

10 DISCOURSE PLANNING: EMPIRICAL RESEARCH AND COMPUTER MODELS

Jerry Andriessen, Koenraad de Smedt and Michael Zock

Chapter prepared for:

A. Dijkstra & K. de Smedt (Eds.), COMPUTATIONAL PSYCHOLINGUISTICS: AI AND
CONNECTIONIST MODELS OF HUMAN LANGUAGE PROCESSING (pp. 247-278).

London: Taylor & Francis, 1996.

© 1996 Taylor & Francis

Nonfinal prepublication copy. Do not quote from this version.

10	DISCOURSE PLANNING: EMPIRICAL RESEARCH AND COMPUTER MODELS	1
10.1	Introduction.....	3
10.2	Some problems and phenomena in discourse	3
10.2.1	Contextualization.....	3
10.2.2	Tailoring the message to the audience	4
10.2.3	Discourse types.....	4
10.2.4	Thematic progression and linearization	4
10.2.5	Reference and cohesion.....	5
10.2.6	Coherence.....	6
10.3	Research on discourse planning	6
10.3.1	Writing research.....	7
10.3.2	Macroplanning and microplanning.....	9
10.3.3	Pragmatic factors in discourse planning	10
10.3.4	Semantic macroplanning	10
10.3.5	Linearization.....	11
10.3.6	Topic and focus.....	13
10.3.7	Rhetorical Structure Theory	14
10.3.8	Conclusion.....	16
10.4	Computational models of discourse planning.....	17
10.4.1	Schemata for discourse planning: McKeown's TEXT	17
10.4.2	Rhetorical relations: Hovy's Structurer	20
10.4.3	Moore & Paris: planning of explanations	24
10.5	Evaluation and conclusion	27
10.6	References.....	28

10.1 Introduction

Discourse, be it written or spoken, consists generally of more than a single sentence. Production of a multisentential discourse requires *planning*, or a series of choices that guide subsequent verbal production. Discourse planning involves the creation and elaboration of communicative goals, and the application of strategies for the selection and organization of content, taking into account the situation and the available linguistic resources. This chapter will focus on aspects of planning that pertain to producing coherent discourse, and on computational models to perform this planning process.

In the introduction we will discuss some of the problems faced by a producer of an extended piece of discourse. The list is not exhaustive, but it captures the most important issues in discourse planning and shows how they are interrelated. In Section 10.3 we will review experimental evidence pertaining to discourse planning. An important part of this discussion will address written rather than spoken discourse. It will be shown that our understanding of the whole process is far from complete. We believe that computational models may help to discover what pieces of the puzzle are lacking and how the different pieces may fit together. Next, in Section 10.4, we will present some of the computational models that have been developed for discourse planning. It should be noted that none of these systems has been developed as a psychological model, hence none of them should be evaluated strictly on that basis. However, it appears that computational work is progressing towards the point where implementation and testing of psychological models of discourse planning will become feasible.

10.2 Some problems and phenomena in discourse

10.2.1 Contextualization

The purpose of discourse production is to perform an act of communication, to realize an intention by linguistic means (Austin, 1962; Searle, 1979). A first problem that a discourse producer has to solve is how to develop intentions into a set of goals realizable in the current context. Bronckart (1985) calls this process *contextualization*. Driven by the intentions of the discourse producer (e.g. *I want to write a letter that leads to an invitation for a job interview*), the contextualization process causes reflection on goals to take into account situational constraints. Many courses on writing (e.g. Flower, 1981) include heuristics for working out a *discourse plan* before actual writing starts. These heuristics include tactics for brainstorming, and making inventories and schematic outlines. Contextualization leads to the creation of an orientation that guides the activities in discourse production. Beginning writers often lack such an orientation. They start a paper without knowing what to include, what to omit, how to provide adequate background information, and how to put into perspective

what is most important. Hence they often provide unnecessary details while leaving out crucial general information (Barnard, Andriessen, Bläcker, & Erkens, 1989).

10.2.2 Tailoring the message to the audience

Early writing is based on experience with oral conversation. Partners in a dialogue provide feedback which indicates their information needs. In contrast, writing is a monologic activity in which a writer must anticipate the interpretation process of a specific audience. The characteristics of such an audience guide the discourse planning strategies needed. Texts written for a lay audience require more general information and explanation of the domain than texts written for experts. The latter may expect more of an in-depth analysis and require less general information. Young writers tend to overlook the demands of a specific audience (e.g. Roussey & Gombert, 1992).

10.2.3 Discourse types

Among the different *types* of discourse, the following four have received most attention: *description*, *narration*, *argumentation* and *exposition*. Each of these discourse types is generally associated with a pragmatic goal and a canonical structure. For example, Toulmin's (1958) model of the structure of *argumentations* describes how an opening statement of opinion needs mentioning of certain supportive evidence, which then crystallizes to a conclusive statement. Such a global structure fosters (but does not prescribe) the organization of the discourse. Different structures are often required for different purposes. A persuasive discourse should minimally include a point of view and some supportive evidence. Further options include counterarguments, which then have to be refuted, as in (1a). If the same information is to be presented in a neutral rather than persuasive manner, arguments for and against a point of view could be presented alongside (1b).

- (1) a. People should go abroad on holidays. Being abroad is important to broaden your mind. Going abroad is costly, but there are usually affordable offers.
- b. Going abroad on holidays has advantages and disadvantages. On the one hand, being abroad broadens your mind. On the other hand, it is costly, even though there are usually affordable offers.

10.2.4 Thematic progression and linearization

Decisions on what to say necessarily involve decisions on *linearization* of information. When a discourse is to convey a complex image or thought, it must be broken down into an ordered set of separate utterances for communicative purposes. The writer must present the content elements in a suitable order and add linguistic cues that enable the reader to re-create the initial whole. For example, a *definition* may be linearized in the following way: One starts by mentioning the general class of an object, then lists some subtypes, and finally describes the functions of the object and its components.

The linear structure of texts shows a *thematic progression* in which the different *themes* or *topics* should be linked without abrupt shifts. Compare the difference between sentences (2b) and (2c) as continuations for (2a) in the following example by Brown & Yule (1985):

- (2) a. The Prime Minister stepped off the plane.
b. Journalists immediately surrounded her.
c. She was immediately surrounded by journalists.

Brown & Yule claim that there is a preference for (2c) as the continuation sentence, rather than (2b). Their explanation is that readers prefer to maintain the same topic. The choice for (2b) would entail a shift of topic.

Moreover, ordering may not only affect ease of processing but also interpretation, as illustrated in example (3) from Levelt (1981).

- (3) a. She married and became pregnant.
b. She became pregnant and married.

10.2.5 Reference and cohesion

A given object may be referred to in different ways, depending on the set of alternatives from which it must be distinguished. If a speaker wants to refer to a big black ball in a situation where the alternative object is a big white ball, the referring expression may be *the black one* or *the black ball* (Levelt, 1989). In other situations the same object may be referred to as *the big one*, *the ball*, or simply the pronoun *it*.

A sentence within a discourse can generally be understood by its links to other sentences, as shown in the following example, taken from Halliday and Hasan (1976, p.14). The meanings of *he* and *so* in (4b) can only be captured by reference to their *antecedents*, which are in (4a).

- (4) a. Did the gardener water my hydrangeas?
b. He said so.

Anaphora, such as *he*, *so*, and *it*, are *cohesive* devices. Halliday and Hasan define *cohesion* as a semantic relationship between two textual elements in which one is interpreted by the other. Clear cohesive ties are essential for the interpretation of discourse. Especially young writers appear to have problems with proper reference (Bartlett, 1984), as evident from sentences like (5).

- (5) John got into an argument with Charlie. Then he hit him and knocked him down.

Another type of cohesion is established by using connectors like *and*, *but*, *then*. Such cue words relate what is about to be said to what has been said before. Furthermore,

they instruct the reader how to link the different pieces of information (temporally, causally, etc.). Still another way of establishing cohesion is to repeat words or semantically related items, as in (6).

10.2.6 Coherence

Example (6) shows a piece of discourse that is cohesive but not *coherent*:

- (6) My daughter works in a library in Amsterdam. Amsterdam has a museum of modern art. Collectors of modern art are often yuppies. Yuppies don't like punks. The punk phenomenon originated in Great Britain in the seventies.

To understand a piece of discourse, the reader or listener must construct a coherent mental representation of that discourse. This requires not only solving problems of reference, as sketched above, but also finding a general frame of interpretation. This frame guides inferences that link different parts of the discourse, based on knowledge of the world. Such a frame is absent in (6). In contrast, example (7), taken from Roberts and Kreutz (1993), is hardly cohesive but still coherent.

- (7) The storm took the vacationers by surprise. The clothes took hours to dry.

The writer relies on the reader's interpretation that the vacationers' clothes got wet from the rain during the storm. The reader is assumed to know that storms usually involve rain which causes clothes to get wet. Furthermore, *storms* and *clothes* occur at the beginning of the sentence, so that inferences related to these concepts can readily be made. Writers and readers are usually both cooperative in handling such inferences.

10.3 Research on discourse planning

Many theories of language production assume some rough distinction between 'preverbal' planning activities and the 'verbal' production of sentences. Preverbal activities include contextualization of the communicative goal, selection and organization of the message. According to the language user framework proposed in Chapter 1, the Conceptualizer component produces preverbal messages. These serve as input to the Formulator or realization component, which prepares the syntactic frame and the word material of the sentences under construction (see Chapters 11, 12 and 13). Only recently, in the eighties, conceptualizing has become an important subject of psycholinguistic research.

In this section, we will take a closer look at some important theoretical aspects of discourse planning. We will start by sketching a picture of current research on written discourse production by novice and expert writers. Next, psycholinguistic theories of discourse planning will be discussed. Since the subject represents a fairly recent branch of psycholinguistics, no complete model or theory can as yet be provided. We will therefore focus on some empirical phenomena that seem to be central to the domain.

Finally, we will present linguistic approaches to discourse structure. Compared to psycholinguistics, linguistics in a sense approaches the problem from the opposite side, by attempting to analyze discourse in terms of its structure rather than in terms of cognitive operations performed over time.

10.3.1 Writing research

Embedded in a tradition of *problem-solving* research (Newell & Simon, 1972; Ericsson & Simon, 1984), many models of writing are based on the analysis of verbal protocols. These protocols are recorded on assignments during which subjects must carry out a particular task while they are simultaneously *thinking aloud* and explain what they are doing. On the basis of the analysis of the resulting text, notes and thinking-aloud protocols, researchers have constructed models of writing as a problem solving activity (Flower & Hayes, 1980; Cooper & Matsushashi, 1983; de Beaugrande, 1984; Bereiter & Scardamalia, 1987). In various ways, these models comprise the problems, processes, and strategies that are supposed to capture the essence of writing. Thus, writing research provides general descriptions of the processes involved in written-discourse production. A problem in reviewing writing research is that different authors often divide the process and its units in different ways.

The most frequently cited model of writing is that by Flower and Hayes (1981). Though the model is not procedural in nature, it can be used as a framework that describes at a high level the activities going on during composition. According to Flower and Hayes, writing involves three interacting processes: *planning*, *translating*, and *reviewing*. Here we will only discuss their notion of planning. Planning involves the retrieval of knowledge from memory and its organization according to the goals of the writer. The planning process is constrained by the writer's knowledge as well as by the writing context. Its main output is a *text plan*, i.e., something like an outline. The resulting plan does not need to correspond to the final surface form of the text, because it may be vague, quite incomplete, and diverse, yet it is often precise enough to guide the discourse producer in the complex task of writing (Flower & Hayes, 1984). *Expert planning* can be distinguished from *novice planning* by four features (Hayes & Flower, 1986):

1. During planning, expert writers include an initial task representation and a body of goals that guide and constrain their efforts to write.
2. This body of goals can be represented as a hierarchical structure, including top-level goals, plans and subgoals.
3. The network of goals is a dynamic structure: it is built and developed and sometimes radically restructured at the top levels while the writer composes and responds to new ideas or to the text. Modifying writing goals may be essential for good writing.
4. Experts tend to develop far more elaborated networks with more connections and integration among goals than novices.

Other authors characterize beginning and expert writing in terms of two qualitatively different production modes, respectively called *knowledge-telling* and *knowledge-*

transforming (Bereiter & Scardamalia, 1983, 1987). *Knowledge-telling* involves text generation through primarily linear processes. The writing of a prototypical knowledge-teller is based on an initial task representation, which signals a relevant discourse type, which triggers a highly canonical schema (narrative, persuasive, etc.). The task representation also provides topic associations that act as probes to retrieve content from memory. Because discourse is generated as a direct consequence of this retrieval process, the coherence of the produced texts is supposedly directly related to the organization of topical information in memory (Scardamalia & Paris, 1985; McCutchen & Perfetti, 1982). What is crucially lacking in knowledge telling is purposeful reflection on the content and form of the discourse. A knowledge teller engages in sentence-to-sentence operations, primarily guided by local topic associations (*what to say next?*). Instruction in awareness of discourse functions does not change this behaviour (Scardamalia & Paris, 1985).

The second production mode, *knowledge-transforming*, can be characterized by an inclusion in the writing process of reflective operations that transform intentional, structural, and gist representations. These operations correspond to the restructuring in adult planning as characterized by Flower and Hayes (see item 3 above). Analysis of thinking-aloud protocols has shown that mature writers plan by globally working through a writing task at an abstract level before working through it at a more concrete level. During the text production process, problems are tackled both at the level of content (*what do I mean?*) and at the level of form (*how do I say it?*). Reflection on both levels during composition leads to the transformation of content and form, giving rise to new thoughts.

A psycholinguistic theory of discourse planning must therefore account for the fact that experts of discourse production perform problem solving in at least two domains: at the ideational level (content determination and organization of ideas) and the rhetorical level (determination of linguistic forms according to communicative goals). The two domains mutually interact, whereby ideas give rise to linguistic planning, while the resulting linguistic forms provoke further reflection on ideas. Expert writing is not a one-shot process, but involves reflection and revision on all relevant aspects of the assignment. Therefore, planning structures seem to be required which can be adapted and modified on the fly. Planning processes appear to take place at all levels of production, from the construction of pragmatic plans to the preparation of articulatory sequences. To what extent the various levels are autonomous or interdependent remains an open question (Fayol, 1991). Finally, there is a fundamental difference between expert and novice writing with respect to the nature of planning. Planning by beginners is opportunistic and driven by local constraints, while expert planning is strategic: the writer's goals determine the generation and organization of content.

The general distinction in terms of writing activities of beginning and expert writers is useful. As we shall see in Section 10.4, some computational models embody characteristics of beginners' writing, including essentially linear processes guided by a highly canonical schema. Some aspects of expert writing, including explicit modeling of intentions and hierarchical planning, have also been subjects of computational

modelling. However, what is meant by dynamic restructuring has yet to be specified. The nature of the processes involved, and the information sources that serve as their inputs and outputs, are as yet unclear. It is generally acknowledged that the way the task is initially represented is a crucial factor in determining subsequent activities, but it is unknown what this initial task representation exactly contains. Furthermore, it is even unclear what a discourse model should look like (see Chapter 9 for a more detailed discussion of this issue). Proposals from various sources incorporate the individual's goals, socio-cultural conventions, abstract (hierarchical, propositional) representations of content, or a *hearer model* or *user model* expressing the speaker/writer's ideas about the hearer/reader (e.g. van Dijk & Kintsch, 1983; Bereiter & Scardamalia, 1987; Levelt, 1989; Hermann & Grabowski, 1994). To avoid an extensive discussion concerning the nature of the discourse model, we will focus on the *processes* of planning and come back to the structural issue only later in Section 10.4, when we discuss computational modelling.

10.3.2 Macroplanning and microplanning

Psycholinguistic notions of discourse planning are generally based on the spoken mode. Moreover, the data often involve dialogues, where the specification of consecutive discourse actions is highly dependent on the direct interaction with the hearer. In monologues, such as in a lecture or a news story, planning may be a more conscious activity, the resulting discourse plan more elaborate, and its execution better controlled (Van Dijk & Kintsch, 1983). In the analysis of dialogues, the focus has been more on individual utterances than on the discourse structure as a whole. Thus, less consideration has been given to higher-level speaker goals which underlie multiple, purposefully interrelated utterances (Redeker, 1992; Paris, 1991). Nevertheless, many insights from psycholinguistic approaches to discourse are clearly relevant to text writing as well.

From the psycholinguistic perspective, Levelt (1989) distinguishes between *macroplanning* and *microplanning*. *Macroplanning* is a hierarchically structured activity which involves the elaboration of some global communicative goal into a series of subgoals, and the retrieval of relevant information instrumental for realizing each of these subgoals. Van Dijk and Kintsch (1983, p. 266) distinguish between pragmatic goals (e.g. *I want you to take my advice*) and their semantic specification (e.g. *I don't want you to go to Nigeria*). *Microplanning* assigns the right propositional shape to the information, as well as the perspective (topic, focus) from which the speaker views the situation and by which the speaker guides the addressee's focus of attention. The output of microplanning has been described by Van Dijk and Kintsch (1983) as a *micro speech act*, whose definite selection depends on local pragmatic coherence constraints and features of the actual local context.

In the following paragraphs, we will focus essentially on macroplanning processes. We will discuss the planning processes from several viewpoints. It should be noted that no serial order is implied in the execution of these processes, since the flow of planning activities in discourse generation involves repeated, recursive and maybe even

simultaneous execution of several processes (see, e.g., Goldman-Eisler, 1968; Butterworth, 1980; Matsuhashi, 1987).

10.3.3 Pragmatic factors in discourse planning

Discourse production is guided by a number of pragmatic factors, on which speakers and listeners implicitly agree during communication. Cooperation is a basic ingredient for the establishment of coherence, i.e., the recognition of the fact that different pieces of a discourse are somehow related. A three-year old child describing a picture scene seems to move from detail to detail in a more or less random way. As she inspects the picture for details to be announced, salient features that catch her attention are reported immediately. Associations that come to her mind may sometimes lead to sidetracking, distracting the flow of speech by details of her personal experience. Such discourse can only be understood by a very cooperative listener who knows the person. Deutsch and Pechmann (1982) examined the way speakers select information for making reference to objects. Speakers describing arrays with many objects do not simply move from one object to the other, pointing out all the details. Rather, they mention only a few objects, and the speaker relies on the addressee's cooperation to pose further questions if the referent cannot fully be identified. Especially younger children tend to exploit such cooperativeness (Levelt, 1989). Speakers try to establish the mutual belief that the object reference is understood well enough for the current purposes (Clark & Wilkes-Gibbs, 1986). The source of coherence is therefore not the discourse itself, but has to be found in the interaction between speaker and addressee.

Important cooperative principles for interaction have been formulated by Grice (1975) as *maxims*, e.g. *be polite*, *be concise*, and *be as clear as possible*. Another important pragmatic factor in the establishment of coherence is *presupposition* (Seuren, 1985). Presupposition can be defined as the logical assumptions underlying utterances. Thus after hearing *Martians appeared again last night*, the hearer may assume that (according to the speaker) Martians had already appeared before.

10.3.4 Semantic macroplanning

After our discussion of pragmatic goals, we will now turn to the macroplanning of their semantic content. A well-known concept that might figure as a plan for translating intentions into content subgoals is the *schema*. Schemata are structured packets of generic knowledge that furnish much of the content needed to interpret, explain, predict, and understand events (Mandler, 1984; Graesser, Singer, & Trabasso, in press; see also Chapter 9). Besides acting as a filter for determining what information is relevant given some discourse goal, they serve as a device for the organization of content.

There is some empirical evidence supporting the importance of schematic knowledge in narrative writing. According to Trabasso, Van den Broek and Suh (1989), a *narrative* is based on a schema with different components (*setting*, *event*, *internal response*, *goal*, *attempt*, and *outcome*) which are supposed to be causally connected.

Recognizing the different components of the narrative in (8) is left as an exercise to the reader.

- (8) a. It was winter.
b. Mary wanted to surprise her mother.
c. She went to the shop and bought a sweater.
d. Her mother was very pleased.

According to Trabasso and Nickels (1992), coherence in narration is achieved when people are able to relate everyday knowledge to the protagonists' behaviour in order to infer their goals and plans according to the narrative schema. The content and the structure of a narration is the result of an interaction between a person's model of physical and psychological causation (e.g. wearing warm clothes causes people not to be cold in winter) and the events to which it is applied.

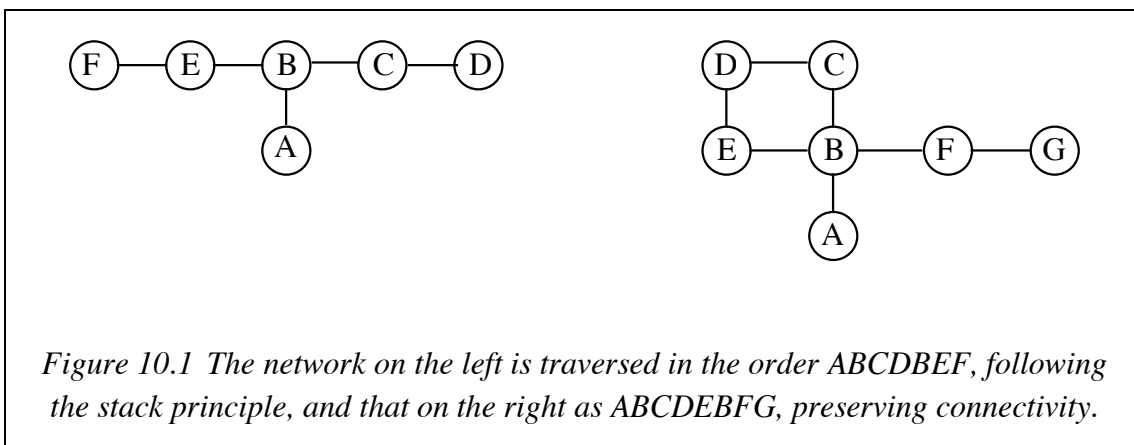
By analyzing the presence and nature of goal-plan structures from the perspective of each character in stories by children of several ages, Trabasso and Nickels (1992) were able to show how children from three years onward progressively move from simple descriptions of states to stories consisting of actions and later to explanations of actions carried out according to a goal plan. Using a sentence selection assignment, involving local planning of next sentences, Andriessen (1991) showed that coherence, defined in terms of intentional and purposeful action of the story characters, was related to the subjects' (10-12 years old) proficiency in reasoning about their decisions and to the quality of their revisions of sentences.

A familiar background knowledge structure (such as a schema) may be easily retrieved from memory for use in discourse production. A well-organized representation solves many problems of selection and organization of discourse, allowing the main focus of the discourse producer to be on what to say next. In Section 10.4.1, we will see a specific use of schemata in computational models of discourse generation. However, we should bear in mind that, no matter how useful schemata may be, they do not work in all situations. In particular, the writer's representation of the content to be expressed may not be organized well enough to fit in a single schema (Andriessen, 1991, 1994).

10.3.5 Linearization

In addition to content selection, there is the problem of determining in what order the different content elements will be presented. To deal with this problem of linearization, speakers apply a number of principles, such as mentioning causes before results, or earlier events before later ones. In several experiments, Levelt (1981, 1982a, 1982b, 1989) studied such principles in the following way. Subjects were asked to orally describe spatial grid-like networks which were put on the table in front of them. These networks consisted of differently colored dots, connected by horizontal and vertical arcs (see Figure 10.1). The subjects were asked to start their descriptions at a node indicated

by an arrow, and to proceed so as to enable the hearer to correctly draw the network on the basis of their tape-recorded description.



Analysing these results, Levelt distinguished between *content-related* and *process-related* determinants of the ordering of information. The content-related determinants derive from the so-called principle of natural ordering, in this case a spatial ordering. Linear spatial structures seem to have a natural order, imposed for the listener's sake: the connective sequence of loci from source to goal (Klein, 1979, 1982). It is also known that preserving chronological order is one of the earliest rhetorical skills in children (Clark, 1970). Process-related determinants of linearization concern the complexity of information and the bookkeeping abilities of the speaker. Levelt discusses the principle of *connectivity*, which predicts that a speaker will go over a pattern as much as possible without lifting the 'mental pencil'. Speakers rarely violated the connectivity principle for string-like patterns. Connectivity is a general ordering principle in perception and memory. Ehrich and Koster (1983) found a high degree of connectivity in the description of play furniture arrangements in a doll house.

Linde and Labov (1975) studied apartment descriptions. The subjects described their own apartments in terms of 'imaginary tours which transform spatial lay-outs into temporally organized narratives' (1975, p. 924). The narrative tour begins at the front door, just as it would if the interviewer were to arrive for the first time at the apartment. Levelt (1989) notes that these descriptions conform to a second process-related principle of linearization, the stack principle, which states that speakers always tend to return to the last node in the waiting line. Levelt's final principle is that of minimal load: when confronted with alternative branches, speakers prefer continuations which involve the least memory load. In other words, do the simplest thing first.

Put briefly, linearization of discourse has been found to follow certain cognitive principles: preserve natural order, continue the path as long as possible, return to the last digression point, and minimize memory load. While these ideas are attractive, they have only been studied in well structured domains and tasks. In other words, they merely determine what to say next.

10.3.6 Topic and focus

Elements of a message (the output of the conceptualization process) usually fulfill certain thematic roles (e.g. actor, source, goal, beneficiary). Thematic roles differ in their importance, saliency, and/or centrality to the discourse producer. Certain roles can be put in the (mental) foreground, others in the background. This perspective is what distinguishes, e.g., the sentences in (9).

- (9) a. Mary bought the book from John.
 b. The book was bought from John by Mary.

The saliency of certain discourse elements varies over time. What the discourse is about at each moment in time is called the *discourse topic*. Normally, an utterance will relate to the discourse topic (as we have seen, this will promote coherence), but sometimes a speaker or writer may want to change the topic. Such changes need to be explicitly marked (Grosz & Sidner, 1985). The fragment of information in the center of attention (often containing new information on a topic) is called the *focus*.

Investigating focus, Grosz (1977) implemented a set of mechanisms for the interpretation of definite noun phrases in a computer program that participated in a dialogue about a task. These mechanisms bring entities into focus as the discourse moves to a subtask, and move the main task back into focus when the subtask is completed. For doing so, the program uses a stack of focus spaces, containing the entities that discourse participants focus on during a specific discourse segment. The focusing techniques allow the correct prediction of the anaphoric referent of a definite noun phrase such as *the screw* when the screw in a wheelpuller has been brought into focus. Appelt and Kronfeld (1987) use the mechanisms for focusing in the generation of referring expressions. McKeown (1982, 1985) and McDonald (1983) adapt the focusing algorithms to generate pronouns in text. In Section 10.4.1 we will see a specific use of these focusing techniques in computational modeling.

Semantic microplanning consists of deciding what to say (in the immediate context) and from what point an event is viewed (perspective). It should be noted, however, that thematization pertains not only to microplanning, as it involves also what theme to select in the clause to follow. As this choice goes beyond the immediate or local context, thematization is also part of macoplanning, i.e., what to say in the global context.

Local planning has been investigated in the context of argumentative discourse by Andriessen, Coirier, Roos, Passerault and Bert-Erboul (in press). In a sentence selection paradigm, an initial and a final statement of an argumentative text were presented, and the subjects were asked to insert six arguments in between. The first and last statements expressed contrastive points of view concerning a topic (e.g., *The car is very practical* and *So, the train is more practical than the car is*, respectively). The polarity of the arguments was varied (in favour of or against the car or train) as well as the means of presentation of the sentences to be selected (6 out of 24, presented at once, or 6 times 1 out of 4). In this way, it was investigated whether the subjects (10-14 years old)

preferred arguments in favour of the first sentence (indicating local planning) or the last sentence (indicating global planning). It appeared that subjects produced better argumentative sequences, containing more kinds of arguments in a more plausible order, when all possible arguments were presented to them at once. When they received consecutive groups of four sentences, local planning predominated. The tendency for subjects of this age to continue earlier themes was also observed in narrative sentence selection (Andriessen, 1991).

10.3.7 Rhetorical Structure Theory

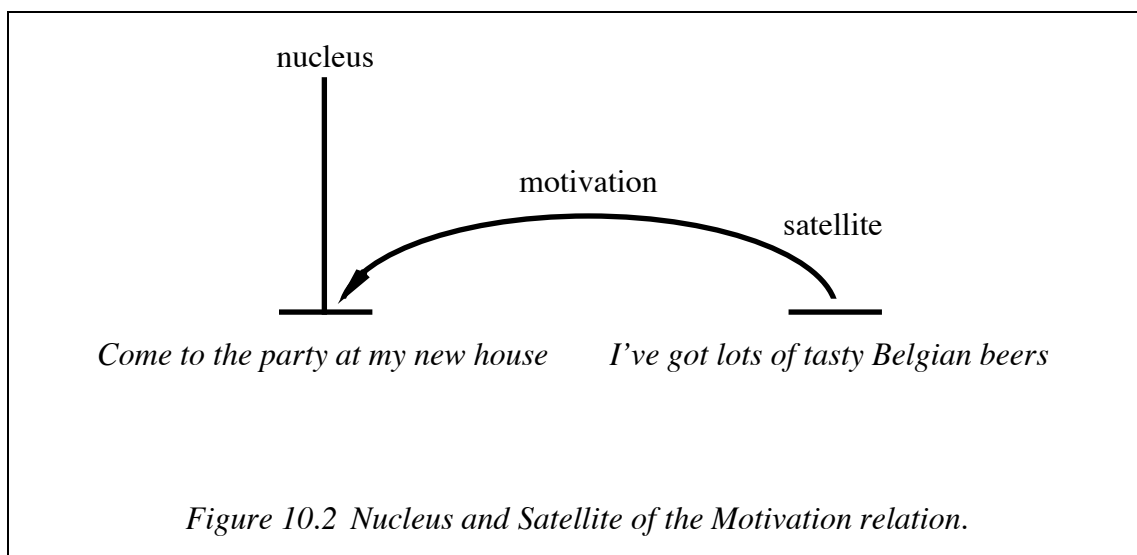
Complementary to observations about how people deal with discourse production tasks are descriptive linguistic approaches. These analyze the result of discourse production in terms of the functions of its components (for an overview see Maier & Hovy, 1991).

An important line of research in this area is Rhetorical Structure Theory (RST), which has inspired several computational modelling approaches for discourse generation (see Sections 10.4.2 and 10.4.3). The goal of RST is to describe text organization in terms of *rhetorical relations*, such as *purpose*, *enablement*, *circumstance*, *background*, *motivation*, etc. Rhetorical relations indicate which role a given part (segment) plays with respect to the whole.

On the basis of an analysis of a wide variety of texts, Mann and Thompson (1987) derived some 25 rhetorical relations. Texts can be characterized by these relations at different levels: a relation may hold between its basic elements (clauses) as well as between larger chunks (paragraphs). Relations are typically signalled by special cue words. For example, the PURPOSE relation is usually signalled by phrases such as *in order to*, *so that*, etc.

The rhetorical relation is embedded into the fundamental unit of RST, the *schema* (not to be confused with McKeown's notion of schema, see below). A schema is composed of three elements: a *nucleus*, a *satellite* and a *relation*, where the relation specifies the satellite's role with regard to the nucleus. An example is the MOTIVATION schema that can be applied to (10). RST relations are conventionally depicted as arcs between the nucleus and the satellite, e.g. in Figure 10.2.

- (10) a. Come to the party at my new house. (nucleus)
 b. I've got lots of tasty Belgian beers. (satellite)



A rhetorical relation is defined in terms of its *effects* and *constraints*. The effects specify the result a given relation shall have on the hearer (in terms of communicative goals). For example, providing a MOTIVATION may stimulate the hearer for some action, by increasing the hearer's desire. The constraints specify under which conditions a given relation holds or may be used. For example, MOTIVATION is only applicable if the nucleus expresses an action. An example of the definition of this relation is provided in (11), based on Moore and Paris (1993).

(11) <i>relation name:</i>	MOTIVATION
<i>constraints on nucleus:</i>	Presents an action (unrealized with respect to the nucleus) in which the hearer is the actor.
<i>constraints on satellite:</i>	None.
<i>constraints on N+S combination:</i>	Comprehending the satellite increases the hearer's desire to perform the action expressed in the nucleus.
<i>effect:</i>	The hearer's desire to perform the action presented in the nucleus is increased.

Schemata are unordered: satellite and nucleus can appear in any order in the schema. Furthermore, they can be used recursively: a text fragment (or *text span*) serving as the nucleus or satellite of one schema may itself be decomposed into a nucleus and satellite, using another schema. A text can thus be represented as a tree structure. In order to avoid uncontrolled growth, control strategies or constraints are needed that dictate when a satellite should appear or not, how often, and when it should be expanded as some other schema.

RST addresses coherence in terms of rhetorical relations. While this approach has successfully inspired many models of discourse generation (see Sections 10.4.2 and 10.4.3), some weaknesses have become apparent. A first limitation is that RST relations

have very weak semantics, merely indicating what general effects a given relation will have upon the hearer. Unlike an outline showing the topical structure of discourse and the writer's flow of thought, an RST tree discloses little of what a text is about. An RST-tree is like a macrostructure stripped of its content. For a complementary line of research where topic trees are built up bottom-up by using domain-specific and background knowledge, see Zock (1986).

A second problem is that the chosen relations should be cognitively basic and reliably applicable to any discourse segment. Otherwise, the list of possible relations could grow indefinitely long and become very complex. Sanders, Spooren and Noordman (1992) present a taxonomy of discourse relations in terms of cognitive primitives, such as the polarity of the relation and the semantic or pragmatic character of the link between the units. As an alternative, Knott and Dale (1994) investigate the use of explicit linguistic markers as a basis for the classification of coherence relations. The validity of these approaches in the context of discourse generation has not yet been shown.

A third problem with RST is its assumption that between two segments only one relation may hold. This is clearly wrong: many of them can signal more than one relation and can do so in a single token (Shiffrin, 1987). For example, (8) can be analyzed in terms of other relations, such as background or evidence. A more principled proposal to deal with different kinds of relations is the Parallel-Components Model (Redeker, 1992), based on the idea that multiple relations exist between utterances:

1. The ideational structure conveys the meaning of the discourse;
2. The rhetorical structure expresses a hierarchy of intentions;
3. The sequential structure signals coordination and subordination of discourse segments.

For example, a causal relation (ideational) may be used as evidence (rhetorical) for a claim or argument. This relation may constitute a structurally coordinated or subordinated segment (sequential) in the discourse.

10.3.8 Conclusion

While writing research up to now has provided a general characterization of discourse production, the relations between process and product are as yet very unclear. One of the main future tasks is to characterize strategies for discourse generation at several grain sizes (discourse and planning units) and in several domains of problem solving (rhetorical and ideational), and also to specify how these interact during planning. Computational approaches may be heuristically applied in simulations of discourse planning, in order to organize and observe the different factors involved in discourse, thus helping in the construction of a coherent and complete theoretical framework.

10.4 Computational models of discourse planning

So far, no computational psycholinguistic models are available that cover all aspects of discourse production discussed in the previous section. To date, computational approaches have especially addressed *structural* aspects of the discourse generation problem. No attempt has been made to simulate the actual writing process. The aim of most computational models is to construct working systems dealing with discourse structure in terms of schemata for canonical discourse types (e.g. McKeown, 1982, 1985), schemata for rhetorical relations (RST), speaker intentions and pragmatic constraints (e.g. Hovy, 1988; Jameson, 1990; Moore & Paris, 1993), and focus constraints (e.g. McCoy & Cheng, 1991). Issues concerning the interaction between topic selection and organization have not been addressed. Most systems simply take a message representation and find a way of expressing it.

In the discussion of the models that will follow, we will point out which processes are included and what kind of data are generated. Our selection of models is based not only on their current theoretical relevance in the field, but also on the clarity of their description and their potential for generating further research. Their inspiration comes from linguistics and Artificial Intelligence, rather than from psycholinguistics. This being so, no evaluation based on psycholinguistic criteria is attempted here. In the final section, we try to place these models in the virtual space of ‘things to do’.

10.4.1 Schemata for discourse planning: McKeown’s TEXT

TEXT (McKeown, 1982, 1985) was one of the first systems to automatically produce paragraph length discourse. The system was built as a front end to a naval data base. By communicating with the system, the user can get information about ships and weapons. McKeown analyzed texts that people produced to identify, describe, and compare objects. This analysis showed that people tended to reach a certain discourse goal by providing the same kind of information in a stereotypical way and in a rigid order.

These discourse strategies are typically composed of *rhetorical predicates*. They describe the relations (of similar grain size as RST relations) holding between two text units. Some examples of predicates are the following:

- IDENTIFICATION: identify the object as a member of some generic class or provide distinguishing attributes; e.g. *This beer is a Belgian beer.*
- CONSTITUENCY: present the constituents of the item; e.g. *This beer contains pure malt.*
- ATTRIBUTIVE: present properties of the object being defined; e.g. *This beer is dark brown.*

The combination of predicates appearing in texts with the same discourse structure is identified as a discourse strategy and can be formally represented as a *schema* (not to be confused with an RST schema). For example, a strategy where the CONSTITUENCY predicate is prominent is represented in the schema in (12), where {} indicates optionality, / indicates alternatives, + indicates ‘at least once’, * ‘any number of times’, and ; means ‘either’.

- (12) {Identification}
 Constituency
 Attributive*/Cause-Effect*
 {Depth-identification/Depth attributive
 {Particular-Illustration/Evidence}
 {Comparison; Analogy} }+
 {Attributive/Explanation/Attributive/Analogy}

Considerable freedom exists within a schema; as can be seen from the example, portions may be omitted or repeated when necessary. Each entry in the schema can be filled by an instantiated predicate or a full schema with the same name. This flexibility allows schemata to be embedded, which provides for a hierarchical account of text structure. Each schema could be associated with one or more discourse goals. For example, the *constituency schema* could be used in order to *define* or to *describe* a concept.

Incorporating these schemata, TEXT can answer three types of requests made by users: requests for a definition of an object (*define*), for a description of an object (*describe*) and for the comparison of two objects (*compare*). When asked to define or describe an object, TEXT chooses between two strategies. According to the quantity of information available in the database it will choose either the CONSTITUENCY schema (which details subparts) or the IDENTIFICATION schema (which gives defining characteristics). To generate the content of a response, TEXT follows the steps defined by the selected schema. The predicates of the schema dictate what kind of information to look for in the database.

Suppose the user requested a definition by asking *What is a guided missile?* Based on the user's question, TEXT would select a relevant subset of the knowledge base. For the current example, the system would select attributes, relations, subordinates and superordinate information of the notion *guided missile*. How the knowledge base is traced in order to determine what to say next is dictated by the *process strategy* (Paris & McKeown, 1987) which follows the structure of the knowledge base closely. Next, according to the discourse goal (define, describe, or compare) and the amount of information available in the relevant knowledge pool, a schema is chosen. Walking through the schema, TEXT instantiates the rhetorical predicates by using information from the selected subset of the knowledge base. When applied to a knowledge base on guided missiles, the constituency schema may lead to the text in (13).

- (13) (IDENTIFICATION) A guided projectile is a projectile that is self-propelled.
(CONSTITUENCY) There are two types of guided projectiles in the ONR database, torpedoes and missiles.
(IDENTIFICATION) The missile has a target location in the air or on the earth's surface.
(IDENTIFICATION) The torpedo has an underwater target location.
(EVIDENCE) The missile's target location is indicated by the DB attribute DESCRIPTION and the missile's flight capabilities are provided by the DB attribute ALTITUDE.
(EVIDENCE) The torpedos underwater capabilities are provided by the DB attributes under DEPTH (for example, MAXIMUM OPERATING DEPTH).
(ATTRIBUTIVE) The guided projectile has DB attributes TIME TO TARGET & UNITS, HORZ RANGE & UNITS and NAME.

One principle of linearization is that one should avoid side-tracking. TEXT accounts for this principle by using *focus* rules, which choose the information that ties in best with the text produced so far. McKeown takes the focus rules introduced by Sidner (1983) as a starting point and reorders the three basic focus moves as follows:

1. Change focus to a recently introduced element.
2. Maintain current focus.
3. Return to the previous focus.

It should be added that is not only an additional means of determining content, but also a means of controlling surface form (pronouns), as can be seen in example (14), in which the topic shifts in different ways (14a–c) according to the three rules, respectively.

- (14) John is a good friend of mine. He told me that he was looking for a flat.
- a. It shouldn't be too expensive.
 - b. He looked at the adds.
 - c. I know him for more than 20 years.

Schemata, as used in TEXT, have a number of interesting features. First of all, they are easy to build and use. Second, they may be defined for each type of paragraph to be generated by a specific application. For each clause typically appearing in such a paragraph, a predicate is incorporated into the schema that represents the type of information in the clause. To use a schema, the conditions of use of the predicates must be evaluated (taking into account focus), the appropriate material in the data base should be found, and the relevant material must be passed on to the realization component. Schemata are equivalent to what Levelt (1989) calls macroplans. Third, besides being useful on the macrolevel, schemata are also useful on the microlevel. The process strategy and the focusing rules adhere to the principles for linearization discussed in Section 10.3.5. In fact, Paris and McKeown point out that their process

strategy resembles the one identified by Linde and Labov (1975; see also above) for apartment descriptions.

Unfortunately, schemata also have a number of shortcomings. One limitation on their use is the fact that they do not specify the role of each part with regard to the whole. Schemata merely describe what comes next. In this respect, they are equivalent to what Bereiter and Scardamalia (1987) call knowledge telling (see Section 10.3.1). Regardless of the number of optional and repeating predicates, the same question will invariably produce the same kind of answer, irrespective of the user's expertise or interest. The instantiation of the predicates in schemata is only driven by what is found in the knowledge base. This indicates that McKeown's communicative goals are not properly contextualized. The complex planning required for knowledge transforming (see Section 10.3.1) is therefore far beyond the capabilities of TEXT. From another perspective, a schema can be viewed as the result of a compilation process where the rationale for all the steps in the process has been compiled out (Moore & Swartout, 1991). Because of this compilation, schemata provide an efficient but inflexible way to produce multisentential texts for achieving generic discourse purposes. A more flexible approach to planning will be discussed in the next section.

10.4.2 Rhetorical relations: Hovy's Structurer

As we have seen in Section 10.3.7, RST is a descriptive theory of the organization of natural language texts. An RST description of a text is a hierarchical structure (consisting of clauses, sentences, paragraphs) that characterizes the text in terms of basic rhetorical relations holding between the parts of the text. The definition of each RST relation includes constraints on the two entities being related and on their combination, as well as a specification of the effect which the speaker attempts to achieve on the hearer's beliefs. Because RST provides an explicit connection between the speaker's intention and the rhetorical means to achieve those intentions, RST offers a more flexible approach to planning (Hovy, 1991; Moore & Paris, 1993) than the use of McKeown's schemata.

In order to be applicable to text generation, RST relations must be implemented in a discourse planner. Hovy (1988) was the first to operationalize a subset of RST relations into plans, by representing them as NOAH-like plan operators (Sacerdoti, 1977). Operators are named after their corresponding RST relation, e.g. SEQUENCE, a simplified example of which is given in (15) below.

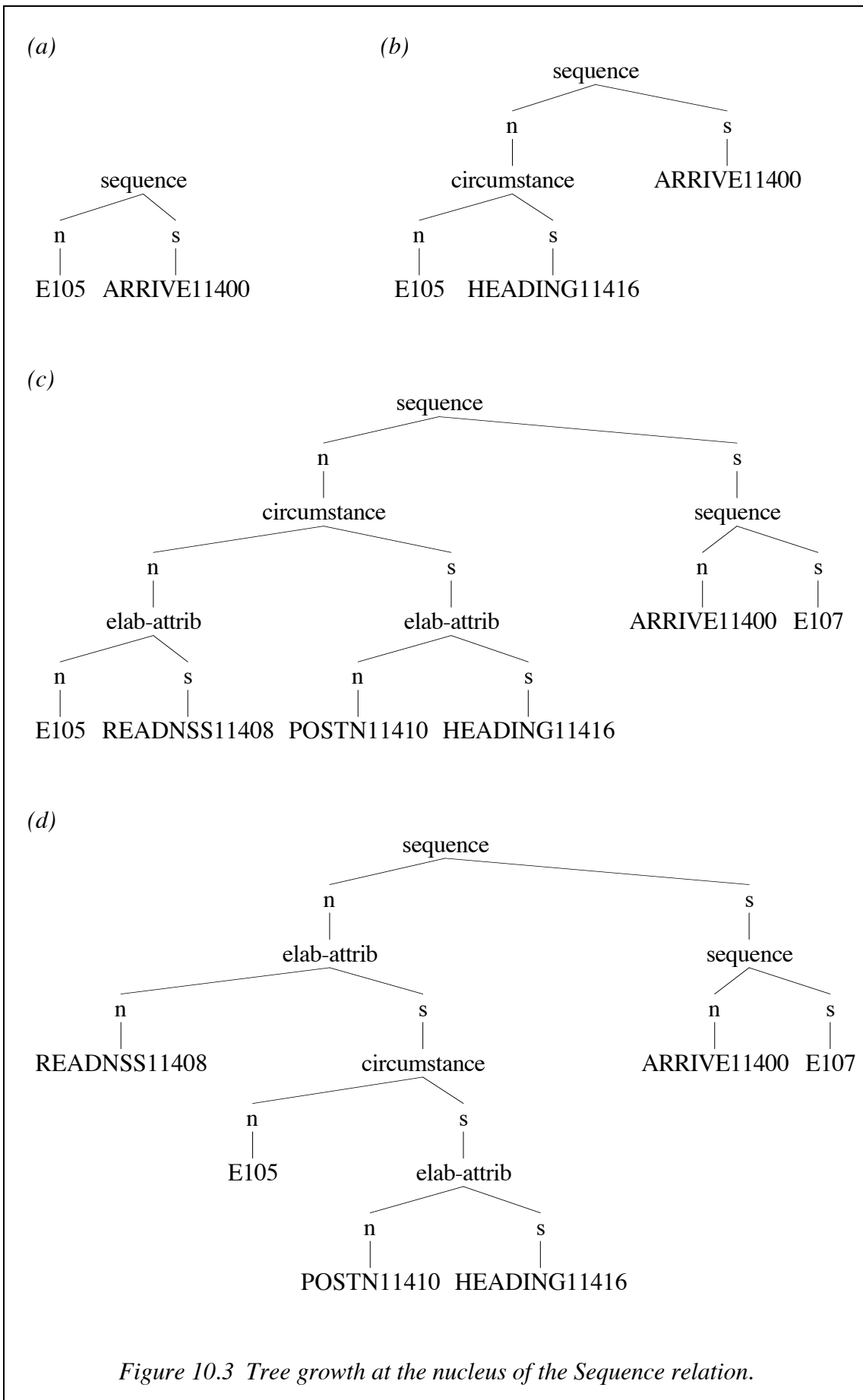
(15) Results	(SEQUENCE-OF ?PART ?NEXT)
N + S requirements/subgoals	((NEXT ACTION ?PART ?NEXT))
Nucleus requirements/subgoals	(TOPIC ?PART)
Satellite requirements/subgoals	(TOPIC ?NEXT)
Nucleus growth points	(CIRCUMSTANCE-OF ?PART ?CIRC)
	(ATTRIBUTE-OF ?PART ?VAL)
	(PURPOSE-OF ?PART ?PURP)
Satellite growth points	(ATTRIBUTE-OF ?NEXT ?VAL)

	(DETAILS-OF ?NEXT ?DETS)
	(SEQUENCE-OF ?NEXT ?FOLL)
Order:	(NUCLEUS SATELLITE)
Relation-phrases:	(“” “then” “next”)

The intended effect of the RST relation is mapped into the *results*-field of the operator, while the constraints of the RST relation are mapped into *requirements/subgoals*, which are treated as semantic preconditions based on the knowledge of the hearer. So-called *growth points* are included which signal appropriate spots for conveying additional material. The inclusion of growth points was motivated by an extensive analysis of relevant texts and interviews with domain experts. Below we will describe how all this is put to work in discourse planning.

Plan operators of the kind we just described are called *relation/plans* in Hovy’s (1991) text planner or Structurer. The domain Hovy uses is a naval application in which the Structurer, together with Penman, a surface generator (Mann & Matthiessen, 1985; Penman Natural Language Generation Group, 1989), are part of a larger system that presents database information about U.S. Navy vessels to a user by means of maps, tables, and text (Arens, Miller, Shapiro & Sondheimer, 1988). The database consists of a network of assertions about entities and actions. When a goal is posted (by a host system), the Structurer tries to find a relation/plan whose results-field matches this goal. The output of the Structurer is a hierarchical structure, called the paragraph tree (Figure 10.3), which contains the discourse relation/plan as the top goal, and retrieved data base elements at the bottom leaves.

Consider an example (based on Hovy, 1991, 1993), in which the user asks for the next position of a particular vessel (called Knox) in the data base. To the Structurer, this goal is represented in the following way: (BMB SPEAKER HEARER (POSITION-OF E105 ?NEXT)), which can put in plain language as: *Achieve the state in which the hearer believes that it is the intention of the speaker that they mutually believe that the event E105 is followed by some other event.* In what follows, we will simplify this goal by leaving out the BMB-part. The Structurer thus starts with this goal, simplified as (POSITION-OF E105 ?NEXT) which matches the results-field of the relation/plan SEQUENCE, shown in (15). In the match, ?PART is bound to E105, and with this binding, the Structurer begins searching for an appropriate nucleus, as the core of the message to be expressed. To accomplish this, it searches for input entities in the database that match the first requirement, which is the combined nucleus and satellite requirements (i.e. line 2 in (15)). The database input contains the information that the arrival of the vessel (ARRIVE11400) is the next action. This becomes bound to ?NEXT, which then becomes the satellite of the SEQUENCE relation/plan (line 4 in (15)), while E105 is the nucleus (Figure 10.3a).



Next, the growth points are considered. Suggestions for additional input material related to the nucleus are considered in the Nucleus growth point field: these call for circumstances, attributes, and purpose. These act as subgoals the planner must try to achieve. A similar set is associated with the satellite. The first growth point to be considered (i.e. line 5 in (15)) is (CIRCUMSTANCE-OF ?PART ?CIRC). This appears to match the results-field of the CIRCUMSTANCE relation/plan (not shown). In the same way as for the SEQUENCE relation/plan, a match is sought for the variable ?CIRC. In this case the data base provides the heading of the ship (HEADING11416). The found CIRCUMSTANCE relation between E105 and HEADING11416 thus fulfills the growth-point goal of the original SEQUENCE nucleus, which causes the tree to grow at this point. The nucleus E105 is moved down to become the nucleus of the newly formed CIRCUMSTANCE relation, where HEADING11416 becomes the satellite. The whole CIRCUMSTANCE relation then replaces the original nucleus in the SEQUENCE (Figure 10.3b).

Put briefly, the propagation of remaining growth points eventually leads to further growth of the paragraph tree (Figure 10.3c). The whole process stops when no satisfiable goals remain posted or the input is exhausted, regardless whether at the end some growth points may remain unsatisfied. After adding the relation's characteristic cue words (line 12 in (15)) or phrases to the appropriate input entries and setting the appropriate syntactic constraints, the tree structures are transmitted to Penman for surface generation (16a). Notice that one of the *relation-phrases* in the last line of (15) is used.

Text planning with RST has enjoyed several enrichments. Consider for example the surface form in (16a), which looks rather odd. For example, the repeated use of the pronoun *it* does not seem natural. A speaker or writer would probably have used the words *the ship* or *the vessel*. Furthermore, with regard to coherence, the text introduces first the circumstance (condition) of the Knox and then enumerates a sequence of events. While this may be structurally appropriate, it fails to group semantically related material concerning the direction: *to head SSW*, and *to be en route to Sasebo* (see Maybury, 1992, pp. 80–81).

- (16) a. Knox, which is C4, is en route to Sasebo. Knox, which is at 18N 79E, heads SSW. It arrives on 4/24. It then loads for 4 days.
- b. With readiness C4, Knox is en route to Sasebo. It is at 18N 79E, heading SSW. It will arrive on 4/24. It will load for 4 days.

To overcome some of these problems, Hovy and McCoy (1989) enriched RST in order to promote coherence, by using *discourse focus trees* (McCoy & Cheng, 1991). In this way, the initial text could be improved to (16b). To this end, the focusing rules used by TEXT (see Section 10.4.1) were extended by representing the topics in the discourse as nodes in a tree, which is built up and traversed as the discourse progresses. McCoy and Cheng identified some general constraints for hopping from one node to the next, according to the conceptual types of the nodes. For example, if the current focus is on

an object, the next focus may be one of its attributes or actions. Hovy & McCoy describe how an RST paragraph tree and a focus tree can be constructed in parallel. During the expansion of a node in the RST discourse structure, the Structurer disregards questions with respect to the ordering of the growth points, collecting all the potential candidate relations and their associated data base inputs. Each candidate relation is then checked against the currently legal focus shifts in the Focus Tree.

To recapitulate, Hovy's Structurer transforms RST into a text planner which allows more versatile planning than the schemata in TEXT, thanks to the use of growth points. The use of independent relation/plan operators explicitly links intentions and rhetorical relations in a hierarchy. Each RST relation/plan is simultaneously a basic rhetorical operator (microplan) which can be incorporated into a schema as well as a generalized schema (macroplan) for building a specific type of paragraph. Furthermore, Hovy (1993) argues that a text planner where RST relation/plans are combined with intracause planning rules (such as those proposed by Appelt, 1985) and focusing rules (McCoy & Cheng, 1991) offers fine grained control over smaller spans of text than schemata. This also allows control over various syntactic aspects, such as relative clauses, the use of the tense, and the combining of several clauses into a single sentence.

10.4.3 Moore & Paris: planning of explanations

Hovy's text Structurer orders the inputs from the domain according to the constraints on the RST relations. It looks for some coherent way to organize the text so that all of the information in the input is included according to the requirements of the plan operators. The lack of a distinction between content selection and organization ignores the possibility that the same content domain may be used for different goals, which may mandate different items to be selected. Conversely, the availability of content may affect the discourse production strategy. For instance, the decision to present an example depends, among other things, upon the speaker's knowledge of such an example.

Furthermore, Hovy's system lacks higher order goals that allow the system to explain why it behaves as it does. The rhetorical relations serve at the same time as communication plans and as discourse structuring relations. In other words, communication goals can only be stated directly in terms of rhetorical relations. Moore and Swartout (1991) argue for a separation of *intentional* and *rhetorical* relations as well as for the primacy of communicative intentions. This is motivated by the fact that there is no one-to-one mapping between intentions and rhetorical relations. Moore and Pollack (1992) present example (17) to illustrate this point:

- (17) S: (a) Come home by 5:00.
 (b) Then we can go to the hardware store before it closes.
 H: (c) We don't need to go to the hardware store.
 (d) I borrowed a saw from Jane.

At the informational level, utterance (17a) is a CONDITION for (17b). Getting to the hardware store before it closes depends on H's coming home, but at the intentional level S may be trying to increase the ability of H to perform the act described in (17b). It is thus an ENABLEMENT, if S believes that H does not realize that the store closes early tonight. On the other hand, S may be trying to motivate H to come home early, say because S is planning a surprise party for H (MOTIVATION). H's reaction (17c-d) requires further motivation.

Example (17) shows that intentions and rhetorical relations do not map one-to-one. In particular, it illustrates that a generation system cannot simply rely on the information to be conveyed, while disregarding the speaker's underlying intentions. It is only on the basis of the intentions underlying (17a) and (17b) that the speaker can decide how to subsequently respond to (17c) and (17d).

Moore and Paris (1993) describe a Text Planner that constructs explanations, in the context of a prototype expert system called Program Enhancement Advisor (PEA), which gives advice to beginning LISP programmers. The Text Planner is based on the intentions (goals) of the speaker at each moment of discourse production and finds the linguistic means available for realizing these intentions. We will describe some details of this Text Planner, first examining the goals represented in the text plan and subsequently the operation of the operators.

Moore and Paris distinguish two types of goals to be reached by the discourse producer: communicative goals and linguistic goals. *Communicative goals* represent the speaker's intentions to affect the beliefs or goals of the hearer. Given a goal representing the speaker's intention, the planner tries to find the linguistic resources available for achieving that goal by posting *linguistic goals*. The latter lead to the generation of text and are of two types: speech acts and rhetorical goals. *Speech acts*, such as INFORM or RECOMMEND, map straightforwardly into utterances that form part of the final text. *Rhetorical goals*, such as MOTIVATION and CIRCUMSTANCE (compare RST relations) cannot be achieved directly but must be refined into one or more subgoals, which may be further communicative goals or speech acts.

Plans are utilized by the same style of hierarchical planner as Hovy's RST-based Structurer. The plan language provides *operators* which implement both general and specific strategies. The effect of an operator is defined in terms of a communicative or linguistic goal; constraints on the operator are listed as conditions which should be true for the operator to have the intended effect; furthermore, an operator specifies a nucleus (the most essential subgoal) and satellites (additional subgoals). As examples, operators for RECOMMEND and MOTIVATION are given in (18) and (19), respectively.

- (18) NAME: **recommend-enable-motivate**
 EFFECT: (GOAL ?hearer (DO ?hearer ?act))
 CONSTRAINTS: (Nucleus)
 NUCLEUS: (RECOMMEND ?speaker ?hearer ?act)
 SATELLITES: (((COMPETENT ?hearer (DONE ?hearer ?act)) *optional*)

(PERSUADED ?hearer (GOAL ?hearer (DO ?hearer ?act))
 optional))

English paraphrase:

To make the hearer want to do an act,
 IF this text span is to appear in the Nucleus position, THEN
 1. Recommend the act
 AND optionally,
 2. Achieve the state where the hearer is competent to do the act
 3. Achieve the state where the hearer is persuaded to do the act

(19) EFFECT: (PERSUADED ?hearer (DO ?hearer ?act))
 CONSTRAINTS: (AND (STEP ?act ?goal)
 (GOAL ?hearer ?goal)
 (MOST-SPECIFIC ?goal)
 (CURRENT FOCUS ?act)
 (SATELLITE))
 NUCLEUS: (FORALL ?goal (MOTIVATION ?act ?goal))
 SATELLITE: none

English paraphrase:

To achieve the state in which the hearer is persuaded to perform an act,
 IF the act is a step in achieving some goal(s) of the hearer,
 AND the goal(s) are the most specific along any refinement path
 AND the act is the current focus of attention
 AND the planner is expanding a satellite branch of the text plan
 THEN motivate the act in terms of these goal(s).

The planning process begins when a communicative goal is posted, for example *Make the hearer set a cup of coffee*. When a goal is posted, the planner searches its library for all operators whose *effect* field matches the goal. To make this search more efficient, plan operators are stored in a discrimination network based on their effect field. When the plan operator in (18) is selected, it posts its nucleus as a discourse subgoal, in this case RECOMMEND. This goal is defined as a speech act, which maps directly into a specification for the sentence generator. The two satellites, however, require further operators, among which the one in (19), indicating that the communicative goal of persuading can be achieved by using the rhetorical strategy MOTIVATION.

Note the explicit representation of various knowledge sources that are included in the constraints of (19). The first constraint (STEP ?act ?goal) says that there must be some domain goal(s) for which the act is a step in achieving. Satisfying this goal requires the planner to search the expert system's domain knowledge for such goals. The second constraint (GOAL ?hearer ?goal) specifies that if any such domain goal(s) exist, they must be goals of the user. For this, the system must inspect the user model. The last two constraints refer to the evolving text plan. They state that the operator can

only be used if the act is in focus and that a satellite branch of the current text plan is expanded.

Planning proceeds by selecting one operator among those whose constraints are satisfied. Once a plan operator has been selected, it is recorded in the current plan node and the others are recorded as untried alternatives. The planner then expands the plan by posting its nucleus and required satellites as subgoals to be refined. For each subgoal, candidate operators are again selected and the planning process is repeated.

The planner maintains an agenda of pending goals to be satisfied. New subgoals for the nucleus and satellites are collected in a list, which is added to the front of the agenda. In this way, the text plan is built in a depth first manner (see Chapter 2). When a speech act is reached, no further elaboration of the plan at that point is necessary, but instead, the system constructs a specification that directs the realization component PENMAN which formulates the utterance. As the planner examines each of the arguments of the speech act, new goals may be posted as a side effect. Whether optional satellites are expanded depends on a parameter representing *verbosity* (terse or verbose) in the system. In terse mode, no optional satellites are expanded, whereas in verbose mode, each satellite is checked against the user model, and those that add to the user's knowledge are expanded.

Summing up, Moore and Paris extended the work on RST-based planning in various principled ways. While using plan-based operators resembling those of Hovy's Structurer, they also take the important distinction between communicative and linguistic goals as a new point of departure. The Text Planner described by Moore and Paris has a set of 150 operators that can answer questions like *Why?*, *Why...conclusion?*, *Why are you trying to achieve goal?*, *Why are you using this method?*, *Why are you doing act?*, *What is a concept?*, *What is the difference between concept1 and concept2?*, *Huh?* etc. The planner is implemented into several knowledge base systems and two intelligent tutoring systems (Rosenblum & Moore, 1993). Work in progress on the system includes, among other things, the use of focus and the implementation of dialogue management strategies. The problem of phrasing utterances for different types of users and situations is also being investigated (Paris, 1991).

10.5 Evaluation and conclusion

Computational models for coherent text generation can only be evaluated according to psycholinguistic criteria when they incorporate explicit mechanisms for selecting and organizing what to say, as well as devices to translate these intentions into semantic and pragmatic plans. When, in addition, mechanisms are included for monitoring and revising both ideas and text, thereby allowing answers to questions about the reasons for producing a piece of discourse, the status of the model can be evaluated with respect to theories of writing. When neither is the case, we can only evaluate the internal characteristics of the model. For example, we can investigate what types of discourse the system is able to generate. Of course, such an exercise may still generate interesting ideas concerning the mechanisms involved in discourse planning.

Despite the relative lack of empirical evidence, computational work is in progress, yielding interesting results with respect to several aspects of discourse planning. We have focussed our discussion on three approaches (McKeown's TEXT, Hovy's Structurer, and Moore and Paris' Text Planner), but many more are worth mentioning. For example, the work by Dale (1992) on the generation of referring expressions, the work of Endres-Niggemeyer, Maier and Sigel (1994) on developing an expert system for abstracting and that by Jameson (1990) on the selection of information, thereby dealing with conversational maxims. It seems that computational approaches will soon be in a position to test specific questions about discourse planning. It is unfortunate that relevant testable hypotheses based on psycholinguistic models are lacking.

The systems that we discussed show a progress from more or less prepackaged plans to dynamic text structuring to a more explicit link between communicative goals and linguistic means. As a consequence, the output of computational text planning steadily improves. In particular, progress has been made with respect to the tailoring of texts to specific users and the coherence of multisentence paragraphs. It remains to be seen whether these approaches model the richness that characterizes human discourse planning. The interaction between communicative goals and linguistic goals in Moore & Paris' Text Planner is at least reminiscent of the interaction between the ideational and rhetorical domains which characterizes knowledge transforming (cf. 10.2.1).

Furthermore, it seems hard to move from a descriptive theory of text structure to a theory of multisentence paragraph planning. Text planning systems may solve problems, but they may do so in a different way than human writers do. Moreover, an open-ended task such as discourse planning may offer a wealth of qualitatively different strategies for problem solving. Only a few of these strategies have been studied so far. If current systems were used as a testbed for the evaluation of alternative strategies for planning discourse, some new insights could be expected.

Currently prominent text planners crucially hinge on the use of RST-inspired planning operators. The status of these operators is questionable, not only from a psychological perspective. Hovy (1993) discusses an attempt to assemble a core library of discourse relations. It is assumed that these should at least be described in intentional, structural, semantic and rhetorical terms. However, since the context of use generally determines their nature and specificity, a closed set of such operators may, in fact, not exist (Grosz & Sidner, 1985; Polanyi, 1988).

Dealing with the flexibility that characterizes human discourse generation may require better theories of planning, text coherence, and knowledge representation than are currently available. Computational research is gradually moving from the descriptive to the procedural phase. That is still a long way from modelling strategies.

10.6 References

Andriessen, J. E. B. (1991). *Minimal strategies for coherent text production*. Doctoral dissertation. Utrecht: Instituut voor Sociaalwetenschappelijk Onderzoek Rijksuniversiteit Utrecht (ISOR).

- Andriessen, J. E. B. (1994). Episodic representations in writing and reading. In G. Eigler & Th. Jechle (Eds.), *Writing: Current trends in European research* (pp. 71–83). Freiburg: Hochschulverlag.
- Andriessen, J. E. B., Coirier, P., Roos, L., Passerault, J. M., & Bert-Erboul, A. (in press). Thematic and structural planning in constrained argumentative text production. In H. VandenBergh & G. Rijlaarsdam (Eds.), *Current trends in writing research: What is writing?* Amsterdam: Amsterdam University Press.
- Appelt, D. E. (1985). Planning English referring expressions. *Artificial Intelligence*, 26, 1–33.
- Appelt, D. E., & Kronfeld, A. (1987). A computational model of referring. *Proceedings of the 10th International Joint Conference on Artificial Intelligence, Milan* (pp. 640–647).
- Arens, Y., Miller, L., Shapiro, S. C., & Sondheimer, N. K. (1988). Automatic construction of user-interface displays. *Proceedings of the 7th Conference of the American Association for Artificial Intelligence, St. Paul* (pp. 808–813).
- Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon.
- Barnard, Y. F., Andriessen, J. E. B., Bläcker, T., & Erkens, G. (1989). Fostering reflection on organizational functions of text segments during expository paper composition. *3rd EARLI Conference, Madrid*.
- Bartlett, E. J. (1984). Anaphoric reference in written narratives of good and poor writers. *Journal of Verbal Learning and Verbal Behavior*, 23, 540–552.
- Beaugrande, R. de (1984). *Text production: Towards a science of composition*. Norwood, NJ: Ablex.
- Bereiter, C., & Scardamalia, M. (1983). Does learning to write have to be so difficult? In I. P. Y. A. Freedman (Ed.), *Learning to write: First language, second language* (pp. 20–33). New York: Longman.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Bronckart, J. P. (1985). *Le fonctionnement des discours: Un modèle psycholinguistique et une méthode d'analyse*. Neuchâtel: Delachaux et Nestlé.
- Brown, G., & Yule, G. (1985). *Discourse analysis*. Cambridge: Cambridge University Press.
- Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 155–175). London: Academic Press.
- Clark, E. V. (1970). How children describe events in time. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in Psycholinguistics*. Amsterdam: North Holland.
- Clark, H. H., & Wilkes-Gibbs, D. L. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Cooper, C. R., & Matsuhashi, A. (1983). A theory of the writing process. In M. Martlew (Ed.), *The psychology of written language* (pp. 3–39). New York: Wiley.
- Dale, R. (1992). *Generating referring expressions*. Cambridge, MA: MIT Press.

- Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*, 159–184.
- Ehrich, V., & Koster, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes*, *6*, 169–195.
- Endres-Niggemeyer, B., Maier, E., & Sigel, A. (submitted). How to implement a naturalistic model of abstracting: Four core working steps of an expert abstractor.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Fayol, M. (1991). From sentence production to text production: Investigating fundamental processes. *European Journal of Psychology of Education*, *6*, (2), 101–120.
- Flower, L. F. (1981). *Problem-solving strategies for writing*. New York: Harcourt Brace Jovanovich.
- Flower, L., & Hayes, J. R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive Processes in writing* (pp. 39–58). Hillsdale, NJ: Erlbaum.
- Flower, L., & Hayes, J. R. (1981). Plans that guide the composing process. In C. H. Frederiksen & J. F. Dominic (Eds.), *Writing: The nature, development and teaching of written communication* (Vol. 2, 39–58). Hillsdale, NJ: Erlbaum.
- Flower, L., & Hayes, J. R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication*, *1* (4), 120–160.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (in press) Constructing inferences during narrative text comprehension. *Acta Psychologica*.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.) *Syntax & Semantics 3: Speech Acts* (pp. 41–58). New York: Academic Press.
- Grosz, B. (1977). *The representation and use of focus in dialogue understanding* (Technical report 151). Menlo Park, CA: SRI International.
- Grosz, B. J., & Sidner, C. L. (1985). Discourse structure and the proper treatment of interruptions. In A. Joshi (Ed.), *Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles* (Vol. 2, pp. 832–839). Los Altos, CA: Morgan Kaufmann.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hayes, J. R., & Flower, L. (1986). Writing research and the writer. *American Psychologist*, *41*, 1106–1113.
- Hermann, T., & Grabowski, J. (1994). Pre-terminal levels of process in oral and written language production. In U. Quasthoff (Ed.), *Aspects of oral communication*. Berlin: De Gruyter.
- Hovy, E. H. (1988). *Generating natural language under pragmatic constraints*. Hillsdale, NJ: Erlbaum.

- Hovy, E. H. (1991). Approaches to the planning of coherent text. In C. L. Paris, W. R. Swartout, & W. C. Mann (Eds.), *Natural language generation in artificial intelligence and computational linguistics* (pp. 83–102). Boston: Kluwer Academic Publishers.
- Hovy, E. H. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence*, 63, 341–385.
- Hovy, E. H., & McCoy, K. F. (1989). Focusing your RST: A step towards generating coherent mutisential text. *Proceedings of the Annual Conference of the Cognitive Science Society, Ann Arbor, MI* (pp. 667–674).
- Jameson, A. (1990). *Knowing what others know. Studies in intuitive psychometrics*. (Technical report 90–10). Nijmegen: Nijmegen Institute for Cognition and Information, University of Nijmegen.
- Klein, W. (1979). Wegauskunft. *Zeitschrift für Literaturwissenschaft und Linguistik*, 9, 9–57.
- Klein, W. (1982). Local deixis in route directions. In R. J. Jarvella & W. M. Klein (Eds.), *Speech, place and action: Studies in deixis and related topics*. Chichester: Wiley.
- Knott, A. & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18, 35–62.
- Levelt, W. J. M. (1981). The speaker's linearization problem. *Philosophical Transactions Royal Society London*, 295, 305–315.
- Levelt, W. J. M. (1982a). Linearization in describing spatial networks. In S. Peters & E. Saarinen (Eds.), *Processes, beliefs, and questions*. Dordrecht: Reidel.
- Levelt, W. J. M. (1982b). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella & W. M. Klein (Eds.), *Speech, place and action: Studies in deixis and related topics*. Chichester: Wiley.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language*, 51, 924–939.
- Maier, E., & Hovy, E. H. (1991). Organising discourse structure relations using metafunctions. In H. Horacek & M. Zock (Eds.), *New concepts in natural language generation* (pp. 69–86). London: Pinter.
- Mandler, J. M. (1984). *Stories, scripts and scenes: aspects of schema theory*. Hillsdale, NJ: Erlbaum.
- Mann, W. C., & Matthiessen, C. (1985). A demonstration of the NIGEL text generation computer program. In R. Benson & J. Greaves (Eds.), *Systemic perspectives on discourse* (pp. 50–83). Norwood, NJ: Ablex.
- Mann, W. C., & Thompson, S. A. (1987). Rhetorical structure theory: Description and construction of text structures. In G. Kempen (Ed.), *Natural language generation: New results in artificial intelligence, psychology and linguistics* (pp. 85–95). Dordrecht: Nijhoff (Kluwer).

- Matsuhashi, A. (1987). Revising the plan and altering the text. In A. Matsuhashi (Ed.), *Writing in real time* (pp. 197–223). Norwood, NJ: Ablex.
- Maybury, M. T. (1992). Communicative acts for explanation generation. *International Journal of Man-Machine studies*, 37, 135–172.
- McCoy, K. F., & Cheng, J. (1991). Focus of attention: Constraining what can be said next. In C. L. Paris, W. R. Swartout, & W. C. Mann (Eds.), *Natural language generation in artificial intelligence and computational linguistics* (pp. 103–124). Boston: Kluwer Academic Publishers.
- McCutchen, D., & Perfetti, C. A. (1982). Coherence and connectedness in the development of discourse production. *Text*, 2, 113–139.
- McDonald, D. (1983). Description directed control: its implications for natural language generation. *Computers & Mathematics*, 9, 111–130.
- McKeown, K. (1982). The TEXT system for natural language generation: An overview. *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics, Toronto* (pp. 113–120).
- McKeown, K. R. (1985). *Text generation: Using discourse strategies and focus constraints to generate natural language text*. Cambridge: Cambridge University Press.
- Moore, J. D., & Paris, C. L. (1993). Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19, 651–694.
- Moore, J. D., & Pollack, M. E. (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics*, 18, 537–544.
- Moore, J. D., & Swartout, W. R. (1991). A reactive approach to explanation: Taking the user's feedback into account. In C. L. Paris, W. R. Swartout, & W. C. Mann (Eds.), *Natural language generation in artificial intelligence and computational linguistics* (pp. 3–48). Boston: Kluwer Academic Publishers.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Paris, C. L. (1991). Generation and explanation: Building an explanation facility for the explainable expert systems framework. In C. L. Paris, W. R. Swartout, & W. C. Mann (Eds.), *Natural language generation in artificial intelligence and computational linguistics* (pp. 49–82). Boston: Kluwer Academic Publishers.
- Paris, C. L., & McKeown, K. R. (1987). Discourse strategies for descriptions of complex physical objects. In G. Kempen (Ed.), *Natural language generation: New results in artificial intelligence, psychology and linguistics* (pp. 97–115). Dordrecht: Nijhoff.
- Penman Natural Language Generation Group (1989). *The PENMAN user guide*. Information Science Institute, University of Southern California, Marina Del Rey, CA.
- Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12, 601–638.
- Redeker, G. (1992). *Coherence and structure in text and discourse*. Unpublished manuscript, Tilburg University.

- Roberts, R. M., & Kreuz, R. J. (1993). Nonstandard discourse and its coherence. *Discourse Processes* 16, 4, 451–464.
- Rosenblum J., & Moore, J. (1993). Participating in instructional dialogues: Finding and exploiting relevant prior explanations. In P. Brna, S. Ohlsson, & H. Pain (Eds.), *Proceedings of the 5th World Conference on Artificial Intelligence in Education, Edinburgh* (pp. 145–152).
- Roussey, J-Y., & Gombert, A. (1992). Ecriture en dyade d'un texte argumentatif par des enfants de huit ans. *Archives de Psychologie*, 60, 297–315.
- Sacerdoti, E. (1977). *A structure for plans and behavior*. Amsterdam: North Holland.
- Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Towards a taxonomy of coherence relations. *Discourse Processes*, 15, 1–35.
- Scardamalia, M., & Paris, P. (1985). The function of explicit discourse knowledge in the development of text representations and composing strategies. *Cognition and Instruction*, 2, 1–39.
- Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.
- Seuren, P. A. M. (1985). *Discourse semantics*. Oxford: Blackwell.
- Shiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Sidner, C. L. (1983). Focusing in the comprehension of definite anaphora. In J. M. Brady & R. C. Berwick (Eds.), *Computational models of discourse* (pp. 267–330). Cambridge, MA: MIT Press.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the Narration. *Discourse Processes* 15, 249–276.
- Trabasso, T., Van den Broek, P., & Suh, M. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, 12, 1–25.
- Van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Zock, M. (1986). *Le fil d'Ariane ou les grammaires de texte comme guide dans l'organisation et l'expression de la pensée en langue maternelle et/ou étrangère*. Paris: UNESCO.