**Natural language generation systems:
a comparative study**

Athanasios Daradoumis

Report LSI–90–47

NATURAL LANGUAGE GENERATION SYSTEMS:
A COMPARATIVE STUDY

Athanasios Daradoumis
Facultad d´ Informática de Barcelona
Pau Gargallo, 5. Barcelona 08025. **SPAIN**

ABSTRACT

The development of natural language generation systems has been one of the most considerable subjects of research in Artificial Intelligence during the last decade. This paper investigates several important research projects in natural language generation developed up to date. Its aim is to provide not only a complete presentation of some of the most significant work in the fields of AI and Computational Linguistics, but also a broad, comparative study of some generation systems. For this purpose, we identify particular techniques in each system which contribute strongly to the quality of the resulting natural language text. Comparison of their individual skills then reveals their performance capabilities and limitations as well as common points, differences and possible advantegeous combinations of these techniques that could lead to even more effective systems.

June 15, 1990

INDEX

1. INTRODUCTION

The use of natural language as a communication means between people and computers requires first of all that the computer analyzes and understands the user's message and store it in the form of an internal representation. Then the system determines (and plans) what to include in the response to this message which results in an internal semantic (or deep structure) representation of what is to be said. Finally, this semantic representation -- the plans -- is realized as a text in the user's desired language.

The first task is the job of a **natural language analyzer**, the second of a **text planner** (or **reasoning module**) and the third of a **realization** (or **linguistic**) **component**.

The first priorities in this research were the development of techniques for analyzing NL inputs, for storing them in the form of internal representations and evaluating them. In the initial stages, less attention was paid to the problem of how NL could be generated from evaluated internal representation structures particularly since human users can correctly interpret communicatively inadequate and ungrammatical utterances on the basis of their world knowledge and contextual information.

In many NL systems, the techniques employed were poorly grounded in semantic and pragmatics. For example, predetermined text schemas (or templates) were instantiated according to the results of their evaluation. In the process, the structure of all responses was predetermined and knowledge about the particular dialog situation was rarely utilized. By doing a direct verbalization of portions of the knowledge base, the quality of the utterances was to a great extent dependent on the structuring of the knowledge [Mann et al. 1982].

With increasing complexity of the NL systems and knowledge bases, the disadvantages of the methods for generating language became evident. Since these disadvantages were often extremely prejudicial to the acceptance of NL systems, the development of general procedures for generating communicatively adequate responses started to show up. In particular, beginning in about 1982, there has been a virtual explosion in the quantity of research being done in the field of natural language generation, and a complete review of all it could well fill a book.

Therefore, NL generation systems have been developed to produce well-constructed texts in the style of a newspaper (e.g., [Danlos 1985], [Kukich 1984]), while others have been used as generation components of on-line systems (e.g., [Jacobs 1985a], [Busemann 1984], [Buchberger and Horacek 1984]) or to produce text in an extended dialog system [Koit and Saluveer 1985].

Whatever is the objective of a generation system, the generation process involves two main problem areas, *which are not wholly independent* :

(1) **deciding what to say:**

(a) *determining what goals the utterance is to achieve* and

(b) *planning what information content and rhetorical force will best meet those goals given the context.*

(2) **deciding how to say it:**

(c) *realizing the specified information and rhetorical intent as a grammatical text.*

These two tasks give rise to the following components of a generation system: A **text planner** (or **reasoning module**) deals with the **what to say** question and returns a *semantic* (or *conceptual*) *representation* of the information that is to be conveyed to the user. A **linguistic generation component** deals with the **how to say it** question; it takes the semantic representation as input and translates it into grammatically correct text (i.e., producing a stream of morphologically specialized words as output).

Not all generation systems in existence today deal with both problem areas (i.e., are composed of both a *text planner* and a *linguistic component*). The majority of them is concerned only with the *linguistic generation component* -- as a natural consequence of the fact that the ultimate and important goal of a generator is to produce natural language text -- designed to accept as input the semantic representations, "plans", developed by another program and consequently has had to be sensitive to the limitations and varying approaches of the present state of the art in conceptual representation.

Relatively few generation research projects address both *"what to say"* and *"how to say it"* questions, and although the generation systems developed for this purpose give the image of complete systems consisting of both a planning and linguistic components, major emphasis has been given on the first one, perhaps because the text produced by these systems is in English, a language with relatively few inflections; on the other hand, languages with a rich inflectional system (e.g., roman or germanic languages) require a much more sophisticated linguistic component able to integrate decisions on lexical, syntactic and morphological operations.

Furthermore, the implementation of both components by these systems gave rise to an interesting argument: whether the two questions *"what to say"* and *" how to say it"* should be treated independently the one from the other or not. Though some researchers have relied on a model of language production which divides processing into the above two stages (e.g., [McKeown 1985]), others have clearly oposed to that (e.g., [Appelt 1985a]). It seems that even though this modularity of a generation system is both theoretically and practically attractive, the "what-how" distinction has lass merit, especially when it is examined in the light of a theory of language based on planning.

Whatever is the method the generation process follows, a language generation system must be capable of producing utterances or text with the same fluency as that of a human speaker or writer. Unfortunately, no NL generation system yet developed can be said to have this high-level performance. In order to achieve this long-term goal, it is not enough to know only about the syntactic and semantic rules of the language; one must also bear in mind that language behavior is part of a coherent *plan* and is directed toward satisfying the speaker´s or writer´s goals.

Concluding, a NL generation system should be capable of taking general decisions about what to accomplish, what knowledge to use; this yields a body of material to convey. Then it should be able to organize (plan) the given body of material and supplement it as needed, with evidence, contrast and other supporting material. It should be capable of using correctly and appropriately such syntactic devices as conjuction and ellipsis; it should be competent at fitting its utterances into a discourse, using pronominal references where appropriate, choosing syntactic structures consistent with the changing focus, making suitable lexical choices, and giving an overall feeling of coherence to the discourse.

Furthermore, other problem areas should be also considered. These include the *knowledge representation* which contains the information to be described, *interpretation* of the user´s questions and *user modelling:*

Knowledge representation and content is important since it limits what the generation system is able to talk about. The system would have a large knowledge base of basic concepts and commonsense knowledge unless an extensive inferencing component is available. Facility for *interpreting* a user´s questions should be also provided which is the job of a natural language analyzer. The system should equally be able to express the same idea from multiple perspectives, or to present it to users with varying backgrounds or goals; it must reason about the user´s state and the ways the produced text affects that state. A *model of the user* (or *identification of the audience*) is therefore quite necessary.

In other words, a complete generation system, which incorporates a comprehensive linguistically principled **grammar**, a **knowledge base** with a sophisticated, commonsensical, conceptual view, and a **text planner** that can make use of **models of the audience** and the **discourse**, must be used so that *versatility* and *creativity* be possible in machine-generated texts [Shapiro 1987].

In this paper, we first present briefly some of the earliest work done in generation (section 2). In particular, in section 2.1, we discuss several early generation systems which only dealt with the linguistic component, while in section 2.2 we present work that concentrated on the development of planning formalisms for NL genaration. Section 2.3 discusses two early generation systems (PROTEUS and KDS) that make use of both a planning and a linguistic component and comparison between them reveals how a possible combination of some of their particular techniques could produce better text than either system alone. A brief overview of later work in generation systems is presented in section 3, which could equally serve as a complete presentation of bibliography in natural language generation up to date.

In the following sections, we describe in some more detail later work on generation systems which we classify according to whether they have primarily focused on **generation based on planning** or on **text organization and generation.**

In particular, in section 4, we present, discuss and criticize a significant work on planning and generation done by Appelt who developed a theory of language generation based on planning and embodied this theory in a computer system called KAMP which reasons from speaker's goals to produce a plan which in turn is refined until an English sentence is completely specified. Then, in section 5, we examine other considerations on planning and generation based on a distinct work done in this area by Mellish and Evans, who start with a non-linear plan structure produced by an AI planning program and generate natural language text explaining how to execute the plan.

In section 6, we pass to a distinct area that treats the subject of text organization (or text structure) and generation. More specifically, we discuss and criticize an important text generation theory and method developed by McKeown who was concerned with how to produce text in response to a given communicative goal (such as *define, desrcibe* and *compare)*. She developed a computational model of discourse strategies and focus rules to guide the generation process in its decisions about *what to say next.* These techniques have been incorporated in a computer system, called TEXT, which generates paragraph-length responses to questions about database structure.

Furthermore, Mann and Thompson have done a considerable research in text generation which resulted in the development of a more general theory of text organization, called *Rhetorical Structure Theory (RST)* and is presented in section 7. Finally, based on the work done by Danlos on text generation in Romance languages, we discuss, in section 8, the use of alternatives other than focus for a more correct treatment of language production based on a set of specific knowledge rather than on rules of general nature.

Concluding, we present here the particular accomplishments and contributions of various approaches to the generation research, by identifying the special techniques and principles that each approach considered as well as their limitations, and possible improvements that might be achieved if additional facilities could be incorporated. At the same time, a comparative study among different generation systems is done. This constitutes the most important part of this paper, since it reveals disadvantages of one technique used in one system in favor of another: how the limitations and inconviniences of one technique have led to the development of another more general and powerful. Finally, we look at the possibility of how different approaches could be combined so that we produce even more effective generation systems.

2. EARLY GENERATION SYSTEMS

It was quite a long time before the problem of language generation began to receive the attention it deserves. Until the early 1980s generation had not been perceived by the larger research community as an important problem to be working on. Yet, the majority of work done during 1970s has been concentrated on problems in the realization phase of the generation process (the *linguistic component*); a *planning facility* as well as *textual organization* were brought into the process only in late 1970s (or early 1980s) when competence one needed in producing more complex text had as result that generation be considered as an important research area.

In the following subsections we present several generation systems distinguishing them as to whether they have focused on the linguistic or the planning component respectively. Finally, we concentrate on the discussion and comparison of two important early generation systems (PROTEUS and KDS) that use both a planning and a linguistic component.

2.1 Linguistic Components

Several early NL generation systems (e.g., [Friedman 1969]), were designed more for the purpose of testing grammars than for communication. The earliest language generation systems designed for communication depended upon ad hoc strategies that produced reasonable behavior in predictable situations. An example of such a language generation system was SHRDLU of [Winograd 1972]. SHRDLU produced language by having a large set of templates with variables that could be instantiated appropriately in different instances. These templates, combined with a number of heuristics about question-answering, reference, and pronominalization, enabled the system to produce dialogs that sounded quite natural, given the simplicity of the technique.

Many application-oriented systems still kept on using the *template-instantiation* approach, since they operated in highly restricted domains in which some generality could be sacrificed in the interest of simplifying the generation component that is only perripheral to their primary concerns. Some well-known expert systems that have employed this technique are the explanation component of MYCIN of [Scott et al. 1977], and the explanation component of Swartout´s digitalis therapy advisor [Swartout 1981].

Although generation systems of this simple but effective sort had become quite important in the early rule-based expert systems, generation researchers, however, started to be interested in more complex texts than the context-free presentation an expert system´s rules could motivate.

In the early 1970s, some research was done to extend the simple approach of instantiating patterns to more general grammar-based approaches. One of these earliest systems was that of

[Simmons and Slocum 1972]. Their system accepted an input in internal representation language -- a *semantic network* -- and used an ATN grammar to drive the entire generation process. Content had already been specified in the semantic network formalism which also specified the words and syntactic structures to be used. Their system generated single sentences and **not** texts. While their system produces different sentences for the same semantic net, they did little work on the reasons for using one of those sentences as opposed to any of the others. Moreover, since the system embodied no notion of the way an utterance fits into a discourse, other than pattern matching with the user's question, it could perform only the simplest generation of definite references. As such, it was designed purely as a question-answering system that never took the initiative in a dialog.

[Goldman 1975] developed a generation system, called MARGIE, that also used an ATN formalism to handle syntactic processes; however, he concentrated his research effort on developing a process to select particular words and idioms for a sentence. For this reason, he assumed a knowledge representation called *conceptual dependency* [Schank 1975] that was based on a very small number of predicates. The primary problem that he addressed then was that of finding a good lexical choice that could describe a concept that was encoded in the internal representation as relationships between a large number of semantic primitives. His solution was to use a *discrimination net* to filter possible lexical choices. Nevertheless, Goldman did not address the questions of *why* to choose one word over another; moreover, since his generator was designed as a question-answerer that produced responses with only a *single* question for a discourse context, it suffered from most of the deficiencies of the Simmons and Slocum generator.

ATN was also independently adapted by [Shapiro 1982] whose generator is the most elaborate of the group. Shapiro's ATN design is particularly enlightening as he uses a sophisticated intentional network formalism to encode the computational state of the underlying knowledge representation. His ATN performs a kind of "parsing" that amounts to the construction of an assessment of the steps that must be taken to satisfy the intended communicative goals of the program. The construction of this assessment is done in terms appropriate for directing the generation of a text.

The general and most important problem with using this approach to drive the entire generation process is that it forces a certain ordering of decisions about the surface structure of the sentence, that must be followed by the system. For example, the active-passive decision must be made at the start, even though more important decisions regarding the overall form of the content and the lexical choices might seem more appropriate to consider first. Thus a syntax-driven generator might be awkward to design or would involve considerable backtracking to redesign the sentence structure when new constraints are found. Given the high cost of this operation, it is not striking to see that Shapiro's ATN never backs up, but it carefully proceeds by incremental refinement and the posting of constraints. However, the difficulty of decoupling perception from action still remains as the most significant deficit of the ATN designs, especially when viewed as a planner.

2.2 Planning Components

The early work in AI on the relationship of plans and goals to language was done not in the area of generation, but rather in the area of understanding. [Bruce 1975] carried out some of the initial research that set the stage for true speech-act planning. He was based on the fact that the task of understanding a sentence is not only a process of recovering its meaning, but also of interpreting the speaker´s intentions in producing it. Later work in the understanding of simple stories [Schank 1977] has recognized the need for reasoning about the underlying intentions of agents. Furthermore, [Wilensky 1978] developed a model in which agents could plan various actions, including asking and telling. The model was used to understand goal-based stories.

[Meehan 1977] extended these early ideas about planning to the design of a system that produced simple stories. Meehan´s system was not a language-generation program since it did not produce any actual English sentences. What it did was to compose formal descriptions of sort stories about different agents who made plans to achieve their goals and could be frustrated by various situations and events. The agent´s plans included actions of telling, asking and persuading.

[Cohen 1978] was interested in the interaction between planning and generation. He addressed the problem of planning speech acts in response to a user´s question. Part of the problem involved determining which speech act (e.g., inform, request, etc) was most appropriate for the response. The implemented system, OSCAR, was capable of selecting speech acts, specifying the agents involved, and the propositional content of the act, but it did not produce actual English output. In later work, [Cohen 1981, 1984] proposes the use of the planning formalism for deciding upon appropriate referential descriptions.

Further work in the planning area has been done by [Cohen and Perrault 1979], [Allen and Perrault 1980], [Perrault and Allen 1980], [Allen 1983, 1984], and [Cohen and Levesque 1985] on analyzing and recognizing intentions from NL utterances. This work enabled a number of intention-recognition heuristics and a rigorous speech act theory to be developed.

The utterance planning research, as well as such other speech act planning work as that of Allen, Cohen, and Perrault, was tested in domains that are fundamentally task-oriented -- e.g., as in performing some cooperative problem-solving task or assisting a customer at an information booth. This work leaves open the question as to whether planning and problem-solving techniques are also useful in less well-structured domains. [Hobbs and Evans 1980] examine the goal structure in a "small talk" dialog and conclude that in fact they are. The goals that arise are of different nature --

more social goals are involved for which formal description is difficult. Nevertheless, careful examination reveals that there is the same goal directed behavior as it is apparrent in the more well-structured domains.

2.3 Two early Generation Systems (*PROTEUS* and *KDS*) that use both a Planning and a Linguistic Component

PROTEUS was an early system implemented by [Davey 1979] within the context of a tic-tac-toe game and provided a running commentary on the game. The capabilities of the system span, in some sense, both the planning and linguistic component, though what needs to be said is given by the moves made by game participants. Nevertheless, the system incorporates the rhetorical principle that one should put in a text only the most salient information in a situation, (e.g., missed opportunities or forks), while leaving the other information to be communicated implicitly by inference.

PROTEUS first analyzes tic-tac-toe moves in terms of threats, blocks, etc., providing thus input to a planning facility that selects the best level of description for each move (e.g., "block" vs "fork"). The planner then groups the moves two or three at a time into sentences according to what game-level relationship seems to provide the best description of their motivation. A *systemic grammar* [Hudson 1971] and a realization (linguistic) facility then takes the described and grouped moves, works out the details of their form as English sentences, and produces the words of the text.

Thus, Davey´s system can use the content and rhetorical structure to influence the surface text, but does this only on the basis of domain-dependent concepts. An important aspect of his program is the capability for omitting details from the text based on dependencies between known facts: some facts dominate others in the choice of what to say.

[Mann and Moore 1979, 1981] and [Mann 1982], on the other hand, developed a system called Knowledge Delivery System (KDS) which could produce a multi-sentential text providing instructions about what to do in case of a fire alarm. The system first extracts the relevant knowledge from the notation given to it, and divides that knowledge into small expressible units, called *fragments* . A kind of planner is then responsible for selecting the representational style of the text and also for imposing the gross organization onto the text according to that style. Propositions identified by a kind of user´s model as known to the user, are not expressed. The system relies on the interesting *hill-climbing* techniques which have three responsibilities: to compose complex fragments (sentences) from simple ones, to judge relative quality among the units resulting from composition, and based on those judgement to repeatedly improve the set of fragments so that it is of the highest overall quality. Finally, a very simple *surface sentence maker* creates the sentences of the final text out of fragments.

Comparing the two systems we note that **they have very little in common**. In particular,

(1) KDS, using the hill-climbing techniques, can obtain sentence scoring and a quality-based selection of how to say things; PROTEUS has no counterpart.

(2) PROTEUS has a sophisticated grammar, a systemic grammar, while the *surface sentence maker* of KDS is a quite simple counterpart.

(3) PROTEUS has only a dynamic, redundancy-based knowledge filtering, while the filtering in KDS is based on an explicit model of the user's knowledge which identifies principally static, foreknown information which consequently is not expressed.

(4) PROTEUS takes a great advantage of the rhetorical principle that puts in the text only the most salient information in a situation; KDS does not use it.

In spite of the differences between the two systems, some of their principal techniques appear to be compatible, and the combination of them would surely produce better text than either system alone.

For example, the hill-climbing and sentence scoring techniques of KDS could be both introduced into PROTEUS and give it great improvement, especially in avoiding the long awkward referring phrases that PROTEUS produced. The system could detect the excessively long constructs and give them lower scores, leading to choice of shorter sentences in those cases.

On the other hand, the systemic grammar of PROTEUS could be introduced into KDS to great advantage. Moreover, as in PROTEUS, KDS could also take advantage of the most salient information principle that it would immediately allow acquisition of the basic propositions and could consequently improve the text significantly.

A final remark can be made on the exchange of knowledge domains between the two systems: The tic-tac-toe domain would fit easily into KDS, but the fire-crisis domain would be too complex for PROTEUS, because PROTEUS lacks the necessary text-organization methods.

In conclusion, the early generation systems discussed in this section share the primitive state of the computer-based discourse-generation art as it was in the late 1970s (or in early 1980s). Their processes are primarily devoted to activities that go without notice among literate people. The deeper linguistic and rhetorical phenomena usually associated with the term "discourse" are hardly touched. These systems make little attempt at coherence, and they do not respond in any way to the coherence (or lack of it) which they achieve. Presupposition, topic, focus, theme, the proper role of inference, implicature, ellipsis, anaphora, direct and indirect speech act performance and other relevant concepts all go unpresented. Even worse, the underlying conceptual representation is extremely ad hoc and idiosyncratic, severely limiting the posibilities for using general knowledge of the semantics of the language.

Despite these deficiencies, these systems produce relatively smooth readable text. They have been significant principally as collections of methods which served as parts of more competent and general-purpose systems of the future. For example, the use of the systemic grammar, a user´s model, or means for text evaluation are techniques that have been taken into account or could be considered in later or future generation systems.

In the next section, we discuss the great explosion in the quantity and quality of research done in the fields of NL generation and computational linguistics, beginning in about 1982. This research resulted not only in the commencement of a competent performance by a computer but also in the development of a computational theory of the human capabilities for language (i.e., *versatility* and *creativity*), and the processes that people employ to enable them to be *versatile* and *creative* .

3. AN OVERVIEW OF LATER WORK IN GENERATION SYSTEMS

In this section we present a sketch of recent research in NL generation that gave rise to the development of more competent and general-purpose generation systems than in the past. Our purpose is to outline this work in generation and orient it according to the **planning**(*what to say*)/ **linguistic**(*how to say it*) division. We recall that there is no clean line between the two activities. However, we assume this division, since by classifying work which is directly related to one (or both) of these two problem areas, it helps us to build up a bibliography relevant to each area. As we note, there is relatively less work done in the planning than in the linguistic area.

Starting with **the area of planning**, one of the most important works on *generation based on planning* has been done by [Appelt 1985a] (the KAMP system). His approach covers both problem areas with more emphasis in the first one. Dealing with the issue of *text organization*, [McKeown 1985] (the TEXT system) has also been concerned with both problem areas, her research however, has been focused on issues concerning the content and planning of the generated text. Further extension of this work has been done by [Paris and Mckeown 1986]. Text organization has also concerned [Weiner 1980] and [Goguen et al. 1982] although they have primarily focused on explanations.

The most elaborated work on text organization (planning) has been done by [Mann 1984,1988], [Mann and Thompson 1985, 1986, 1987] and [Mathiessen and Thompson 1986] (the Rhetorical Structure Theory). Earlier, [Mann 1983a, 1983b] (the PENMAN system) and [Mann and Mathiessen 1985] (the NIGEL grammar) developed a generation system with both a planning and linguistic component. [Koit and Saluveer 1986] (the TARLUS system) dealt with the generation of natural language text in a dialog system whose components enable it to solve the problems of deciding "what to say", and then proceed to a treatment of the "how to say it" question.

Two of the most recent works devoted to the planning approach to language generation have been done by [Houghton 1989] and [Mellish and Evans 1989]. Finally, other work related to planning, as we also mentioned in section 2.2, has been done by [Cohen and Perrault 1979], [Allen and Perrault 1980], [Perrault and Allen 1980], [Cohen 1981, 1984], [Allen 1983, 1984], and [Cohen and Levesque 1985] who developed a rigorous logical framework to encode basic notions such as intention and reference.

Passing to the **linguistic area**, we see that there is quite a lot of work done on this component, in addition to the work mentioned above by Appelt, McKeown, Mann and Mathiessen, and Koit and Saluveer.

One of the researchers that has particularly concentrated on the linguistic generation component and done a considerable amount of work is [McDonald 1980, 1983, 1984, 1985] as well as

[McDonald and Conklin 1982] and [McDonald and Pustejovsky 1985] (the MUMBLE system, the TAG formalism and examination of several psycholinguistic phenomena).

Grammars have always played a significant role in generation. Besides the ATN grammar that was the first one to be used in generation, other more flexible and effective formalisms have been widely used, each one playing a distinct role in the linguistic component. Thus, the *systemic grammar* [Halliday 1976,1985] and [Patten and Ritchie 1986] has been used by [Mann and Mathiessen 1985] to control the generation process, while the *functional unification grammar (FUG*) of [Kay 1979, 1984, 1985a, 1985b] has been only one of the many components that contributed to the design of the sentence. [Ritchie 1986] and [Lancel et al. 1986] have also been concerned with the problems and applications of the FUG formalism. Particularly, the TELEGRAM grammar developed by [Appelt 1983] and the realization component written by [Bossie 1981] for the generation system of [mcKeown 1985] have been succesful paradigms of the use of the FUG grammars in the linguistic generation component. *Lexical functional grammars* [Bresnan 1984] are close neighbors to FUGs, though not actually employed in generation yet.

Furthermore, [McDonald and Pustejovsky 1985] applied the *tree adjoining grammar* of [Joshi 1985], [[Kroch and Joshi 1985] and [Vijay-Shankar and Joshi 1985] to their generation theory of surface structure as an intermediate representation. Other generation researchers, besides McDonald, that have independently chosen to interpose a level of explicitly linguistic representation between the levels of the "message" (the semantic representation) and the words of the text have been [Kempen and Hoenkamp 1982], [Jacobs 1985b, 1986, 1987] (the KING system) and [Jacobs and Rau 1984] (the ACE knowledge representation framework), [Busemann 1984] (the SUTRA system), and [Buchberger and Horacek 1984] (the VIE-GEN system).

The last two systems, working in German, have developed quite elaborated linguistic generation components due to the complexities of the rich inflectional system of the German language. [Luckhardt 1986] has also focused on the German language, but discusses generation as part of mechanical translation. The Japanese generator of [Ishizaki 1984] could in some sense be regarded as a mechanical translation too. However, it is actually working from completely conceptual representation.

[Danlos 1984, 1985, 1986, 1987a, 1987b, 1989], [Danlos and Namer 1988] and [Lancel et al. 1988] have done an important research on generation of texts in Romance languages. Danlos has made use of a particular grammar, called *"discourse grammar"* that integrates conceptual and linguistic decisions, as well as of a specific lexicon, called *"lexicon-grammar"* [Gross 1986], which describes the syntactic properties of terms of the language under use.

Finally, an important tool used by various generation researchers has been the *phrasal lexicon* of [Becker 1975] that proved to be the best solution to the problems they faced on lexical choice of complex and idiomatic terms. One of the most important works in this area has been done by [Hovy 1985, 1986a, 1987a, 1987b, 1988, 1990] (the PAULINE system) who focuses on the form that a generator´s linguistic knowledge should take. He finally presents a model where all linguistic knowledge, grammatical and lexical should be included in the lexicon. [Kukich 1984] developed a system, called ANA, that organizes a large phrasal lexicon as a set of production rules in which semantic and linguistic knowledge are used in a highly integrated fashion during the generation of the text. PHRED, a generation system developed by [Jacobs 1985a], is also centered on the use of preformed, productive phrases. Since it was constructed to share a linguistic knowledge base with its counterpart component PHRAN (a NL analyzer), it eliminates the redundancy of having separate grammars and lexicons for input and output.

It is not in the scope of this paper to present a detailed discussion of all these systems. It would have been actually impossible. Rather, we present some of those as the most representative and contributive works in NL generation research. We concentrate more on two of the most important areas in which NL generation is applied: generation based on planning and text (or discourse) organization and generation. This is done because a considerable work on these two areas has initiated the first steps that really opened wide roads of further research and development in the generation field.

For this reason, a comparative and evaluative study of these systems, rather a descriptive one, is done in order to identify the important issues addressed and developed so far and what else remains to be done so that we obtain more competent and more general generation systems. We begin first by discussing generation based on planning.

4. LANGUAGE GENERATION BASED ON PLANNING: THE KAMP SYSTEM

[Appelt 1985a]

A significant work on planning and generation has been done by Appelt who developed a theory of language generation based on planning. To illustrate this theory, he examined the problem of planning referring expressions in detail. A such, a theory based on planning makes it possible for one: to account for noun phrases that refer, that inform the hearer of additional information, and that are coordinated with the speaker´s physical actions to clarify his communicative intent; The theory is embodied in a computer system called **KAMP** -- which is an acronym for **K**nowledge **A**nd **M**odalities **P**lanner -- which plans both physical and linguistic actions, given a high-level description of the speaker´s goals. In other words, the system´s purpose is to plan and produce one or more utterances that satisfy multiple goals.

KAMP is a hierarchical-planning system that uses a nonlinear-representation of plans, called a *procedural network* [Sacerdoti 1977]. This approach regards decisions about *"what to say"* and *"how to say it"* as two facets of the **same** overall process, and *recognizes the interactions between them.* The planning of an action appropriate for a given situation necessitates consideration of several crucial elements: *different kinds of goals that are satisfied by the utterances* (physical, knowledge-state, discourse and social goals); *the knowledge of the hearer; general knowledge about the world; linguistic knowledge* (the constraints imposed by the syntax of the language). The utterance planner can integrate these diverse knowledge sources to arrive at a plan involving abstract specifications of speech acts and then, finally, produce English sentences. Instead of regarding the hearer as the mere consumer of a message, the utterance planner treats him/her as an active participant in the communication process.

Appelt was principally faced with the problem of explanation of behavior and the aim of his research was, at least in part, to explain how particular mental states account for particular actions. In other words, how holding a particular set of beliefs and intentions results in an agent´s making a particular utterance. Therefore, he proceeded by constructing a formal, computational theory (along with a system that embodies this theory) in which it is possible to represent mental states and provide an explicit mechanism whereby these states determine actions.

Nevertheless, to provide a complete computational theory of speech acts and the use of natural language is a very high level and complex task, and this research falls far away of this ultimate goal. The principal significant **contributions** until now have been the establishment of a framework for such a complete theory, and an initial examination of the major problems in representing knowledge about knowledge, planning, the axiomatization of illocutionary acts, surface speech acts, and focusing actions so the accounts for all them can be unified and incorporated into a single computational model.

The KAMP system is **not** intended to be a cognitive model **nor** a psychological model of human behavior, although it may reflect some of the aspects of human language processing.

In particular, Appelt considers only the planning of concept activation actions in which the speaker and hearer **mutually know** some facts about the object of the concept activation at the time of the speaker´s utterance. An adequate formal theory of concept activation actions accounts for the following phenomena: (1) how speakers reason about mutual knowledge to arrive at a description of the object; (2) how speakers use non-linguistic means (e.g., pointing) to contribute toward the satisfaction of a goal to activate a concept; and (3) how speakers plan noun phrases that satisfy goals in addition to reference.

The logic that is used for the formal description of referring actions is based on the possible-worlds-semantics approach to reasoning about knowledge, intention, and action of [Moore 1977, 1985]. At this point, an interesting argument could be to use a **theory of belief** rather than a **theory of knowledge** as the basis of an adequate theory of language planning. Although this is a valid point, an adequate theory of belief is difficult to formalize, because once one admits the possibility of holding beliefs that are not true of the world, a theory of belief revision and truth maintenance is required. Anyway, it is true that Moore´s logic of knowledge can be transformed into a logic of belief by approximately weakening the axioms. However, without addressing all the problems associated with belief and justification, one has really accomplished little else besides changing the word **Know** to **Believe**. Because a detailed study of reasoning about belief was beyond the scope of Appelt´s research, he finally considered only axioms using **Know** as a first approximation to the best theory.

Adapting KAMP from a general-purpose hierarchical planner to a language planner involved **techniques** or capabilities such as: the axiomatization of various linguistic and physical actions (illocutionary acts, surface speech acts, concept activation, describing and pointing) in terms of the possible-worlds-formalism, the development of *action summaries* descriptions which summarize the preconditions and effects of actions, and the use of special procedures called *plan critics* to examine the plan globally and to determine interactions between proposed actions. The result of incorporating these capabilities into KAMP is a system capable of producing English sentences with complex referring expresions as part of an agent´s plan.

Furthermore, the linguistic capabilities of KAMP have been considerably extended by the use of a functional unification grammar, called TELEGRAM [Appelt 1983], that served as a coherent grammatical representation and a clean interface between the linguistic and planning components of the generation system, while still permitting adequate communication between them.

Employment of *action subsumption* to construct utterances that satisfy multiple goals requires the axiomatization of a great diversity of action types: surface speech acts, concept activation actions, and focusing. In addition, another important advance represented by KAMP is a means of coordinating physical and linguistic actions that enables the planning of pointing actions to communicate the intention to refer.

KAMP has proven to be a useful tool for the investigation of planning referring expressions and utterances in general, and promises to be useful in developing a speech-act theory to account for many aspects of natural language use. In this initial study, Appelt has emphasized his work on a broad class of issues instead of a few in depth; hence a fairly large number of problems, while *touched upon* , remain as topics for future research.

The division of actions into levels of absttraction is crucial to the logical separation of communicative acts from the utterances that realize them. One major area of research is the formalization of the relationship among actions in the abstraction hierarchy. More specifically: intentions behind surface speech acts should be recognized; the basic idea is that the inference of intentions, given a speech act and some knowledge about the speaker, can be classified according to several patterns. Also, the problem of a more rigorous formal description of action subsumption is needed to be completely resolved, since the treatment of action subsumption by KAMP is one of the least well formalized aspects of the entire system.

In addition to the illocutionary acts such as inform or request, there are many other illocutionary act types and their relation to surface speech acts, including *commit, declare and express* , that have not been incorporated into KAMP´s axioms or action summaries. Enlarging the repertoire of linguistic actions in KAMP will require fundamental research in representation of knowledge about concepts such as permission and obligation in addition to examination of the speech acts themselves.

KAMP´s analysis of referring expressions is extremely limited since KAMP is only capable of referring to objects about which the speaker and hearer share (mutual) knowledge at the time the utterance is planned, and for which the speaker intends the hearer to pick out the individual to which he/she is referring. Because speakers frequently use descriptions that are **not** mutually believed with the intention that the hearer match the description to objects in the world to pick out the intended one, this is a large gap that needs to be filled.

KAMP is capable of planning and producing **short** multisentence discourses, but there is no attempt to reason about the discourse as a whole. The discourse is coherent, because sentences are part of the same overall plan, rather than from being the realization of any discourse-level actions. The development of such discourse-level planning operators is an important area for future research. Since a good discourse planning theory was not available, Appelt did not deal with *global* *focusing* actions. On the other hand, he was able to axiomatize the local changes resulting from *immediate focusing* [Sidner 1982] or *centering* [Grosz et al. 1983].

Moreover, in the case of complex goals involving a large number of propositions that the hearer is supposed to believe, Appelt´s process becomes inadequeate, and some type of higher-level organization is needed. One particularly promising type of higher-level organization of text is found in the form of *discourse schemata* of [McKeown 1985] -- see section 6 -- and the *rhetorical structure shemata* of [Mann and Mathiessen 1987] -- see section 7. Such discourse schemata could be possibly treated by a planner as very high level actions. However, further research is needed to demonstrate that this assumption is true. Planning such rhetorical actions makes it necessary to explicitly represent the speaker´s and hearer´s beliefs about concepts such as the topic of the discourse and the relationship between propositions in a discourse. This analysis is high speculative, but appears to hold promise that an utterance planning sytem will eventually be able to treat extended discourse succesfully.

Finally a very important concern has been to make KAMP theory useful in the design of practical computer systems. Given that a great deal of deduction is necessary for KAMP, it should at least quarantee that the axioms and the deduction system are as efficient as possible. However, a lot of improvement is needed to be done in the current deduction system. The current axiomatization of the possible-world-semantics of the **Know** operator makes very wasteful use of antecedent rules that cause a large number of irrelevant inferences to be made during the course of a proof. The current implementation does not take advantage of well-known techniques such as structure sharing and indexing that could be used to reduce some of the computational effort required. One solution to the problem could be the development of a new axiomatization of semantics that does not depend on forward chaining rules, but is still structured so that rules can be indexed and found quickly when needed.

The efficiency issues and the computational costs associated with KAMP almost certainly prevent the practical application of its problem solving approach to language generation in the near future. However, the theoretical ideas are important, since it seems from the examination of dialogs that reasoning processes similar to those modeled by KAMP must be followed by speakers when they produce utterances. It remains a topic for future research to determine how the ideas presented in this project can be applied at a reasonable cost.

Concluding, KAMP is evidence that a general planning mechanism, suitably modified, is a feasible design for a system that plans communicative acts and generate sentences. Although many problems remain to be solved, the basic system design can provide a foundation from which future research in natural language generation can proceed.

5. OTHER CONSIDERATIONS ON PLANNING AND GENERATION

In the previous section Appelt focused his research on how speakers **plan** referring expressions that can be coordinated with physical actions and that may satisfy multiple goals. His system KAMP is based on a theory of speech acts and has to be capable of general reasoning about agent´s beliefs and intentions. In other words, the task of the generation system is to reason from speaker´s goals to produce a plan which in turn is refined until an English sentence is completely specified.

A distinct work done in the area of planning and generation is the development of a system by [Mellish and Evans 1989] which starts with a non-linear plan structure produced by an AI planning program and generates natural language text explaining how to execute the plan. Examination of this work revealed further advantages and deficiencies from the use of the domain of plans in natural language generation.

In particular, providing explanations is a practical reason for considering ways of generating natural language from plans. There are also theoretical reasons why plans are a good domain for studying natural language generation: Although there may be a great deal of material in a given plan, there is a kind of agreement among planning researchers on what short of information a plan is likely to contain. Thus, it is possible that interesting general principles about producing explanations of plans can be formulated, independently of the domains in which the plans are produced. This property, of providing a relatively formally defined and yet domain-independent input, **makes plans very attractive from a natural language generation point of view.**

Nevertheless, although the domain of plans seems *a priori* to provide rich structure that a natural language generator can use, in practice a plan that is generated without the production of explanations in mind rarely contains the kinds of information that would yield an interesting natural language account. Indeed, deficiencies of a planner and/or its plans can explain some of the deficiencies presented in natural language texts. Some problems stem from the use of plan operators not designed with text generation in mind, and can be solved within the scope of the planning system. More serious are the problems that arise because of deficiencies in the planning methodology itself.

More specifically, the planner´s representation may play a significant role in the coherence of the text which depends considerably by the prudent use of pronouns. To use pronominalization, one needs to be able to determine that the same domain object is being mentioned several times and this is accomplished only by a type of plan representation that accurately supports the representation of domain objects. A planner that cannot make use of properties of a general action results in having its representation of actions at a less finer level of *granularity* than that required to produce good text. Moreover, plans may only encode very weak information about causality and temporal relationships. Again, this problem can be thought of as a mismatch between the granularity of the representation used for planning and that needed to exploit the facilities of the natural language.

The effect of the granularity problem can be reduced by allowing the plan generator to provide deeper information about the internal structure of actions and states through domain-dependent rewrite rules -- to talk about repeated actions, for instance. However, the more one relies on domain-dependent rewrite rules for good text, the less one can claim to have a domain-independent basis for generating text from plans.

Domain-dependent rewrite rules can be regarded as a way of adorning the planner´s output to match up better with the requirements for generation. The basic *conceptual framework* of the plan is, however, something that cannot be changed unless the generator itself is to start doing some of the planning (as Appelt´s generator is actually doing). Assuming that there is some point in distinguishing the planner from the generator, the generator is therefore sometimes faced with a mismatch between self-evident concepts in the conceptual framework of the planner and those concepts that can be expressed simply in natural language.

In fact, since the planner regards all primitive actions as essentially instantaneous, it is in some sense incorrect to express the planner´s recommendations in this way *in all cases* . In other words, AI planners have a restricted view of the world that is hard to match up with the normal semantics of natural language expressions. Thus constructs that are primitive to the planner may be only clumsily or misleadingly expressed in natural language. It might therefore be suggested that a generator working from plans could and should always try to convey the plan semantics accurately, even if this involves an excessively detailed and unnatural prose.

Furthermore, the justification structure built by the planner -- its *explanatory power* -- could also help one to explain why the given actions, with the ordering described, are the right ones to achieve the plan´s goal. Unfortunately, the notion of "plan" adopted in many cases only makes reference to the succesful actions, even though the plan generator may have spent a lot of time exploiting other posibilities that did not work out. As a consequence, the generated texts, for instance, may only tell one *what to do* and not *what not to do* . It might therefore be appropriate, in future work, to consider natural language generation based on the *trace* of a planning system, rather than on the final results.

Similarly, in many other texts produced by a generation system based on plans, the reader is told **what** *to do* but is given no illumination as to **why** *things have to be done in this way* . Unfortunately, although in principle every plan is justified by earlier actions achieving the preconditions of later actions, many plans do not contain this information in a useful form. Such problems are, of course, similar to those that [Swartout 1983] encountered in producing explanations from expert systems. Moreover, one cannot necessarily expect machine-generated plans to come at the right level of detail to be really useful to a human being. This is because a plan with a rich explanatory power could be very large, and it would be well beyond the state of the art for such a plan to be produced automatically. In addition to it, such a plan would undoubtably contain a lot of information that could be redundant or obvious to a human reader and hence of no interest whatsoever.

Finally, a generation system may be conditioned by *arbitrary planner restrictions* : As is typical with application programs, most planners have particular features that represent non-standard or new approaches to certain situations. This fact means that any natural language generator using plans as input must customize itself somewhat to the peculiarities of the particular planner it is working with. For instance, one problem peculiar to a planner´s representation language could be the manner in which preconditions are specified.

There are also other problems in natural language generation from plans, not entirely due to deficiencies in the plans one is working on. One of these may result from *relying on plan structure* , i.e., using only information that the planner itself understands, in order to build a domain-independent system to generate text from plans. Thus, one could think that the plan abstraction hierarchy would be a useful source of information about how the text should be organized. However, there seem to be general differences between the kind of abstraction useful to a planner and the kind useful to a text generator. This fact makes clear that the planner´s abstraction hierarchy alone is not fully adequate for text generation. Whether we can device general principles for producing alternative decompositions of plans more suitable for text generation remains an open research area.

Furthermore, the raw material we can gain from plans has the tendency to lead to *repetitiveness* in the text. Even if we managed to enrich the plan representations suitably, however, the generator would still be deficient when the input really is uniform. In particular, the uniformity of the text output often leads to unwanted ambiguities, simply because of the lack of variation in the stylistic devices used. Our structure-building should not be based solely on *local* patterns in the message; although, we can gain some improvement in our system by having the choice of structure-building rules be determined partially by some random factor, the problem of repetition can only be solved by a *global* planning of the text.

In conclusion, plans produced by an AI planner could be unsatisfactory for the natural language generation task. Therefore, we must specify what we would like a plan to look like and contain, within the general constraint that a planning system could reasonably produce it, so that to produce better generation-from-plans systems. Another main area of research besides planning is the development of discourse stategies for text generation that we examine in the next section.

6. A TEXT GENERATION THEORY AND METHOD: THE TEXT SYSTEM

[MCKeown 1985]

In this section, we discuss and criticize an important text generation theory and method developed by McKeown whose work was the first real attempt in the history of generation research to deal with principles of discourse structure and thus provide a good theory and base for text organization. For this purpose, she developed a computational model of discourse strategies and focus rules to guide the generation process in its decisions about *what to say next.* Furthermore, we spend a great deal discussing the addition of a user model to McKeown´s generation system which is viewed as a fundamental prerequisite to the overall improvement of the system.

McKeown was concerned with how to produce text in response to a given communicative goal. She primarily focused her research on determining *what to include in the text* and *how to organize this information* so that it can be easily understood. The approach she has taken towards text generation is based on two fundamental hypotheses about the production of text: (1) that how information is organized (stored) in a system´s knowledge base and how a person describes that information need not be the same, and (2) that people have preconceived notions about the ways in which descriptions can be achieved.

Therefore, one major **contribution** of this work is a characterization of four different common text structures or "**schemata**" : *identification, constituency, attributive* and *contrastive* schemata. A computational model of these discourse strategies has then been developed and used to guide the generation process in its decisions about *what to say next* . This model is actually based on a previous analysis of natuarally occuring texts and represent strategies that can be used for three communicative goals: *define, describe,* and *compare* .

The schemata are characterized (composed) in terms of smaller functional units, called *rhetorical techniques* (or *predicate semantics*), each of which can then be instantiated against particular substructures in the system´s knowledge base. (What constitutes a legal instantiation is specified in the system and would vary with the type of knowledge base).

Another **contribution** of this work is McKeown´s extension of **focus-movement rules** of [Sidner 1982] for NL generation. McKeown uses these rules both to decide among alternative possible expansions of a schema and to influence syntactic realizations (e.g., the choice of active or passive voice, the choice between a pronoun or a definite descriptor). Finally, McKeown has implemented this model in **TEXT**, a system which generates paragraph-length responses to questions about database structure.

In particular, generation processing in TEXT was based on a division of two stages. The first stage -- the **strategic component**, completely determines the content and discourse structure (or text organization) and the output of this stage (the "message") is then passed to the second component -- the **tactical component**, that uses a FUG grammar to translate the message into English. This distinction allowed McKeown to focus on the problems and processes of the strategic component -- deciding what to say and how to organize it effectively.

Use of backtracking to provide cooperation between the strategic and tactical components

Although McKeown assumes a separation between the processes of deciding what to say and how to say it --unlike [Appelt 1985a] who found no separation between them -- a control structure which allows for **backtracking** between the tactical and strategic components would also be possible. Backtracking would thus give the system an additional constraint -- functioning in the same way as the focus constraints -- on *when* to include a proposition in the text. If a proposition could not be appropriately expressed at a particular point in the text, control would return to the strategic component which would decide what else it should say at this point.

In order to allow for such a control strategy, a minor change would be needed in the TEXT system flow of control. Instead of waiting until the text "message" is completely constructed before invoking the tactical component, the tactical component would be invoked for each proposition as it is added to the message. Backtracking, as it is used in an ATN mechanism to control the construction of the "message", could be exploited in TEXT to allow for new decisions in the text planning process to be taken when a proposition could not be appropriately translated by the tactical component.

Concluding, the approach that McKeown has actually taken clearly specifies how the planning of the text *influences* the realization of a "message" in natural language. On the other hand, backtracking would allow for processes that produce the surface expression to *influence* the planning of the discourse. *McKeown's division of the processes, however, does allow one to focus on textual organization, an issue which Appelt has not addressed.*

The use of schemata and focus rules in text organization and realization

We are now going to examine the particular techniques and tools McKeown considered for text generation as well as the performance capabilities and limitations of the TEXT system. At the same time, we discuss what possible improvements could be considered for the development of a more competent and general system.

The use of *schemata* (or *rhetorical strategies*) and a *focusing mechanism* provides a computationaly tractable method for determining what to include in a text and how to organize it. Moreover, the generation method, as described by McKeown, not only provides techniques for determining high-level choices about order and content, but it also provides information that can be used by the tactical component to make decisions about various surface level choices.

The use of schemata to encode knowledge about structure embodies a computational treatment of *rhetorical strategies* that can be used to guide the generation process. This reflects the hypothesis made above, that the generation process does not simply trace the knowledge representation to produce text, but instead uses communicative strategies that people are familiar with. This has the consequence that the same information can be described in different ways for different discourse purposes.

Therefore, schemata identify common means for effectivelly achieving certain discourse goals. They capture patterns of textual structure that are frequently used by a variety of people (and thus do not reflect individual variation in style) to succesfully communicate information for a particular purpose. Thus, they describe the norm for achieving given discourse goals, although they do not capture all the means for achieving these goals. Since they formally capture means that are used by people to create understandable texts, they can be used by a generation system to produce effective text.

Schemata do not function as a complete **grammar** of text in general. However, they do serve as a text grammar for the system since they describe all possible text structures that TEXT can generate.

McKeown has further shown that schemata could be **recursive**, describing text structure at many levels, and thus being much like a hierarchical plan for text. Schema recursion can be achieved by allowing each rhetorical strategy (predicate) in a schema to expand to either a single proposition (e.g., a sentence) or to a schema (e.g., a text sequence). A text generated by applying schemata recursively will be tree-structured, with a sub-tree occurring at each point where a predicate has been expanded into a schema. Propositions then occur at the leaves of the tree.

The schemata developed for the TEXT system encode structures suitable for description of **static** information. Other text types will require different kind of schemata and probably different kinds of predicates as well. Description of processes involving *cause and effect, reasoning involved in explanation, narrative* and *temporal* or *spatial relations* are all examples of different text types that will require additional examination of text to determine commonly used means of description and explanation. This may require the use of different communicative techniques as well as different combinations of these techniques. It is unclear in what ways strategies for generating these types will be like the discourse strategies developed in TEXT. Strategies for describing processes, cause and effect, and temporal sequences may well include directives that specify how to trace the underlying knowledge, unlike McKeown's major hypothesis that textual organization need not mirror the organization of the knowledge being described. It might be possible that these two different perspectives on textual strategy could and would be merged. However, exactly how this integration can come about remains to be seen.

One influence on the final structure of the text which was **not** taken into account in TEXT, is a **model of the user** (knowledge about the user's beliefs and goals). This information could be taken into account both in the initial selection of a schema and in deciding between alternatives within a schema (thus, a proposition could be selected if it was determined to be most appropriate for the given user). The choice among several alternatives at a schema choice point is currently determined by the *focus constraints* alone, which select the proposition that ties in most closely with the previous discourse. Conditions on focus provide the means for using tests based on the system's beliefs about the user to select a *predicate* most satisfactory for the user, given what he already knows. Analysis of other ways in which information about a particular user affects the level of detail needed in a response has been made by [Paris 1984].

Furthermore, although TEXT is capable of performing recursion in some instances, **full recursion** is not currently implemented. The use of full recursion would mean the system would have the ability to provide different responses to the same question which vary by level of detail. The development of full recursion, as well as the question of *when* recursion is necessary are related to the incorporation of a **user model**. It is necessary to investigate the influences on level of detail [Paris 1984], some of which will be part of a user model, and incorporate these as conditions on recursion. In general, in order to develop a comprehensive theory on recursion a full user-model must be developed [Moore 1977, 1985], [Rich 1979], [Allen and Perrault 1980] and [Allen 1983].

Finally, schemata can be used to aid in the **interpretation** of natural language by determining the discourse goals of the user when interacting with the system. This information will be useful both in responding to individual questions and anticipating follow-up questions.

Focus of attention has been used in TEXT to constrain the choice of what to say next. The choices are constrained in two ways: (1) by **global focus**, and (2) by **immediate focus**. When deciding what to say next, the system need only consider information that is relevant to the question currently being asked, as opposed to all information in the knowledge base. This information is contained in the *relevant knowledge pool* and determines the system's *global focus* . After every generated utterance, the system is further limited by a set of *immediate focus* constraints on what it can say next.

Using global and immediate focus proves to provide a computationally tractable approach to the problem of deciding what to say next. Furthermore, it guarantees the production of a semantically cohesive text, while the interaction of focus constraints with the schemata allows for the construction of a greater variety of the paragragh-length responses to questions about database structure.

Focus information is also available for the tactical component to use in making decisions about relatively sophisticated linguistic devices. In the TEXT system, tests for pronominalization and varying syntactic structures (active, passive and there-insertion) on the basis of focus information have been implemented. Other syntactic constructs -- not actually implemented in TEXT -- which have been shown to be related to focus include: *it-extraposition* (e.g., *It* was Sam who left the door open), *left-dislocation* (e.g, *Sam* , I like *him*), and *topicalization* (e.g., *Sam* I like).

Several other linguistic effects which were neither implemented in the TEXT tactical component and can be achieved through the use of focus information include: *Parallel sentence structure* (used when focus remains the same from one sentence to next), *subordinate sentence structure* (used to combine two propositions into a single complex sentence, when focus shifts in a proposition to an item just introduced into conversation in the last proposition), and *semantic markers* or *textual connectives* (used to provide the needed link, when there are no referential links from one sentence to the next).

In the TEXT system as currently implemented, **global focus** remains unchanged through the course of an answer. For longer sequences of text, global focus may shift. Its implementation is dependent upon the implementation of full recursion which in turn requires the development of a user-model. That is, where more detail is required focus shifts to the details is presented. The implementation of shifting focus would require capabilities for expansion and stacking of focus in addition to the full development of the recursive mechanism.

Finally, [Sidner 1982] presents the use of *focus sets* : She notes that a discourse need not always focus on a single central concept. A speaker may decide to talk about several associated concepts at once and yet the resulting discourse is still coherent. She calls this type of phenomenon *"co-present foci"* . In TEXT, no account of the use of foci implicitly associated with focus was made, since this requires use of general world knowledge and an inferencing capability. A more manageable, but equally necessary, task that must eventually be addressed is the specification of conditions and influences on *when* to shift focus to associated items rather than to explicitly mentioned items. Such a specification would likely be part of an extension to the preference ordering on immediate focus shifts. More complex structures may be needed for some other situations: For example, a speaker may introduce an item into conversation, but specify that he/she will continue to talk about it at a later point [Joshi and Weinstein 1982].

*Further **limitations** of the implemented system*

(1) The lack of an **inferencing capability**. This means that the TEXT system is only capable of talking about information wich is explicitly encoded in the knowledge base. This fact makes the knowledge representation and content an important issue in generation research in general, since it limits what the generation component is able to talk about unless an extensive inferencing component is available. The inclusion of an inferencing capability would allow the system to generate additional information from what is known which might be appropriate in different situations.

(2) An important limitation of the TEXT system is the lack of specific information about the particular users of the system. This information could be used to vary detail (for example, not all users need to know the same information about a description or identification of something), and include a finer determination of relevance (for example, if the user already knows about a description of an object, the information the system uses to specify this description should be compared against this existing knowledge, requiring the system to elaborate further on the description it should make), so that to be able to tailor responses for different individuals. Currently, the system assumes a static, casual and naive user and gears its responses to a level appropriate for that type of person. Inclusion of an **extensive user-model**, which can represent and infer information about different types of users, would allow for a significant improvement in the quality of text produced.

Although such a user-model was not implemented in TEXT, an analysis was done on the effect of previous discourse on the generation of responses. This analysis indicates how a generation system can make use of the previous dialog to tailor its responses to the current user and thus improve the quality of the response generated.

Adding a user model to the generation system

The addition of a user model to the generation system is needed for extensions in the use of discourse structure, relevancy and dictionary for generation:

(1) As regards **discourse structure**, a user model is necessary for the use of full schema recursion (when recursion is necessary and how recursion is achieved), i.e., for determining when more detail is necessary. This requires an extensive examination of the effect of the user´s knowledge state on the need for more detail, a research on exactly what information about the users can be determined from the discourse, and an analysis of how that information can be derived from the discourse.

Segmentation of the discourse is another issue that involves decisions about where sentence and paragraph boundaries should occur. The delineation of paragraphs within a text depends in part on the amount of information presented about a given topic and it therefore seems to be closely related to the recursive use of schemata.

(2) As regards **relevancy** of information, a user model could be used in determining which information to include in the relevant knowledge pool initially, as well as in expanding the relevant knowledge pool to include detail that is necessary for the current user. In these cases, knowledge of what the user knows and believes can help in determining exactly what information is relevant for the current user and in determining how much information is necessary given what he already knows. If relevance continually changes as the text is constructed, then a more sophisticated theory would likely interact with the recursive mechanism and global focus shifts.

For both discourse structure and relevancy, existing work on formalisms and methods for resoning about knowledge and belief such as [Moore 1977, 1985], [Allen and Perrault 1980], [Allen 1983] and [Appelt 1985a] would be helpful. However, particular attention must be addressed to the problem of how much to say. Other techniques for determining relevancy might rely on an account of salience [McDonald and Conklin 1982].

(3) As regards the **dictionary** (or **lexicon**) used for generation, if the dictionary has access to a user model and pragmatic information about lexical choice, its designer would have to do less work in selecting appropriate lexical items for the translation of the "message". Moreover, research on reasoning about referential choice [Cohen 1981], [Appelt 1985a], could further increase the sophistication of this component.

Practicality of the generation system

An important question is now logically following: How practical is the TEXT system for use in a real-world application? The system as it stands is not ready to be used in a real life situation. Faster response time is obviously one problem that must be solved before the TEXT system can be considered practical. Practicality, however, involves more than just questions of speed. Two very important issues should be also considered:

(1) The system should appropriately address the **needs of the user** and to do so, it must be augmented so that it can avoid *repetition* of information within a single session, and so that it can provide *shorter* or *longer* answers.

The problem of when to omit information because of its presence in previous discourse and when to repeat it, is not a simple question. Some issues involved in this problem are discussed during the analysis of previous discourse: tracking the *discourse history* involves remembering what has been said in a single session with a user and using that information when generating additional responses. The discourse history can be used to avoid repetition within a single session, and more importantly, to provide responses that contrast with previous answers.

In order to make decisions about the length of an answer (text), i.e., about how much information is sufficient for answering a particular user, an explicit model of the user´s beliefs is necessary. In cases where a user is clearly unfamiliar with the terms involved, a more detailed explanation is necessary. When it is clear that the user understands the concepts involved, longer text is unnecessary and cumbersome.

(2) The system should **cover all the questions** that a user might want to ask. In the case of TEXT, three classes of questions are addressed: requests for definitions, requests about available information, and questions about the differences between entities. Other classes of questions were not covered by TEXT, because the typical knowledge base used in natural language database systems did not encode the kind of information needed to answer such questions. These include: *questions about the system´s capabilities* , such as "Can the system handle ellipsis?", *questions about system processes* , such as "How is manufacturer´s cost determined?", and *questions involving counterfactuals* , such as "If John Brown were promoted to system analyst, what would his salary be?", etc.

Although the generation principles developed in the TEXT system were used for the tasks of providing definitions, describing information, and comparing entities, these principles can be applicable to all generation tasks. That is, the schemata capture discourse strategies in terms of text structure, and their representation does not rely on domain (or system) dependent concepts. Similarly, the focus constraints describe general preferences for what should be appropriately said next and apply to all situations in which coherent text must be produced. Thus, the methods for generation developed for the TEXT system are not specific to the database application and could be adapted for systems where generation of descriptions of static information is required.

Computer assisted instruction systems provide a good example of where generation of language could be enhanced by taking into account the best techniques for presenting information (descriptions or explanations). The generation of explanations in **expert systems** is another area where communicative techniques could be used to improve the quality of text output.

Concluding, significant **improvements** to the text can only be made by incorporating into the generation system facilities such as, *inferencing, varying detail,* finer determination of *relevance,* means for *evaluating* its own text, and a *user´s model* .

Despite its limitations a significant improvement in the quality of computer generated multi-sentential text was achieved by the TEXT system through the use of text structuring techniques to guide the generation process, and an account of focusing to ensure discourse coherency. Future research on issues concerning the content and organization of the generated text is needed in order to overcome the limitations of TEXT and develop a more general and competent system. Another energetic attempt on the central issue of text organization has been made by Mann and Thompson (see next section).

7. A MORE GENERAL THEORY OF TEXT ORGANIZATION

One of the most central issues of text generation involves text organization. Text has parts, arranged systematically and its organization is essential to its function. [Mann 1984,1988] and [Mann and Thompson 1985, 1986, 1987] addressed important questions concerning the nature of text organization or structure, the parts in which a text is arranged and the principles of arrangement. The result of this research has been the development of a theory of text organization, called **Rhetorical Structure Theory (RST)**.

RST is a *structural account* of the nature of discourse, characterizing the kind of units that can make up an extended text and the relationships between them. At the same time, RST is a *functional theory* , with structural units characterized according to the role they play in the forwarding of a speaker´s goals.

The development of this theory was the second attempt after McKeown that dealt with generation tasks which involved the need for extensive text organization. Its purpose was to classify different discourse strategies in a manner more general than that of McKeown. For this reason, a comparison between the two approaches is worthwhile in order to see the differences between them as well as other considerations, improvents and difficulties that result from realizing a more general theory of text organization and generation.

RST defines four classes of objects: **(1)** *Schemas* are configurations of text composed of two spans, one called the *nucleus* and the other the *satelite* . They are defined in terms of either a single relation or two relations. **(2)** *Schema Applications* are instances of schemas which correspond to the schema definitions according to a set of application conventions. **(3)** *Text structures* are compositions of schema applications. **(4)** *Relation definitions* determine how relations are defined and all of the above objects rest on these definitions of relations.

The kind of conventions presented above are all that is needed to analyze a text. The analyst finds spans for which the relations hold. When they combine into schema applications, such that those applications form a tree, then the tree is an RST structure.

The most significant differences between the text structure accounts of McKeown´s system TEXT and RST are:

Size of the largest structural unit: In TEXT the scale of schemas is the whole text. Texts are specified as whole units, whose structural possibilities are prespecified rather than created for the individual instance. Thus, TEXT has a simple schema-follower that is well documented. On the other hand, RST text structures are compositions of various schema applications. RST has a structure-building process, as one of several processes involved in developing the pre-grammatical plan for a text.

Dual use of TEXT schemas for: (1) what constitutes an adequate answer (*what to say*), and (2) text structure (*text planning*). RST separates these. Indeed, since TEXT relies on predetermined text structures based on study of prior texts, it does not need to reason extensively about the need for particular text elements, nor about how to structure them. *Its schemas encode both the task and the method for accomplishing the task* . This produces a workable but rigid text structuring method (*narrow task*) with the major advantages of being implementable and effective in its designed domain.

On the other hand, RST separates text structure building from other processes (e.g., processes for goal, knowledge and audience identification -- done before the building process -- as well as processes for lexical and syntactic choices, sentence scope or other grammatical relations -- done after the building process). This has the obvious advantages of flexibility and greater generality (*text diversity*), but produces a corresponding need to implement and coordinate a diversity of processes.

Because the RST approach creates structure based on the **immediate needs** of the text under construction, rather than past texts, it must reason much more about the user´s state and the ways that text affects that state. In other words, the procedure is based on *diverse and detailed appeals to the model of the user* . Moreover, the procedure contains *test conditions* determining the need for each RST schema.

Another characteristic of the predetermined text structure in TEXT is that the *optionality of elements is built in* at the whole-text level, there is an *absence of relations,* and the *order* of elements is *prespecified.* On the contrary, RST schemas are distinguished by their explicit nuclearity, its free ordering, and by the fact that any schema can be used to decompose any span.

In fact, the input to the procedure is a goal and the control structure is top-down and recursive, proceeding by progressive refinement to the single clause level. Moreover, the use of some schemas, e.g., Conditional, involve simple goal decomposition, whereas use of others, e.g., Evidence, involve addition of goals.

Finally, TEXT recognizes more detail and diversity in its predicates than RST does. TEXT is specifying structure with a much finer grain than RST. The reasons of TEXT functioning like this are the specific benefits TEXT is obtaining from this finer grain:

(1) *Selectivity that is useful for particular tasks* (e.g., it provides a superordinate class information for the Identification schema).

(2) *Selectivity on what to say when.*

(3) *Accomodation of details of search* (e.g., attributes with and without values).

(4) *Basis for selecting grammatical forms* for expressing the information.

We note that none of these benefits relate to the structure-building functions in RST´s approach; that is, all of these functions are located somewhere outside of RST structure-building. This could be interpreted in either two ways: that RST´s present level of detail is sufficient for structure building, but needs to be supplemented to support other processes, or that RST´s present level of detail is insufficient. However, evidence for any decision is at present unavailable.

Of all these differences described above, the crucial one is the difference between **structure derived from previous texts** (in TEXT) and **structure derived from goal pursuit** (in RST). These represent two competing and deep influences on language: *prior patterns or conventions*, and *immediate needs*. The first type includes fixed text and fill-in-the-blanks text as the simplest cases. The second includes many kinds of responsiveness to present needs beyond the sort represented by RST. Neither of these influences is sufficient by itself in a mature account of text structure or construction. The interactions between them are responsible for a major part of the complexity of language and communication.

8. FINAL COMMENTS ON NL GENERATION BASED ON PLANNING, ON TEXT ORGANIZATION AND ON LINGUISTIC ISSUES

In this section, we rely on the work done by [Danlos 1985] on text generation in Romance languages and we discuss the use of alternatives other than focus for a more correct treatment of language production based on a set of specific knowledge rather than on rules of general nature.

Danlos has done a quite extensive work in the field of NL generation. In particular, she was concerned with generating newspaper-style text in Romance languages and concentrated her work on the linguistic component of the generation system. Taking into account the rich inflectional system of Romance languages -- an issue which has not been seriously considered in English -- she had to specify the way of integrating conceptual, linguistic, syntactic and morphological decisions so that to deal with the complex interactions and interdependencies of various issues concerning both the text structure, and the grammatical and morphological aspects of language.

Both Appelt´s and McKeown´s approaches are criticized in contrast to the approach Danlos considered in text generation and especially that of McKeown´s focusing mechanism.

In particular, Appelt´s claim of integrating the speech (linguistic) and physical acts into the same plan, is a justified contradiction of the traditional approach according to which speech acts are the translation of the ideas the speaker wants to express. Nevertheless, when the generator of sentences is not situated in the framework of an oral dialog between two people in presence, one could adopt the traditional approach, that of separating the generation of text according to the questions *"what to say"* and *"how to say it"*. Therefore, it is left to see what happens when Appelt´s perspective on sentences generation (i.e., the kind of action taken in order to accomplish one or more goals) goes beyond the framework of an oral dialog between two people in presence.

Within the scope of natural language interfaces, computer responses should satisfy the goal of informing the user in the best possible "cooperative" way [Grice 1975]. Moreover, [Joshi et al. 1984] state that a speaker can be more "cooperative" in a dialog, if he/she can supply complementary information in an answer to a question in order to prevent possible false inferences. For example, *Is John a university professor? Yes, but he is not still a doctor* . (A simple *yes* could possibly make a hearer falsely believe that John is also a doctor). Similarly, *Is there a train to London? No, but there is a bus*. This means that the system recognizes the user´s goal (to go to London) and knowing that taking a train is not possible, it suggests another means (to take a bus).

McKeown´s focusing mechanism has been criticized by [Danlos 1985] who found several deficiencies in its use for text generation. In particular, the role of focus in **ordering the information** in a text does not always give the desired results. Indeed, use of a general rule such as: *If there is a choice between keeping the focus of the preceding sentence or shifting it to an* *element of the potential focus list, then decide on the second solution,* may result in badly formed texts, especially when one regards causal and/or temporal relations. For example, we suppose that after producing the sentence

(1) John was married to Mary **[focus=John] [potential focus=Mary]**

we would like to add the following information

(2) John was unfaithful to Mary (*cause*) **[focus=John]**

(3) Mary commited suicide (*effect*) **[focus=Mary]**

Therefore, after (1), one has to choose between keeping the same focus -- (1) followed by (2), or shift to a potential focus -- (1) followed by (3). If we make an appeal to the above rule, this one indicates that (1) should be followed by (3) and thus produces the text:

*John was married to Mary. She commited suicide. He was unfaithful to her. **[order (1)-(3)-(2)]**

which is bad in comparison with the following text where (1) is followed by (2):

John was married to Mary. He was unfaithful to her. She commited suicide. **[order (1)-(2)-(3)]**

In fact, this situation is not only specific to the above rule. Rules on focus are supposed to reflect abstract principles of general use. As a consequence, they do not take into account neither the *content* nor the *form* of sentences under question. In general, every decision requires the consideration of several specific factors: pragmatic, semantic, syntactic, lexical, or stylistic. Thus one cannot apply the above rule on issues of meaning (i.e., causal and temporal relations between the events described in (1), (2) and (3)).

Similarly, the role of focus is neither appropriate for deciding on **syntactic choices**, as McKeown considered it in order to choose between active and passive constructs or to take pronominalization decisions. In particular, if we consider the rule: *If the focus is on the AGENT, construct the sentence in the active; otherwise, if the focus is on the OBJECT, construct the sentence in the passive,* the choice between an active or passive construct cannot be made without taking into account the lexical element involved. This is because there are some verbs that do not permit the formation of the passive. For example,

John left home.

*Home was left by John.

Concluding, since the notion of focus reflects an indisputable intuition, one has to look for establishing conditions when applying the focus rules by examining the interactions between focus and pragmatic, semantic, syntactic, lexical and stylistic factors. Nevertheless, the effort needed to develop such a method would be tremendous. McKeown, herself, signals that she has not taken into account the interaction between discourse structures (schemata) and focus constraints, but that undeniably these two phenomena interfere.

Consequently, Danlos proposes that it would be reasonable to avoid using focus rules. Instead, she presents a *"discourse grammar"* that, in a more general manner, can handle the problems of ordering the information, sentence linearization and syntactic constructs. Unlike focus, the notion of a discourse grammar reflects the fact that the automatic treatment of natural language needs a set of specific knowledge and not rules of general nature. Her discourse grammar thus integrates conceptual, linguistic and stylistic decisions (unlike the discourse structures of McKeown and Mann that treat only the ordering of information), and interacts with a *"lexicon-grammar"* and other syntactic and morphological operations in order to offer a more correct treatment of natural language text generation -- correct, in the sense that it guarantees the production of texts which are grammatically and semantically appropriate.

Furthermore, the discourse structure of McKeown determines only the ordering of information independently of any linguistic consideration, and then passes it to a dictionary that decides which lexical items to use. Finally, the output of this module passes to a grammar which takes decisions on syntactic choices. In other words, such a modularity indicates that the ordering of information has always priority over lexical choices which in turn have priority over the choice of syntactic constructs. However, such an order of priorities is not justified.

In fact, let´s consider the case of a situation that describes a suicide, and suppose that the order ACT - EFFECT has been initially indicated by a discourse structure. If that information is then passed to a dictionary, this module should eliminate the verbal form *"to commit suicide"*, since this term cannot be placed after a description of an *act* (*Max jumped from the Empire Building committing suicide). However, the elimination of *"to commit suicide"* is not a good idea: It is desirable to use that verbal form to describe a suicide. For this reason, in this case the priority should be given to the lexical choice and not to the ordering of information. This does not mean, however, that we should always give the priority to lexical choices.

In conclusion, the decisions concerning the choice between lexical terms, ordering of information, segmentation into sentences, syntactic constructions and morphology are all dependent one upon each other, and there is *a priori* no reason to consider them in a fixed order: the priorities can vary from one case to another. Consequently, a generation algorithm should not be modularized but should allow an interaction between these decisions. In particular, it should determine those that must be sacrificed in case of contradiction between them. For example, in front of the choice:

- order of information = ACT - RESULT
- <RESULT> = N_0 to assasssinate N_1

which are incompatible (since the verb "to assassinate" cannot be placed after a description of an act), the algorithm should determine which of the two to sacrifice.

9. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we addressed four principal issues in NL generation which have led to the development of four main areas of research: **Discourse strategies for text organization,** representing common textual organizations found in naturally occurring texts and are used by a generator system to decide what information to include at each point in the text. **Generation based on planning formations**, focusing on the use of planning and reasoning about speaker and hearer beliefs to produce a text. **User modeling**, indicating the audience for whom text is generated. Information about the intended reader of the generated text is often used in conjuction with discourse strategies in order to determine what to say. Finally, an adequate **knowledge representation** is a prerequisite to later processes in text generation. A good representation of content is necessary to support generation before the process of generation begins.

As regards **discourse strategies**, we primarily focused on the TEXT system of McKeown which is capable of generating a variety of texts, since the large number of alternatives in a single strategy does not totally restrict text content. The approach of discourse strategies taken in TEXT, however, is limited in that it does not take into account constraints from the user of the system. Moreover, the semantics of the rhetorical predicates hold only for the database domain.

The Rhetorical Structure Theory of Mann and Thompson, on the other hand, embodies a set of rhetorical schemata for a wide variety of purposes in a manner more general than that of McKeown, but it is not currently fully implemented. They suggest that goal-pursuit techniques will be used to choose between the wide variety of options available within this approach. Thus new results are promised by this work from the interaction between speaker goals and strategies.

The use of **planning formalisms for language generation** is based on the necessity of reasoning about the intended speaker and hearer beliefs when generating language. Here, we focused on Appelt´s development of KAMP system which made two major contributions to NL generation: First, the characterization of language generation as a planning activity, and second, the development of a formal mechanism, relatively different from other generation systems, in which he made use of a logical representation of speaker and hearer beliefs and of a sound method of reasoning about plans and beliefs.

In KAMP´s limited task domain, it is possible to use a detailed model of speaker and hearer beliefs and determine through theorem proving whether the hearer has the information needed to carry out a task. In less well-defined domains such as tutoring or explanation production, it may be difficult to enumerate all relevant speaker and hearer beliefs and to formalize in detail the plans required to produce texts. These points raise the question of whether Appelt´s model for reasoning about belief (in particular, the use of a theorem prover) is the best approach. Finally, other work in this area that use similar formalisms, has been done by [Kronfeld 1986] whose system produces referring expressions -- noun phrases -- that accurately reflect certain distinctions in speaker intentions and the effect of the expressions on the hearer. Also, [Pollack 1986] presents a model of plan inference that distinguishes between the beliefs of the questioner and the respondent. By representing plans as mental concepts, separating out beliefs of the questioner and respondent about plans, Pollack´s model can infer invalid plans as well.

Research on **user modeling**, i.e., incorporating the effect of the reader (user) on text generation, is still in early stages as indicated by the lack of a unified approach to the problem. Recent work on this area has been done by [McCoy 1986] who identifies how to respond to different user misconceptions (i.e., incorrect user beliefs) by making use of discourse strategies. Also, [Paris 1985] identifies and uses different discourse strategies to produce descriptions of physical objects to users who range from domain novice to domain expert. [Hovy 1985, 1986a, 1986b] shows how information about the reader can influence the style of generated text, both in its emphasis (i.e., what is included or left out) and phrasing (i.e., how to say it).

Finally, most of the work done in user modeling dealt primarily with the effect of the user on selection of content. However, one would expect that how one chooses to realize a text (how to say it) would be influenced by the expected audience as well. Two works that touch on this topic and promise future results are: [Appelt 1985b] examines the influence of mutual beliefs on the production of noun phrases and [Hovy 1986a,b] notes the influence of the hearer-speaker relation (*"speaker affect"*) on overall textual style. Another advance in user modeling may come from the consideration of generation within the context of an ongoing dialog. This will allow the user to provide feedback indicating how well text has been understood.

As regards **knowledge representation**, two main issues arise from the standpoint of generation: The first, *generation adequacy,* concerns the content of a knowledge base, the second, concerns the *form* of the knowledge representation formalism. As we also saw in section 5, when we generate text from a plan structure produced by an AI planning program, the range of possible natural language constructs may be artificially limited by the shallowness of the planner´s representation.

Finally, Danlos notes that a variety of influences on *surface* decisions (e.g., lexical choice) and *conceptual* decisions (e.g, ordering of information) interact in unpredictable ways. Appelt also notes that there is no clean line between the *"what to say"* and *"how to say it"* questions. One major direction in current generation research is investigation into the nature of the interaction that must occur between these two generation components. In sum, language generation is a young but rapidly maturing field that offers a set of tools and approaches for practical application across various topics as well as a range of challenging open problems ready for further exploration.

10. BIBLIOGRAPHY

[Allen 1983] Allen, J.F., *Recognizing intentions from NL utterances* , in *Computational Models of Discourse* (M. Brady and R.C. Berwick, eds.) pp. 107-66, MIT Press, Cambridge, MA, 1983.

[Allen 1984] Allen, J.F., *Towards a general theory of action and time* , AI(23), 123-154, 1984.

[Allen and Perrault 1980] Allen, J.F. and Perrault, C.R., *Analyzing intention in utterances* , AI, 15, 143-78, 1980. Reprinted in *Readings of Natural Language Processing* (B.J. Grosz, K.S. Jones, B.L. Webber eds.), pp. 441-58, Morgan Kaufmann: Los Altos, 1986.

[Appelt 1983] Appelt, D.E., *TELEGRAM: A grammar formalism for language planning* , Proc. of the 8th IJCAI, pp. 595-599, 1983.

[Appelt 1985a] Appelt, D.E., *Studies in NL processing: Planning English sentences* , Cambridge, England: Cambridge, U. Press, 1985.

[Appelt 1985b] Appelt, D.E., *Some pragmatic issues in the planning of definite and indefinite noun phrases.*, in Proc. of the 23rd Ann. Meet. of the ACL, Chicago, IL., pp. 198-203, July 1985.

[Becker 1975] Becker, J., *The phrasal lexicon* , Proc. Theoretical issues in NL Processing (TINLAP-I), 60-64, 1975.

[Bossie 1981] Bossie S., *A tactical component for text generation: sentence generation using a functional grammar,* Univ. of Pennsylvania, TR: MS-CIS-81-5, Philadelphia, PA, 1981.

[Bresnan 1984] Bresnan, J.,*The mental representation of grammatical relations,* MIT Press, Cambridge, MA, 1984.

[Bruce 1975] Bruce, B.C., *Belief systems and language understanding,* BBN Technical Report, nº 2973, 1975.

[Buchberger and Horacek 1984] Buchberger, E. and Horacek, H., *VIE-GEN: A generator for German texts* , in *Natural Language Generation Systems* (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 166-204, 1988.

[Busemann 1984] Busemann, S., *Surface transformations during the generation of written German sentences* , in *Natural Language Generation Systems* , (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 98-165, 1988.

[Cohen 1978] Cohen, P., *On knowing what to say: Planning speech acts* , PhD thesis, University of Toronto, TR 118, 1978.

[Cohen 1981] Cohen, P., *The need for identification as a planned action* , Proc. of the 7^{th} IJCAI, pp. 31-36, 1981.

[Cohen 1984] Cohen, P., *Refering as requesting* , COLING - 84, 207-11, 1984.

[Cohen and Levesque 1985] Cohen P.R., Levesque H.J., *Speech acts and rationality,* in Proc. of the 23^{rd} ACL, 49-59, 1985.

[Cohen and Perrault 1979] Cohen P.R., Perrault C.R., *Elements of a plan-based theory of speech acts,* Cognitive Science, Vol. 3, pp. 177-212, 1979; (Reprinted In RNLP).

[Danlos 1984] Danlos, L., *Conceptual and linguistic decisions in generation* , Proc. of the COLING - 84, pp. 501-504, 1984.

[Danlos 1985] Danlos, L., *Génération automatique de textes en langues naturelles* , Masson, Paris, 1985.

[Danlos 1986] Danlos, L., *A robust syntactic component in generation* , in *Proccedings of the 3^{rd} International Workshop on Text Generation* (Gerard Kempen eds.), Nijmegen, The Netherlands, August, 1986.

[Danlos 1987a] Danlos, L., *The linguistic basis of text generation* , Cambridge University Press, Cambridge, 1987.

[Danlos 1987b] Danlos, L., *A French and English syntactic component for generation* , in *Natural Language Generation: New Results in Artificial Intelligence, Psychology and Linguistics* (Kempen G. eds.). Dortrecht/Boston, Martinus Nijhoff Publishers, 1987.

[Danlos 1989] Danlos, L., *Some pronominalization issues in generation of texts in Romance languages* , in *Procesamiento del Lenguaje Natural* , Actas de la IV Reunión Annual de la SEPLN, Boletín nº 7, Enero 1989.

[Danlos and Namer 1988] Danlos, L., Namer, F., *Morphology and cross dependencies in the synthesis of personal pronouns in Romance languages* , Proc. of the COLING - 88, Budapest, 1988.

[Davey 1979] Davey, A., *Discourse production,* Edinburgh Univ. Press, Edinburgh, 1979.

[Friedman 1969] Friedman, J., *Directed random generation of sentences,* Communications of the ACM, Vol. 12, 1969.

[Goguen et al. 1982] Gogeun J., Linde C., and Weiner J., *Reasoning and natural explanation,* unpublished proceedings of ISI Workshop on Explanation, 1982.

[Goldman 1975] Goldman, N.M., *Conceptual generation* , in *Conceptual Information Processing* (R.C. Schank eds.), Amsterdam: North-Holland, 1975.

[Grice 1975] Grice, H., *Logic and conversation,* in *Syntax and Semantics III: Speech Acts,* (Cole P. anf Morgan J.L., eds.), Academic Press, New York, 1975.

[Gross 1986] Gross, M., *Lexicon-Grammar: The representation of compound words* , in Proc. of the COLING - 86, 11^{th} International Conference on Computational Linguistics, Bonn, 1986.

[Grosz et al. 1983] Grosz B., Joshi A., and Weinstein S., *Providing a unified account of definite noun phrases in discourse* , Proc. of the Association for Computational Linguistics, pp. 44-50, June, 1983.

[Halliday 1976] Halliday, M.A.K., *System and function in language* , Oxford University Press, 1976.

[Halliday 1985] Halliday, M.A.K., *A short introduction to functional grammar* , London: Arnold, 1985.

[Hobbs and Evans 1980] Hobbs J, Evans D., *Conversation as planned behavior,* Cognitive Science, Vol. 4, pp. 349-377, 1980.

[Houghton 1989] Houghton, G., *A computationl model of discourse production* , Ablex, Norwood, N.J., 1989.

[Hovy 1985] Hovy, E., *Integrating text planning and production in generation* , Proc. of the 9th IJCAI, L.A., 848-851, 1985.

[Hovy 1986a] Hovy, E., *Putting affect into text* , Proc. of the 8th Cognitive Science Society Conference, Amherst, MA, 669-671, 1986.

[Hovy 1986b] Hovy, E., *Some pragmatic decision criteria in generation,* in *Proccedings of the 3rd International Workshop on Text Generation* (Gerard Kempen eds.), Nijmegen, The Netherlands, August, 1986.

[Hovy 1987a] Hovy, E., *Interpretation in generation* , Proc. of the AAAI, 1987.

[Hovy 1987b] Hovy, E., *Generating natural language under pragmatic constraints* , in *Journal of Pragmatics* , vol XI, nº 6, 1987.

[Hovy 1988] Hovy, E., *Generating language with a phrasal lexicon* , in *Natural Language Generation Systems* , (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 353-384, 1988.

[Hovy 1990] Hovy, E., *Pragmatics and natural language generation,* AI, 43, 2, pp. 153-197, May 1990.

[Hudson 1971] Hudson, R.A., *English complex sentences,* in North Holland Linguistic Series, Vol. 4, North Holland, London and Amsterdam, 1971.

[Ishizaki 1984] Ishizaki, S., *Generating Japanese text from conceptual representation,* in *Natural Language Generation Systems* , (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 256-279, 1988.

[Jacobs 1985a] Jacobs, P., *PHRED: A generator for N.L. interfaces* , Computational Linguistics, 11, 4, 219-242, 1985.

[Jacobs 1985b] Jacobs, P., *A knoweledge-based approach to language production* , PhD thesis, Berkeley Computer Science Department, TR 86/254, 1985.

[Jacobs 1986] Jacobs, P., *Knowledge structures for natural language generation* , Proc. of the COLING - 86, 1986.

[Jacobs 1987] Jacobs, P., *Knowledge-Intensive natural language generation* , AI, vol 33, nº 3, pp. 325-378, November 1987.

[Jacobs and Rau 1984] Jacobs P., Rau L., *ACE: associating language with meaning,* in Proc. of the 6^{th} European Conference of AI, Pisa Italy, 1984.

[Joshi 1985] Joshi, A., *TAG grammar* , in *NL Processing - Theoretical, Computational and Psychological Perspectives* (Dowty, Karttunen, Zwicky eds.), 1985.

[Joshi and Weinstein 1982] Joshi A., Weinstein S., *Control of inference: role of some aspects of discourse structure-centering,* in Proc. of the 7^{th} IJCAI, Vancouver, BC, August 1982.

[Joshi et al. 1984] Joshi A., Webber B., Weischbedel R., *Preventing false inferences,* in Proc. of the COLING-84, Stanford University, CA, 1984.

[Kay 1979] Kay, M., *Functional grammar* , Proc. of the Berkeley Linguistic Society, 1979.

[Kay 1984] Kay, M., *FUG: A formalism for machine translation* , Proc. of the COLING - 84, Stanford, CA, pp. 75-78, July, 1984. Also in Computational Linguistics, 1984.

[Kay 1985a] Kay, M., *Parsing in FUG* in *Natural Language Parsing* (Dowty D.R., Karttunen L., and Zwicky A., eds.), 251-278, NY: Cambridge, U. Press, 1985, (reprinted in RLNP).

[Kay 1985b] Kay, M., *Unification grammar,* in *NL Understanding and Logic Programming* (Dahl V., and Saint-Dizier P., eds.), pp. 233-40, North-Holland, Amsterdam, 1985.

[Kempen and Hoenkamp 1982] Kempen, G. and Hoenkamp, E., *Incremental sentence generation: Implications for the structure of a syntactic processor* , Proc. of the COLING - 82, Prague, August, 1982.

[Koit and Saluveer 1985] Koit M., and Saluveer M., *Generatring natural language text in a dialog system,* in Proc. of the COLING-86, Bonn, FRG, August, pp.576-580, 1986.

[Kroch and Joshi 1985] Kroch, T. and Joshi, A., *The linguistic relevance of Tree Adjoinjng Grammar* , Univ. of Pennsylvania, Dept. of Computer and Information Science, 1985.

[Kronfeld 1986] Kronfeld, A., *Donnellan's distinction and a computational model of reference,* in Proc. of the 24th Ann. Meet. of the ACL, New York, pp. 186-91, June 1986.

[Kukich 1984] Kukich, K., *Fluency in natural language reports* , in *Natural Language Generation Systems* (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 280-311, 1988.

[Lancel et al. 1986] Lancel J.M., Rousselot F., Simonin N., *A grammar used for parsing and generation* , Proc. of the COLING - 86, Bonn, FRG, August, pp. 536-539, 1986.

[Lancel et al. 1988] Lancel J.M., Otani M., Simonin N., Danlos L., *Sentence Parsing and generation with a semantic dictionary and a lexicon-grammar* , in Proc. of the COLING - 88, Budapest, 1988.

[Luckhardt 1986] Luckhardt H. D., *Generation of sentences from a syntactic deep structure with a semantic component,* in *Natural Language Generation Systems* (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 205-255, 1988.

[Mann 1982] Mann, W., *Text generation* , JACL, 8(2), 1982.

[Mann 1983a] Mann, W., *An overview of the Pennman text generation system* , Proc. of the AAAI, 261-265, 1983. Also in the 21st meeting of the ACL, 74-78, 1983.

[Mann 1983b] Mann, W., *An overview of the Nigel text generation grammar* , in the 21st meeting of the ACL, 79-84, 1983.

[Mann 1984] Mann, W., *Discourse structures for text generation* , in Proc. of the COLING - 84, Stanford University, CA , 1984.

[Mann 1988] Mann, W., *Text generation: The problem of text structure* , in *Natural Language Generation Systems* (McDonald D. and Bolc L. eds.), Springer-Verlag, New York Inc, pp. 47-68, 1988.

[Mann et al. 1982] Mann W., Bates M., Grosz B., McDonald D., McKeown K., and Swartout W., *Text generation: The state of the art and literature* , JACL, 8(2), 1982.

[Mann and Mathiessen 1985] Mann W., and Mathiessen C., *Nigel: A systemic grammar for text generation* , in *Systemic Perspectives on Discourse* (Freedle eds.), Ablex, Norwood, NJ, 1985.

[Mann and Moore 1979] Mann W., and Moore J., *A snapshot of KDS, a Knowledge Delivery System,* in Proc. of the 17th Ann. Meet. of the ACM, pp. 51-52, August 1979.

[Mann and Moore 1981] Mann W., and Moore J., *Computer generation of multi-paragraph English text* , JACL, 7(1), 1981.

[Mann and Thompson 1985] Mann W., and Thompson S., *Assertions from discourse structure* , in Proc. of the 11th Annual Meeting of the Berkeley Linguistic Society, Berkeley, 1985.

[Mann and Thompson 1986] Mann W., and Thompson S., *Relational propositions in discourse* , in *Discourse Processes* , 9(1), 57-90, Jan - March, 1986.

[Mann and Thompson 1987] Mann W., and Thompson S., *Rhetorical structure theory: A theory for text organization* , in *Discourse Structure* (Livia Polanyi eds.), Ablex, Norwood, NJ, 1987.

[Matthiessen and Thompson 1986] Matthiessen C., Thompson S., *The structure of discourse and subordination* , in *Clause Combining in Grammar and Discourse* (Haiman and Thompson eds.), Benjamins, Amsterdam, 1986.

[McCoy 1986] McCoy, K.F., *The ROMPER system: responding to object-related misconceptions using perspective,* in Proc. of the 24th Ann. Meet. of the ACL, New York, N.Y., pp. 97-105, June 1986.

[McDonald 1980] McDonald, D., *Natural language production as a process of decision making under constraint,* PhD thesis, Cambridge, Mass: MIT Press, 1980.

[McDonald 1983] McDonald, D., *NLG as a computational problem: An introduction* , in *Computational Models of Discourse* (M. Brady and R. Berwick eds.), MIT Press, Cambridge, MA, pp. 209-266, 1983.

[McDonald 1984] McDonald, D., *Description directed control: Its implications for NLG* , in *Computational Linguistics* (Cercone eds.), Plenum, NY, pp. 403-424, 1984.

[McDonald 1985] McDonald, D., *Description directed NLG* , Proc. of the 9th IJCAI, Los Angeles, CA, pp. 799-805, 1985.

[McDonald and Conklin 1982] McDonald D., and Conklin E.J., *Salience as a simplifying metaphor for NLG* , Proc. of the AAAI, 75-78, 1982.

[McDonald and Pustejovsky 1985] McDonald D., and Pustejovsky J., *TAGs as a grammatical formalism for generation* , Proc. of the ACL, Chicago, pp. 94-103, July 1985.

[McKeown 1985] McKeown K.,*Text generation* , Cambridge U. Press, Cambridge, UK, 1985.

[Meehan 1977] Meehan, J.R., *TALE-SPIN, an interactive program that write stories,* in Proc. of the 5th IJCAI, pp. 91-98, August 1977.

[Mellish and Evans 1989] Mellish C., and Evans R., *Natural language generation from plans,* Computational Linguistics, Vol. 15, nº 4, December 1989.

[Moore 1977] Moore, R.C., *Reasoning about knowledge and action,* in Proc. of IJCAI, pp. 223-227, 1977.

[Moore 1985] Moore, R.C., *A formal theory of knowledge and action,* in (J.R. Hobbs and J. Moore eds.), *Formal Theories of the Common Sense World (vol. 1),* Norwood, NJ: Ablex, 1985.

[Paris 1984] Paris, C., *Determining the level of expertise of a user in a question-answering system,* Techn. Rept., Columbia University, New York, 1984.

[Paris 1985] Paris, C., *Description strategies for naive and expert users,* in Proc. of the 23rd Ann. Meet. of the ACL, Chicago, IL, pp. 238-45, July 1985.

[Paris and McKeown 1986] Paris C., and McKeown K., *Discourse strategies for descriptions of complex physical objects* , in *Proccedings of the 3rd International Workshop on Text Generation* (Gerard Kempen eds.), Nijmegen, The Netherlands, August, 1986.

[Patten and Ritchie 1986] Patten T., and Ritchie G., *A formal model of systemic grammar* , in *Proccedings of the 3rd International Workshop on Text Generation* (Gerard Kempen eds.), Nijmegen, The Netherlands, August, 1986.

[Perrault and Allen 1980] Perrault C.R., and Allen J.F., *A plan-based analysis of indirect speech act* , JACL, 6(3), 167-182, 1980.

[Pollack 1986] Pollack, M.E., *Inferring domain plans in question-answering,* PhD thesis, Univ. of Pennsylvania, 1986.

[Rich 1979] Rich, E.A., *User modeling via stereotypes,* Cognitive Science, 3, 329-354, 1979.

[Ritchie 1986] Ritchie, G., *The computational complexity of sentence generation using FUG* , Proc. of the COLING - 86, Bonn, FRG, 25-29, August 1986.

[Sacerdoti 1977] Sacerdoti, E., *A structure for plans and behavior,* North-Holland, Amsterdam, 1977.

[Schank 1975] Schank, R.C., *Conceptual information processing* , North-Holland Publishing Company, 1975.

[Schank 1977] Schank, R., Abelson, R., *Scripts, plans goals, and understanding,* Laurence Erlbaum Associates, New Jersey, 1977.

[Scott et al. 1977] Scott C., Clancy W., Davis R., and Shortliffe E., *Explanation capabilities of production-based consultation systems,* JACL, Michrofiche 62, 1977.

[Shapiro 1982] Shapiro, S.C., *Generalized augmented transition network grammars for generation from semantic networks* , JACL, 8(1), 12-25, 1982.

[Shapiro 1987] Shapiro, S.C. (ed), *Encyclopedia of A.I.* , New York, Wiley, 1987.

[Sidner 1982] Sidner, C., *Focusing in the comprehension of definite anaphora* , in *Reading in natural language processing* (Grosz B.J., Jones K.S., Webber B.L., eds.), Morgan Kaufmann Publishers Inc., Los Altos, CA, pp. 363-394, 1986.

[Simmons and Slocum 1972] Simmons, R. and Slocum, J., *Generating English discourse from semantic networks* , Commun. of the ACM, 15, 10, 891-905, 1972.

[Swartout 1981] Swartout, W., *Producing explanations and justifications of expert consulting programs,* Massachusetts Institute of Technology, TR, 1981.

[Swartout 1983] Swartout, W., *XPLAIN: A system for creating and explaining expert consulting programs* , AI (21), 285-325, 1983.

[Vijay-Shankar and Joshi 1985] Vijay-Shankar, K. and Joshi, A., *Some computational properties of Tree Adjoining Grammars* , 23rd Annual Meeting of the ACL, pp. 82-93, 1985.

[Weiner 1980] Weiner, J.L., *BLAH, a system which explains its reasoning* , AI (15), pp. 19-48, 1980.

[Wilensky 1978] Wilensky R., *Understanding goal-based stories,* Yale University, Dept. of Computer Science, TR, nº 140, 1978.

[Winograd 1972] Winograd, T., *Understanding natural language,* Academic Press, New York, NY, 1972.