well-organized knowledge structure because of the recency of the event. Thus the task was assumed to elicit spontaneous organization of content and planning. The subjects were using paper and pencil, and were asked and trained to think-aloud as she wrote. The sessions were video-taped, the think-aloud transcribed verbatim, and the transcripts were analyzed to identify episodes of planning and execution of plans and to represent the content of plans. The protocols were first segmented at syntactic clause boundaries following Winograd (1983).

Planned writing involves setting up plans and executing them. The elaboration of a plan can be conceived of as the organization of content, labels for knowledge structures being high-level goals, and the execution of a plan as the actual writing, i.e., the realization of goals into sequences of surface structures in text. The execution phase consists essentially in producing grammatical sentences to form a text, according to the goals in the plan that controls the actions. Plans, however, can appear at various levels along the hierarchy of processing already mentioned in the analysis of writing as a transformation problem. Initial content-related plans that encompass the whole text to be produced can be considered highest in the hierarchy, together with overall rhetorical plans (e.g. to convince or to trick the reader). More local plans are subordinated to goals that are set earlier, or do not involve new content but lower levels of the writing process (e.g., signalling, syntax). The analyses that follow more specifically concern the extent to which, in writing, the high-level decisions made during planning interact with decisions at lower levels during the execution of plans.

The protocol excerpt from subject SH3 presented in Table 3 corresponds to an important planning episode taking place at the onset of the task. In this episode, the major goals for the overall task are determined and organized. In the verbalizations, the temporal order between goals is often indicated explicitly by temporal adverbs like "first" and "then". Other clues on the relations between goals come from the writing produced during this episode (see the itemization and embedding marks presented between brackets in Table 3).

The excerpt in Table 3 is actually composed of two small episodes. In the first part, from segment 3 to segment 15, the subject jots down the major themes to be included in the text: the nice show or spectacular side of the event, Reagan and the general question of disasters helping to keep political leaders in place, the successful demystification of space, and the esthetics of the event. In the second part, from segment 16 to 26, the subject elaborates these themes and organizes them together. She also adds one new theme (the parallel with an event in Colombia), deletes one (the esthetics) because it seems redundant with the show theme, and includes some rhetorical structures (introduction and conclusion). The indicators of planning in these episodes include both the semantic of the verbalization (especially the future marking of segments 16 to 18 and the modality and interrogative truth value on segment 26), and the abstract properties of the actions, which are "illegal" textual actions since they result in single noun phrases ("live television") or incomplete sentences ("exactly planes explodes"[1]). The resulting plan is presented graphically in Figure 2. The plan is labelled in the figure by a "head" (the thick oval node); the square nodes linked under the plan head represent the content of the plan, a structured set of goals. Numbers in plan heads correspond to segment numbers in the protocol, and can serve to indicate the order in which the planning episodes occurred.

[1]"Exactement avion explose" in the original protocol.

**Table 3**

Initial Planning Episode from Subject Sh3

| Segment no | Segment |
|---|---|
| 3.* | here it's going to be first the **(-) nice show** side **(,) good time** a bit, in a certain way |
| 4. | it's still vague but it's |
| 5.* | after that there's Mr **(-) Reagan** in that |
| 6.* | and the sort of **helping hand** it provides this sort of things to leaders |
| 7.* | because they have the opportunity to put themselves into **spotlight** |
| 8.* | then there is what could I talk about? |
| 9.* | there is the famous **demystification** of **space** |
| 10.* | **(->) successful**, very successful |
| 11. | ok after that |
| 12.* | I have got to try to find one more idea at least |
| 13.* | I have to find something else |
| 14.* | why not the **esthetics** of the thing |
| 15.* | it can change but |
| 16.* | ok so in the **(1) intro**duction I will talk about the **shuttle** as such a little |
| 17.* | then I will **bring in** my **ideas** a little |
| 18.* | ok here it's going to be **(1)** the **show** the **live television** aspect and the **kids crying** |
| 19.* | **(II) Reagan** who puts himself under the spotlight that I will make a parallel **(=)** with **Colombia** when there was nearly going to be a **coup (->)** and then when there was a volcano eruption the guy I don't remember his name it had given him a hand and then finally he stayed there |
| 20.* | and then the school teacher **Mrs McCauliff** she wanted to demystify space |
| 21.* | and and she demystified it very well it's **exactly** like a **plane** it **explodes** |
| 22. | and then **(III)** the esthetics |
| 23. | no |
| 24. | it's not necessary finally |
| 25. | I take this out because if I talk about the show side it's the same thing **~~(esthetics)~~** |
| 26.* | ok in conclusion here why not **conclude** on the **human comedy** |

Note. Segments corresponding to goals are preceded by an asterisk.
Items in bold indicate information written by the subject, items within brackets were not verbalized.

The protocol section presented in Table 4 is an illustration of a transition episode, when the subject is moving, at segments 165 and 166, from the currently active goal from the plan —the intro— to the next item —the show. First, in segment 165, there is an overall evaluation of the introduction, indicating that it is finished and satisfactory. Then, in segment 166, a bringing into focus of the next goal occurs.

The excerpt in Table 4 also contains an illustration of a plan reorganization. From segments 167 to 174, the subject reorganizes the remaining items from the initial plan, and in the course of doing so she provides clear evidences on the motivations behind this reorganization. The subject clearly indicates, in segment 169, that it is the ending of the introduction (on the wondrous aspect of the space adventure) that suggests bringing in the demystification theme and pushing in third place the spectacular aspect of the event. The subject also anticipates a link between the spectacular aspect and the theme on

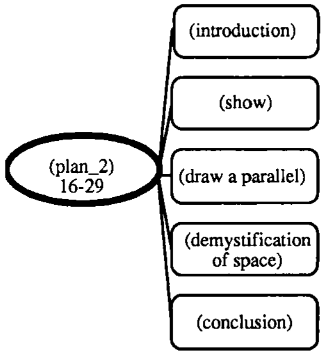



Fig. 2. Initial plan from subject SH3

Reagan, based on the fact that Reagan used to be an actor. Finally the first item in the resulting plan is simply an empty cohesive devise: "on the other hand". All the goals in this passage are identified on the basis of a modality marking ("have to...", "would", "could"). The structure of this reorganization is presented in Figure 3.

A second plan reorganization takes place later on, after the execution of the demystification theme. The protocol section for that episode is presented in Table 5. In segments 280 and 281, the subject indicates she is getting at the next goal in her plan, the spectacular aspect, but she starts taking into account the time and space constraints imposed for the task; this constraint inclusion is indicated in segments 283 and 284. The subject decides that she must drop one of the next major thematic goals in her plan: the spectacular aspect or the theme on Reagan. She weights the advantages of each possibility, as indicated in segments 285 to 288. In consistence with the earlier preparation for the introduction of the spectacular aspect, she chooses to go on with this theme and to drop the anecdote involving Reagan, although she finally tries to insert that in the conclusion by including the goal of having a little punch at Reagan. In some respect, this reorganization is similar to the first one: it is driven by "local" properties or newly introduced constraints.

Figure 4 presents the planning for an important portion of the protocol of a second subject, SH2, who did not reorganize any of his plans along the way. The protocol shows, however, transition episodes during which the subject carefully monitors the execution of the goal sequences in his plans. The protocol represents a very neat top-down application of plans, without corrections. However, as in the case of SH3, subordinated goals are elaborated locally, the initial plan being a simple sequence of top-level goals. The refinement of each of these goals is done after the execution of the preceding one(s).

This partial account of two protocols illustrates certain characteristics of the writing processes that have been observed in other protocols as well: (a) planning is distributed over the whole time span of the task; (b) plans are

**Table 4**

First Reorganization from Subject Sh3

| Segment no | Segment |
|---|---|
| 165. | for the intro it's going not too bad |
| 166. | ok the show side |
| 167.* | I'd better wait |
| 168. | before I put that then |
| 169. | if I finish like that |
| 170.* | I'd be better with ok ok |
| 171.* | I'd put on the other hand |
| 172.* | and then I'd talk about the demystification of space |
| 173.* | and then I'd finish like the demystification of space and yes and so it's it becomes like a plane or like a good show |
| 174.* | and then I could oh yes with the show Reagan he's an actor it could make a link somewhat ok ok |

Note. Segments corresponding to goals are preceded by an asterisk.

shallow and decomposed locally; (c) sub-optimal situations are detected and repaired in context.

In the protocols of SH2 and SH3, planning episodes alternate with execution phases. The initial plans are gradually refined, each goal being expanded, usually through further planning, into a piece of discourse. Plans are kept simple until the point of expansion, and the expansion process leads either to expression, elaboration of a sub-plan or modification of a previous plan. The expansion process can be linked with transition episodes, during which the subject decides on the type of expansion by evaluating text and high-level goals. The plans observed in these protocol as well as others are remarkably simple and shallow: very little hierarchical decomposition appears in the absence of actions. The plans are vague, and we can speculate that subjects would rather not commit themselves to particular objectives or that they cannot anticipate precisely the outcome of further processing (namely the expression process). Such plans are either decomposed locally or reorganized according to local properties of the problem-state. The occurrence of plan reorganizations is consistent with the view of writing as an ill-structured task, in which individuals can redefine relevant characteristics of the task, and where the consequences of moves are not easily anticipated (Simon, 1978). The context in which these reorganizations take place suggests an ability to take into account unexpected situations and local characteristics during the plan execution process; it implies also that low level characteristics of actions can bring about modifications at higher-levels of representation. This process is similar to that of Sacerdoti's (1977) computer model of planning, which includes a correction mechanism based on the interaction between actions recently added to the representation and on the coherence between these detailed actions with the more abstract plans. One question arising from this is whether the ability to elaborate detailed plans is simply a function of

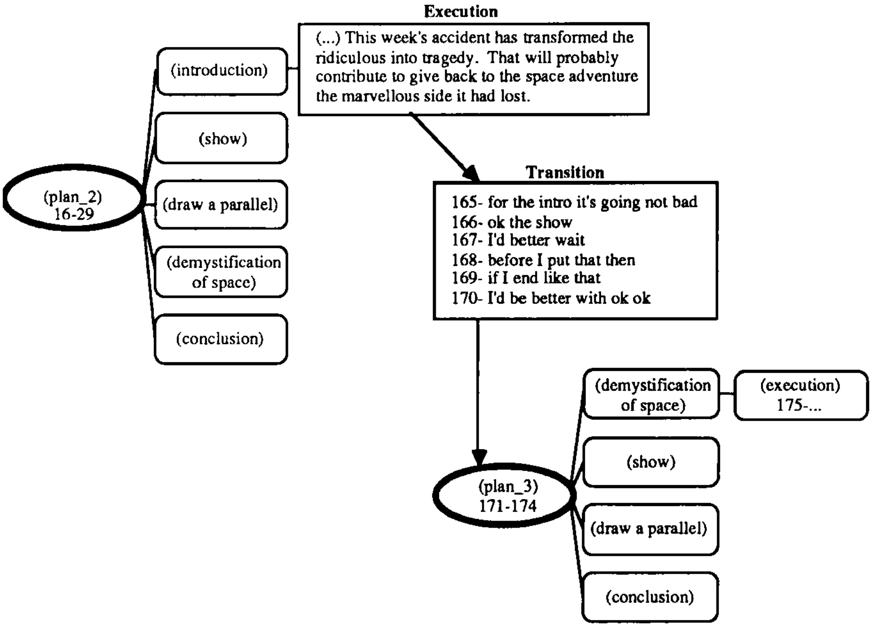



Fig. 3. First reorganization in the protocol of subject SH3

knowledge available on the topic or whether it is a stable characteristic of individuals. Further research should try to determine to what extent depth of planning is related to expertise in writing or familiarity with content domain.

The ATN grammar presented here is a first approximation of a model of planning in writing. Since it postulates only general processes, the model's definition is as yet still rough. The rules of the grammar, however, are meant to be refined according to the distinctions that can be made, among the plans identified in protocols, between different types of planning. Similarly, the encoding permits the gradual inductive construction of a problem-space that cannot be determined by a task analysis alone. For the purpose of inducing the problem-space and the production system for a subject from a protocol, Waterman & Newell (1971) suggest a lexical approach based on listing out knowledge elements and operators that a subject is using. This list can then be incrementally reentered into the analysis procedure.

This can be done rather directly from the propositional representation of the think-aloud protocols. The operators used by the subject can be extracted by listing the act identifiers of all goal statements. Such a list, extracted from two protocols, including the one discussed here, is presented in Table 6. The operators have been grouped tentatively according to certain characteristics of the task. The operators in the first category of Table 6 are somewhat uninformative: they simply refer to the main action relevant to the task, expressing in written form. Sometimes the use of these "empty" operators indicates an *embedding* relation between two themes, since a majority of goals set up by

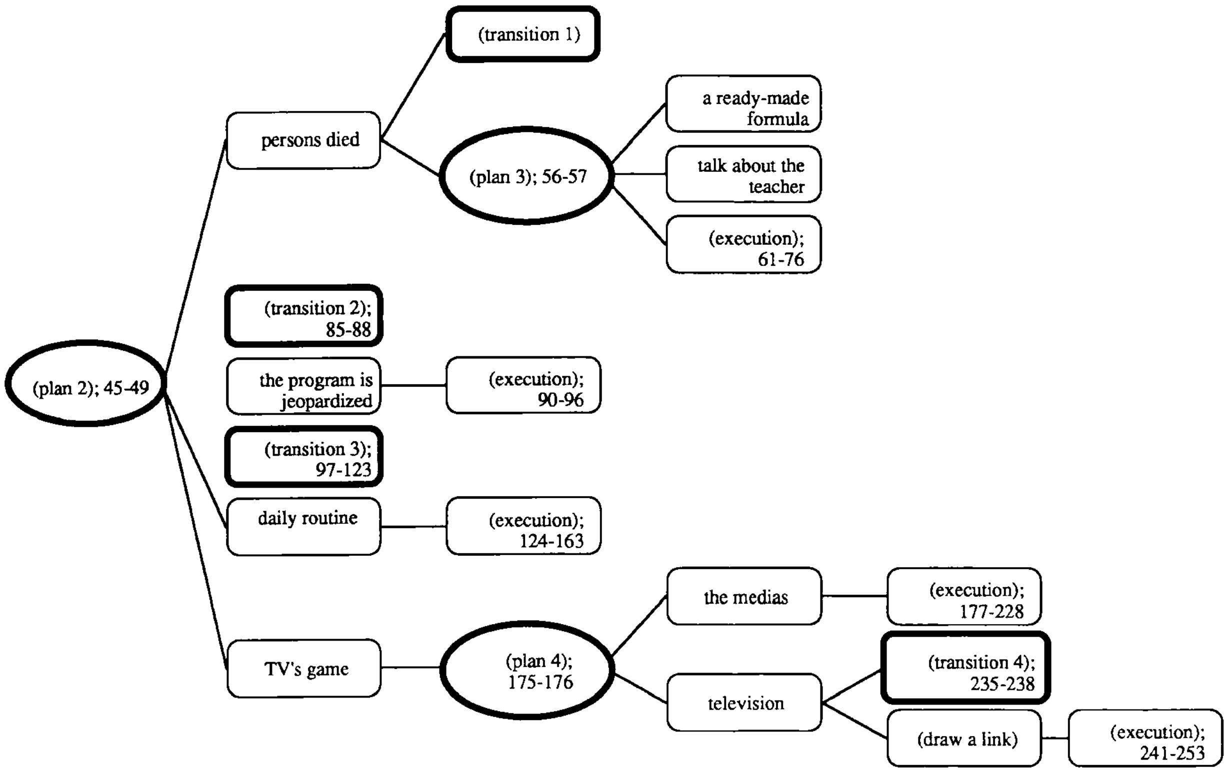



Fig. 4. Planning and execution of plans in the protocol of subject SH2

**Table 5**

Second Reorganization from Subject Sh3

| Segment no | Segment | Comment |
|---|---|---|
| 279. | well I've got about a page | |
| 280. | yes ok the show | |
| 281.* | I will talk about the show | |
| 282.* | it would be easy to get into that | |
| 283. | but then I already have a lot | |
| 284. | I wouldn't be able to | |
| 285.* | either I drop the show or I drop Reagan and his Colombia | |
| 286.* | that would take a long time to explain | |
| 287. | but it's interesting | |
| 288. | the show | |
| 289.* | it would be easy to go on with the show | |
| 290.* | let's take the easy way | Subject crosses out item 3 of plan |
| 291. | and the conclusion after | |
| 292.* | oh I could put my little anecdote in the conclusion the human comedy | Subject draws a link between items 3 ("Reagan") and 5 ("conclusion") |
| 293.* | then finish with a little punch at **Reagan** | Subject writes under item "conclusion" of her plan |
| 294.* | **punch** | |
| 295. | ok the live show on television ok | |

Note. Segments corresponding to goals are preceded by an asterisk.
Items in bold indicate information written by the subject.

writers are achieved by actually writing something. The second category of operators in Table 6 are *structure building* operators. These operators indicate certain structural functions of sections of a text, like opening, ending, reintroducing a theme, or rhetorical functions like explaining and summarizing. The third class of operators in Table 6 are *structure modification* operators that can be applied to either text structures of plan elements. The final category in Table 6 refers to very general activities that have only indirect effects on the state of the problem; for that reason, they are termed *meta-operators.* A further analysis would lead to the specification of the classes of knowledge states that are admissible to any given operator. Waterman and Newell (1971) present a similar table of operator-state associations for cryptarithmetic problems. The list here is incomplete since it contains exclusively the operators that were explicitly verbalized. The analysis of the discourse structures produced could be examined in order to augment the list and to see what actions are associated with the operators.

**Table 6**

List of operators verbalized by SH2 and SH3

| Embedding | Structure building | Structure modification | Meta-operators |
|---|---|---|---|
| say | start with | change | wait |
| write | introduce | edit | put aside |
| mark | bring in | take out | plan |
| put | bring back | start again | re-read |
| talk about | get back at | | get back to |
| attach | use again | | |
| | get to | | |
| | link | | |
| | get away from | | |
| | describe | | |
| | dramatize | | |
| | explain | | |
| | summarize | | |
| | treat | | |
| | conclude | | |
| | end | | |

## Computer-based tools for the analysis of think-aloud protocols

The encoding system presented here is based on a well-defined semantic model that has already been implemented as a computer-based interactive analysis program (CODA). An automatic system is also currently being developed by interfacing the semantic model with a syntactic parser for English (Frederiksen, Décarry & Hoover, 1988). This system can automatically generate semantic representations from sentences by the activation of a set of rules that relate the components of propositions defined by the semantic model with syntactic informations generated by the parser.

The information structures and rules for extracting and representing plans can be thought of as a further layer on this semantic parser. The result would be a modular automatic encoding system based on well-defined analyses of syntax, semantics, and problem solving processes. Although this system has not yet been completely worked out, the process of generating plan structures from propositional representations of protocol statements has been implemented partially. The result of this implementation are encouraging. Further development will involve testing and adapting the syntactic component of the parser for think-aloud discourse.

The development of a parser has a number of motivations, including the need for explicitness and formalization. The increase in explicitness is achieved by computer implementation of the grammars and of the analysis process. This implementation in turn produces tools that facilitate the process of analyzing think-aloud protocols by reducing the time required for the analysis and by increasing the number of protocols that can be handled in a research project. Most of all, the implementation increases the objectivity and the reliability of the coding and enables the comparison of protocol data across conditions in an experimental design.

## Conclusion

In congruence with the ill-structured nature of writing, the encoding tactic described in this paper is based on a deliberately small set of assumptions about planning and other processes in writing. With regards to building a model of writing, the tactic consists of initially hypothesizing a small set of general productions in view of increasing their number and specificity according to the fit to the protocol (see Waterman & Newell, 1971, and Kuipers & Kassirer, 1983).

The tactic is based upon a detailed semantic analysis of the clauses in think-aloud protocols of subjects. While the procedure may appear tedious, work is currently under way to design computer tools to assist and automatize parts of it.

The technique presented provides data on a subject's representation of the task and on his strategies, from which it is possible to specify and test aspects of a model of writing. It can be used, for example, to ascertain the importance of planning and distinguish different types of plans. Since it provides detailed information on the behavior of individuals it is useful generally in investigating individual differences. Furthermore, we can anticipate systematically looking at how expert writers and novices differ with regards to their use of plans. Future research will also focus on the relation between planning and various sources of knowledge, mainly domain-knowledge and writing skills. The verification of hypotheses related to these issues depends on a reliable method for identifying plans from think-aloud protocols.

This work is also relevant to any research activity involving the use and analysis of complex verbal data and natural language protocols. The tactic described in the present paper can be applied, for example, to knowledge acquisition in the development of expert-systems or tutoring-systems, in order to produce representations of expert knowledge and pedagogical strategies that are both explicit and congruent with human problem-solving methods. These characteristics are necessary for ensuring that the reasoning processes are comprehensible (see Clancey, 1984).

Finally, this work can be applied to protocols collected on well-defined problems. The benefits of such an application are two. First, an analysis based on a semantic grammar would provide parallel and independent data that could be compared with data collected using more subjective procedures. Second, the methods could be used initially in the development of more valid subjective coding procedures for protocols from well-defined tasks.

## References

Bhakshar, R. & Simon, H.A. (1977). Problem solving in semantically rich domains: an example from engineering thermodynamics. *Cognitive Science, 1,* 193-215.

Beaugrande, R. de (1981). Design criteria for process models of reading. *Reading Research Quarterly, 16,* 261-315.

Beaugrande, R. de (1982). *Text production: Toward a science of composition.* Norwood, NJ: Ablex.

Breuleux, A. (1988). *L'élaboration et l'exécution de plans dans une tâche de rédaction* [The elaboration and execution of plans in a writing task]. Unpublished doctoral thesis. University of Montréal.

Bracewell, R.J. (1986, May). *Signalling semantic representation: the role of topical patterning in text.* Paper presented at the Fourth International Conference of Teachers of English, Ottawa.

Burtis, P.J., Bereiter, C., Scardamalia, M., & Tetroe, J. (1983). The development of planning in writing. In B.M. Kroll & G. Wells (Eds.), *Explorations in the development of writing.* New York: Wiley.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and Talk.* New York: Academic.

Clancey W.J. (1984). *Acquiring, representing and evaluating a competence model of diagnostic strategy* (HPP Memo 84-2). Palo Alto: Stanford University, Stanford Knowledge Systems Laboratory.

Danks, J.H. (1977). Producing ideas and sentences. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory.* Hillsdale, NJ: Erlbaum.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis; verbal report as data.* Cambridge, MA: MIT Press.

Flower, L.S., & Hayes, J.R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. Gregg, & E. Steinberg (Eds.), *Cognitive processes in writing.* Hillsdale, NJ: Erlbaum.

Flower, L.S., & Hayes, J.R. (1981). Plans that guide the composing process. In C.H. Frederiksen, & J. Dominic (Eds.), *Writing: the nature, development, and teaching of written communication.* Hillsdale, NJ: Lawrence Erlbaum.

Frederiksen, C.H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology, 7*, 371-458.

Frederiksen, C.H. (1986). Cognitive models of discourse analysis. In C. R. Cooper & S. Greenbaum (Eds.), *Written Communication Annual: Vol. 1. Linguistic approaches to the study of written discourse.* Beverly Hills, CA: Sage.

Frederiksen, C.H. ,& Breuleux, A. (in press). Monitoring cognitive processing in semantically complex domains. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skills and knowledge acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Frederiksen, C.H., Décary, M., & Hoover, M.L. (1988) The semantic representation and processing of natural language discourse: Development of a computational model. *Proceedings of the International Colloquium on "Informatique et langue naturelle."* Nantes, France.

Frederiksen, C.H., Donin-Frederiksen, J.D., & Bracewell, R.J. (1987). Discourse analysis of children's text production. In A. Matsuhashi (Ed.), *Writing in real time.* Norwood, NJ: Ablex.

Greeno, J.G. (1978). Natures of problem-solving abilities. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 5. Human information processing.* Hillsdale, NJ: Erlbaum.

Groner, R., Groner, M., & Bischof,W.F. (1983). *Methods of heuristics.* Hillsdale, NJ: Erlbaum.

Hayes, J.R., & Flower, L.S. (1977). Problem solving strategies and the writing process. *College English, 39,* 449-461.

Hayes, J.R., & Flower, L.S. (1980a). Identifying the organization of writing process. In L. Gregg, & E. Steinberg (Eds.), *Cognitive processes in writing.* Hillsdale, NJ: Erlbaum.

Hayes, J.R., & Flower, L.S. (1980b). Writing as problem solving. *Visible Language, 14,* 388-399.

Hayes, J.R., & Flower, L.S. (1983). Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenthal, L. Tamor, & S. Walmsley (Eds.), *Research on writing: principles and methods.* New York: Longman.

Hayes, J.R., & Simon, H.A. (1976). Understanding complex task instructions. In D. Klahr (Ed.), *Cognition and instruction.* Hillsdale, NJ: Erlbaum.

Kintsch, W. (1982). Levels of processing language material: Discussion of papers by Lachman and Lachman and Perfetti. In L.S. Cermak, & F.I.M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ: Erlbaum.

Kuipers, B.J., & Kassirer, J.P. (1984). Causal reasoning in medicine: analysis of a protocol. *Cognitive Science, 8,* 363-385.

Matsuhashi, A. (1987). Revising the plan and altering the text. In A. Matsuhashi (Ed.), *Writing in real time.* Norwood, NJ: Ablex.

Miller, M. & Goldstein, I.P. (1976). *PAZATN: a linguistic approach to automatic analysis of elementary programming protocols* (AI Memo 388). Cambridge, MA: Massachusetts Institute of Technology, Artificial Intelligence Laboratory.

Miller, G.A., Galanter, E., & Pribram, K.H. (1960). *Plans and the structure of behavior.* New York: Holt.

Newell, A., & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Rips, L.J., & Marcus, S.L. (1977). Suppositions and the analysis of conditional statements. In M.A. Just, & P.A. Carpenter, (Eds.), *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum.

Sacerdoti, E.D. (1975). The Nonlinear Nature of Plans. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence.* Thilisi, Georgia, USSR. pp. 206-218.

Sacerdoti, E.D. (1977). *A structure for plans and behavior.* New York: Elsevier.

Scardamalia, M., & Bereiter, C. (1985). Written composition. In M. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed.). New York: Macmillan.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflexive processes in written composition. *Cognitive Science, 8,* 173-190.

Simon, H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology, 7,* 268-288.

Simon, H.A. (1978). Information-processing theory of problem solving. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 5. Human information processing.* Hillsdale, NJ: Erlbaum.

Sussman, G.J. (1975). *A Computational model of skill acquisition.* New York: American Elsevier.

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

Volpert, W. (1982). The model of the hierarchical-sequential organization of actions. In W. Hacker, W. Volpert, & M. von Cranach (Eds), *Cognitive and motivational aspects of actions.* New York: North Holland.

Waterman, D.A., & Newell, A. (1971). Protocol analysis as a task for artificial intelligence. *Artificial Intelligence, 2,* 285-318.

Waterman, D.A., & Newell, A. (1973). PAS-II: an interactive, task-free version of an automatic protocol analysis system. *Proceedings of the Third International Joint Conference on Artificial Intelligence* (pp. 431-445). Menlo Park, CA: Stanford Research Institute.

Winograd, T. (1983). *Language as a cognitive process: Vol. 1. Syntax.* Reading, MA: Addison-Wesley Publishing.

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# PRODUCTION OF ARGUMENTATIVE DISCOURSE: THE TEXTUAL FUNCTION OF STATEMENTS CONSIDERED IMPORTANT BY THE SPEAKER

**Pierre Coirier**

Laboratoire de Psychologie du Langage - U.A. C.N.R.S. 666
Université de Poitiers
Poitiers - France

Suppose after writing a text, the subject is asked to indicate which of the statements in the text he/she considers the most important. What are the textual characteristics of the statements chosen? What is the relationship between them and the typical constituents of the type of text in question? This is the problem that will be examined here.

The notion of importance, whether subjective or objective, has been the topic of various experimental studies on comprehension. Important textual units (sentences, propositions, ...) are generally memorized, recognized, and summarized better than unimportant ones (for a review, see Passerault, 1984). The semantic representation of a text elaborated by its reader is essentially structured around the most important units.
There are two aspects to importance. First, the objective aspect, which expresses the characteristics of the text, e.g. position in the text structure (at the macrostructural level, for example), importance based on thematization, emphasis, etc. Such characteristics are not taken into account automatically, but depend on various cognitive operations involving general or specific linguistic knowledge and acquired textual skills. Thus, we should not speak of objective importance, but of importance linked to the objective properties of the text, that is, textual importance.
Subjective importance, on the other hand, refers to factors linked to the subject's cognitive structures. These factors are independent of the text being read: knowledge of the topic being dealt with (structure of the "situation model", in Kintsch, 1986 for example), typological representations (schemas, superstructures, ...), as well as personal involvement, opinion, interest. The characteristics of the reading situation also fall into this category: perspective, objectives assigned to the task, individual strategies, ... The factors of importance are indeed highly varied.

The overall importance of a given unit can be considered to be the resultant of a complex evaluation process involving the characteristics of the text, the reader, and the reading situation, all of which interact. This evaluation process determines how central the role of the units contributing to the construction of the text's representation will be (Coirier & Passerault, 1988; Passerault & Coirier, 1987).
It would not be legitimate to view the processes involved in comprehension and those involved in production as highly symmetric, but a certain amount of complementarity must be admitted if one takes into account the communicative function of texts. The procedural aspects of importance are likely to provide an approach to this problem. In this perspective, we will assume that when speakers plan their texts, they organize their plan on the basis of the units they consider the most important to the given situation, and that they then place these units in the text in such a way that their status can be identified by the addressee. Various textual characteristics may reflect this process.
The other aspect that interests us pertains to the relationship between importance and type of text. According to Bronckart (1985, p. 8), "To each context category there corresponds a text type (or discourse type) that can be characterized by a specific structure and a specific distribution of linguistic units" (our translation). The results obtained by this author, as well as those observed by Espéret, Coirier, Coquin, and Passerault (1987) in the same perspective indeed confirm the idea that there is a text-context link. But this link is not necessarily a simple one, not just a "distributive" one. If we grant a powerful status to the notion of type of text in the sense of superstructural model, as with the narrative schema, we can expect the distribution of linguistic marks to occur essentially on the most typical constituents of the text type in question. Thus, in the case of a story for example, the characteristic marks should pertain primarily to the essential narrative constituents: complication, resolution, ... **If importance also plays a role in production, then this role ought to be linked above all to the most typical text constituents.**
This problem will be approached here in regards to argumentative discourse, which we will characterize rather globally as a type of discourse in which the speaker's goal is to lead the addressee to share his point of view on a given question. The function of such discourse (or such texts -- we will not distinguish the two categories here) grants it the following typical characteristics:

- In argumentative discourse, a stand is taken. The speaker explicitly or implicitly proposes choices. The utterances that express those choices thus have a particular status in the discourse, both at the textual level -- they are required constituents, as it were -- and at the subjective level -- they represent the choices of the speaker. Take for example the ideological standpoint: **Will the assertions that are congruent with the speaker's standpoint take on the same degree of importance as the incongruent ones?**

- Because a stand is taken, argumentative discourse is necessarily strongly characterized by the enunciative involvement of the speaker, particularly through individual or collective value judgments (see Espéret et al.'s results on this point, 1987). As for the enunciative modality of such assertions (Dispaux, 1984; Pierrault Le

Bonniec, 1974), two levels of argumentative discourse can be distinguised: the value level and the fact level (Grize, 1982). Two typical categories of utterances can thus be distinguised: those formulating obligations (moral, social, ...) or value judgments (right/wrong, happy/unhappy, ...), which generally carry the marks of the prescriptive or axiological modality, and those expressing acknowledgements or facts, or presenting their content as such, with the marks of the factual modality. **What is the relative importance of these two types of assertions?**

- The conception of natural reasoning set forth by Grize (1982) describes it as a "reason-based organization", a "schematization", by means of which the speaker attempts to make his own logico-discursive representation of the referent appear believable to his/her addressee. The most important means of doing so is the "supporting" procedure, that is, the use of an assertion as a support, as evidence, as justification of another assertion in view of making others accept that second assertion; A supports B if A makes one believe B, regardless of the nature of the relationship between A and B. The supporting relation appears as a minimal structure of argumentative discourse, just as the opening-closing pair is the minimal constituent of narrative discourse (Botvin & Sutton-Smith, 1977). **Is the distinction between the supporting element and the supported element associated with a variable degree of importance? Are some supporting modes more important than others?**

The relationship between the subjective importance of utterances and the typical characteristics mentioned above will be dealt with here in the following experimental framework. The subjects were instructed to compose an argumentative text on a given topic. When finished, they were asked to indicate which statements in their text were the most important. In the corpus thus obtained, the frequency at which the above textual characteristics occurred in the statements judged subjectively as important was compared to their frequency in the other statements.

## METHOD

The production task. The task was not a completely free production task, but rather a text "composition" task in which subjects were given a pre-established list of arguments containing 66 statements. After examining these statements, each subject was to choose 15 of the statements to form a text.
The instructions were to state one's opinion on the topic in question by (1) putting the chosen statements in the desired order, (2) linking them with appropriate connectors (a non-comprehensive list of suggested connectors was provided: for, thus, and, since, ...), and (3) adding modal adverbs if required (perhaps, certainly, always,...). Any statement could be expressed in the negative by addition of expressions such as "It is not true that", "never", ...
But -- and this is one of the important clauses in the instructions -- the statements themselves could not be modified. They could only be selected, ordered, linked together, completed, ... without changes in content. To assure adherence to these

instructions, the statements were written on cards that the subjects glued to their answer sheets. Finally, the instructions emphasized the fact that each subject was supposed to take a personal stance and then support it.
Subjects were given 45 minutes to compose their texts.

Determining subjective importance. After the composition task was completed, and without changing the produced text, each subject was asked to designate the three statements he/she considered the most important in his/her text. These statements will hereafter be called the subjectively important statements (SIS) to distinguish them from the other statements (OS).

List of arguments. A list of frequently used arguments on the topic of women, work, social roles, children, etc. was extracted from literature on the status of women. Based on this list, 66 statements were constructed so as to counterbalance the following two variables:

Ideological standpoint. A given statement was classified as egalitarian (E) if it opposed sex discrimination, and inegalitarian (I) if it justified or placed value on sex discrimination. Whenever possible, one egalitarian and one inegalitarian statement were constructed from the same initial content.
Discursive modality. The statement was first put into a prescriptive or axiological form. For the sake of simplicity, these two forms were considered as the same modality, which was called the prescriptive (P) modality. The same content was then reformulated in "factual" form (F).

The following four statement types were thus obtained:
F.I. Maternity leave is often a burden to business.
F.E. Having children is no longer an obstacle in women's careers.
P.I. Working mother, father at home: that's not really a satisfactory solution.
P.E. In order for the sexes to be equal on the job, men must be encouraged to do their part of the housework.

The P/F and E/I features thus defined were submitted to 7 judges (3 males and 4 females) for verification. To be included in the final list of 66 arguments, 5 out of the 7 judges had to agree on those features. For the E/I feature, we obtained 30 E statements, 29 I statements, and 7 unclassified statements; for P/F, there were 27 P's, 33 F's, and 6 unclassified.
However, classification of this sort is not absolute, and would be irrelevant out of context. Moreover, the additions made by the subjects when composing were likely to modify this classification. Thus, the final classification used to analyze the data was based on the case-by-case examination of the texts actually produced.

The subjects. 69 students (38 men and 31 women) who were first-year mathematics majors were tested collectively. A 4-item questionnaire was used after the experiment to classify them as egalitarian or inegalitarian in the concerned field, giving 17 E's and 21 I's for the males, and 25 E's and 6 I's for the females. (1)

Processing the data. As a general rule, the indexes calculated pertain to the corpus as a whole, i.e. all protocols taken together. However, whenever possible, individual protocol calculations were verified to make sure that they did not contradict the general tendencies observed.
The final membership of a statement in the egalitarian or inegalitarian category and in the prescriptive or factual category, of course, took the modifications made by the subjects into account.

## RESULTS

Each of the 69 subjects used an average of 17 statements. The entire corpus contained 1173 statements. Nearly all of the subjects selected 3 important statements in their text, as requested. In all, there were 204 subjectively important statements and 969 other statements.

Ideological standpoint in the statements. Are the statements that express the ideological position of the speaker judged as more important than the others? SIS statements and OS statements were compared as to their congruity/incongruity with the speaker's standpoint, where congruity was considered to exist on a given statement when the subject's standpoint corresponded to the standpoint expressed by that statement (egalitarian or inegalitarian in each case) (Table I below).

**TABLE I.** Ideological standpoint expressed in subjectively important statements and other statements: congruity with the subject's standpoint (percentages in rows).

| | Congruent choices | Incongruent choices | d (congruent/ incongruent) |
|---|---|---|---|
| Subjectively important Statements (SIS) N = 204 | 70.1 | 26.5 | + 43.6 |
| Other Statements (OS) N = 969 | 53.6 | 36.7 | + 16.9 |

The corpus was marked as a whole by the tendency of the speakers to choose statements that were congruent with their personal standpoint. All statements taken together, 56.4% of the choices were congruent. This fact in itself is trivial, although it indicates, on an overall basis, that subjects do not only use congruent statements. The resulting discourse thus left room for counter argumentation, for the presence of opposing arguments.

The salient fact here is that congruity was much more greater on SIS's (d = 43.6%) than on OS's (d = 16.9%): more than twice as much. **The discourse obtained thus allocates a privileged place to the expression of the ideological position of the speaker.**

Discursive modality. Both types of statements taken together (SIS and OS), the discourse as a whole was predominantly factual. Out of the total of 1173 statements, 47.5% were factual, 37.8% prescriptive. To decide whether this predominance is a general tendency in argumentative discourse, other corpus should be studied. (The corpus gathered by Espéret et al., 1987, was 60% to 70% factual on natural argumentation by 7- to 14-year-olds.)
When SIS and OS are considered separately (Table II below), we can see a marked difference between the two.

**TABLE II**. Discursive modality used in subjectively important statements and other statements (percentages in rows).

| | Prescriptive and axiological | Factual | d (P-F) |
|---|---|---|---|
| Subjectively important Statements (SIS) N = 204 | 54.4 | 26.5 | +27.9 |
| Other Statements (OS) N = 969 | 34.4 | 52.0 | -17.6 |

SIS's are prescriptive twice as often as they are factual (d = +27.9%). On the other hand, OS's are predominantly factual (d = -17.6%). **In argumentative discourse, value judgments have more weight than facts.**

Supporting relationships. What role do the subjectively important statements play in the structure of the argumentation? If we look at the consecutive statements in a given protocol, we can distinguish two types of pairs: those containing at least one SIS and those made up solely of OS's. In the 1173-statement corpus (all protocols), there were 1008 pairs (given the initial and final statements and the paragraph changes in each protocol). 324 of these 1008 pairs included at least one important statement, while 684 did not.

1) Were the important statements used more often for support? Supporting relationships may or may not be expressed by means of an argumentative connector. Use of a connector was assumed to be more likely in the case of supporting arguments, which led us to raise the following question: In how many of the pairs with and without SIS's were connectors used?
A wide range of connectors were taken into account: because, thus, since, if, however, even if, although, moreover, etc. (see the Bronckhart's list, 1985).
The results are given below (Table III).

**TABLE III**. Percentage of connected statements.

| | Argumentative link | | % of connected statements |
|---|---|---|---|
| | present | absent | |
| With an SIS<br>N = 324 | 221 | 103 | 68.2 |
| Without an SIS<br>N = 684 | 4265 | 258 | 62.3 |

68.2% of the pairs containing important statements turned out to be linked by a connector. This percentage was 64.3% for the other pairs. This slight difference does not allow us to attribute any specific supporting function to the important statements from a quantitative standpoint (at least not when assessing it by the use of connectors, which is not a simple endeavor!).

2) Supporting modes. Various supporting modes were distinguished here: causality (A because B), finality (A so that B), specification and restriction (A only if B), etc. (See the analysis grid proposed by Apothéloz & Miéville, 1986). The following four main categories were chosen:
Linkage relations (final or causal). Prominent use of the following connectors: thus, for, if, so, since, ...
Refuted opposition relations (or concession). Connectors: however, although, nevertheless, even if ...

Addition relations. Connectors: in addition, moreover, besides, ...
Specification relations. Connectors: in particular, accordingly, in this case, ...

Based on this analysis grid (only applied to pairs with connectors), the distribution of supporting modes on pairs containing important statements was compared to that on pairs without important statements (Table IV below).

**TABLE IV**. Supporting modes observed in the connected pairs with and without a subjectively important statement (SIS) (percentages in rows).

| | Linkage (causality, finality) | Refuted opposition (concession) | Addition | Specification |
|---|---|---|---|---|
| With an SIS N = 221 | 47.1 | 28.9 | 19.9 | 4.1 |
| Without an SIS N = 426 | 34.7 | 32.2 | 28.4 | 4.7 |

Table IV indicates above all that when the pair contained an SIS, nearly half of the relations were of the linkage type (causality, finality, ...). For the non-SIS pairs, however, linkage, refuted opposition, and addition relations were basically equivalent.

3. Were the important statements of the supported or supporting type?. In an "A since B" type relation, A is the supported element and B is the supporting element. In pairs with a clearly oriented relation (the linkage pairs), 65% of the subjectively important statements were supported and 35% were supporting.
**Important statements thus play a role in the argumentative structure: they tend to be supported rather than supporting, and they pertain predominantly to the linkage category.**

## DISCUSSION

The method used in this experiment to study the role of subjectively important statements raises some questions.
One of the problems is that an uncommon production task was used. The speakers did not have complete control over the content of their texts. They could only select the desired statements and then arrange them in the order of their choice. Did this factor have any significant effects on the characteristics of the texts? If so, what effects? The only way to answer these questions would be to apply the same analysis method to a free writing task.

A second important reservation: Did the judgments of importance made after text composition reflect the "text plan" or the text itself, once written? In other words, was this a comprehension variable in some way rather than a production variable? As in one of our other studies (Passerault & Coirier, 1988), we could have asked the subjects to make judgments about the arguments in the list before beginning to compose. But this presents another problem similar to the preceding one: What would be the impact of those judgments on the subsequent production task? What's lacking here is an objective means of measuring the importance of a given textual unit, for a given subject, in a given situation, which would not depend on the final structure of the text. How could this be done? Could we say that importance is only constituted in and by the text itself? This in fact turns out to be false if prior works on subjective importance in comprehension are taken into account; the importance of a unit depends not only on the text, but also on the reader. The relevant parameters describing the readers, as well as the related effects of extralinguistic context parameters, have yet to be precisely determined.

Now that the above reservations have been made, let us briefly summarize the results obtained here. The statements that the writer of an argumentative text considers most important have specific textual functions: being more closely linked to the expression of the speaker's stance, they serve to express value judgments. These subjectively important statements also play a role in organizing precisely that which is typical of this type of text: the support feature. They are the statements that "link" arguments, and they constitute the justified elements rather than the justifying elements.

The three functions studied -- congruity, modality, and support -- were viewed here as independent dimensions. Reality is in fact more complex. It is difficult to imagine, for example, how the discursive modality and the supporting function could be completely dissociated. In an opinion-giving discourse, to take a stance the subject generally must use prescriptive or axiological formulations. Thus, there is automatically some dependency between prescriptive or axiological forms and dominance of the congruent ideological stance. The fact that the important statements are usually congruent may be due to the fact that they are also prescriptive, and vice versa. Likewise, it is "natural" to expect facts to support values, and not the opposite. But at the same time, there is relative dependency at the subjective importance level between the supporting function and the discursive modality. More than just autonomous functions, we are dealing here with a configuration of partially interdependent dimensions, whose structural properties should be closely analyzed. This could lead us to propose a minimal model for ordering information according to the links between the various typical textual constituents: a minimal model of subjective information processing. Accordingly, we could say that the more directly linked a statement is to the function or goal of the discourse, the more important it becomes. In certain cases, it is the statement which announces the author's position that becomes this key statement, but not in all cases. The "best possible argument" to support a shared standpoint can also be considered. We are approaching both Dispaux's (1984) analysis of strategies in the dialogue of strategists, experts, ..., even the deaf! and the propositions made by Grize (1982) on the conditions of acceptation/refutation of schematizations.

It thus appears here that production of argumentative discourse is organized around those few statements which are viewed as important by the speaker. This should be related to the observations made on comprehension. In this respect, further study of protocols based on the analysis of macrostructures could perhaps confirm this organization process. In addition, the following question should be answered: are the statements that are important to the speaker the ones the addressee takes into account? Is there complementarity? In production, the importance of a statement can be marked in a variety of ways, whether implicit (to be inferred from the macrostructure, for example) or explicit (thematization, repetition, degree of expansion, etc.). It would be interesting to list the possible categories, and then, attempt to determine whether or not there is a typology of both producers and understanders that describes their relative acknowledgement of these categories of marks and levels. Few studies (except in classical rhetoric) have dealt with this aspect of "text competency": the ability of a speaker to choose, among a wide range of marks, those which will be the most effective in a given speech interchange situation. The same is true for the ability of the understander to extract the revelant cues out of the "author's way of thinking". This competency in production can be illustrated by examining a composition written collectively, and in comprehension, by analyzing journalists' commentaries on a speech.

Let us now come back to the initial problem set forth in this study, i.e. the statements that are important take on that importance in relation to the intrinsic properties of argumentative discourse: the discursive modality, the standpoint of the speaker and the supporting procedure. **The subjectively important statements are predominantly those which provide the means of translating the discourse goal -- to make others agree with one's value judgments by supporting them -- into a text.**
This is consistent with the idea that the notions of text type and typical constituents should be considered as essential to the analysis of text-context links. Such links are probably not manifested by the simple, homogeneous distribution throughout the text of surface marks according to extralinguistic parameters. Our idea is that in the text-context relationship, the goal of the discourse may be considered not just as an ordinary parameter, but rather as the central monitoring component of production. Our results, of course, do not allow us to go that far. They only point out a possible orientation for further research in this field, as previously stressed by Fayol (1986, p. 1) when he spoke of the role of a "schema activated in long-term memory corresponding to each type of text and guiding the various processing levels" (our translation).

Is the function of a text the central monitoring component? After all, this is due almost directly to the acknowledgment of the goals of all behavior, whether linguistic or non-linguistic, and to the guiding role of such goals in the elaboration of that behavior. It is natural to expect texts to be organized, above all, according to the speaker's aims in a given situation: to inform, justify, explain, entertain, etc. This seems to resemble the central role that Adam (1987) granted to discourse macroacts in his analysis of text sequence types. In the case of an argumentative text, the essential aim -- to make others share one's point of view by means of appropriate justifications -- leads to the maximization of the importance of the statements that fulfill that function, i.e. those statements which are congruent, prescriptive, and supporting.

## NOTES

1. The corpus analyzed here was prepared and gathered by Olivier Varasson as part of his Master's thesis in experimental psychology:

Varasson (O.), **Variations de la production argumentative en fonction de la cible**, Mémoire de Maîtrise, duplicated manuscript, Poitiers.

Other data pertaining to the same corpus are analyzed in:

Coirier, (P.), "Modalité prescriptive ou factuelle d'une production argumentative en fonction du destinataire", Poitiers, duplicated manuscript.

## REFERENCES


ADAM J.M. (1987).- Textualité et séquentialité: l'exemple de la description. **Langue Française, 74**, 51-72.

APOTHELOZ P., MIEVILLE D. (avec la collaboration de J.B. Grize): (1986).- Cohérence et discours argumenté, in M. Charolles (Ed.), **The resolution of discourse**, Hambourg: Buske Verlag.

BOTVIN G.J., SUTTON-SMITH B. (1977).- The development of structural complexity in children's fantasy narratives, **Developmental Psychology, 13**, 377-388.

BRONCKART J.P. (1985).- **Le fonctionnement des discours. Un modèle psychologique et une méthode d'analyse**, Neuchâtel: Delachaux et Niestlé.

COIRIER P., PASSERAULT J.M. (1988).- Interpretative aspects of text summarization: diversification as a function of task goals, **Cahiers de Psychologie Cognitive, 3**,....

DISPAUX G. (1984).- La logique et le quotidien: **une analyse dialogique des mécanismes d'argumentation**. Paris: Les Editions de minuit.

ESPERET E., COIRIER P., COQUIN D., PASSERAULT J.M. (1987).- L'implication du locuteur dans son discours: dicours argumentatifs formel et naturel, **Argumentation, 1**, 155-174.

FAYOL M. (1986).- Les typologies de textes: approche cognitive. Contribution au Colloque sur les Textes et leurs Traitements, Poitiers, 22-25 septembre.

GRIZE J.B. (1982).- **De la logique à l'argumentation**. Genève: Droz.

KINTSCH W. (1986). - Learning from text. **Cognition and instruction**, **3**, n°2, 87-102.

PASSERAULT J.M. (1984). - Niveau, importance relative et rappel des éléments d'un texte: résultats et interprétation. **L'Année Psychologique**, **84**, 251-266.

PASSERAULT J.M., COIRIER P. (1987).- Résumé d'un texte argumentatif: effet de l'accord avec les phrases et de leur importance subjective. **Psychologie Française**, **32**, 307-312.

PASSERAULT J.M., COIRIER P. (1988).- Le marquage de l'implication énonciative dans le discours argumentatif: les effets de l'interlocuteur. Poitiers (document renéoté).

PIERRAULT LE BONNIEC G. (1974).- **Le raisonnement modal: étude génétique**. Paris: Mouton.

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# TOWARDS THE GENERATION OF DISCOURSE UNITS IN CONVERSATION*

Uta M. Quasthoff

Fakultät für Linguistik und Literaturwissenschaft
Universität Bielefeld
Bielefeld, FRG

## 0. INTRODUCTION

Research in the field of text production – as opposed to comprehension – has been neglected for a long time by the relevant disciplines, cognitive psychology and Artificial Intelligence. Standard theoretical reasoning in linguistics is more affiliated to the ideas of modelling verbal production. One should remember that the term `generative grammar' stems from the attempt fo describe language structure by explicating the rules by which an idealized speaker–listener generates (and analyzes) sentences. In comparison to the approaches in psychology and AI, however, linguistic models often are not interested in the *empirical* reconstruction of the generating *process* in itself but use the concept of a production rule on a more abstract level as a descriptive means for the analysis of the structure of the linguistic product. Furthermore, the vast majority of research done on production in linguistics, psycholinguistics, psychology, and AI has not exceeded the sentence level of analysis.

Among the authors who do deal with text or discourse the overwhelming majority is only interested in the *cognitive* processes of planning, realizing and controlling discourse (see also the state–of–the–art–article by Esperet/Piolat, this volume). The few approaches which acknowledge the decisive role of situational or *interactive* factors in discourse planning in general are still lacking empirical evidence for the theoretical

* The research reported here was made possible by two succeeding grants, one given by the Volkswagen–Foundation for the research project "Kognitive und sprachliche Entwicklung am Beispiel des Erzählens in natürlichen Interaktionssituationen" (Freie Universität Berlin) and one given by the German National Science Foundation (DFG) for the project "Generation of narratives in natural and experimental settings: Interaction and language development: GENESIS" (University of Bielefeld).

assumptions (van Dijk/Kintsch 1983, ch. 8) or formulate their models on a level too abstract for empirical testing (Zammuner 1981; for an exception see Herrmann 1982). The empirical validation of AI–models (cf. Meehan 1980, Davey 1979, McDonald 1983, McKeown 1985) naturally is only given in the form of "running" computer programms in a very limited domain of discourse generation.

In the following I will present the outlines of a linguistic model of discourse production at its present state, which is formulated along the lines of the following theoretical claims:

1. empirical testability (possibility of falsification)
2. integration of the cognitive and the interactive constituents of discourse planning
3. integration of global and local planning procedures.

These (meta–)theoretical goals are transformed in the formulation of the model in the following way:

1. The level of abstraction is chosen such that empirical falsification is not ruled out by principle. In other words, a model of *discourse* production will never be open to empirical falsification since observable phenomena will always be specific to certain *types* of discourse. Consequently, the heuristics must be to first explicate the rules which govern the generation of one type of discourse unit (in our case: narratives) and then reach the higher levels of abstraction first by comparison with other types of discourse units and then by careful and empirically based generalization.

2. The separation of cognitive and interactive aspects in the study of verbal behavior is seen as a consequence of disciplinary boundaries and traditions instead of as being induced by the object of description (cf. Quasthoff 1983). Two conceptual decisions in the model relate to this basic assumption: a) the object of description is delimited such that only *discourse units* in conversation (cf. Wald 1978) are dealt with. These are structural chunks in the flow of conversation, marked by initial and closing boundaries, which establish a special variety of the turn–taking–mechanism and are thus clearly distinguished from the surrounding turn–by–turn–talk. It is clear that the modelling of a conversationally constituted discourse structure in particular asks for the integration of interactive aspects into the reconstruction of cognitive processes. b) a central descriptive category is the concept 'plan' (cf. Miller/Galanter/ Pribram 1960; for a recent survey see Esperet/Piolat, this volume). In other words, the goal–oriented advanced *cognitive* organisation of verbal and other interactive behavior is based on the rules and mechanisms of *interaction* and the embedding in interactive context as much as *interaction* is based on *cognitive* planning and processing by the participants.

3. Although the level of descriptive explicitness differs with respect to global and local aspects of discourse planning, this is a matter of descriptive economy rather than a conceptual deficiency. The structural description of discourse units in terms of the 'relational' and the 'informational structure' integrate the most global schematic characteri-

stics of stories in conversation and the propositional, lexical, syntactic and phonetic aspects of single utterances in one hierarchical structure. This entire structure is built into a representation of the global planning procedure including all the "pragmatic" aspects.

In the following I will first give an outline of the descriptive format, which will have to be very brief and sketchy and then add some empirical evidence on the basis of narrative production data by children.

## 1. PLANNING DISCOURSE UNITS IN INTERACTION

Let me first give you a schematic representation of the basic ideas of the descriptive approach which serves purely illustrative purposes (see fig. 1 below).

### 1.1. The situational boundedness of the generation process

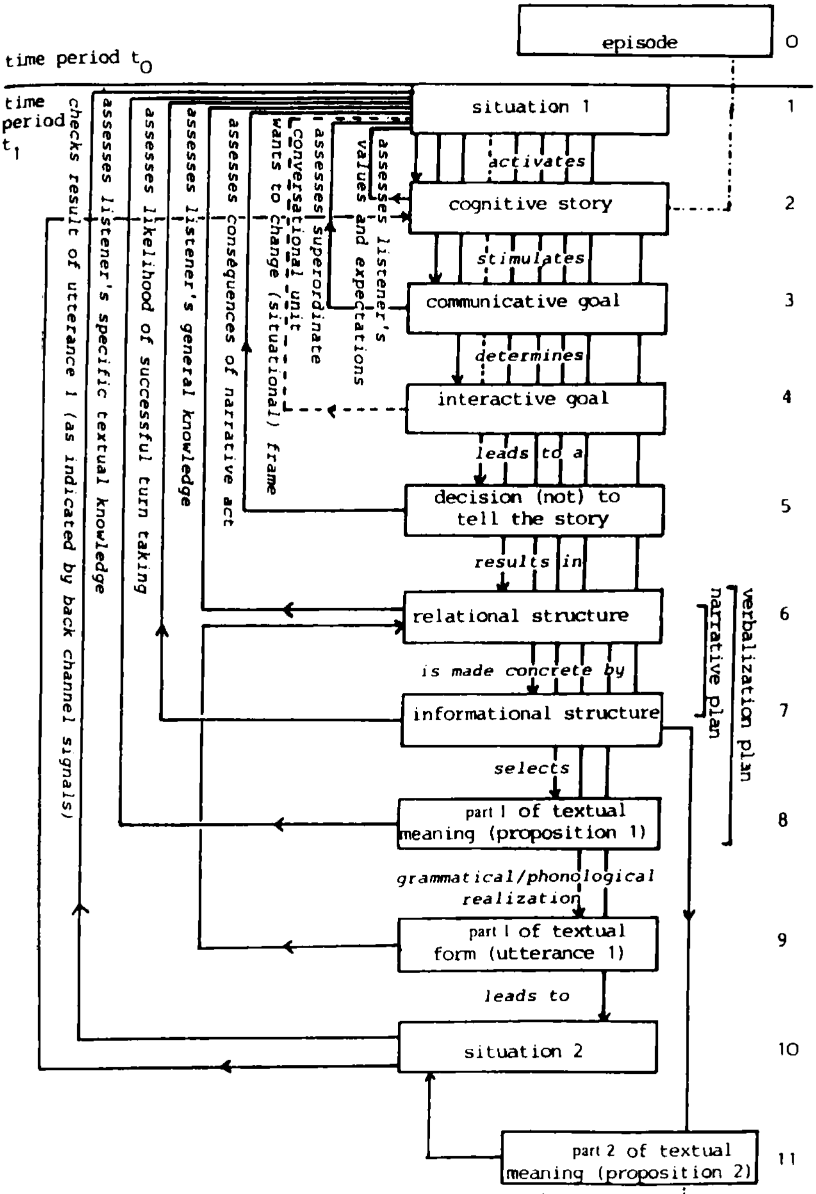
The most important factor of the production model and the most conspicuous aspect of the diagramm is the basic role of the "situation", which constitutes all the steps of the planning procedure. The term "situation" covers all the traditionally called "pragmatic" aspects of language use (social context, spatio–temporal surroundings, personal and social relations between speaker and listener, their mutual knowledge, and the like). Some situational aspects are specified as names of the feedback loops connecting "situation" with the other steps of the production schema. Each and every aspect of narrative production (or, for that matter,of discourse production in general) depends on, is influenced by and influences many pragmatic factors.

By "cognitive story" we wish to refer to whatever mental representation of the real–world episode is retrieved (or reconstructed) from memory at the time of the narration. This reflects the distinction between what the narrator recalls about what happened and what really happened; it also reflects the distinction between what the narrator recalls and what he/she actually tells.

As far as the distinction between "communicative" and "interactive goal" is concerned, we are well aware of the fact that this terminology is a bit awkward; for lack of anything better, we will use these terms in the following sense: Communicative functions rely on the *content* of a narrative, whereas interactive functions rely on its linguistic *form*, i.e. the type of discourse pattern (for example, "report" vs "replaying narrative", cf. Quasthoff 1979, 1980, 1986).

Obviously, some functions of conversational narratives – like argumentation or self–aggrandizement – depend primarily on *what* is told in the particular story. These are the functions that are subsumed under the heading of "communicative functions". On the other hand, the very act of "replaying" conversational narration (Goffman 1974) (no matter what the subject is) may serve the function of creating an atmosphere of intimacy and thus help the narrator to express and promote his view of the interactional relation,

**Fig. 1:** Production schema for conversational narratives: (– – –), variable processes; (———), necessary processes; (– · – · –), processes not included in narrative production.



namely "We're engaged in an informal context with a relatively close personal relationship" (Quasthoff 1979). Therefore, we call this latter function of a conversational narrative and the corresponding intention of the narrator "interactive".

The intended function(s), the cognitive story, and their appropriateness must be constantly checked against one another and the situation; this finally results in a decision to tell a particular story (or to refrain from doing so) and to do this by use of a particular discourse pattern. Once these decisions have been made, information about the episode that is accessible in the memory store is retrieved (and missing links are reconstructed). This recall is guided by narrative schemata that are probably culture-specific. The process of recalling finally results in an informational structure which is embedded in a relational net, the details of which cannot be given here.

This complex semantic structure then is serialized following the constraints of the pattern structure: i.e., one information unit at a time is selected, verbalized, and finally uttered. Of course, even the planning of the linguistic form of the utterance(s) is dependent on all the preceding steps in the production schema. With the realization of the first utterance and listener reactions, a new situation is created so that the planning circle starts all over again.

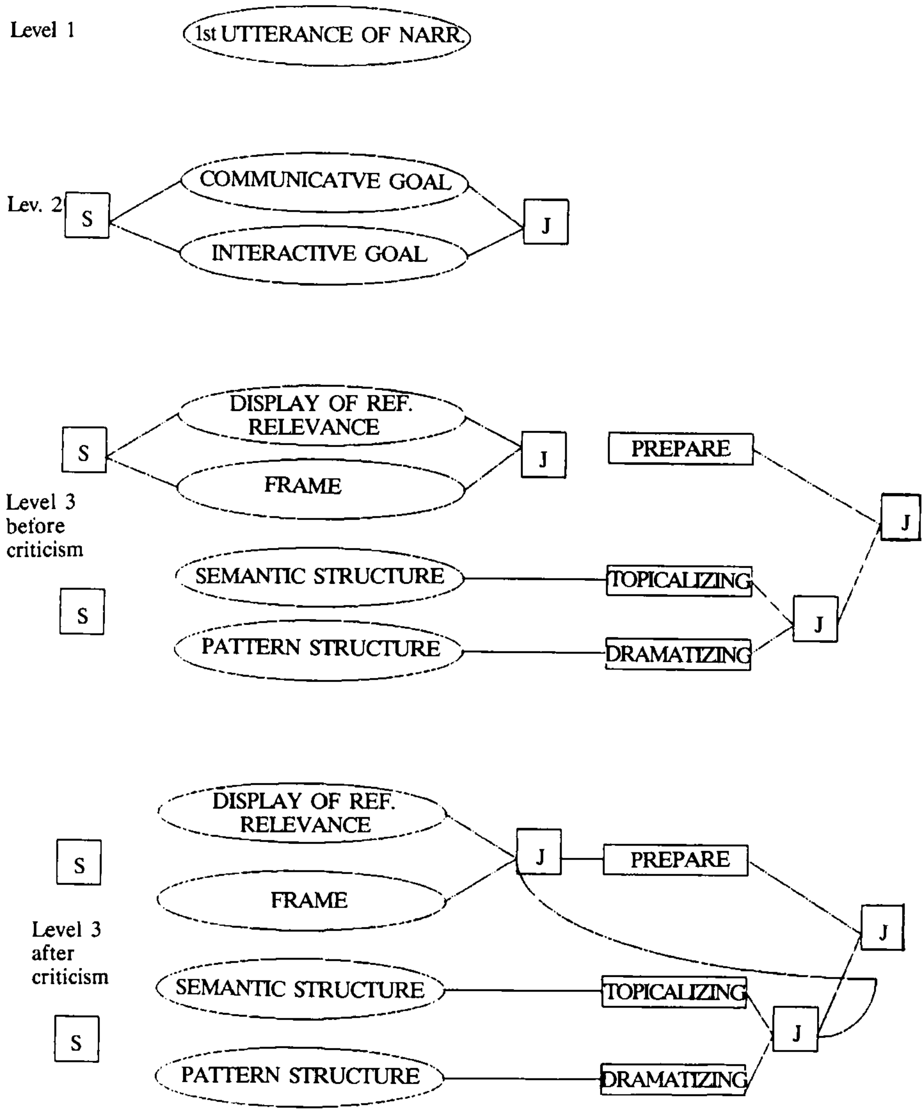
This illustration of the basic ideas of our descriptive approach leaves many important aspects unexplicated and does not yet provide a consequent representation of the planning *process* as a procedure of sequentially organized steps. In order to achieve a higher degree of formal explicitness in this structural description of a process I have made use of a format developed in AI to represent plans of behavior (Sacerdoti 1977).

### 1.2. Towards a formal representation

For space reasons I can make only a few remarks as to the general principles of Sacerdoti's plan representation NOAH and the specific application to our descriptive problem:

- NOAH operates hierarchically in that it first explicates the most abstract notion of a plan by giving the ultimate goal and then, at successive steps, specifying action–goal relations with an ever increasing degree of detailedness.
- The different levels shown in the diagramm (fig. 2) represent the different states of the problem solver. In comparison to the illustrative representation of fig. 1, due to lack of space only some of the higher levels of abstraction are represented here.
- The nodes labeled S and J respectively indicate a split in the plan and a join where the separate paths come together. Obligatory linear ordering is indicated by horizontal sequence in the graphs.
- NOAH uses so called 'critics' which evaluate the plan and deal with interactions among subplans. The "criticism to resolve conflicts" (see level 3, fig. 2), e.g., handles constraints in the ordering of operations with respect to the split paths of the plan.

**Fig. 2: Representation of narrative generation according to NOAH (Sacerdoti 1977)**
**(Square boxes denote operators, boxes with rounded ends denote goals)**



The specific formulation of our narrative plan in analogy to NOAH relies partly on empirical findings as to the conversational embedding of the initiation and conduction of narrative discourse units, which will be presented in the next section. These findings and their formal representation shown in fig. 2 above differentiate the unsatisfying crudeness in the explication of the situational boundedness as was suggested in fig. 1.

## 2. EMPIRICAL EVIDENCE

### 2.1. The data

The empirical data used to differentiate or modify the theoretical conceptions presented above are production data of the following kind: We worked with approxiamately 20 children in each of the following age groups: 5, 7, 10, and 14. Each age group was divided into two groups of 10 children. The first group related the incident in a conversational setting which the subjects did not perceive as being part of the experiment, whereas the second group recounted the same incident following the explicit request of the experimenter. I will refer to the first type of story telling situation as the "informal situation" and to the second type as the "formal situation".

The 10 children of each age group and situational type were recorded in institutions such as day-care centers or schools. The children were provided with a cover story that explained out presence in the institution and the reason for the technical equipment as well as the recording that we made in the "official" part of the experiment. During the preparations for the experiments the children witnessed a staged incident (the same for all age groups and situational types) that they perceived as being a natural event. The "plot" of the incident is as follows: a male student trips over the cord of a tape recorder causing it to fall off the table onto the floor. He does not want to tell his supervisor because he is afraid of being shouted at again. His co-worker, a female student, argues that they have to tell the supervisor or they will jeopardize the experiment. She finally decides to tell him and thus tattles from the perspective of the male student. The supervisor, in fact, does get very angry at the student, but, in the course of the argument it turns out that not much harm has been done anyway since the tape recorder was a broken one that had been brought by mistake. While the supervisor goes to look for a new tape recorder, the male student accuses the female student of unneccessarily tattling.

In the informal situation each of the 10 children then was asked to take part in an experimental interview (the "official" part of our experiment). While the child was waiting for the experimenter to show up, he/she found him/herself waiting in a room with a person who was well known to the child, usually one of the care-takers at the day-care-center, and who was supposedly in charge of guarding the cassette recorder. If the child did not volunteer to tell about what had just happened, the listener, who of course had not witnessed the incident, used subtle elicitation devices.

With the exception of the staged incident, the same procedure was followed on the two following days with the same children, using, of course, different kinds of "official" interviews for each of the three days. This meant that each child told about the same real world incident on three subsequent days to three different listeners who (from the childs perspective) knew nothing about the incident. As far as the chiild was concerned these activities did not belong to the experimental tasks, which only began when the official experimenter entered the room (after the child had told the incident).

In the formal situation there is no "waiting room phase". Instead the experimenter immediately asks the child to tell about something that he or she has experienced and explicitly asks about "what just happened in the other room". Please note that in this, as in the other condition, the child thinks that the listener of the story has not been informed about the details of the incident.

## 2.2. *The conversational embedding of narrative discourse units*

The data of our informal situation have a special quality in comparison to most data collected in the laboratory or to entirely spontaneous data: The contextual conditions of story−telling in conversation are especially transparent because normally the child does not volonteer to tell the story spontaneously nor does the listener give explicit promptings. Instead, the adult listener governs the conversation with the child such that a recounting of the event becomes expectable. In other words, the data give us en excellent chance to study what we calll the "interactive achievement of reportability", i.a., the sequential implicatures which call for a narrative discourse unit in the specific replaying discourse pattern (cf. Hausendorf/ Quasthoff 1989).

In accordance with relevant principles of Conversation Analysis we claim that there are certain global JOBS which have to be done in order to organize conversational interaction, and that the participants work jointly − using different pragmatic DEVICES − to get the jobs done. Utterances and other interactive − non−verbal or para−linguistic − FORMS have to be analyzed in terms of the device−bound functions which they perform in getting the interactive work done. Jobs are defined as tasks for the interactive dyad whereas devices and forms are narrator− or listener−specific.

At least for the data of our informal situation the following formulation of JOBS, DEVICES, and FORMS seems to cover all the variants of actual conversational embeddings/ preparations of a particular story (see tables 1 to 3 below. For a more detailed recent description see Hausendorf/ Quasthoff 1989).

Table 1: **JOBS**

| | |
|---|---|
| 1. | Display of referential/ formal relevance |
| 2. | Topicalizing |
| 3a | Elaborating |
| 3b | Dramatizing |
| 4. | Closing |
| 5. | Transition |

Table 2: **LISTENER DEVICES**

| | | | |
|---|---|---|---|
| 1. | – Local, temporal, personal or objectoriented reference to setting of event<br>– expression of need for information / lack of knowledge | | |
| 2. | – Establishing of "eventability" and "reportability" | | |
| 3a | – content–oriented questions (local, temporal, personal or objectoriented)<br>– expression of need for more information / lack of knowledge<br>– explicit request to tell the story | | |
| 3b | – evaluative appreciation of 2 | | |
| | – orientation towards | visual and acoustic experiencing<br>emotional involvement<br>unusualness | with resp.<br>to event |
| 4. | – evaluation of unusualness of event<br>– evaluation of particular actions / persons | | |
| 5. | Coda | | |

Table 3: LINGUISTIC FORMS

1. What were you doing *just now in the other room*?

2. The cassette recorder fell down

3a Who broke the recorder? Couldn't they use the r. anymore?
Tell me what happened!

3b Oh really?
Did you see it happen?

4. That's really something because these rècorders are terribly expensive

5. Do you think it's O.K. that B told K ("tattled")?

The complete version of the narrative plan according to NOAH integrates these conceptualizations of the interactive conditions of story–telling and the schematic aspects of story structure, which I will deal with in the next section.

2.3. Narrative production as evidence for the relational structure

The explication of the schematic structure of conversational narratives used in our model is strictly confined to typical or invariant *content* elements of such stories. Formal characteristics are dealt with in the so called pattern structure. The semantic structure is represented in a kind of hierarchically organized relational net, which is also integrated in the narrative plan of the NOAH – version.

This structural description of content elements provides the possibility of representing not only meaning elements of particular texts but also structural elements of the real world episode itself – defined as meaning elements uttered by at least one child in our sample.

This kind of structural description of the incident served in our project as a *tertium comparationis* for the semantic analyses of our data. A version of the relational structure of the incident was developed before the systematically comparable production data described above were available. When we looked at these data, using the structural description of the entire episode as an analytic tool, we found the description inadequate to the data in one respect. It turned out that practically all the children who mentioned the incident at all referred to (at least) the initiating event and the final result. In other words, if the initiating event was mentioned spontaneously, there was a high probability that the final result was also mentioned, whereas the mentioning of other content elements could not be expected with the same degree of predictability.

On the basis of this finding we reformulated the structural description moving the initiating event and the final result to a higher hierarchical level of the description.

The interesting aspect about this revision is not so much the empirical finding which is in principle an accordance with findings one might expect on the basis of, e.g., story grammar approaches (cf. among many others Stein/Glenn 1979). The interesting aspect is of a methodological nature: It concerns the way in which empirical evidence (or falsification) for schematic structural descriptions can be provided by production instead of by comprehension data.

This brief outline of an attempt to describe the interactively founded planning of discourse units in conversation in an empirically testable way had to be extremely sketchy. Hopefully it was explicit enough to at least point to what is left to be done!

## 3. REFERENCES

Davey, Anthony. 1978. Discourse production: a computer model of some aspects of a speaker. Edinburgh (Edinburgh University Press).

van Dijk, Teun A./ Kintsch, Walter. 1983. Strategies of Discourse Comprehension. New York etc. (Academic Press).

Esperet, E./ Piolat, A.. 1986. "Production: planification et controle." In this volume.

Goffman, Erving. 1974. Frame Analysis. An Essay on the Organization of Experience. New York etc. (Harper & Row).

Herrmann, Theo. 1982. Sprechen und Situation. Eine psychologische Konzeption zur situationsspezifischen Sprachproduktion. Berlin etc. (Springer).

McDonald, D. "Natural Language Generation as a Computational Problem. An Introduction." In: Brady, D./ Berwick, R. (eds): Computational Models of Discourse. Cambridge, Mass. (MIT Press).

McKeown, Kathleen R.. 1985. Text Generation. Using discourse strategies and focus constraints to generate natural language text. Cambridge etc. (Cambridge University Press).

Meehan, James Richard. 1980. The Metanovel. Writing Stories by Computer. New York, London (Garland Publishing).

Miller, George A./ Galanter, Eugene/ Pribram, Karl H.. 1960. Plans and the Structure of Behavior. New York (Holt, Rinehart & Winston).

Quasthoff, Uta M.. 1979. "Eine interaktive Funktion von Erzählungen." In: Soeffner, Hans-Georg (ed): Interpretative Verfahren in den Sozial- und Textwissenschaften. Stuttgart (Metzler): 104-126.

Quasthoff, Uta M.. 1980. Erzählen in Gesprächen. Linguistische Untersuchungen zu Strukturen und Funktionen am Beispiel einer Kommunikationsform des Alltagss. Tübingen (Narr).

Quasthoff, Uta M.. 1983. "Text Connection as Mental Plan and Interactional Structure." In: Ehlich, K./ van Riemsdijk, H. (eds): Connectedness in Sentence, Discourse and Text. Proceedings of the Tilburg Conference held on 25. and 26. Jan. 1982. (Tilburg Studies in Language and Literature 4). Tilburg University: 237-257.

Quasthoff, Uta M.. 1986. "Narrative Discourse Patterns as a Matter of Perspective." In: Pragmatics and Linguistics. Festschrift für Jacob L. Mey, ed. by Sonne, H. and Johansen, J.D.. Odense (Odense University Press): 151-162.

Hausendorf, H./ Quasthoff, Uta M.. 1989. "Ein Modell zur Beschreibung von Erzählerwerb bei Kindern." In: Ehlich, Konrad/ Wagner, Klaus (eds): Erzähl-Erwerb. Bern etc. (Peter Lang): 89-112.

Sacerdoti, E.D.. 1977. A Structure for Plans and Behavior. Amsterdam (Elsevier North-Holland Inc.).

Stein, Nancy L./ Glenn, Christine G.. 1979. "An analysis of story comprehension." In: Freedle, R.O. (ed): New Directions in Discourse Processing (vol.2). Norwood, N.J. (Ablex).

Wald, Benji. 1978. "Zur Einheitlichkeit und Einleitung von Diskurseinheiten." In: Quasthoff, U.M. (ed): Sprachstruktur - Sozialstruktur. Zur linguistischen Theorienbildung. Kronberg/Ts. (Scriptor).

Zammuner, Vanda L.. 1981. Speech Production: Strategies in Discourse Planning: A theoretical and empirical enquiry. Hamburg (Buske Verlag).

TEXT AND TEXT PROCESSING
G. Denhière and J.-P. Rossi (Editors)



# STRATEGIES IN DISCOURSE PRODUCTION: COMPUTATIONAL MODELS

Vanda Lucia ZAMMUNER*

University of Padova
Dipartimento di Psicologia dello Sviluppo e Socializzazione
Via B.Pellegrino, 26, 35137 Padova, Italy

On the assumption that computational models can be crucial tests of a theory, this paper reviews a few recent Artificial Intelligence language generation models and programs that address issues relevant for a theory of discourse production and that are or might be psychologically plausible, discussing the extent to which they converge to support psychological concepts relevant for discourse production.

## 1. HUMAN ACTIONS AND PROBLEM SOLVING

I am not a very good cook, I am usually busy, and I have a small house. Yet, now and then, I concoct dinner parties for about a dozen friends, and they come back the next time around. How do I manage? I use strategies, of course (but I shall not tell you my secret!).

The idea that, given a problem, a set of limitations, and a desired final state, we can form or use strategies to find a solution - presumably as old as mankind - has been intensively exploited in the last three decades within the human information processing approach (see Estes, 1978; Anderson, 1980; Loftus and Schooler, 1985). The concept is that, given a task environment, the information processing system - the problem solver - conceptualizes the problem in terms of a problem space that can be searched actively for a solution, using heuristics such as means-ends analysis. The basic psychological characteristics of the human information processing system - its adaptability, its capacity to modify its own behaviour over time by

---

* This work was partially done while the author was working at the University of Leiden, Faculty of Social Sciences, Unit of Experimental Psychology, with a NATO-CNR fellowship.

learning, its flexibility, as well as its resource limitations (e.g. processing, knowledge, memory) - define broad boundaries on its behaviour. The latter is also constrained by the structure of the problem space, e.g. in terms of which moves (strategies) are legal, and it depends on the adequacy of the problem solver's representation of critical features of the task environment (Simon, 1979, 1978).

A crucial question is the following: are the mechanisms that govern problem solving for simple, well-structured problems (e.g. the Tower of Hanoi puzzle) applicable to less-structured and semantically rich domains? Among the most demanding, ill-structured intellectual problems are those called of "composition", such as composing a painting, a piece of music, or a novel. To these we can add the 'normal' text/discourse - a 'problem' that is the focus of this paper. Here, much effort is involved in defining and developing the problem itself, not only in solving it (Greeno, 1978). As in simpler problems, production must be guided by an understanding of the requirements of a good solution to enable the problem solver to produce possible solutions and components, and there has to be a continuous evaluation of the current attempts at solution. However, the relevant criteria are complex and involve very abstract concepts. According to Simon (1978), between well- and ill-structured problems there is a continuum. However, for the latter it is more complex and difficult to decide whether the goal has been achieved, to define which information is needed to solve the problem and where it can be found, and there is no simple "legal move generator" to establish alternative possibilities at each step in the solving process. Indeed, we know that even in the solution of simple and well-structured problems different alternative strategies can be equally effective. In the attempt to explain how ill-structured problems are solved the information processing approach has been recently extended to several, broader task domains (e.g. Brodie, Mylopoulos & Schmidt, 1983). A most interesting development within Cognitive Science deals with the complex task of language generation.

As Esperet and Piolat (1986) point out, current psychological models of discourse production - including my own hypotheses on the role of strategies in discourse planning (Zammuner, 1981) - leave open many questions, one of the most foundamental being their validity. Empirical research is an obvious and necessary step to verify a theory (e.g., Zammuner, 1990, 1989, 1987, 1982). Another useful strategy, at this point, is to consider models of language use that claim some degree of psychological plausibility and are (or might be) computationally effective. This approach shares the perspective (e.g. Bara, 1984; Castelfranchi, 1984; Job and Tabossi, 1984) that computational models can be crucial tests of our theories because of their high level of formalisation and explicitness, and the very fact that, when they are implemented as computer programs, we can verify their output adequacy. However, given the state of the art in both fields and the lack, still, of a serious attempt at a transfer and integration of knowledge within and across these domains (and given the serious space limitations), the matching (and the validation) of models can be at present only partial and indirect.

## 2. COMPUTATIONAL MODELS OF LANGUAGE GENERATION

### 2.1 Recent trends in Artificial Intelligence planning models

Artificial Intelligence (AI) research on problem-solving and planning has been very large. While the beginning work mainly dealt with simple tasks performed by a single agent that executed simple actions, about a decade ago programs that dealt with more complex problems started to be implemented. Among the first applications of planning models to natural language processing (NLP) there was work on dialogue about plan-based activities (e.g., Grosz, 1977), on story understanding and production (e.g., Wilensky, 1978; Schank & Abelson, 1977), and on the planning necessary to produce single speech acts (e.g., Cohen & Perrault, 1979. For a review of work on NLP see Carbonell & Hayes, 1983; Bara and Guida, 1984a; Danlos, 1985). Most of this work was concerned with the representation of real-world knowledge and its application to the *understanding* of natural language input. As Bara and Guida (1984b) suggest, a possible taxonomy of NLP models reflects the three main aspects involved in the study of language: 1. cognitive models, basically concerned with the conscious and unconscious mechanisms underlying the use of language, that treat it only as a medium for expressing mental states; 2. pragmatic models, concerned with language as a medium for communication, and thus with the interpersonal reasons for its use; and 3. linguistic models that analyse the structure of language, trying to capture its regularities and justify the relationship between structure and meaning. In practice, the distinction is not clear cut and models often overlap at least to some extent. Also, most A.I. models are not concerned with the extent to which their procedures resemble human mechanisms (see also Winograd, 1983). As we would hope, in very recent times researchers have started to model language *generation* (LG), explicitly trying to consider and integrate cognitive, pragmatic and linguistic components, with the purpose of modelling and accounting for human competence and performance. Needless to say, these attempts are still very partial, i.e. restricted to narrow task domains in relation to which they model a small set of goals, plans and strategies, with reference to very limited sets of knowledge representations. However, each of the models which I shall briefly present addresses some of the issues that are relevant for a theory of *discourse* production and, taken altogether, they suggest a less partial account of this complex task.

## 2.2 Conversation as a Planned Activity

An early attempt at modelling conversation is by Hobbs and Evans (1980). They posit a *planning mechanism* that allows for plans to be developed and *modified* during the execution of the conversation itself. The planning process is a procedure that relies on *"causal axioms"* to derive a plan with a *sequence of actions* that will enable achievement of a *goal*. Causal axioms, a subset of "beliefs" about the world and about the mechanism itself, include *conversational strategies*. It is these that provide a link between goals and actions. In a conversation domain, the planning begins with *high-level conversational goals* (the focus is on image and coherence goals, such as to project a favorable image or hold the floor for continuation of own turn) and relies on its causal axioms, including the conversational strategies, to generate a plan whose actions are utterances, gestures and other typical conversational moves (e.g. gaze direction). Because conversational plans are executed in a number of steps and may meet with unforeseen problems, the mechanism needs to work in association with a monitor and a debugger. The *monitor* tries to relate inputs from other participants in a conversation to the conversational plan, in order *to alter* the plan or *check* how effective it is in achieving its goal(s). If the monitor learns *new information* that contradicts its expectations, the *debugger* singles out which causal axioms happen not to be true, accounts for this by carrying out a deeper *search of the knowledge base* to find factors not previously considered, and passes the information to the planning mechanism so that it may generate a *repair plan*. It is these two components that enable *flexible* plans.

The definition of the planning mechanism - in terms of goals, actions, causal axioms, unspecified means for choosing among the options presented by the causal axioms, and the planning mechanism itself - reflects related research findings. However, to my knowledge, this model has not been implemented and the rather subjective micro-analysis of a conversational fragment - reported by the authors to show how the mechanism works - in the main fails to make explicit the relation between the posited planning components and the actual conversation, introducing at the same time actions, goals, intentions and strategies that are not defined explicitly anywhere. Nonetheless, the model is interesting in that it underlines the relationship between goals, plans and strategies in the conversational domain, it stresses the concept of ongoing monitoring in a conversation, and takes into account the possibility of conscious moves and the availability of different means (i.e., both verbal and non-verbal) to obtain conversational goals.

2.3 Conversational Coherence

A more elaborated theory of conversation, written as an ATN grammar and designed for future computer implementation, is proposed by Reichman (1981). Her model tries to explain the semantic/logical coherence relations that utterances have to one another and that define well-formed discourse. These relations are explained in terms of *"context spaces"*, i.e., abstract structures characterized by slots that are *functionally* defined, capturing both *implicit* and *explicit* discourse information. A discourse is assumed to be the result of *conversational moves* - 12 such moves are discussed - that serve a speaker's communicative *goals* (e.g. "support", "interrupt", "challenge"). The focus in on utterances whose role is to begin a new communicative act (or speaker's goal) that results in a topic shift, and not on utterances that simply constitute an embellishment or continuation of a current discourse unit. The goal of a speaker in his/her conversational move reflects the functional relation of the current utterance to preceding utterances. When utterances fulfill a single goal - equated here with a single conversational move - they lie in a single context space. It is in relation to this unit that a speaker chooses the next conversational move. Just as a sentence conveys more information than that expressed by each of the specific words, a context space carries more information than that associated with the specific set of utterances that form it. Interesting features of the context space structure include an *internal representation* of the set of utterances that lie in the context space; a marker reflecting the *influential status* of these utterances at any given point in the discourse; a pointer to the *preceding discourse* context spaces in relation to which the present context space was developed; a specification of what *type of relation* was involved. Furthermore, there is a *goal slot* for each context space that specifies the speaker's goal in developing it - e.g. the value of this slot might be "support" or "challenge". A given goal may be performed in different ways and it refers to some preceding part of a conversation. Therefore, all context spaces also have a *"contextual-functional" slot*, itself a structured entity composed of (a) annotations of the specific method by which the corresponding goal has been achieved (e.g. "modus-ponens", "flat denial", "emotive-flat-rejection"); (b) information about the previous discourse part in connection to which this relation holds. (Other interesting aspects of the context space structure cannot be discussed here). The grammar models generation of discourse involving two or more participants; however, the updating of the relevant discourse context, on the basis of the preceding speaker's conversational moves, is not separate for each speaker.

I can only mention that Reichman's theory both differs from, and integrates, previous notions and models in linguistics, psychology, ethnomethodology, sociolinguistics, and A.I. (including some of the work quoted here). This model explains, among other things, surface linguistic phenomena such as the use of *referring expressions* in spontaneous discourse - by showing, for instance, that pronominal references are limited to elements currently in "high focus" - and the use of specific *connectives* - such as *so*, *by the way* and *because* - that *signal* that a context space boundary has been reached and specify the type of shift about to take place. Furthermore, each conversational move is formalized in terms of an associated set of *preconditions* that specify its "appropriateness" at any given point in time, and a set of *effects* that enables participants to set up expectations and available options for subsequent discourse development. Although the conversational excerpts used for discussing the model and grammar are, basically, argumentative, Reichman claims that the abstract context space deep-structure can account for all discourse forms, on the ground that different discourse genres have different thematic developments and not different structures.

As my brief presentation hopefully shows, this model is extremely interesting and it captures many complex discourse phenomena. Among its theoretical limits, are the following. Participants in a conversation presumably construct discourse models that integrate several discourse units at any given time (usually with much abstraction); especially for complex contents, the formed representation may at times be conflicting, biased, partially wrong, or it may involve open-ended relations, etc. Thus, to update the conversation model only in terms of the unit under discussion and the preceding unit currently most influential - i.e. the relevant context space-pair - is not sufficient; the updating mechanism should operate separately for each participant. The model assumes that speakers are very consequential in their choices - e.g. once a speaker concedes an argument as invalid, he or she cannot re-use it in challenge of an opponent claim. It is fairly obvious, instead, that people often argue in a rather different fashion, because they are stubborn, forget what has already been said, re-negotiate previously reached agreements, etc. For Reichman, there is usually only one function achieved by a given, coherent set of utterances (and thus by a conversational move and associated goal). This characterization is too rigid and does not take into account that a move (or strategy) that performs a function at one level might be associated with a different function at the same or at a different level, and be subservient to different goals. The definition itself of which are the basic goals and conversational moves in a conversation would require much discussion. Another important point is that the updating of a discourse model cannot be done by a mechanism that does not possess knowledge of a speaker's higher goals, such as overall purpose for (a part of) a discourse. For instance, the knowledge should include the possibility of creating new plans and strategies if the overall goal does not seem to be near to being achieved. Finally, while an approach in terms of a strict hierarchical procedural organization indeed facilitates the formulation of a coherent and integrated theory of discourse, it does not seem psychologically plausible.

### 2.4 Strategies in the Generation of Answers

The issues of how a system determines what information has to be included in a text and how that information has to be organized to achieve its communicative goals are both complex and central for a theory of discourse production. McKeown's model, implemented as a computer program (1985, 1983), deals with these problems in a restricted domain, i.e. *answers to questions* about database structure. As for Reichman, the assumption is that people have predefined, *stereotypical* ideas about how they can achieve a goal as well as how they can integrate these means in order to form an acceptable text. The model is also similar to Reichman's in that the representation of the *current focus of attention* constrains what information needs to be considered, i.e. is relevant, in deciding what to say next, constraining also which choices can be made when the chosen discourse strategy allows for alternatives. The author correctly stresses that a generation system needs to formulate reasons for selecting between different options pertaining to what to say, when to say it, and how to say it. Previous work, as we know, mainly dealt only with this third class of options. In this model, the discourse content and structure are determined by a processing stage that is called *"strategic"*, while a *tactical"* component uses a grammar and dictionary to realize in English the message produced by the strategic component. The strategic component determines the answer to an input question as follows. Given a knowledge base, 1. it determines which information is relevant and sets it off as the *"relevant knowledge pool"*. 2. It retrieves a set of possible *schemata* associated with a specific question-type: this is the step at which the association of *discourse strategies* with *discourse goal* is achieved. For instance, if the goal is "to define" an object - i.e. the question implies a "request for definitions" - the possible schemata are "identification" and "constituency"; if the goal is "to describe", the schemata are "attributive" and "constituency"; and, if it is "to compare", the schema is "compare and contrast". These schemata are different *"rhetorical predicate" patterns* that reflect those preconceived ideas people have about how to obtain a given goal. The predicates (e.g. identification, analogy, particular-illustration) are the basic units of the discourse strategies used to answer a question. They characterize the predicating acts a speaker may use and delineate the structural relation between propositions in a text - her analysis of these predicates borrows from previous work.

A given schema guides the generation process by acting as a *text plan*. Selection of a specific schema is dictated by the information available to answer the question. Thus, semantic information *interacts* with information about discourse structure to determine the text structure. It should be noted that *schema recursion* is possible and is achieved by allowing each predicate in a schema to expand into a single proposition (e.g. a sentence) or a schema (e.g. a text sequence). So, for instance, when using an identification schema, the identification predicate can expand into an identification schema, where the speaker provides a definition of a previously mentioned obiect *A* in relation to the question object *B* and then, using an analogy schema, points out the similarities and differences between *A* and *B*. 3. The choice of a strategy then *guides* the selection of propositions from the relevant knowledge pool. 4. The content thus determined is passed to the tactical component that uses a functional grammar to transform the message into English. At this level, *focus of information* is used to make appropriate *surface choices* such as referring expressions, active vs. passive constructions, and sentential connectives (e.g. the connective *for example* is chosen in association with the rhetorical predicate "particular-illustration").

The strenght of this model is that it incorporates explicitly in the planning mechanism knowledge about discourse structure; it deals with several strategies associated with different discourse goals; and it uses a large number of "rhetorical predicates". Although the model assumes that decisions made in the lowest level of the generation process can influence decisions about what to say, its strategic and tactical components cannot interact. Other weak points are the following. The system cannot evaluate its own production nor make inferences on the basis of its knowledge; it does not have a model of the hearer; it is restricted to a very narrow domain of generation, and, of course, it is not really creative (for other critiques to this model see Danlos, 1985). The last works I shall discuss here have tried to address some of these crucial issues.

## 2.5 Pragmatically Motivated Generation

In a program developed by Hovy and Schank (1984) we find due consideration of pragmatic elements - not taken into account by previous models - that have a central role in communication. Namely, knowledge about social and psychological aspects that are relevant in discourse production, such as the relative *social statuses* of speaker and hearer, knowledge of *different types* of hearer and of their interests and 'simpathies', and knowledge of *communicative goals.* To produce texts that are adequate for different hearers the generation of language is governed by a set of *strategies.*

The starting assumption of this model is that when a speaker produces two different versions of an episode or story, the differences are not due to different underlying *representations* of the story, but to *different goals.* These reflect pragmatic considerations and influnce the text generation at four *decision levels*: story and sentence content, sentence structure, and word choice. Pragmatic criteria, however, are not sufficient, and *syntactic* considerations also have to be taken into account. The production of texts in which pragmatic and syntactic constraints are both taken into account is achieved by a complex mechanism that consists of four parts, each having a different *functional* role in the generation process, i.e. a speaker's world knowledge and representation of the story, knowledge of the hearer, a set of strategies, and generation routines. The interplay is between a speaker's goals and his/her knowledge of his/her hearer. The program implementing the model is simplified in that it assumes the hearer to have the same general world knowledge as the speaker and it models only positive cooperative conversations - thus the only goals (and associated strategies) modelled at present are to create and maintain the hearer's sympathy and interest. Hearers - specifically, an IRA terrorist, the terrorist's wife, a British soldier, and an American person - are defined in terms of the world knowledge and concept definitions that exist in the generator. The final output is the result of the semantic representation, plus speaker's goals and knowledge of the specific hearer to be addressed.

The strategies used to obtain the speaker's goals *select and order the concepts* to be said as well as various *sentence parts*, control the *amount* of generated detail/information and specify which *words* to use. For instance, when the goal is *to create interest*, one strategy selects concepts and uses words that are linked with strong emotions rather than being neutral for the hearer; if the hearer already knows about a certain fact, the strategy dictates a simple reference to it; otherwise, the fact is fully described; in a sentence, the element unknown to the hearer is put in focus position; etc. To achieve the goal of *creating sympathy*, likely strategies include talking about topics the hearer is sympathetic with, and placing in focussed position the aspect of the topic the hearer is sympathetic with. The strategies, on the basis of pragmatic knowledge, guide the generation process at the four hierarchical levels mentioned (topic choice, etc.) using expansion rules based on this hierarchy.

A few examples. When speaking to the IRA terrorist, a suitable topic is the part of the story dealing with the shooting of a British soldier; when talking to the wife, the first chosen topic would be related to the wounding of a girl and her later discharge from hospital. At the next level of sentence content, the focus for the first example - as sentence subject - would be on the soldier; for the second example, on the girl; and so forth (a program that generates similar kinds of text but which focusses on linguistic issues is proposed by Dunlos, 1985).

Some of the issues not addressed in this model have been taken up by Hovy who has implemented a program called PAULINE (Planning And Uttering Language In Natural Environments). The main concern is still with the pragmatic decision criteria that inform discourse in the real world, the points of the generation process at which they become relevant, and the strategies that are necessary to produce texts in accordance with them. A first improvement is that the planning and the production components interact *only* at critical decision points (e.g. topic choice, etc., mentioned for the previous model), in such a way that the planner does not need to contain explicit syntactic knowledge, and the generator does not have explicit goal-related information (see Hovy, 1985). This feature models the fact that, in normal discourse, not all decisions are planned before an utterance starts to be produced. Textual plans (the only one so far implemented is *to convince* a hearer) are formalized as *sets of injunctions*, similar to McKeown's schemata, that are used to guide decisions and instructions about *possible topics*. Associated with these plans are *"suggestions"* on how to choose a topic, such as "consider the topic if the hearer and hearer agree over its connotation", or "minimize difference between the speaker's desire/interpretation and reality, if it is small". These, in fact, are rules that express *a set of conditions* and an associated *action* to be performed if the conditions are met. Thus a plan is formalized as a production system.

In determining *what* to talk about, the system follows an *"affect rule"* - a rule based on a reformulation of Grice's cooperative principle as a "principle of payment". This rule states that, when the goal is "to convince" the hearer that some concept has an affective value (good or bad) and there is a suitable topic of conversation, *"enhancers"* and *"mitigators"* can be used to manipulate the text to achieve the desired effect (e.g. for a good effect, the speaker should talk about good topics using enhancers, and bad topic with mitigators). We are not surprised at finding that sentences of the form "not only X but Y" are enhancers, together with such connectives as *also* and phrases such as *what's more*; *however*, *merely*, and *accidentally* are mitigators. Constraints are associated on how enhancers and mitigators can be used in relation to sentence content. In general, a speaker has several *techniques* available to emphasize or evade topics she likes or dislikes, such as being selective about sentence topics, selecting verbs with appropriate predicate forms, or selecting appropriate stress (see Hovy, 1986a).

This approach is further elaborated (Hovy, 1986b) by enlarging the type of *pragmatic concerns* the system can deal with, and by positing *intermediate-level goals* that have a *"rhetorical"* nature, such as "simplicity", "formality" and "detail" (each of which may be characterized by one of two values, e.g., formal or informal). These goals achieve their results by the *characteristic effects* they cause in the text *content and form*. Associated with each rhetorical goal and its values are *strategies* that enable the system to achieve its rhetorical goal, itself activated on the basis of the topic, affects and pragmatic values given to the generator for its conversation. Thus PAULINE produces *stylistically* appropriate texts under various pragmatic circumstances from a *single representation*. In this model, in sum, rhetorical planning and generation are processes that interact at crucial decision points. In other words, the program integrates a top-down and a bottom-up approach, in that partially realized syntactic options represent available choices that are selected on the basis of rhetorical criteria, leading in turn to further planning of the text.

The specific conversation topic modelled in different pragmatic situations is the outcome of a hypothetical election between Carter and Kennedy as presidential candidates. For instance, PAULINE may be a Carter supporter that is speaking formally at a debate where the audience is presumed to support Kennedy, or it may simply be informing a neutral acquaintance of the election outcome in an informal fashion. The generator sets its rhetorical goals (e.g. level of formality and detail) using rules again conceptualized as production systems, i.e. as a set of conditions and associated action.

Although not devoid of problems, this last computational model, presented here all too briefly, is the one that seems to capture best much of the already existing knowledge on discourse 'products' and processes. Indeed, it shows that, in order to account for the various decisions that a speaker needs to make in order to produce texts that are both pragmatically appropriate and linguistically acceptable, it is necessary to posit a rather complex architecture whose different levels may interact and influence each other.

## 3. A FEW CONCLUDING COMMENTS AND A SUGGESTION

Each of these models to some extent purports to be a theory of discourse production but they were mainly developed only with reference to other computational models. Therefore, the theoretical validation they can provide for non-computational models, as I said earlier, can only be indirect and partial. Not discounting these limits, I think that to some extent they converge to support a number of general concepts to be found also in non-computational models (e.g. Zammuner, 1981). In producing a discourse, we have communicative goals and in order to achieve them we plan strategically our actions. In a given context of production the relevant goals may be hierarchically ordered and be of different nature (i.e. rhetorical, social, etc.). Plans and strategies are formed and/or activated according to a representation of the requirements of the production context, and on the basis of knowledge (preconceived ideas, stereotyped schemata, etc.) of how goals can be achieved in a given context.

Plans and strategies have to be monitored during implementation, may have different hierarchical status, and may be subject to change if the context's evaluation so requires. During text production, each participant in a conversation forms a discourse model, i.e. a representation of the universe of discourse so far created, that is biased by his/her own discourse/context goals and tries to consider other participants' (inferred) goals. This model is continuously updated. The scope of implemented strategies may be quite varied: from the ordering of information and choice of text topics to choice of words, sentence focus and intonational stress. Thus, a given utterance may result from the activation of more than one strategy and often it reflects, at the linguistic level, the strategy chosen. In turn, language itself on the one hand constrains which local strategies are legal, on the other hand it offers a wide range of means to execute a strategic choice. In sum, we can say that strategies are quite essential in linking goals to (discourse) actions, but, the less well-structured is the problem at hand, the greater the complexity of strategic choice, the less defined the range of necessary and/or available strategies, and the greater their possibility of failure. Thus, the range of available strategies, their interaction with the problem structure, the circumstances for their use, the influence of previous experience or of motivational factors, the possibility of strategy transfer across domains or tasks, and so forth, are still key open issues for ill-structured tasks.

My review of recent computational models of language generation that explicitly deal with the notion of strategies (and plans and goals) was not meant to be exaustive, nor a discussion of the 'technical' aspects of these models - which would be quite necessary in order to assess their relative merits. My purpose was simply to provide a beginning attempt at crossing boundaries from the point of view of a psychologist interested in discourse processing. The crossing-over of disciplinary boundaries is difficult and sometimes even dangerous, because of major differences in the adopted theoretical and practical perspective, interests, technical language and level of formalisation. Nonetheless it is necessary. Recently, in reading the literature on discourse processing, I have had the impression that much of the current work overlaps to a greater or minor extent. This overlap is not itself negative, but should be acknowledged and defined precisely in order to lead to productive future research. The blame can be divided equally among psychologists, linguists, sociolinguists, A.I. scholars, and so forth, as well as among those working within a given field. My suggestion is that a good strategy would consist in a serious attempt at integrating different research approaches and their results. I am aware that this particular problem is not easy to solve, and that the constraints and resource limitations are many. I suggest it as a global strategy. Each of us may work out intermediate level goals, and associated strategic plans.

## REFERENCES


Anderson, J.R. (1980). *Cognitive Psychology and its Implications*. San Francisco, W.H. Freeman & CO.

Bara, B.G. (1984). Artificiale e no. *Storia e Critica della Psicologia*, 5: 43-56.

Bara, B.G. & Guida, G. (Eds.) (1984a). *Computational Models of Natural Language Processing*. Amsterdam, Elsevier Science Publishers B.V. (North-Holland).

Bara, B.G. (1984b). Competence and performance in the design of natural language systems. In B.G. Bara & G. Guida (Eds.), *Computational Models of Natural Language Processing*, pp. 1-7. Amsterdam, Elsevier Science Publishers B.V. (North-Holland).

Brodie, M.L., Mylopoulos, J. & Schmidt, J.V. (Eds.). *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*. New York, Springer-Verlag.

Carbonell, J.G. & Hayes, P.J. (1983). Tutorial on Artificial Intelligence: Natural Language Processing. *Proc. of the 8th Int. Joint Conf. on A.I.*, , pp.1-39, Karlsruhe.

Castelfranchi, C. (1984). Psicologia e metodo simulativo. *Storia e Critica della Psicologia*, *5*: 57-68.

Cohen, P.R. & Perrault, C.R. (1979). Elements of a plan-based theory of speech acts. *Cognitive science*, *3*: 177-212.

van Dijk, T.A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York, Academic Press.

Danlos, L. (1985). *Génération automatique de textes en langues naturelles*. Paris, Masson.

Esperet, E. & Piolat, A. (1986). Production: planning and control. Paper presented at the Workshop on Texts and Text Processing, Poitiers (F), 22-25 September

Estes, W.K. (Ed.) (1978). *Handbook of Learning and Cognitive Processes, Vol.5*. Human Information Processing. Hillsdale, Lawrence Erlbaum Associates.

Greeno, J.G. (1978). Nature of problem-solving abilities. In Estes, W.K. (Ed.), *Handbook of Learning and Cognitive Processes, Vol.5*. Human Information Processing, pp 239-270. Hillsdale, Lawrence Erlbaum Associates.

Grosz, B. (1977). The representation and use of focus in a system for understanding dialogue. *Proc. of the 5th Int. Joint Conf. on A.I.*, M.I.T., pp.67-76.

Hobbs, J.R. & Evans, D.A. (1980). Conversation as planned behavior. *Cognitive Science*, *4*: 349-377.

Hovy, E.H. (1985). Integrating text planning and production in generation. *Proc. of the 9th Int. Joint Conf. on A.I.*

Hovy, E.H. (1986a). Putting affect into text. *Research Report*, Yale University, Dept. of Computer Science.

Hovy, E.H. (1986b). Some pragmatic decision criteria in generation. *Research Report*. Yale University, Dept. of Computer Science.

Hovy, E.H. and Schank, R.C. (1984). Language generation by computer. In Bara & Guida (Eds.), *Computational Models of Natural Language Processing*. Amsterdam, Elsevier Science Publishers B.V. (North-Holland)..

Job, R. & Tabossi, P. (1984). Umana o artificiale, purchè sia intelligenza. *Storia e Critica della Psicologia*, 5: 69-89.

Loftus, E.F. & Schooler, J.W. (1985). Information processing conceptualizations of human cognition: Past, present and future. In Ruben, B.D. (Ed.), *Information and Behavior, vol.1*, pp. 225-250. New Brunswick, Transaction Books.

McKeown, K.R. (1985). Discourse Strategies for generating natural-language text. *Artificial Intelligence, 27*: 1-41.

McKeown, K.R. (1983). Focus constraints on language generation. *Proc. of the 8th Int.Joint Conf. on AI.*, pp.582-587, Karlsruhe.

Reichman, R. (1981). *Plain Speaking: A Theory and Grammar of Spontaneous Discourse.* Technical Report no. 4681, Bolt, Beranek & Newman, Inc.

Schank, R. & Abelson, R. *Scripts, plans, goals, and understanding.* Hillsdale, Laurence Erlbaum Associates.

Simon, H. A. (1978). Information-processing theory of human problem solving. In Estes, W. K. (Ed.), (1978). *Handbook of Learning and Cognitive Processes, Vol.5. Human Information Processing*, pp. 271-295. Hillsdale, Lawrence Erlbaum Associates.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology, 30*: 363-396.

Wilensky, R. (1978). Understanding goal-based stories. *Report 140.* Yale University, Dept. of Computer Science Research.

Winograd, T. (1983). *Language as a Cognitive Process.* Reading, Addison-Wesley Publishing Co.

Zammuner, V.L. (1990). Strategies in the expression of attitudes toward abortion. In M. Charolles, S.Fisher & J. Jayez (Eds.), *Représentations discoursives.* Nancy, Presses Universitaires de Nancy (in press).

Zammuner, V.L. (1989). Strategies speakers use to achieve coherence. In M.Charolles (Ed.), *The Resolution of discourse*, pp. 88-113. Hamburg, Buske Verlag.

Zammuner, V.L. (1987). Pragmatic factors and strategies in discourse production. In A.Bertuccelli Papi & J.Verschueren (Eds), *The Pragmatic Perspective. Selected Papers from the 1985 International Pragmatics Conference*, pp. 255-276. Amsterdam, J.Benjamins.

Zammuner, V.L. (1982). Semantic discontinuities as text production strategies. In A.Flammer & W.Kintsch (Eds.), *Discourse Processing*, pp. 137-150. Amsterdam, North-Holland.

Zammuner, V.L. (1981). *Speech Production. Strategies in Discourse Planning.* Hamburg, Buske Verlag.

This Page Intentionally Left Blank

# AUTHOR INDEX

## SUBJECT INDEX

This Page Intentionally Left Blank