AUTOMATED DISCOURSE GENERATION USING DISCOURSE STRUCTURE RELATIONS

Eduard H. Hovy

Information Sciences Institute
of the University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292-6695

Tel: 310-822-1511
Fax: 310-823-6714
Email: HOVY@ISI.EDU

Abstract

This paper summarizes work over the past five years on the automated planning and generation of multisentence texts using discourse structure relations, placing it in context of ongoing efforts by Computational Linguists and Linguists to understand the structure of discourse. Based on a series of studies by the author and others, the paper describes how the orientation of generation toward communicative intentions illuminates the central structural role played by intersegment discourse relations. It outlines several facets of discourse structure relations as they are required by and used in text planners — their nature, number, and extension to associated tasks such as sentence planning and text formatting.

In *Artificial Intelligence* 63, Special Issue on Natural Language Processing, 1993.

This work was partially supported by the Rome Air Development Center under RADC contract FQ7619-89-03326-0001.

1 Introduction

Every day, people produce thousands of words of connected discourse from complicated internal knowledge for little-understood reasons. Today, after three decades of work on natural language processing, computers are beginning to approach this capability. Computational studies such as [Appelt 85] have established the power of viewing language generation as a goal-driven and hence essentially planning process (in contrast to analysis, which is input-driven and essentially inferential). This perspective leads to the construction of text planners and sentence generators that govern the selection and assembly of material into coherent grammatical text in service of the speaker's communicative goals.

An important issue under this perspective is the nature of text plans. What text plans there are, whether they can be thought of as implementing a multisentence grammar, what information they contain, and what kind of discourse structure they are assembled into, are all questions for which answers are required before a thorough understanding of discourse is possible. Though none of these questions has been fully answered to date, several interesting new results have come to the fore over the past five years on the role of discourse structure relations, an important aspect of text plans which help make up and give structure to coherent discourse.

This paper focuses on discourse structure and discourse structure relations as seen from the text planning perspective. It can serve as a survey of what has been done recently and a pointer to where research can fruitfully be performed. After arguing in Section 2 that without an understanding of discourse structure, communication is unlikely to succeed, the paper outlines various theories of discourse structure, linguistic and computational. Section 3 describes an early computational attempt, the first of several similar efforts, to plan discourse structure automatically by dynamically constructing a tree of interclause operators or relations. These attempts' general requirements for discourse structure are summarized in Section 4. Section 5 then presents four primary aspects of discourse structure relations that arise, regardless of particular theory of discourse structure, when they are employed to plan discourse automatically. Finally, Section 6 describes the effects of discourse structure relations on related tasks such as sentence planning and text formatting.

As an initial assumption, we take it that discourse is goal-oriented: people communicate for a reason. Though these goals do not always decompose into a structure of increasingly specific subgoals — think of interacting with a 4-year-old, joking in a supermarket line, reminiscing around a fire — enough of them do to make the traditional Artificial Intelligence planning approach, namely goal decomposition, rewarding. Discourses that admit such an analysis are typically informative messages such as annual reports and encyclopedia entries, instructions, explanations, and other collaborations toward some purpose — the kinds of conversations we want to have with computers in any case.

In this paper, we discuss only monologic discourse; the additional issues that are required for multi-party discourse are still at early stages of study.

2 Discourse Structure

2.1 The Problem

No account of language that stops at the sentence level is adequate. Neither are programs that communicate solely on the sentence level. But moving "up" to the paragraph level has proven a difficult matter — you cannot simply string together sentences. Of the $n!$ permutations possible for n sentences, usually only a handful of them make semantic sense, and often their meanings differ quite radically. For example, on being told that

1. Zurab and Maria had a fight last night.
2. Maria was found dead this morning.

you are fully within your rights to assume that the fight somehow caused Maria's death, and that Zurab was the perpetrator. The juxtaposition of these two sentences in the null context combines with world knowledge that a fight can cause a death to license the inference of Zurab's guilt. However, both prior and subsequent knowledge can block that inference and/or cause others to be made, especially when aided by cue words, as in:

a1. Maria was diagnosed with cancer some months ago.
a2. Zurab and Maria had a fight last night.
a3. (And then) Maria was found dead this morning.

b1. Maria was diagnosed with cancer some months ago.
b2. She was found dead this morning.
b3. (And) Zurab and Maria had a fight last night.

c1. Zurab and Maria had a fight last night.
c2. Maria was found dead this morning.
c3. (And) she had been diagnosed with cancer some months ago.

When the discourse is not properly structured, numerous things go wrong. To ensure correct communication, the interlocutors need to understand how individual clauses relate to each other. Discourse structure is the matrix in which clauses are embedded and which, aided by cue words, permits or blocks implicit inferences. Several discourse phenomena signal discourse structure, including clause juxtapositioning, pronoun and other reference use, quantifier scoping, focus shifts, tense, and aspect.

Determining the interactions due to sentence juxtaposition can be a significant problem. Unfortunately, there are no grammars of paragraph structure, no general linguistic theories of the parts of speech of discourse and inference. But people do assemble sentences into well-structured multisentence texts in a principled way. What principles do they use? How do the principles relate to inferences? What basic elements govern discourse structure?

The key insight for solving these questions is the notion of text coherence. Following [Mann & Thompson 88], we define coherence as follows:

> A discourse is coherent if the hearer knows the communicative role of each portion of it; that is, if the hearer knows how the speaker intends each clause to relate to each other clause.

In other words, a discourse is coherent and will succeed only if it is properly structured: if (i) segments properly reflect communicative intentions, and (ii) interrelationships among segments are properly expressed, enabling the hearer to recognize them, draw the appropriate inferences, and build up the desired structures. Any person or system producing multisentence discourse must therefore confront the problem of discourse structure, which can be posed as a set of questions:

- Since the discourse under discussion is goal-based: How do the speaker's communicative intentions give rise to the discourse?
- Since communication succeeds only if the hearer participates: How can the speaker guide the hearer's inferences? Or: how can the speaker take precautions against undesired inferences?
- Since we are interested in computer-based generation: By what process can a computer plan an effective communication?

All the key notions have now been introduced: text coherence, discourse segments, intersegment relationships, communicative intentions, and hearer inferences.

2.2 Theoretical Antecedents: Descriptions of Discourse Structure

The question of what makes discourse coherent has been studied from several perspectives. Within Computational Linguistics and Natural Language Processing work on monologic discourse[1], two major approaches can be identified: the formalist and the functionalist perspectives. As it turns out, the theories being developed in these two perspectives are largely complementary, and in fact they seem to be converging, hopefully toward a unified model of general (single- and multi-person) discourse.

Following typical *formalist* analyses, such as [Kamp 81], the argument goes as follows: discourse exhibits internal structure, where structural segments encapsulate semantic units that are closely related. Typically, the theories are used to explain pronominalization and quantifier scoping effects. The theories tend to concentrate on the development of formalisms for and formal properties of discourse segments and the discourse structure itself (that is, the "scaffolding" that supports the text), which usually is a tree of some form. The theories tend to be weak on the actual contents of the structure, such as the precise interrelationships between segments and the communicative purposes of the discourse. Some of the more influential formalist work is Discourse Representation

[1]With regard to dialogue, research has focused on cooperative plan-based endeavors such as tutoring and interactive explanation. As a result, many discourse generation ideas are shared with work on plan recognition [Kautz 87, Hobbs et al. 88, Charniak & Shimony 90]. Several research efforts are investigating the nature and role of participants' beliefs and intentions [Pollack 86, Cohen & Levesque 90, Grosz & Sidner 90, Lochbaum 91], and much effort is focused on the types of plans that underlie this type of discourse (see [Litman 85, Lambert & Carberry 91, Ramshaw 91]). Most of these theories postulate several levels of plans, each level handling a distinct phenomenon (discourse management, domain knowledge, etc.).

Theory (DRT) [Kamp 81], and that of [Polanyi 88, Reichman 85, Cohen 83, Heim 83]. Extending beyond dialogue-length discourse, [Van Dijk 72] discusses large-scale text organization and defines the notion of macro-structures and [Rumelhart 72] develops the idea of story grammars.

The *functionalist* argument goes as follows: discourse exhibits internal structure, where the segments are defined by communicative purpose. The theories tend to concentrate on the goals of the speaker and on the ways these goals are reflected in the discourse structure, often as interrelationships between segments (see [Levy 79]). Often, such interrelationships are viewed as reflecting plans of one sort or another which serve the interlocutors' communicative goals. The theories are strong on the particular intersegment relations and their use as operators in planning algorithms; they tend to be weakest on the precise form of the discourse structure. This approach has a fairly long history as well; researchers going back to Aristotle [Aristotle 54] have recognized that in coherent text successive pieces of text are related in a relatively small set of particular ways. Hobbs [Hobbs 78, Hobbs 79] produced a set of relations organized into four categories, which he postulated as the four types of phenomena that occur during communication. Other categorizations of typical intersentential relations were developed by [Grimes 75, Shepherd 26, Dahlgren 88, Mann & Thompson 88, Martin 92], to name a few.

A combination of the formalist and functionalist ideas is embodied in the theory of discourse developed by [Grosz & Sidner 86]. This theory describes a three-way parallel analysis of discourse into the (formalist) segmentation of the utterances, the (functionalist) structure of interlocutor intentions, and the attentional state (an additional record of the referentially available objects).

2.3 Computational Antecedents: Generating Coherent Text

Early computational systems working with multisentence text simply ignored the issue of text structure per se. Generators followed "guided consumption" strategies for deciding what material to include and how to organize it, such as hill-climbing (KDS) [Mann & Moore 81] or proceeding according to the organization of the domain semantics (e.g., TALESPIN [Meehan 76] and PROTEUS [Davey 79]). Early multisentence analyzers either used predefined large-scale knowledge structures that spanned the relevant content of the text, such as scripts (SAM [Cullingford 78], FRUMP [DeJong 79], BORIS [Dyer 83]), or else dynamically built up structures using rules particular to the purpose, such as the argument structure work of [Birnbaum et al. 80] and [Sycara 87].

One of the first text generators that took discourse structure into account explicitly was TEXT [McKeown 85]. The system contained schemas, predefined representations of a stereotypical paragraph structures which acted as templates to mandate the content and order of the clauses in a paragraph; coherence was achieved by the correct nesting and filling-in of a schema. TEXT used four schemas – Identify, Describe, Compare&Contrast, and Attributive — to generate short texts describing various naval objects such as submarines. An example schema is shown in Figure 1. Each schema part is defined in terms of a rhetorical predicate, which specifies what type of material may fill that part by providing semantic attributes the material must contain. Considerable freedom exists within a schema; schemas may nest within others and where permitted portions may be omitted or repeated as necessary to handle the material to be conveyed. This variability was further extended by [Paris 93], who developed methods of switching between schemas depending on their appropriateness to various levels of the hearer's knowledge.

IDENTIFICATION

1. Identification (class & attribute/function)
2. {Analogy/Constituency/Attributive/Renaming}+
3. Particular-illustration/Evidence+
4. {Amplification/Analogy/Attributive}
5. {Particular-illustration/Evidence}

Example:

Eltville (Germany) 1) An important wine village of the Rheingau region. 2) The vineyards make wines that are emphatically of the Rheingau style, 3) with a considerable weight for a white wine. 4) Taubenberg, Sonnenberg and Langenstuck are among vineyards of note.

Figure 1: The IDENTIFICATION schema in TEXT, [McKeown 85].

Though schemas remain a clear and popular method of generating multisentence texts today (see for example [Rambow & Korelsky 92]), their utility is limited because of their essential shortcoming: the lack of representation of the purpose of each part in the whole. Without such information, the system cannot replan any portion of its text in the case that a portion should not communicate successfully, and cannot motivate why it said what it said. This shortcoming is crippling to any system that must be able to assemble its text dynamically and then reason about it, such as interactive explanation generators or documentation generators (see Section 5.1.4).

In order to address this shortcoming, a method of dynamically assembling coherent discourses from basic building blocks had to be developed.

2.4 Planning Text Structure Dynamically

The planning of multisentence paragraphs by computer requires both a sound theory of text organization and an algorithm that can make efficient use of it. For text generation, an influential theory of text structure is Rhetorical Structure Theory (RST) [Mann & Thompson 88, Mann & Thompson 86], which, based on a study involving some hundreds of paragraphs (ranging over advertisements, scientific articles, letters, newspaper texts, and others), postulates that a set of approximately 25 relations suffices to represent the relations that hold within normal English texts. The theory holds that the relations are used recursively, relating ever smaller segments of adjacent text, down to the single clause level; it assumes that a paragraph is only coherent if all its parts can eventually be made to fit into one overarching relation. Most relations have a characteristic English cue word or phrase which informs the hearer how to relate the adjacent clauses; larger blocks of clauses are then related similarly, so that eventually the role played by each clause can be

RELATION NAME: ELABORATION

CONSTRAINTS ON NUCLEUS: none

CONSTRAINTS ON SATELLITE: none

CONSTRAINTS ON THE NUCLEUS AND SATELLITE COMBINATION: The Satellite clause presents additional detail about the situation or some element of subject matter which is presented or inferable from the Nucleus clause in one of the following ways (Nucleus listed first) [set :: member; abstract :: instance; whole :: part; process :: step; object :: attribute; generalization :: specific].

EFFECT: The reader recognizes the situation presented in the Satellite as providing additional detail for the Nucleus.

LOCUS OF THE EFFECT: Nucleus and Satellite

Figure 2: The RST relation ELABORATION, [Mann & Thompson 88].

determined with respect to the whole. Most relations contain two parts, a Nucleus (the major, central material) and a Satellite (the ancillary, qualifying, material). For example, the ELABORATION relation is given in Figure 2.

To address some of the shortcomings of schemas, the author and colleagues have over the last five years carried out an investigation into the compositional planning and generation of multisentential paragraphs. In the first attempt, the author operationalized some relations from Rhetorical Structure Theory as plans and created a text structure planner by simplifying a top-down incremental refinement system patterned on the AI planner NOAH [Sacerdoti 77]. The structurer planned coherent paragraphs in several domains to achieve communicative goals for affecting the hearer's knowledge in some way. It operated after some application program such as a data base or expert system and before the sentence generator Penman [Hovy 90c, Penman 89, Mann & Matthiessen 83]). From the application program, the structure planner accepted one or more communicative goals along with a set of clause-sized input entities that represented the material to be generated. It assembled the input entities into a tree that embodied the paragraph structure, in which nonterminals were RST relations and terminal nodes contained the input material. It then traversed the tree, specifying sentence boundaries and various aspects of syntactic phrasing, and submitted the annotated input entities to Penman to be generated a sentence at a time. A short review of the structure planning process appears in the next section.

This experiment uncovered a set of issues that had to be addressed before a powerful general-purpose theory of the automated production of discourse could be developed. Section 5 describes four major issues, including further studies that were carried out, and outlines remaining work to be done.

3 A First Attempt: Text Structuring Using RST

The first experiment in dynamic text structure planning involved developing a paragraph structure planner and applying it to several domains, including an expert system [Hovy 88], a code development system [Hovy & Arens 91], and a multimodal database information display system [Hovy 90a, Arens et al. 88]. This paper contains examples from the latter, the Integrated Interfaces system, a multimodal presentation program that uses maps, tables, and paragraphs of text to answer users' requests for the display of information from a data base of naval information about ships' deployments. In the example, the display agent furnishes the text structure planner with a set of six related semantic entities, along with a goal to achieve the state in which both the system and the user mutually know about the principal entity (at least; implicit is the planner's freedom to incorporate as many additional entities as mandated by the coherence requirements of its relations). After rewriting the input into a standard form (called here input entities, and shown in Figure 3), the structurer proceeds to plan a paragraph, producing the tree shown diagrammatically.

The hardest task in developing the structurer was understanding how to operationalize RST relations. Simultaneously, they had to enforce coherence by capturing the desired hearer inferences, expressing the speaker's communicative goals, and guiding the planning process. By treating Nucleus and Satellite requirements as semantic preconditions on material to be conveyed and by introducing so-called growth points of subgoals permitted by coherence, RST relations were formulated as relation/plan operators. Since Nucleus and Satellite requirements depended on the hearer's knowledge, and since growth points had to be formulated as structurer subgoals, the plans' effects and requirements were best represented in terms of the communicative intent of the speaker and the beliefs of the interlocutors. Suitable terms for this purpose are provided by the formal theory of rational interaction being developed by, among others, Cohen, Levesque, and Perrault, such as the basic modal operators `BEL` and `BMB` from [Cohen & Levesque 85].

In the structurer's operationalized relations, then, each relation/plan has two primary parts, a *Nucleus* and a *Satellite*, and recursively relates some unit(s) of the input, or another relation (cast as Nucleus), to other unit(s) of the input or another relation (cast as Satellite). A simple relation/plan, SEQUENCE, is shown in Figure 4. The term `(BMB x y P)` stands for *P follows from X's beliefs about what x and y mutually believe*. To admit only properly formed relations, the Nucleus and Satellite fields contain requirements that independently have to be matched by characteristics of the input, and another field contains requirements relating Nucleus and Satellite material. In addition, since the Nucleus and Satellite material is usually expanded upon in typical domain-specific ways (see the discussion in [Conklin & McDonald 82]), possible paths of expansion are contained in *growth points*: collections of goals that suggest the inclusion of additional material in appropriate places in the text. Determining the contents of growth points is a major task; in the example Navy domain, for instance, not only were dozens of paragraphs analyzed, but the Navy expert responsible for producing them was interviewed and taped over a period of three days.

On finding (an) RST relation/plan(s) whose effects include achieving (one of) the system's communicative goal(s), the structure planner searches for input entities that matches the requirements holding for its Nucleus and Satellite. If fulfilled, the planner then considers the growth points of the relation/plan. It tried to achieve each newly activated growth point goal by again searching for appropriate relation/plans and matching their Nucleus and Satellite requirements to the input,

```
        (GOAL (BMB SPEAKER HEARER (SEQUENCE-OF E1 ?NEXT))


((ENROUTE E1)              ((POSITION P1)            ((SHIP K1)
 (ACTOR E1 K1)              (HEADING P1 H1)            (NAME K1 KNOX)
 (DESTINATION E1 S1)        (LATITUDE P1 79)           (READINESS K1 C1))
 (NEXT-ACTION E1 A1)        (LONGITUDE P1 18))        ((PORT S1)
 (LOCATION E1 P1))         ((HEADING H1)               (NAME S1 SASEBO))
((ARRIVE A1)                (COURSE H1 195))          ((DATE T1)
 (ACTOR A1 K1)             ((LOAD L1)                  (DAY T1 24)
 (TIME A1 T1))              (ACTOR L1 K1)              (MONTH T1 4))
 (NEXT-ACTION A1 L1))       (STARTTIME L1 T2)         ((DATE T2)
((READINESS-STATUS C1)      (ENDTIME L1 T3))           (DAY T2 25)
 (NAME C1 C4))                                         (MONTH T2 4))
                                                      ((DATE T3)
                                                       (DAY T3 28)
                                   |                   (MONTH T3 4))
                                   |
                               SEQUENCE
                                /     \
                               /       \
                           CIRC      SEQUENCE
                           /   \        /  \
                       ELAB    ELAB   A1  L1
                       /  \     /  \
                     E1   C1   P1  H1

    Knox, which is C4, is en route to Sasebo. It is at 79N 18E heading
    SSW. It will arrive on 4/24, and will load for four days.
```

Figure 3: Communicative goal and Navy data base assertions provided to the structurer as input (top), resulting paragraph structure tree (left branches of the tree are Nuclei, right branches, Satellites), and corresponding text (bottom).

Figure 4: The RST relation/plan SEQUENCE

```
Name: SEQUENCE

Results:
  ((BMB SPEAKER HEARER (SEQUENCE-OF ?PART ?NEXT)))

Nucleus requirements/subgoals:
  ((BMB SPEAKER HEARER (TOPIC ?PART)))

Satellite requirements/subgoals:
  ((BMB SPEAKER HEARER (TOPIC ?NEXT)))

Nucleus+Satellite requirements/subgoals:
  ((NEXT-ACTION ?PART ?NEXT))

Nucleus growth points:
  ((BMB SPEAKER HEARER (CIRCUMSTANCE-OF ?PART ?CIR))
   (BMB SPEAKER HEARER (ATTRIBUTE-OF ?PART ?VAL))
   (BMB SPEAKER HEARER (PURPOSE-OF ?PART ?PURP)))

Satellite growth points:
  ((BMB SPEAKER HEARER (ATTRIBUTE-OF ?NEXT ?VAL))
   (BMB SPEAKER HEARER (DETAILS-OF ?NEXT ?DETS))
   (BMB SPEAKER HEARER (SEQUENCE-OF ?NEXT ?FOLL)))

Order: (NUCLEUS SATELLITE)
Relation-phrases: ("" "then" "next")
Activation-question:
  "Could ~A be presented as start-point, mid-point, or end-point
   of some succession of items along some dimension? -- that is,
   should the hearer know that ~A is part of a sequence?"
```

The contents of this relation/plan can be paraphrased as follows: The plan, when used successfully, guarantees that both speaker and hearer will mutually believe that the relationship `SEQUENCE-OF` holds between two input entities (that is to say, that one entity follows another in temporal, ordinal, or spatial sequence). That is the contents of the RESULTS field. To ensure proper ordering and focus, one input entity is bound to the variable `?PART` in the NUCLEUS REQUIREMENTS field and the other to the variable `?NEXT` in the SATELLITE REQUIREMENTS field. No other semantic requirements hold on the input entities individually. There is, however, the requirement that they be semantically related by some kind of sequential link (in the current domain, the temporal relation `NEXT-ACTION`), as stated in the NUCLEUS+SATELLITE REQUIREMENTS field; that is, that `?PART` does in fact precede `?NEXT`. Suggestions for including additional input material related to the nucleus are contained in the NUCLEUS GROWTH POINTS field: these call for circumstantially related material (time, location, etc.), attributes (size, color, etc.) and purpose. They are stated in terms of mutual beliefs in order to act as subgoals that the planner must try to achieve. A similar set is associated with the Satellite. The typical order of expression in the text is Nucleus first and the Satellite, using either no cue word, "then", or "next".

recursively, adding successfully instantiated relations to the paragraph tree structure. The planning process bottoms out when either all of the input entities have been incorporated into the tree or no extant goals can be satisfied by the remaining input entities. The tree is then traversed in a depth-first left-to-right manner, adding the relations' characteristic cue words or phrases to the appropriate input entities and appropriate syntactic constraints on realization, and transmitting them to Penman to be generated as English sentences.

This experiment was an early step toward the eventual ability to plan coherent discourse dynamically. Capturing the internal organization and rhetorical dependencies between clauses in the text, the paragraph structure tree enables some powerful reasoning about the text. For example, since it contains the derivation of each part of the paragraph, one knows the role each clause plays with respect to the whole, and thus can identify and repair mistakes. In addition, when the text structure is known, various important syntactic aspects can be determined; note in the example text the following:

- Expression of the Satellites of the ELABORATION relation as relative clauses: `Knox, which is C4...` instead of, say, `Knox is C4. It is en route....` In English, this is the standard realization for the ELABORATION Satellite.

- Use of the future tense in the final sentence. Since information provided by the data base was always based on the present time, anything that appeared in the Satellite of a temporal SEQUENCE relation had to be in the future.

- Linkage of the last two clauses into a single sentence. Deciding to link clauses is easily done when a paragraph structure is available; the complexity of each subtree can readily be determined by counting the number of subnodes, and appropriate sentence-building decisions made.

4 General Requirements for Discourse Structure

As illustrated by Zurab and Maria, successful communication ensues only if the speaker and hearer are aware of the structure of their discourse. However, as mentioned earlier, the nature of discourse structure is still being debated. No existing theory or description, RST included, has enough descriptive power to support all the needs of text planners. Whether formalist or functionalist, each theory addresses some phenomena better than others.

From the rather specific perspective of text planning, however, the descriptions of discourse used by various text planners are quite similar, a fortunate fact that enables one to synthesize a relatively neutral working definition. This common working definition also conforms with the core descriptions of the various theoretical accounts of discourse, despite their other differences.

Surveying the text planning systems of several researchers for a variety of domains (aside from the author's text structurer, EPICURE [Dale 88], the EES text planner [Moore & Swartout 90, Moore 89, Paris 90], TEXPLAN [Maybury 90], EDGE [Cawsey 90], SPOKESMAN [Meteer 90], PIT [Kreyss & Novak 90], POPEL [Reithinger 91], JOYCE [Rambow & Korelsky 92] and others) and taking as far as possible into account the theoretical work of [Grosz & Sidner 86, Asher 92, Polanyi 88] and the work on intention recognition [Allen & Perrault 80, Litman 85, Pollack 86, Lambert & Carberry 91], the following general assertions about the structure of plan-based English discourse can be formulated:

1. **Discourse:** A discourse (a text) is a structured collected of clauses. By their semantic relatedness, clauses are grouped into segments; the discourse structure is expressed by the nesting of segments within each other according to specific relationships. A discourse can thus be represented as a tree structure, in which each node of the tree governs the segment (subtree) beneath it. At the top level, the discourse is governed by a single root node; at the leaves, the basic segments are single grammatical clauses.

2. **Purpose:** Each discourse segment has an associated purpose, which (following [Grosz & Sidner 86]) we call the Discourse Segment Purpose (DSP) and represent at each node of the tree. Each DSP is a communicative goal of the speaker. In a successful discourse, the contents of each segment achieve its DSP. Each segment can thus be seen as a step in a plan to achieve the overall communicative purpose of the discourse.

3. **Coherence:** A discourse is only communicatively successful if it is mutually coherent, i.e., if the speaker's and hearer's beliefs agree about how each segment relates to its neighbors (and thus to the whole). Coherence is enforced by the constraints of intersegment discourse structure relations, which are discussed in Section 5.2.

4. **Discourse segment:** A discourse segment S is represented by a tuple (*name, purpose, content*), where:
   - The *name* is a unique identifier for the segment.
   - The *purpose* is one or more communicative goals the speaker has with respect to the hearer's mental state (the DSP)
   - The *content* is either:
     - an ordered list of discourse segments, together with one or more intersegment discourse relations that hold between them (either there is a relation between every two adjacent segments in the list, or a relation holds among all the segments in the list simultaneously); or
     - a single discourse segment; or
     - the semantic material to be communicated (usually statable as a single clause in English). This material often takes the form of a set of knowledge base assertions or data base facts.

5. **Discourse structure:** A discourse structure D is a discourse segment which is not contained in any discourse segment and all of whose leaves (the innermost segments) contain semantic material to be communicated. It is the matrix in which clauses are embedded which permits or blocks implicit inferences.

In most computational applications, the discourse is a tree; this is of course not the general case, since discourses may include interruptions and other discontinuities. The RST based paragraph trees of the first and subsequent applications (Sections 3, 5.4, 6.3) can be reformulated to conform to this definition by the addition of explicit communicative goals to each relation's branch (i.e., to each segment); for presentational clarity, however, this has not been done in this paper.

5 Four Central Aspects of Discourse Structure Relations

The initial attempt with RST based discourse structure planning established that is is possible to dynamically construct coherent paragraph-length discourses in a variety of domains using RST and similar relations as plan operators. Simultaneously, it opened up a set of issues that had to be addressed before robust discourse planning and generation could become a reality, illustrating the effects that discourse structure relations can have on a wide range of phenomena, from tense and aspect selection to focus and theme development. In this section we describe four major aspects of text planning, all repeatedly found in the text generation literature, which centrally involve discourse structure relations:

1. Text plans — content and format: The operationalized RST relations themselves were quickly found inadequate, especially in their inability to capture communicative intent. Text planners switched to using a new kind of plan, one keyed on intentionality.

2. A collection of relations: Intersegment discourse relations are still however required to structure the discourse. An ongoing effort to collect and taxonomize a core corpus of relations is described.

3. Predefined structures (schemas): In spite of the utility of text plans and discourse relations, predefined structures remain necessary to control the combinatorics of longer texts.

4. Controlling planning by focus shift: Being able to juxtapose clauses coherently did not mean being able to make them flow successfully. Discourse relations and focus shift rules work together to co-constrain the possibilities.

Though these issues have been addressed in subsequent studies by the author and others, none have been fully resolved. Taken together, however, the current state of text planning work represents a significant advance over what was known about the automated planning and generation of discourse five years ago.

5.1 Text Plans: Content and Format

Since the first attempts with RST-based text structure planning, the nature of text plans has been an issue. What kinds of plans are needed to generate coherent text? How do they relate to discourse structure relations? Text planning is evolving its own types of plans and its own brand of planning.

5.1.1 AI Planners and Text Planners

By the standards of the most advanced AI research planners today, text planners are not very sophisticated. To perform their two major functions of content selection and organization, most of them use a variant of the basic top-down successive refinement algorithm such as employed in, say, NOAH [Sacerdoti 77], without employing critics. The input goals contain the instruction to communicate some central portion(s) of information, and the final low-level actions are direct calls to a sentence generator, in appropriate order. The resulting plan serves to act simultaneously as discourse structure and as the plan for achieving the desired communicative goal(s). Thus the discourse is simultaneously a linguistic construct and a plan of action. The equivalent of preservation goals, such as the (not hierarchically decomposable) goals to imbue the text with specific stylistic qualities in order to achieve pragmatic effects such as clarity and formality, are not yet handled by text planners (although see the work of [Green 92, Hovy 88]).

5.1.2 Text Plans vs. Intersegment Relations

The dual nature of the planner's output — simultaneously a communication plan and a linguistic structure — has led to much confusion and remains unresolved. In Rhetorical Structure Theory, relations are structural entities that reflect underlying semantic and interpersonal relationships between the discourse segments. In the RST structurer, the relations *themselves* were viewed as plans — the operators that guided the planner's search through the space of inputs. The structurer's goals were all directly related to its relations, thereby limiting it to a "rhetorical" goal language, planning to achieve goals such as "create an elaboration between the current material and some additional material" (see for example the goal in Figure 3). A similar line of argument can be found in [Levy 79]. Later work [Moore & Swartout 90, Moore & Paris 91] argued that using discourse structure relations as goals erroneously conflates "rhetorical" (i.e., structural) information with intentionality. Using RST-like relation/plans to control the selection of material is, they claimed, artificial; more natural is to select material on the basis of communicative intentions. Therefore, as described in the next section, Moore, Paris, and Swartout developed a set of text plans they considered "intentional", such as the plan RECOMMEND, which decomposes into a set of user actions appropriate to some task. These plans were utilized by the same style of hierarchical decomposition planner as the RST structurer.

As a result of these claims, several questions arose in the research community: What information should a text planner properly use? What information should appear at the branch points of a discourse structure? Is there a real difference between "intentional" and "discourse-structural" information? If so, are both types needed, and how do they interact?

These questions are still being debated; see [Moore & Pollack 93]. Neither the initial RST-based approach nor the later experiments are wholly satisfactory in this regard. Certainly, for the selection and overall organization of material in plan-based hierarchicalizable discourse, text plans should somehow express the speaker's communicative intentions. But for several other aspects of text construction, as for example described in Sections 5.4, 6.1, 6.2, and 6.3, practical experience has shown the need for linguistically attuned structural information of the kind embodied in RST.

Unfortunately, no-one has succeeded in outlining precisely what makes a text plan intentional or not. The difference lies neither in the role played during planning — hierarchical decomposition occurs with both types of plan — nor in the role played within the discourse structure — a branch node governing subportions of the discourse. To the extent that a difference does exist, however, the dilemma is resolved when one recognizes that the two types of object — intentional plans and discourse relations — perform different functions and hence are both needed *simultaneously* to govern discourse. To determine what material to include and to provide the overall structure of the discourse, intentional plans are most appropriate; within this framework, it is the function of discourse relations to ensure textual coherence, prevent unintended inferences, govern sentence formation, tense, pronominalization, and focus shift, as described in subsequent sections of this paper. To see this, note that the same communicative purpose can be achieved in many ways; for example, the (intentional) goal to PROVE clause (1) can be achieved using several (discourse) relations with clause (2):

> CAUSE: "(1) *He knows how to deal with red tape because* (2) *he lives in Moscow.*"
> CIRCUMSTANCE-LOCATION: "*Living in Moscow, he knows how to deal with red tape.*"
> SEQUENCE-TIME: "*After he went to live in Moscow, he knew how to deal with red tape.*"

In general, some text genres tend to be more intentional, such as explanatory discourse, while others tend to be more structural, such as encyclopedia entries (for a discussion, see [Maier & Hovy 91]). In the former, almost every clause is governed by a separate intention, while in the latter, large portions of the text serve a single discourse intention (often, DESCRIBE) and are organized under a considerable tree of discourse structure relations. Texts generated by TEXT [McKeown 85] and the RST structurer are both of this type. Texts generated by PEA [Moore 89] and TEXPLAN [Maybury 90] are explanations, with a rich subgoal structure. To accommodate both types, the definition of discourse segments in Section 4 associates both intentions and structural relations with each discourse segment.

Differentiating the two types of object into intentional plans and structural relations may correspond with the distinction made in [Austin 65] between sentences with perlocutionary effect, such as persuading or motivating, and those with illocutionary effect, such as elaborating, identifying, or describing, though, as Maybury's attempt to do so shows, this distinction is unfortunately hampered by the vagueness of the notions of perlocution and illocution and imprecision in plans' and relations' definitions [Maybury 90].

5.1.3 Text Plan Formalism

The contents and formalism of text plans have evolved in many ways since the early RST structurer's relation/plans of Figure 2. Moore, Paris, and Swartout defined for PEA, the text planner of the Explainable Expert System (EES), plans that included, in addition to Effect, Nucleus, and Satellite fields, also a field for constraints — the facts (within the system's knowledge base or user model) that had to be true about the data before the plan could be applied. In addition, they annotated Satellite subgoals as mandatory or optional. The same formalism was used by Reithinger [Reithinger 91]. Maybury further elaborated text plans, adding preconditions of two kinds, essential and desirable.

The effectiveness of these additions to the basic plan format is discussed at length in [Moore 89, Maybury 90]. Based on the above work, as well as on the EDGE planner [Cawsey 90], the planners of [Kreyss & Novak 90] and [Rösner & Stede 92], and the more structurally oriented text representation in SPOKESMAN [Meteer 90], one can define a text plan P as a tuple (*name, effects, constraints, preconditions, decomposition*), where:

- The *name* is a unique identifier of the segment.

- The *effects* are one or more communicative goals that the plan achieves, if properly executed. Since these goals pertain to the speaker's desire with respect to the hearer's state of knowledge, opinion, goals, and similar structures, they are phrased in terms of the hearer's mental state.
- The *constraints* are facts in the knowledge base or the user model that must hold before the plan may be used.
- The *preconditions* are facts in the knowledge base or user model that should hold for felicitous communication. If they are violated, the hearer may be confused. As mentioned above, the planner in a dialogue situation may be given the ability to ignore the preconditions, trusting the hearer to request help when communication fails; in such cases, the planner should mark the affected preconditions to facilitate repair.
- The *decomposition* is an ordered list of subgoals to be achieved. Each subgoal may be flagged as optional, in which case the planner can ignore it under appropriate conditions, depending on the planner's sophistication: at the minimum, it can simply ignore the subgoal if instructed to produce terse text; being more sophisticated, it may reason about various contributing factors, such as the balance of material within the discourse structure so far or the level of detail of the indicated material). The order of subgoal segments within this list must respect the coherence requirements of discourse structure relations. Subgoals are generally of two types:
  - communicative intentions on portions of knowledge base contents, which can be achieved by other text plans (for example, a PERSUADE may call for a MOTIVATE or a DESCRIBE), and
  - "primitive" Speech Acts on clause-sized knowledge base entities, such as INFORM, ASK, and ORDER, which are achieved by the sentence generator.

An example of Maybury's plan formalism appears in Figure 5; note that the subgoals in the DECOMPOSITION field are ordered and, unless explicitly flagged, mandatory, and that planning proceeds along the HEADER fields, not the EFFECTS — that is, subgoals are achieved by plans whose HEADER fields match; the EFFECTS are simply for updating the hearer model).

5.1.4 Example Text Plans

The Explainable Expert System text planner [Moore 89] is an advanced attempt at text planning with backtracking, using a partial hearer model and marking in the discourse structure all assumptions made about the hearer's knowledge. The plan library of EES contains almost 100 plans at various levels of detail, all supporting the informative actions one needs to explain the behavior and data of expert systems (a full list appears in an appendix of [Moore 89]). Judging by name and content, these plans range from intentional (including for example INFORM, RECOMMEND, INFORM-AND-PERSUADE, PERSUADE-BY-MOTIVATION) to structural, RST-like (including SEQUENCE-STEPS, CONTRAST, ELABORATE-OBJECT-ATTRIBUTE). Two example EES text plans appear in the boxes in Figure 6, together with a discourse fragment in which they are used. In

NAME:	Extended-description
HEADER:	Describe(S, H, entity)
CONSTRAINTS:	Entity?(entity)
PRECONDITIONS	
ESSENTIAL:	KNOW-ABOUT(S, entity) $\wedge$ WANT(S, KNOW-ABOUT(H, entity))
DESIRABLE:	$\neg$ KNOW-ABOUT(H, entity)
EFFECTS:	KNOW-ABOUT(H, entity)
DECOMPOSITION:	Define(S, H, entity) *optional*(Detail(S, H, entity)) *optional*(Divide(S, H, entity)) *optional*((Illustrate(S, H, entity)) $\vee$ Give-Analogy(S, H, entity))

(S and H stand for Speaker and Hearer respectively. Describe, Define, Detail, Divide, Illustrate, and Give-analogy are communicative intentions.)

Figure 5: Text plan *Extended-Description* from [Maybury 90].

the example, the goal to persuade the user is matched by the effect of the plan PERSUADE-BY-MOTIVATION; since its constraints are met, its Nucleus goal is posted on the discourse structure. This goal is in turn matched by several plans, including MOTIVATE-ACT-BY-MEANS, whose constraints are satisfied, and whose Nucleus and Satellite subgoals are consequently posted. The Nucleus subgoal, being an INFORM, is directly achievable by the sentence generator, which produces the sentence shown; the Satellite subgoal (means) is matched by a MEANS plan, which causes the generation of the cue word "by" and eventually gives rise to further text.

5.2 A Library of Discourse Structure Relations

5.2.1 The Problem: Which Relations? How Many?

Given the evident need for discourse structure relations, one of the central problems confronting discourse work is the construction of a core library of such relations, defined in a general enough way to be of common use. Since that they have variously been described as essentially intentional, structural, semantic, "rhetorical", and have been variously estimated at maximally two in number and estimated to number in the tens of thousands, this is not a straightforward task.

At the heart of the problem is their intended use. Is it better to think of relations as basic tree-building operators (for which one needs only two, Dominate and Precede), as resembling closed-class syntactic classes (i.e., mirroring Subject, Direct and Indirect Object), as open-class

```
NAME:          PERSUADE-BY-MOTIVATION
EFFECT:        (PERSUADED H (GOAL H (DO H ?act)))
CONSTRAINTS:   (AND (GOAL S ?g)
                    (GOAL H ?g)
                    (STEP ?act ?g))
NUCLEUS:       (FORALL ?g (MOTIVATION ?act ?g)))
SATELLITES:    ()
```

```
NAME:          MOTIVATE-ACT-BY-MEANS
EFFECT:        (MOTIVATION ?act ?goal)
CONSTRAINTS:   (AND (GOAL S ?goal)
                    (GOAL H ?goal)
                    (STEP ?act ?goal))
NUCLEUS:       (INFORM S H (GOAL S ?goal))
SATELLITES:    ((MEANS ?goal ?act))))
```

(PERSUADED USER (GOAL USER (DO USER REPLACE-1)))

(MOTIVATION REPLACE-1 ENHANCE-READABILITY)

(INFORM SYSTEM USER ENHANCE-READABILITY)
"I'm trying to enhance the readability of the program"

"by"
(MEANS REPLACE-1 ENHANCE-READABILITY)

(INFORM SYSTEM USER APPLY-1)
"applying transformations that enhance readability"

(BEL USER (STEP REPLACE-1 APPLY-1))

Figure 6: Example text plans and discourse structure fragment from the EES planner, [Moore 89].

semantic relations (embodying all the possible semantic relations), or as something somewhat more limited (mirroring semantic case relations such as Agent, Patient, and Beneficiary)?

In available attempts at listing relations, the intended purpose determines the nature and number identified. Approaching the problem of discourse structure from several intellectual subfields, various researchers have produced somewhat more extensive lists of intersegment relations — from philosophers (e.g., [Toulmin 58]) to linguists (e.g., [Quirk & Greenbaum 73, Halliday 85, Martin 92]) to computational linguists (e.g., [Hobbs 79, Mann & Thompson 88]) to Artificial Intelligence researchers (e.g., [Schank & Abelson 77, Dahlgren 88]). Typically, their lists contain between five and fifty relations, and they argue that (at least) tens of interclausal relations are required to describe the structure of English discourse; one can call this the *Profligate Position.*

On the other hand, some researchers, (e.g., [Grosz & Sidner 86, Polanyi 88, Kamp 81]) prefer not to identify a specific set of such relations. They argue that trying to identify the "correct" set is a doomed enterprise, because there *is* no closed set; the closer you examine intersegment relationships, the more variability you encounter, until you find yourself on the slippery slope toward the full complexity of semantics proper. Though they do not disagree with using relationships between adjacent text segments to provide meaning and enforce coherence, they object to the notion that some small set of relations describe English discourse adequately. As a counterproposal, Grosz and Sidner define two basic relations, DOMINANCE and SATISFACTION-PRECEDENCE, which carry intentional (that is, goal-oriented, plan-based) but no semantic import, and suffice to represent tree-like nature of discourse structure. One can call this the *Parsimonious Position.*

5.2.2 Collecting and Taxonomizing the Relations

While the parsimonious relations may satisfactorily represent discourse structure for purposes of analysis, practical text generation experience, such as [McKeown 85, Hovy 88, Moore & Swartout 90, Paris 90, Rankin 89, Cawsey 90, Maybury 90, Dobeš & Novak 91], has shown that they are insufficient and that planners need considerably more information of rhetorical and semantic nature to ensure successful communication. For example, when generating the following two clauses

"His car was much admired because it was a red Ferrari."

the speaker needs to know which semantic interrelationship to express: it is the semantic relation of causality that provides the appropriate linking word and much of the structural/realizational information (had the interclausal relationship been temporal coincidence, the cue word would have been "when"; had it been elaboration, the second clause would have been subordinated to the first in a relative clause "His car, which was...", and so on).

Accordingly, in 1989 the author started collecting intersegment relations that are expressive enough to satisfy the requirements of text planning systems while avoiding an unbounded ad hoc collection of semantic relations. Over 350 such relations from approximately 30 researchers in various fields were collected and taxonomized; see [Hovy 90b]. Subsequently, in joint work, over 50 additional relations in other sources were found and an improved taxonomization, consisting of about 70 relations, was produced. A new text planner constructed at USC/ISI and its partner institute IPSI in Germany contains three taxonomies of approximately 120 relations [Hovy et al. 92].

The core set of relations, organized into a taxonomy, are reproduced in the Appendix; the sources, definitions, and taxonomization procedure is described in more detail in [Hovy & Maier 92].

Given the semantic overlaps of many of the relations, a natural taxonomy suggested itself, in which one dimension is constrained in the number of relations and the other unconstrained (the more a relation is specified to distinguish it from others, the more its semantics are enhanced, and the lower it appears in the hierarchy). Though the unboundedness at the bottom places one on the slippery slope toward having to deal with the full complexity of semantic meaning, there is no reason to fear such complexity. The terms are well-behaved and subject to a pattern of organization which makes them manageable: all the pertinent information about discoursal behavior is captured near the top; each relation inherits from its ancestors all necessary processing information, such as cue words and realization constraints, and adds its unique peculiarities, to be used for inference (in parsing) or for planning out a discourse (in generation). Increasing differentiation of relations, continued until the very finest nuances of meaning are separately represented, need be pursued only to the extent required for any given application.

The top-level differentiation of relations into three basic kinds (see Figure 12) is motivated on linguistic and semantic grounds. As discussed in [Halliday 85], two clauses can be related in at most three different ways simultaneously — semantically, interpersonally, and presentationally (what Halliday calls the metafunctions of language: ideational, interpersonal, and textual):

1. *Well* (presentational), *frankly* (interpersonal), *earlier* (semantic) *I had a wonderful time....*
2. *Fortunately* (interpersonal), *second* (presentational), *it seems that....*
3. *Consequently* (semantic), *in conclusion* (presentational), *we see that....*

Discourse structure relations exist for each of these three classes (though frequent linkages can cause confusion; for example, TemporalSequence which is semantic and PresentationalSequence which is presentational are both cued by the words "first", "second", "finally", etc.). A discourse segment representation must be able to maintain three intersegment relations simultaneously. A similar partitioning of discourse relations is discussed in [Mann & Thompson 88].

Of course, there is no guarantee that the relations collected are indeed the "right" and only ones. Their strongest support is that they are the amalgamation and synthesis of the efforts and proposed terms of several investigations in different fields, including actual attempts to construct working text planners and discourse analyzers. When different interclausal relations are proposed, we expect that the hierarchy will grow primarily at the bottom, and that the ratio of the number of relations added at one level to the number of relations added at the next lower level will be low, for all levels. This accords with our experience when compiling the hierarchy: halfway through this study, the topmost tiers had essentially been established, and almost all new relations found were simply specializations of existing ones.

We are continuing the collection and taxonomizing of relations, as well as collecting precise, formal definitions for them, such as those of [Ivir et al. 80, Hobbs 79, Hobbs 90, Sanders et al. 92, Martin 92, Lascarides & Asher 91].

5.3 Schemas

In planning, scripts or macro-operators are useful compilations of plan structures formed out of oft-repeated stereotyped plans [Fikes et al. 72, Schank & Abelson 77]. Similarly, fossilized discourse structures that represent formulaic texts (such as encyclopedia entries and business reports) are called schemas [McKeown 85].

From several attempts to plan longer texts, it became clear that systems without some explicit representation of the structure of longer spans of text than single paragraphs are not feasible in practise. There is simply too much variability in text plans or discourse structure relations; as plan and relation libraries grow, the number of possible texts grows alarmingly (as one would expect, given the plasticity of language). So, as argued in for example [McKeown 85, Mann 87, Rambow 90, Mooney et al. 90], one should capture the idiosyncratic regularities of discourse structure, which may depend on genre, domain, or even simply custom, in schemas and use them as frozen plans by simple schema instantiation. Where additional structuring is required — when no frozen plan exists to achieve the communicative intention — then discourse structure plans and intersegment relations can be used.

When using a schema, one foregoes the ability to reason about the function and interrelation of each portion of the text. One can however replace some of this information back into schemas, essentially formulating them as fossilized discourse structures, thereby gaining a homogeneity of representation with text plans that simplifies the planning process. Since both schemas and text plans specify the nature and order of the material to be communicated, it is possible to view text plans operationally as mini-schemas. One way of unifying the representation of text plans and schemas was outlined in [Hovy 90a]. By treating any text structuring operator — schema or text plan — as an ordered list of mandatory communicative subgoals, the effect is that of a schema. The planner simply constructs a portion of the discourse for each subgoal without reasoning about the interrelatedness of portions. By instead treating the subgoals as a list of suggested possible communications, the effect is that of planning using text relation/plans. The planner must perform additional reasoning to determine why the material satisfying various subgoals should be included and how it relates overall to ensure textual coherence. Thus, as shown in [Hovy 90a], by treating the growth point goals in RST relation/plans as *injunctions* that specify the type and order of additional material to include, rather than as *suggestions* to do so, a text plan acts as a schema. Of course, some growth point goals can be made required and others optional, enabling plans simultaneously to incorporate both fixed structural options whose relationship with the remainder is not explicitly specified (i.e., act as schemas), as well as inferentially motivated patterns that are developed dynamically. This treatment has been adopted in some form or another by most text structure planners and some schema appliers; the schema planner TEXT [McKeown 85] and the EES and TEXPLAN planners, for example, label some subgoals optional. This hybrid approach combines the complementary strengths of schemas and plans.

Several open issues remain. As yet no representation for schemas captures well the underlying semantic and rhetorical interrelations of the parts. Also, when growth point goals are treated as suggestions for additional growth, two problems are immediately introduced: Which growth point goals should be considered? And in what order should new growths be added to the discourse? It is easy to think of criteria for controlling the inclusion, but difficult to formalize them adequately;

for some candidates see [Hovy 90a]. One criterion, however, has been studied to some degree. This is the effect of focus shift on discourse structure.

5.4 Focus Shift

In any plan of action, the sequence of steps may be fixed or not, depending on the underlying interrelationships among their contents. As illustrated by NOAH, ordering requirements cannot all be precompiled into plans, and some additional process has to exercise additional control.

The position is the same in text planning, when the order of relations/plans' subportions is free. To ensure coherence, and to direct the reader's inferential attention, the material must be developed in an appropriate order and with appropriate signals. An important ordering consideration is focus, which we define as the locus of the principal inferential effort needed to understand the text[2]. Consider the example texts and corresponding RST discourse structure in Figure 7. The three ELABORATIONS providing the balloon's features — its color, size, and heat-reflecting ring — are joined by a JOIN relation, which is defined in RST to be multinuclear and thus imposes no order on its parts. As illustrated in the texts, however, the three parts are not interchangeable; text (2) is more connected since it places the two clauses about color together (and in fact these two clauses could well have been been conjoined using "and").

Linguistic and computational investigations reveal strong constraints on what material may occupy the focus position as a text progresses. Three so-called focus shift rules expressing these constraints were formulated by Sidner [Sidner 83] (see also [Grosz 77, Grosz 81]). These rules are however not sensitive to discourse structure, and when used for text generation more specific rules are needed. For the TEXT generator, for example, McKeown had to add an additional focus shift rule [McKeown 85]. Later, McCoy and Cheng generalized the linear operation of focus shift rules using a construct called a Focus Tree, which represents a focused concept at each node with as its branches all possible topic continuations [McCoy & Cheng 88, McCoy 85].

In an attempt to overcome the underdetermination of RST discourse structures (such as the text variations allowed the tree in Figure 7), the author and Prof. Kathleen McCoy from the University of Delaware described the parallel use of Focus Trees and RST discourse structures to co-constrain the order of clauses [Hovy & McCoy 89]. In this approach, the text structure planner constructs an RST paragraph structure and a Focus Tree in tandem. During the expansion of a node in the RST discourse structure, the structurer disregards questions of ordering the growth point subgoals and simply tests *all* the growth point goals active at that node, collecting all the potential candidate relations and their associated clause-sized input entities that can be included at that node in the discourse structure. Each candidate relation is then checked against the currently allowed focus shifts in the Focus Tree and invalid candidates are simply removed from consideration. Thus the underdeterminedness introduced by not specifying the order of communicative subgoals in the relation/plan is handled by the specifications of focus shift. However, though this procedure can help

[2] See [Hovy & Lavid 92]. Severe terminological confusion surrounds the issue of focus, theme, and given; we take focus here in the sense of the Prague School [Daneš 74] and [Halliday 67, Fries 81] to mean a privileged element of the clause that usually appears in its latter, high-informational, portion. It is closely related, but not identical to, the notion of New [Prince 81].

```
                |
                |
           ELABORATION
             /       \
            /         JOIN
     balloon      /   |   \
              red     |    silver
                     CAUSE
                     /   \
                 carry   large
```

(JOIN relation is multinuclear; no order of branches is implied.)

1. At last John saw the balloon. It was bright red. Because the balloon was designed to carry people, it was large. It had a silver circle at the top to reflect heat.
2. At last John saw the balloon. It was bright red. It had a silver circle at the top to reflect heat. Because the balloon was designed to carry people, it was large.

Figure 7: RST structure and two possible texts. Example adapted from [McKeown 85].

significantly to prune the search space, occasionally it can be too powerful, prohibiting any further paragraph structuring when no allowable focus moves remain. In such cases it is sometimes possible to invert the current RST relation's default order, thereby producing a thematically marked but still coherent and well-focused text. For example, in Figure 8, paragraph structure (a) is allowed by RST constraints by simply adding the ELABORATION relation before the CIRCUMSTANCE in the leftmost branch. However, since the material in `C1`, the ELABORATION Satellite, is semantically directly related to a portion of `E1`, the Focus Tree requires that the `C1` clause be generated contiguously with the `E1` clause. To avoid failure, the RST structure is made acceptable to the Focus Tree criterion by inverting the ELABORATION relation, reordering the `C1` clause to precede the `E1` clause. According to RST, an inverted ELABORATION relation is possible but must be linguistically marked, and the resulting text, with a marked dependent clause, is shown as paragraph (b).

6 Three Text Planning Tasks Involving Discourse Structure Relations

The previous section described four central aspects of the nature of discourse structure relations. This section describes three distinct text planning tasks in which discourse structure relations play a role:

1. Casting of syntactic roles: An important sentence-level planning task is the assignment of material to syntactic classes within a sentence.

```
            (a)                              (b)
             |                                |
            SEQ                              SEQ
           /   \                            /   \
       ELAB     SEQ                   ELAB-1     SEQ
      /   \    /   \                  /   \     /   \
   CIRC   C1  A1   L1               C1  CIRC   A1   L1
  /    \                               /    \
E1    ELAB                           E1     ELAB
      /  \                                  /  \
    P1    H1                              P1    H1
```

```
(a) Knox is en route to Sasebo.  It is at 79N 18E heading SSW.  It
    is C4.  It will arrive on 4/24, and will load for four days.

(b) With readiness C4, Knox is en route to Sasebo.  It is at 79N 18E
    heading SSW.  It will arrive on 4/24 and will load for four days.
```

Figure 8: (a) Another version of the Navy text, treating growth points in free order, and (b) using Focus Trees during structure planning to ensure proper focus shifts.

2. Concept aggregation: Another planning task involving discourse relations is the compacting of text by aggregation.

3. Text formatting: Several discourse structure relations achieve their communicative purposes presentationally using text formatting devices such as itemized lists, headings, and footnotes.

6.1 Discourse Relations and the Casting of Syntactic Roles

When the constraints imposed by content to be communicated, discourse structure, and focus, are merged together during planning, the text begins to take shape. However, its final form is still not fully specified. One of the major remaining tasks is the scoping of information into sentence components and the subsequent assignment of such units to syntactic classes. For example, the final SEQUENCE segment in Figure 3 has at least the following realizational alternatives:

(a). `It will arrive on 4/24 and will load for 4 days.`
(b). `It will arrive on 4/24.  It will load for 4 days.`
(c). `After arriving on 4/24, it will load for 4 days.`

and, on the noun phrase level, the first ELABORATION relation has at least:

(d). `Knox, which is C4, is en route.`
(e). `Knox is en route and it is C4.`
(f). `Knox is en route.  It is C4.`

How to plan the sentence? How even to know when realizational alternatives exist, without performing some grammar-based inspection of the material to be generated? Beyond focus, any solution must take several additional issues into account, including the complexity of the remainder of the discourse substructure, the desired overall style of the text (such as a general preference for simple or complex sentences), the rhythm of sentences (long alternating with short, as suggested in numerous books on good style, such as [Shepherd 26]).

Although much more research remains to be done on this problem, intersegment discourse relations provide certain amount of help, either by indicating where alternatives of realization exist or by suggesting candidate syntactic realization forms. Situations in which different sentence scopings exist can often be recognized by characteristic configurations of the discourse structure. The ELABORATION relation provides a simple example: Since it always holds between a clause constituent (such as the actor of a process) and another clause (some attribute of the actor), the Satellite (the attribute) can be realized as a relative clause to the Nucleus (the process containing the constituent), as long as the Nucleus is not itself a subtree in the discourse. In fact, this is the standard realization in English.

A study by Scott and de Souza [Scott & De Souza 90, De Souza et al. 89] of the use of several RST relations in both English and Brazilian Portuguese proposed a set of heuristics to govern sentence formation, including:

1. A Satellite can only be embedded in its Nucleus.
2. Embedding can be realized as an adjective, appositive NP, PP, or relative clause, in this order of preference.

3. Embedding can occur in the leftmost nuclear clause with the same focus value.
4. Satellites in a JOIN within an ELABORATION should be embedded, provided there are no, or else more than one, remaining clauses.
5. Coordination occurs only between elements of JOIN, SEQUENCE, and CONTRAST relations.
6. The more shared parameters between clauses, the more they should be coordinated.
7. Prefer coordinating NPs over PPs over Vs or VPs.
8. Sentences should contain no more than 3 clauses.
9. Sentences should contain at most one level of embedding.
10. Embedding should occur before coordination and before focus transformations.

Forms of some of these heuristics have been implemented in several text planners.

Within noun phrases, the problem of delimiting and organizing content involves three major issues. The first issue relates to pronominalization. It is widely accepted that pronominalization is sensitive to segmental boundaries, at least on the relatively major level; see for example [Björklund & Virtanen 89], or the analyses of conversations by Passoneau, which suggest that discourse referents are available for pronominalization in the local context only [Passoneau 91]. Studies by [Levy 84, Marslen-Wilson et al. 82] indicate that explicit referring expressions (say, a full noun phrase instead of a pronoun) help indicate discourse segment boundaries. The availability of the discourse structure as a tree of intersegment relations, in which segments manifest themselves as subtrees, enables the development of sophisticated pronominalization strategies. Exactly which segment boundaries permit pronominalization, however, remains an open question.

A related case occurs when material in a dependent clause can be realized instead within the noun phrase proper (as an adjective, say). Again from Figure 3, "Knox, which is C4,..." could have been realized as "the C4 Knox..."; in Figure 8, we deemed the clause-sized "Being C4, Knox..." (which was realized by default) unacceptable, preferring the realization "With readiness C4, Knox...". Determining the optimal syntactic class of material depends, among other things, on the balance of the paragraph structure tree, on focus, and on the stylistically desired density of information in the noun phrase.

6.2 Aggregation Guided by Discourse Relations

An important sentence-level planning task involves the compacting of material to be communicated. Often, the detailed representations used within data bases and expert systems result in redundant or verbose text unless some kind of aggregative planning takes place. Aggregation uses the fact that information units, represented by the domain system as separate individuals, are often generated in the text as a group sharing pertinent features, and can therefore be abbreviated. For example, the Integrated Interface data base represented each ship separately, but could decide to display several ships moving together. Without rules for syntactically grouping the ships into a single clause or portion of a clause, the text was of poor quality:

```
MEKAR-87 takes place in the South China Sea from 10/20 until 11/13.
Knox, Fanning, and Whipple are participating.  Knox arrives on 10/20.
It leaves on 10/31. Fanning arrives on 10/20.  It leaves on 11/13.
Whipple arrives on 10/29.  It leaves on 11/13.
```

It is easy to invent aggregation rules to improve the text. It turns out, however, that by formulating some rules in terms of discourse structure one can significantly reduce the complexity of the aggregation process. If aggregation is performed without discourse structure structure planning, the aggregator has to inspect every pair of input elements for each aggregation rule it has, an order n^2 operation per rule for n elements, while if aggregation is performed after structuring, the aggregator need only inspect the pairs of elements within the discourse segments that directly contain the material to be generated, a reduction to (typically) two or three elements. In the example, the paragraph structure involves three parallel ELABORATION relations; see Figure 9(a). To improve this text, the following three aggregation rules were applied:

1. *If two instances of the same RST relation emanate from a single Nucleus, then merge the two instances into one relation, and merge their Satellites into the same leaf node* — see Figure 9(b).

2. *If several instances of the same RST relation appear in a* JOIN*, then promote the relation, and* JOIN *the respective Nuclei and Satellites together* — see Figure 9(c).

3. *If input elements A and B within the same leaf node of the discourse structure contain the same action, the same ending date or time, and the same location, and they contain different actors, then merge the elements* — see Figure 9(d).

The result generated was:

```
MEKAR-87 takes place in the South China Sea from 10/20 until 11/13.
Knox, Fanning, and Whipple are participating.  Knox and Fanning
arrive on 10/20.  Whipple arrives on 10/29.  Knox leaves on 10/31.
Fanning and Whipple leave on 11/13.
```

The general problem of aggregation for fluent text involves many non-structural issues as well; see for example [Van Dijk & Kintsch 83, Hovy 87, Dale 88]). But having access to the discourse structure enables one to begin addressing this problem in a realistic way; see [Horacek 92, Dalianis & Hovy 93].

6.3 Discourse Relations and Text Formatting

This section describes a preliminary study that illustrates how, with suitable extensions, text planning with discourse structure relations can be broadened to include some text formatting[3]. Little written discourse — certainly no journal or conference papers, reports, or overhead transparencies — is generated completely without formatting devices, whether they be simple headings, section

[3]This work was done by the author and Dr. Yigal Arens of USC/ISI.

```
                 (a)
                  |
                 ELAB
                /    \
               /      \
           ELAB        \
          /    \        \
         /      \        \
      ELAB       \        \
     /    \       \        \
    /      \       \        \
ELAB-PT-WHOLE   SEQ     SEQ     SEQ
     /  \       / \     / \     / \
  mekar part   Ka Kl   Fa Fl   Wa Wl
```

```
              (b)
               |
               |
              ELAB
             /     \
            /       \
ELAB-PT-WHOLE       JOIN
     /  \          /  |  \
    /    \        /   |   \
 mekar  part     /    |    \
              SEQ   SEQ    SEQ
              / \   / \    / \
            Ka Kl  Fa Fl  Wa Wl
```

```
                 (c)
                  |
                  |
                 ELAB
                /    \
               /      \
     ELAB-PT-WHOLE    SEQ
          /  \       /   \
         /    \     /     \
      mekar  part  JOIN    JOIN
                  / | \   / | \
                Ka  Fa Wa Kl Fl Wl
```

```
                 (d)
                  |
                  |
                 ELAB
                /    \
               /      \
     ELAB-PT-WHOLE    SEQ
          /  \       /   \
         /    \     /     \
      mekar  part  JOIN    JOIN
                  /   \   /   \
              Ka,Fa  Wa Kl  Fl,Wl
```

Figure 9: (a) Original paragraph structure. (b) After rule 1: merging same relations. (c) After rule 2: merging relations in lists. (d) After rule 3: merging noun phrases.

names, and occasional italicized portions, or more sophisticated itemized lists, footnotes, indented quotations, and boldfaced terms.

Why? The reason is clear: each such formatting device carries a distinct meaning, and writers select the device that best serves their communicative intent at each point in the text.

A more interesting question is: How? That is, how do writers know what device to use at each point? How is device selection integrated with the discourse production process in general? Can the two processes be automated — can a text production system be made to plan not only the content and structure of the text but also the appropriate textual formatting for it?

The answer is yes, and this section describes an experiment that demonstrates this ability.

6.3.1 Textual Devices

In the course of work on multimedia communication [Hovy & Arens 90, Arens & Hovy 90], we noticed an interesting fact: not only are the text layouts and styles (plain text, itemized lists, enumerations, italicized text, inserts, which are called here *textual devices*) used systematically to convey information, but it is possible to define their communicative semantics precisely enough for some of them to be used in a text planner. What's more, the systematicity holds across various types of texts, genres, and registers of formality. It is found in books, articles, advertisements, papers, letters, and even memos. The information these devices convey supplements the primary content of the text.

Though manuals of style (such as [CMS 82, APA 83, Van Leunen 79]) may seem relevant, they contain little more than precise descriptions of the preferred forms of textual devices in fact. We therefore classified textual devices into three broad classes — *Depiction*, *Position*, and *Composition* — and tried to provide functional descriptions of them. In all three cases, their communicative function is to delimit a portion of text for which certain exceptional conditions of interpretation hold. The following are some general uses of these devices (more detail appears in [Hovy & Arens 91]):

- 1. **Depiction:** selection of an appropriate letter string format.
  - *Parentheses:* text is tangential to the main text.
  - *Font switching:* text has special importance (new term, of central importance, foreign expression) when the surrounding text is not italicized.
  - *Capitalization:* text string names (identifies) an entity.
  - *Quotation marks:* text was written by another author, or some non-literal, special meaning is intended.
- 2. **Position:** Repositioning of text blocks.
  - *Inline:* non-distinguished normal case.
  - *Offset* (horizontal repositioning): text was authored by someone else.
  - *Separation* (vertical repositioning): text addresses a single point (a paragraph) or identifies subsequent text (headings or titles).
  - *Offpage:* text provides explanatory material (appendix, footnote).

- **3. Composition:** imposition of an internal structure on the text.
  - *Itemized list:* set of (maximally paragraph-length) discourse objects on the same level of specificity with respect to the subject domain, each more than a clause (e.g., this list of textual devices).
  - *Enumerated list:* set of (maximally paragraph-length) discourse objects on the same level of specificity with respect to the domain, which are ordered along some underlying dimension, such as time, distance, importance.
  - *Term definition:* pair of texts separated by a colon or other delimiter, in which the first names a discourse object and the second defines or explains it (e.g., this item on term definition).

Selecting appropriate textual devices relies on the author's ability to accurately characterize the meaning expressed by the specific portion of text as well as its relationship to the surrounding text (after all, the same sentence can properly be a footnote in one text and a parenthesized part of the text proper in another). Thus (ignoring such issues as textual prominence and style), the problem has three parts: the underlying semantic content to be communicated, the discourse structure, and the textual devices available. With respect to semantics, we took a standard approach (namely, using frame-like representation structures that contain terms from a well-specified ontology), and to define the communicative semantics of textual devices, we employed an extension of RST.

6.3.2 Extending the Structurer: An Experiment in Layout Planning

The RST text structure planner was used to plan and generate paragraphs of text about procedures to be followed by air traffic controllers, using representations from the ARIES system [Johnson & Harris 90, Johnson & Feather 91], an automatic programming project. In one example, the structurer was activated with the goal to describe the procedure to be followed by an air traffic controller when an aircraft is "handed over" from one region to the next. The underlying representation for this example consisted of a semantic network of 18 instances, defined in terms of 27 air traffic domain concepts and 8 domain relations, implemented as frames in the Loom knowledge representation system [MacGregor 88]. The structure planner built the paragraph tree shown in Figure 10.

Though the form of the text closely mirrors that of the actual Air Traffic Control Manual [ASA 89], the differences in formatting are significant; and these differences make the manual much more readable. The manual contains headings, term definitions signaled by italicized terms, enumerated lists, and so forth. After studies of instructional texts (including recipes, school textbooks, and manuals for cars, sewing machines, and video players) conducted at USC/ISI and the University of Nijmegen [Vossers 91, Arens et al. 92], we concluded that certain textual formatting devices are highly correlated with specific configurations of the underlying text structure tree. For example, a series of nested SEQUENCEs, such as appears in Figure 10, is usually realized in the text as an enumerated list. Exceptions occur (in general) only when the individual items enumerated are single words (in which case the whole list is realized in a single sentence) or when there are few enough of them to place in a paragraph in-line (though usually in this case the keywords *first*, *second*, etc., are added).

```
                    COND
                  /      \
        make-handoff   ELAB-PROCESS-STEP
                            /     \
                  relay-info       SEQ
                                  /   \
                            give-1     SEQ
                                      /   \
                                give-2     give-3
```

```
When making a handoff, the transferring controller relays information
to the receiving controller in the following order.  He gives the
target's position.  He gives the aircraft's identification.  He gives
the assigned altitude and appropriate restrictions.
```

Figure 10: Discourse structure and text for Air Traffic Control domain.

On the assumption that one can capture most of the reasons for using such formatting devices as enumerations on the basis of RST alone, we augmented the text plan SEQUENCE in order to include explicit text formatting commands and adapted the structure planner accordingly. For the formatting commands we used LaTeX forms such as `\begin{enumerate} \item \end{enumerate}` [Lamport 86]. Although our implementation was done within the framework of our specific generation technology, we believe a similar augmentation could be performed with most if not all the text planners being developed at this time. The resulting tree (with formatting commands indicated) and the resulting text, generated by Penman and run through LaTeX, is shown in Figure 11.

6.3.3 Semantics of Textual Devices

Despite its rather extreme simplicity, however, the example demonstrates that to the extent one can characterize textual formatting devices in terms of configurations within the discourse structure, one can plan appropriate formatting commands of several types. Some textual devices with structural definitions are:

- *Enumeration:* As described in the example above, the text structure relation SEQUENCE can generally be formatted as an enumerated list. The enumeration follows the sequence of the relation, which is planned in expression of some underlying semantic ordering of the items involved, for example time and location.

- *Itemization:* The textual structure that relates a number of items without any underlying order is the RST relation JOIN, which can be realized by an itemized list (unless the items are small enough to be placed into a single sentence).

- *Appendix, footnote, and parentheses:* These are three devices that realize the same textual relation, namely BACKGROUND. They differ in the amount of material included in the relation's Satellite.

```
                    COND
                  /      \
    make-handoff     ELAB-PROCSTEP
                          /     \
             relay-info       SEQ-1
                                /   \
("\begin{enumerate} \item" give-1)   (SEQ-2 "\end{enumerate}")
                                       /   \
                  ("\item" give-2)   ("\item" give-3)
```

When making a handoff, the transferring controller relays information to the receiving controller in the following order.

1. He gives the target's position.
2. He gives the aircraft's identification.
3. He gives the assigned altitude and appropriate restrictions.

Figure 11: Augmented discourse structure and text for Air Traffic Control domain.

- *Section title or heading:* This device realizes the textual relation IDENTIFICATION, which links an identifier with the body of material it heads. A section or subsection is appropriate when the IDENTIFICATION is combined with a SEQUENCE chain that governs the overall presentation of the text.

The utility of discourse structure relations for specifying the communicative semantics of text formatting devices is a somewhat unexpected bonus. However, two limitations should be borne in mind: unstudied stylistic factors also play a role, and the representational power of current theories of discourse structure is still very limited; for some textual devices, no discourse relation has been identified by discourse linguists (for example, the Quotation device realizes the linguistic relation PROJECTION), and others work on a level too detailed for text coherence theories, since they operate on individual words within a clause.

7 Conclusion

As natural language processing systems become more powerful, they increasingly address the complexities of multisentence discourse. Without a good understanding of how discourse really works, however, no successful communication is possible; too much is missed if sentences are considered individually alone. From the perspective of language generation, discourse structure plays a central role throughout the text planning process, from helping organize the speaker's communicative intentions and specifying what material to include, to constraining how to cast it, how to ensure that it is presented in an understandable, coherent, and linguistically appropriate way, and how to format it.

As discussed in this paper, a full understanding of the nature of discourse is impossible without a clear description of the form and role of intersegment discourse relations, which form the backbone of discourse structure. With regard to these relations, this paper outlines the following topics:

- the relationship between intentional plans and structural relations,
- the underlying similarity of relation/plans and schemas,
- the assembly of a taxonomy of discourse structure relations,
- the relationship between discourse relations and focus,
- the effect of discourse relations on the syntactic casting of material,
- the aggregation of material under discourse relations,
- the communicative semantics of text formatting devices in terms of discourse structure relations.

The studies described here all address some aspect of the problems of discourse structure. Starting with schemas and the RST-based text structure planning, a considerable amount has been learned in the last decade, though much work remains to be done before text planning under communicative intent and text structuring using intersegment discourse relations are understood. However, the availability of a crude discourse structure (in the form of a tree constructed from discourse relations) as a central construct with which to work makes the task of addressing these questions and evaluating the answers a great deal easier than it was a decade ago, when it was often difficult even to formulate the problems.

Few of the studies described here constitute the final word on the subject. They serve as signposts to further areas to explore. However, taking into account the magnitude of the problem of discourse, the enterprise of text planning and discourse analysis has come a long way in a short time. It is not unreasonable to expect the flexible planning and generation of coherent, high-quality multi-page texts in limited domains within the next five years. The new developments are a challenge and an invitation, promising an interesting decade of the nineties!

Acknowledgments

Many people have contributed to the work described here, directly and indirectly, particularly:

For Section 5.1: Thanks to Dr. William Mann (retired from USC/ISI).

For Section 5.1.2: Thanks to Dr. Cécile Paris (USC/ISI), Prof. Johanna Moore (University of Pittsburgh), and Dr. William Swartout (USC/ISI).

For Section 5.2: Thanks to the ISI Text Planning Group: Dr. Cécile Paris (USC/ISI), Dr. Julia Lavid (University Complutense of Madrid), Ms. Elisabeth Maier (IPSI Institute, Darmstadt), Mr. Vibhu Mittal, USC, and Mr. Richard Whitney, USC/ISI, and to visitors Mr. Giuseppe Carenini

(IRST, Trento) and Mr. Thanasis Daradoumis (University of Barcelona). Also thanks for discussions to Dr. Gisela Redeker (Free University of Amsterdam), Dr. Ted Sanders (University of Utrecht), and Dr. Wilbert Spooren (Tilburg University).

For Section 5.4: Thanks to Prof. Kathleen McCoy (University of Delaware) amd Dr. Julia Lavid (University of Madrid).

For Section 6.1: Thanks to Prof. Donia Scott (Brighton University) and Dr. Dietmar Rösner (University of Ulm).

For Section 6.2: Thanks to Mr. Hercules Dalianis (University of Stockholm).

For Section 6.3: Thanks to Dr. Yigal Arens, to Dr. William Swartout for the idea of using LaTeX, to Mr. Richard Whitney for implementation help, and to Dr. Lewis Johnson for ARIES (all from USC/ISI).

The author also wishes to acknowledge the comments of two anonymous reviewers and to Prof. Barbara Grosz (Harvard University).

8 Appendix

This section contains the top levels of the discourse structure relations collected and merged in several studies performed by the author and colleagues, as described in Section 5.2. Over 500 relations from approximately 35 researchers in various fields were collected and taxonomized in three parallel hierarchies totaling approximately 120 relations; see [Hovy 90b, Hovy et al. 92]. The core set of relations is shown in Figure 12. Details about its sources, definitions, and taxonomization procedure can be found in [Hovy & Maier 92].

The classification into three parallel hierarchies is motivated by appealing to factors central to text planning: the types of information required to define and use the relations and the resulting types of illocutionary and perlocutionary effects that the relations have in the discourse.

8.1 Semantic Relations

Semantic relations are defined as those that hold between adjacent segments of material that expresses some experience of the world about us and within our imagination. For example, in:

> *"Ben poured coffee into the cup. When next he looked, he saw that it had been drunk."*

the temporal relationship between the two clauses is cued by the word "when" and by the referential identity of "Ben and "he" and "coffee and "it". The semantic sequentiality of the second clause after the first is given by the fact that Ben's discovery could only occur *after* he poured the coffee into the cup. The interclausal relation SEQUENCE must be specified in terms of the underlying temporal relationship between the events mentioned in the two clauses — a fact about the world.

Given their nature, the use of semantic relations can be determined by the presence of the material related in a system's factual knowledge base. In many instances, relations can be mapped onto knowledge base constructs; for example, the GENERAL-SPECIFIC subtype of ELABORATION can be mapped onto IS-A or CONCEPT-INSTANCE links in conventional knowledge representation formalisms. No explicit reference to a user model or any other external source of knowledge is required.

8.2 Interpersonal Relations

Interpersonal relations are defined as those holding between adjacent segments of material in which the author attempts to affect the addressee's beliefs, attitudes, desires, etc. The perlocutionary effects achieved by these relations are to convince, enable, motivate, give evidence, interpret, or evaluate.

The definitions of interpersonal relations all necessarily involve the addressee's knowledge, beliefs, or attitudes toward the propositional content of the text. For example, in:

> *"The new Tech Report abstracts are now in the journal area of the library near the abridged dictionary. Please sign your name by any that you would be interested in seeing."* (from [Mann & Thompson 88])

			ObjectAttribute (9)
		ElabObject (1)	ObjectFunction (3)
			Set-Member (3)
		ElabPart	Process-Step (5)
	Elaboration (12)		Whole-Part (8)
		ElabGenerality	Genl-Specific (15)
			Abstr-Instance (14)
		Identification (10)	
		Restatement (11)	Summary (4)
		Location (6)	
		Time (8)	
		Means (4)	
	Circumstance (4)	Manner (4)	
		Instrument (1)	
		ParallelEvent (3)	
		SeqTemporal (6)	
	Sequence (6)	SeqSpatial (1)	
		SeqOrdinal (3)	
Semantic (1)			VolCause (1)
		C/RVol (1)	VolResult (2)
	Cause/Result (17)	C/RNonVol (1)	NonVolCause (1)
		Purpose (8)	NonVolResult (2)
		Condition (9)	
	GeneralCondition (1)	Exception (3)	
		Equative (6)	
		Contrast (16)	
	Comparative (1)	Otherwise (8)	
		Comparison (3)	
		Analogy (4)	
	Interpretation (3)	Evaluation (3)	
	Enablement (10)	Background (4)	
Interpersonal (1)	Antithesis (7)		Solutionhood (1)
		Support (2)	Evidence (10)
	Exhortation	Concession (7)	Justification (4)
		Qualification (2)	Motivation (7)
	LogicalRelation	Conjunction (6)	
Presentational (2)	PresentationalSeq (1)	Disjunction (3)	
	Join (7)		

Figure 12: A taxonomy of discourse segment relations. The number associated with each relation indicates the number of different researchers who listed the relation and may be interpreted as a vote of confidence in it.

the enabling relation that holds between the two sentences concerns the addressee's knowledge and desire to express his or her interests in certain Tech Reports. It is not possible to define the interclausal relationship used without reference to the addressee. This essential aspect of interpersonal relations is reflected in the RST definitions [Mann & Thompson 88] of, say,

- EVIDENCE:
  The reader's comprehending the satellite increases his belief of the nucleus.
- MOTIVATION:
  Comprehending the satellite increases the reader's desire to perform the action presented in the nucleus.

Other interpersonal relations, such as INTERPRETATION and EVALUATION, are defined in terms of the goals and intentions of the author.

Since the use of interpersonal relations is predicated mainly on the interests, beliefs, and attitudes of the addressee and/or author, relations of this type are usually defined in a computer system with respect to a user model.

8.3 Presentational Relations

Presentational relations are defined as those holding between adjacent segments of text that are not meant to be directly related semantically or interpersonally, but whose relationship exists solely due to the juxtaposition imposed by the nature of the presentation medium.

Typically, the "linear" nature of language enforces the use of relations for presentational purposes; examples are CONJUNCTION and PRESENTATIONALSEQ. For example, the latter is used as follows:

> *"There are a number of criteria for distinguishing Ranges from Goals:* **First**, *the Range cannot be probed by* do to *or* do with, *whereas the Goal can.* **Second**, *since nothing is being 'done to' it, a Range element never can have a resultative Attribute added within the clause, as a Goal can...* **Next**, *the Range cannot be a personal pronoun, and it cannot normally be modified by a possessive.* **Finally**, *a range element (other than one with an 'empty' verb like* have *or* do*) can often be realized as a prepositional phrase and under certain conditions it has to be....*
> (from [Martin 92], text formatting removed)

The text makes no claim about the semantic orderedness of the sentences enumerated.

Most collections of intersegment discourse relations indiscriminately intermix explicitly presentational relations with semantic and interpersonal ones. This is probably due to the fact that *all* intersegment relations play some presentational role in text, which causes a certain amount of confusion. However, for most relations the presentational function is not primary, and when one is aware of this distinction, the problem is greatly reduced. One major remaining source of difficulty is the SEQUENCE family, since in English the same cue words and other textual markers are used to signal presentational sequence as semantic sequence. We solve the problem by creating the purely presentational relation PRESENTATIONALSEQ.

A further reason for distinguishing the three classes is their difference in illocutionary force. All the semantic relations are expressed by the single illocutionary act DESCRIBE, while the interpersonal relations are expressed by various perlocutionary acts, including CONVINCE, MOTIVATE, and JUSTIFY. The consequences of this difference on the design of text planning systems are outlined in [Maier & Hovy 91].

References

[Allen & Perrault 80] Allen, J.F. and Perrault, C.R. 1980. Analyzing Intention in Utterances. *Artificial Intelligence* 15 (143–178).

[Appelt 85] Appelt, D.E. 1985. *Planning English Sentences.* Cambridge: Cambridge University Press.

[APA 83] *Publication Manual of the American Psychological Association* (third edition). 1983. Washington: American Psychological Association.

[Arens et al. 88] Arens, Y., Miller, L., Shapiro, S.C. & Sondheimer, N.K. 1988. Automatic Construction of User-Interface Displays. *Proceedings of the 7th AAAI Conference*, St. Paul. Also available as USC/Information Sciences Institute Research Report RR-88-218.

[Arens & Hovy 90] Arens Y., and Hovy, E.H. 1990. How to Describe What? Towards a Theory of Modality Utilization. *Proceedings of the 12th Annual Conference of the Cognitive Science Society (487*–494).

[Arens et al. 92] Arens, Y., Hovy, E.H, and Vossers, M. 1992. Describing the Presentational Knowledge Underlying Multimedia Instruction Manuals. In *Intelligent Multimedia Interfaces*, M. Maybury (ed) (forthcoming).

[Aristotle 54] Aristotle. The Rhetoric. In *The Rhetoric and the Poetics of Aristotle*, W. Rhys Roberts (trans). 1954. New York: Random House.

[ASA 89] ASA, Inc. 1989. *Airman's Information Manual.* Aviation Supplies and Academics.

[Asher 92] Asher, N. 1992. *Abstract Object, Semantics and Anaphora.* Unpublished manuscript, University of Texas at Austin (forthcoming).

[Austin 65] Austin, J.L. 1965. *How to do Things with Words.* Oxford: Oxford University Press.

[Birnbaum et al. 80] Birnbaum, L.A., Flowers, M., and McGuire, R. 1980. Towards an AI Model of Argumentation. *Proceedings of the 1st AAAI Conference*, Stanford, (195–198).

[Björklund & Virtanen 89] Björklund, M. and Virtanen, T. 1989. Variation in narrative structure: A Simple Text vs. an Innovative Work of Art. Presented at the 16th International Systemics Congress, Helsinki.

[Cawsey 90] Cawsey, A. 1990. Generating Communicative Discourse. In *Current Research in Natural Language Generation*, Dale, R., Mellish, C., and Zock, M. (eds), Boston: Academic Press (75–102).

[Charniak & Shimony 90] Charniak, E. and Shimony, S.E. 1990. Probabilistic Semantics for Cost Based Abduction. *Proceedings of the 8th AAAI Conference* (107–111).

[CMS 82] *The Chicago Manual of Style* (thirteenth edition). 1982. Chicago: University of Chicago Press.

[Cohen 83] Cohen, R. 1983. *A Computational Model for the Analysis of Arguments.* Technical Report CSRG-151, University of Toronto.

[Cohen & Levesque 85] Cohen, P.R. and Levesque, H.J. 1985. Speech Acts and Rationality. *Proceedings of the 23rd ACL Conference*, Chicago (49–59).

[Cohen & Levesque 90] Cohen, P.R. and Levesque, H.J. 1990. Rational Interaction as the Basis for Communication. In *Intentions in Communication*, Cohen, P., Morgan, J., and Pollack, M. (eds). Cambridge: MIT Press.

[Conklin & McDonald 82] Conklin, E.J. and McDonald, D.D. 1982. Salience: The key to the Selection Problem in Natural Language Generation. *Proceedings of the 20th ACL Conference*, Toronto (129–135).

[Cullingford 78] Cullingford, R. 1978. *Script Application: Computer Understanding of Newspaper Stories.* Ph.D. dissertation, Yale University.

[Dahlgren 88] Dahlgren, K. 1988. *Naive Semantics for Natural Language Understanding.* Boston: Kluwer Academic Press.

[Dale 88] Dale, R. 1988. Generating Referring Expressions in a Domain of Objects and Processes. Ph.D. dissertation, University of Edinburgh.

[Dalianis & Hovy 93] Dalianis, H. and Hovy, E.H. 1993. Aggregation in Natural Language Generation. *Proceedings of the Fourth European Workshop on Language Generation*, Pisa, Italy, 1993.

[Daneš 74] F. Daneš. 1974. Functional Sentence Perspective and the Organization of the Text. In *Papers on Functional Sentence Perspective*, F. Daneš (ed.). The Hague: Mouton.

[Davey 79] Davey, A. 1979. *Discourse Production.* Edinburgh University Press.

[DeJong 79] DeJong, G.F. 1979. *Skimming Stories in Real Time: An Experiment in Integrated Understanding.* Ph.D. dissertation, Yale University.

[De Souza et al. 89] De Souza, C.S., Scott, D.R. and Nunes, M.G.V. 1989. Enhancing Text Quality in a Question-Answering System. Unpublished manuscript, Pontificia Universidade Católica de Rio de Janeiro.

[Dobeš & Novak 91] Dobeš, Z. & Novak, H-J. 1991. From Constituent Planning to Text Planning. *Proceedings of the Third European Workshop on Language Generation*, Judenstein, Austria (46–54).

[Dyer 83] Dyer, M.G. 1983. *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension.* Cambridge: MIT Press.

[Fikes et al. 72] Fikes, R.E., Hart, P.E. and Nilsson, N.J. 1972. Learning and Executing Generalized Robot Plans. *Artificial Intelligence* 3(4) (251–288).

[Fries 81] Fries, P.H. 1981. On the Status of Theme in English: Arguments from Discourse. *Forum Linguisticum* 6 (1–38).

[Green 92] Green, S. 1992. A Basis for Formalization of Linguistic Style. *Proceedings of the 30th Annual Conference of the Association of Computational Linguistics*, Newark (312–314).

[Grimes 75] Grimes, J.E. 1975. *The Thread of Discourse.* Mouton: The Hague.

[Grosz 77] Grosz, B.J. 1977. The Representation and Use of Focus in Dialogue Understanding. Technical Report 151, SRI International, Menlo Park, CA.

[Grosz 81] Grosz, B.J. 1981. Focusing and Description in Natural Language Dialogues. In *Elements of Discourse Understanding.* Joshi, A., Webber, B., and Sag, I. (eds), Cambridge: Cambridge University Press.

[Grosz & Sidner 86] Grosz, B.J. and Sidner, C.L. 1986. Attention, Intentions, and the Structure of Discourse. *Journal of Computational Linguistics* 12(3) (175–204).

[Grosz & Sidner 90] Grosz, B.J. and Sidner, C.L. 1990. Plans for Discourse. In *Intentions in Communication*, Cohen, P., Morgan, J., and Pollack, M. (eds). Cambridge: MIT Press.

[Halliday 67] Halliday, M.A.K. 1967. Notes on Transitivity and Theme in English. *Journal of Linguistics* 3 (37–81), 4 (179–215).

[Halliday 85] Halliday, M.A.K. 1985. *An Introduction to Functional Grammar.* Baltimore: Edward Arnold Press.

[Heim 83] Heim, I. 1983. File Change Semantics and the Familiarity Theory of Definiteness. In R. Bauerle, Ch. Schwartze, and A. von Stechow (eds), *Meaning, Use, and Interpretation of Language.* Berlin: De Gruyter.

[Hobbs 78] Hobbs, J.R. 1978. Why is Discourse Coherent? Technical Note no. 176, SRI International, Menlo Park.

[Hobbs 79] Hobbs, J.R. 1979. Coherence and Coreference. *Cognitive Science* 3(1) (67–90).

[Hobbs 90] Hobbs, J.R. 1990. *Literature and Cognition.* CSLI Lecture Notes no. 21.

[Hobbs et al. 88] Hobbs, J.R., Stickel, M., Martin, P. and Edwards, D. 1988. Interpretation as Abduction. *Proceedings of the 26th ACL Conference*, Buffalo (95–103).

[Horacek 92] Horacek, H. 1992. An Integrative View of Text Planning. In *Aspects of Automated Natural Language Generation*, R. Dale, E. Hovy, D. Rösner, O. Stock (eds). Heidelberg: Springer Verlag Lecture Notes in AI number 587 (1–31).

[Hovy 87] Hovy, E.H. 1987. Interpretation in Generation. *Proceedings of the 6th AAAI Conference*, Seattle (545–549). Also available as USC/Information Sciences Institute Research Report ISI/RS-88-186.

[Hovy 88] Hovy, E.H. 1988. Planning Coherent Multisentential Text. *Proceedings of the 26th ACL Conference*, Buffalo (163–169).

[Hovy 90a] Hovy, E.H. 1990. Approaches to the Planning of Coherent Text. In *Natural Language in Artificial Intelligence and Computational Linguistics*, Paris, C.L., Swartout, W.R. and Mann, W.C. (eds). Boston: Kluwer (83–102).

[Hovy 90b] Hovy, E.H. 1990. Parsimonious and Profligate Approaches to the Question of Discourse Structure Relations. *Proceedings of the Fifth International Workshop on Text Generation*, Pittsburgh.

[Hovy 90c] Hovy, E.H. 1990. Natural Language Processing at ISI. *The Finite String* 16(4) (37–42).

[Hovy & Arens 90] Hovy, E.H. and Arens, Y. 1990. Allocating Modalities In Multimedia Communication. Presented at AAAI Spring Symposium on Knowledge-Based Human-Computer Communication, Stanford University.

[Hovy & Arens 91] Hovy, E.H. and Arens, Y. 1991. Automatic Generation of Formatted Text. *Proceedings of the 8th AAAI Conference*, Anaheim.

[Hovy & Lavid 92] Hovy, E.H. and Lavid, J.M. 1992. Focus, Theme, Given, and Other Dangerous Things. Working paper.

[Hovy & Maier 92] Hovy, E.H. and Maier, E. 1992. Parsimonious and Profligate: How Many and Which Discourse Structure Relations? *Discourse Processes* (forthcoming).

[Hovy & McCoy 89] Hovy, E.H. and McCoy, K.F. 1989. Focusing your RST: A Step toward Generating Coherent Multisentential Text. *Proceedings of the 11th Cognitive Science Conference*, Ann Arbor (667–674).

[Hovy et al. 92] Hovy, E.H., Lavid, J., Maier, E., Mittal, V., and Paris, C.L. 1992. Employing Knowledge Resources in a New Text Planner Architecture. In *Aspects of Automated Natural Language Generation*, R. Dale, E. Hovy, D. Rösner, O. Stock (eds). Heidelberg: Springer Verlag Lecture Notes in AI number 587 (57–72).

[Ivir et al. 80] Ivir, V., McMillan, D. and Merz, T. 1980. S-Relators. Unpublished manuscript, University of Zagreb.

[Johnson & Harris 90] Johnson, W.L. and Harris, D. 1990. Requirements Analysis Using ARIES: Themes and Examples. *Proceedings of the 5th Knowledge Based Software Engineering Conference* (79–84). Tech Report no. RL-TR-91-11, Rome Laboratory.

[Johnson & Feather 91] Johnson, W.L. and Feather, M. 1991. Using Evolution Transformations to Construct Specifications. In *Automating Software Design*, Lowry, P. and McCartney, J. (eds), AAAI Press.

[Jullien & Marty 89] Jullien, C. and Marty, J-C. 1989. Plan Revision in Person-Machine Dialogue. *Proceedings of the European ACL Conference*, Manchester (153–160).

[Kamp 81] Kamp, H. 1981. A Theory of Truth and Semantic Representation. In *Formal Methods in the Study of Language*, Groenendijk, J.A.G., Janssen, T.M.V. and Stokhof, M.B.J. (eds), Mathematical Centre Tracts (vol 136), Amsterdam, (277–322).

[Kautz 87] Kautz, H. 1987. *A Formal Theory of Plan Recognition.* Ph.D. dissertation, University of Rochester.

[Kreyss & Novak 90] Kreyss, J. and Novak, H.-J. 1990. The Textplanning Component PIT of the LILOG System. *Proceedings of the COLING Conference*, Helsinki (431–433).

[Lambert & Carberry 91] Lambert, L. and Carberry, S. 1991. A Tripartite Plan-Based Model of Dialogue. *Proceedings of the 29th Annual Conference of the Association of Computational Linguistics*, Berkeley (47–54).

[Lamport 86] Lamport, L. 1986. LaTeX *User's Guide and Reference Manual.* Reading: Addison-Wesley.

[Lascarides & Asher 91] Lascarides, A. and Asher, N. 1991. Discourse Relations and Defeasible Knowledge. *Proceedings of the 29th ACL Conference*, Berkeley.

[Levy 79] Levy, D.M. 1979. Communicative Goals and Strategies: Between Discourse and Syntax. In T. Givon (ed), *Discourse and Syntax.* Academic Press (183–210).

[Levy 84] Levy, E. 1984. *Communicating Thematic Structures in Narrative Discourse: The Use of Referring Terms and Gestures.* Ph.D. dissertation, University of Chicago.

[Litman 85] Litman, D. 1985. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues.* Ph.D. dissertation, University of Rochester.

[Lochbaum 91] Lochbaum, K.E. 1991. An Algorithm for Plan Recognition in Collaborative Discourse. *Proceedings of the 29th Annual Conference of the Association of Computational Linguistics*, Berkeley (33–38).

[MacGregor 88] MacGregor, R. 1988. A Deductive Pattern Matcher. *Proceedings of the 6th National Conference on Artificial Intelligence AAAI-88*, St. Paul (696–701).

[Maier & Hovy 91] Maier, E. and Hovy, E.H. 1991. Organizing Discourse Structure Relations using Metafunctions. In *New Concepts in Natural Language Generation: Planning, Realization, and Systems*, H. Horacek (ed), London: Pinter (to appear).

[Mann 87] Mann, W.C. 1987. Text Generation: The Problem of Text Structure. USC/Information Sciences Institute Research Report RR-87-181.

[Mann & Matthiessen 83] Mann, W.C. and Matthiessen, C.M.I.M. 1983. Nigel: A Systemic Grammar for Text Generation. USC/Information Sciences Institute Research Report RR-83-105.

[Mann & Moore 81] Mann, W.C. and Moore, J.A. 1981. Computer Generation of Multi-Paragraph English Text. *Computational Linguistics* 7(1) (63–89).

[Mann & Thompson 86] Mann, W.C. and Thompson, S.A. 1986. Rhetorical Structure Theory: Description and Construction of Text Structures. In *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*, Kempen, G. (ed), Boston: Kluwer (279–300).

[Mann & Thompson 88] Mann, W.C. and Thompson, S.A. 1988. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text* 8(3) (243–281). Also available as USC/Information Sciences Institute Research Report RR-87-190.

[Marslen-Wilson et al. 82] Marslen-Wilson, W., Levy, E., and Tyler, L.K. 1982. Producing Interpretable Discourse: The Establishment and Maintenance of Reference. In *Speech, Place and Action*, R.J. Jarvella and W. Klein (eds), New York: John Wiley (339–378).

[Martin 92] Martin, J.R. 1992. *English Text: System and Structure.* Amsterdam: Benjamins (forthcoming).

[Maybury 90] Maybury, M.T. 1990. *Planning Multisentential English Text Using Communicative Acts.* Ph.D. dissertation, Cambridge University. Also available as RADC Technical Report 90-411.

[McCoy 85] McCoy, K.F. 1985. *Correcting Object-Related Misconceptions.* Ph.D. dissertation, University of Pennsylvania.

[McCoy & Cheng 88] McCoy, K.F. and Cheng, J. 1988. Focus of Attention: Constraining What can be Said Next. In *Natural Language in Artificial Intelligence and Computational Linguistics*, Paris, C.L., Swartout, W.R. and Mann, W.C. (eds). Boston: Kluwer (103–124).

[McDermott 81] McDermott, D.V. 1981. Artificial Intelligence meets natural stupidity. In *Mind Design*, Haugeland, J. (ed), Cambridge: MIT Press (143–160).

[McKeown 85] McKeown, K.R. 1985. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text.* Cambridge: Cambridge University Press.

[Meehan 76] Meehan, J. 1976. *The Metanovel: Writing Stories by Computer.* Ph.D. dissertation, Yale University.

[Meteer 90] Meteer, M.W. 1990. *The Generation Gap: The Problem of Expressibility in Text Planning.* Ph.D. dissertation, University of Massachusetts at Amherst.

[Mooney et al. 90] Mooney, D.J., Carberry, S., and McCoy, K.F. 1990. The Basic Block Model of Extended Explanations. In *Proceedings of the 5th International Workshop on Text Generation*, Pittsburgh.

[Moore 89] Moore, J.D. 1989. *A Reactive Approach to Explanation in Expert and Advice-Giving Systems.* Ph.D. dissertation, University of California in Los Angeles.

[Moore & Paris 89] Moore, J.D. and Paris, C.L. 1989. Planning Text for Advisory Dialogues. *Proceedings of the 27th ACL Conference*, Vancouver (67–75).

[Moore & Paris 91] Moore, J.D. and Paris, C.L. 1991. Discourse Structure for explanatory Dialogues. Presented at the AAAI Fall Symposium on Discourse, Pacific Grove.

[Moore & Swartout 90] Moore, J.D. and Swartout, W.R. 1990. Dialogue-Based Explanation. In *Natural Language in Artificial Intelligence and Computational Linguistics*, Paris, C.L., Swartout, W.R. and Mann, W.C. (eds), Boston: Kluwer (3–48).

[Moore & Pollack 93] Moore, J.D. and Pollack, M.E. 1993. A Problem for RST: The Need for Multi-Level Discourse Analysis. Squib in *Computational Linguistics* 18(4).

[Paris 90] Paris, C.L. 1990. Generation and Explanation: Building an Explanation Facility for the Explainable Expert Systems Framework. In *Natural Language in Artificial Intelligence and Computational Linguistics*, Paris, C.L., Swartout, W.R. and Mann, W.C. (eds), Boston: Kluwer (49–82).

[Paris 93] Paris, C.L. 1993. *The Use of Explicit Models in Text Generation.* London: Francis Pinter.

[Passoneau 91] Passoneau, R.J. 1991. Getting and Keeping the Center of Attention. In *Challenges in Natural Language Processing*, R. Weischedel and M. Bates (eds) Cambridge: Cambridge University Press.

[Penman 89] *The Penman Documentation.* 1989. 5 unpublished volumes. USC/Information Sciences Institute, Marina del Rey.

[Polanyi 88] Polanyi, L. 1988. A formal Model of the Structure of Discourse. *Journal of Pragmatics* 12 (601p-638).

[Pollack 86] Pollack, M.E. 1986. A Model of Plan Inference that Distinguishes between the Beliefs of Actors and Observers. *Proceedings of the 24th Annual Conference of the Association of Computational Linguistics*, Chicago, (207–214).

[Prince 81] Prince, E.F. 1981. Toward a Taxonomy of Given-New Information. In P. Cole (ed), *Radical Pragmatics*. Academic Press (223–255).

[Quirk & Greenbaum 73] Quirk, R. and Greenbaum, S. 1973. *A Concise Grammar of Contemporary English*. New York: Harcourt Brace Jovanovich Inc.

[Rambow 90] Rambow, O. 1990. Domain Communication Knowledge. *Proceedings of the 5th International Workshop on Text Generation*, Pittsburgh (87–94).

[Rambow & Korelsky 92] Rambow, O. and Korelsky, T. 192. Applied Text Generation. *Proceedings of the Applied Natural Language Processing Conference*, Trento, Italy.

[Ramshaw 91] Ramshaw, L.A. 1991. A Three-Level Model for Plan Exploration. *Proceedings of the 29th Annual Conference of the Association of Computational Linguistics*, Berkeley (39–46).

[Rankin 89] Rankin, I. 1989. *The Deep Generation of Text in Expert Critiquing Systems*. Licentiate thesis, University of Linköping, Sweden.

[Reichman 85] Reichman, R. 1985. *Getting Computers to Talk Like You and Me*. Cambridge: MIT Press.

[Reithinger 91] Reithinger, N. 1991. *Eine Parallele Architektur zur Inkrementellen Generierung Multimodaler Dialogbeiträge*, Ph.D. dissertation, University of the Saarland.

[Rösner & Stede 92] Rösner, D. and Stede, M. 1992. Customizing RST for the Automatic Production of Technical Manuals. In *Aspects of Automated Natural Language Generation*, R. Dale, E. Hovy, D. Rösner, O. Stock (eds). Heidelberg: Springer Verlag Lecture Notes in AI number 587 (199–214).

[Rumelhart 72] Rumelhart, D.E. 1972. Notes on a Schema for Stories. In *Representation and Understanding*, Bobrow, D.G. and Collins, A. (eds). New York: Academic Press.

[Sacerdoti 77] Sacerdoti, E. 1977. *A Structure for Plans and Behavior*. Amsterdam: North-Holland.

[Sanders et al. 92] Sanders, T.J.M., Spooren, W.P.M.S., and Noordman, L.G.M. 1992. Towards a Taxonomy of Coherence Relations. *Discourse Processes* 15 (1–35).

[Schank & Abelson 77] Schank, R.C. and Abelson, R. 1977. *Scripts, Plans, Goals, and Understanding*. Hillsdale: Lawrence Erlbaum Associates.

[Scott & De Souza 90] Scott, D.R. and De Souza, C.S. 1990. Getting the Message across in RST-Based Text Generation. In *Current Research in Natural Language Generation*, Dale, R., Mellish, C., and Zock, M. (eds), Boston: Academic Press (47–74).

[Shepherd 26] Shepherd, H.R. 1926. *The Fine Art of Writing*. New York: The Macmillan Co.

[Sidner 83] Sidner, C.L. 1983. Focusing and Discourse. *Discourse Processes* 6 (107–130).

[Sycara 87] Sycara, K. 1987. *Resolving Adversarial Conflicts: An Approach Integrating Case-Based and Analytical Methods*. Ph.D. dissertation, Georgia Institute of Technology.

[Toulmin 58] Toulmin, S. 1959. *The Uses of Argument*. Cambridge: Cambridge University Press.

[Van Dijk 72] Van Dijk, T.A. 1972. *Some Aspects of Text Grammars*. The Hague: Mouton.

[Van Dijk & Kintsch 83] Van Dijk, T.A. and Kintsch, W. 1983. *Strategies of Discourse Comprehension.* New York: Academic Press.

[Van Leunen 79] Van Leunen, M.-C. 1979. *A Handbook for Scholars.* New York: Alfred Knopf.

[Vossers 91] Vossers, M. 1991. Automatic Generation of Formatted text and Line Drawings. Master's thesis, University of Nijmegen, The Netherlands.