

Bolt Beranek and Newman Inc.



AD A131003

Report No. 5338

Discourse and Problem Solving

Diane Litman

July 1983

Prepared for:
Defense Advanced Research Projects Agency

DTIC FILE COPY



83 08 02 04 1

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER BBN Report No. 5338	2. GOVT ACCESSION NO. AD-A131 003	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DISCOURSE AND PROBLEM SOLVING		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER BBN Report No. 5338
7. AUTHOR(s) Diane Litman		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0378
9. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt Beranek and Newman Inc. 10 Moulton Street Cambridge, MA 02238		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, VA 22217		12. REPORT DATE July 1983
		13. NUMBER OF PAGES 48
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Distributed Artificial Intelligence, Computational Linguistics, Natural Language Understanding, Discourse, Planning, Goal Recognition, Tractable, Communicative Goals, ARGOT		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) > This report proposes a plan-based natural language system that incorporates knowledge of both plan and discourse structure of task-oriented dialogues. An initial representation of communicative (discourse) actions is discussed, in particular how to incorporate knowledge of legal moves as action effects rather than grammars. The subtle differences implicit in various surface realizations are also examined,		

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (cont'd.)

as well as the structure of these communicative actions in actual dialogues. It is suggested that both local and global discourse structures are necessary, (although analysis of the latter has been emphasized here). It is also suggested that planning models need to be extended to include two agent plan execution. Finally, a model of the goal recognition process is presented. Communicative and task knowledge work in parallel, one source dynamically taking control over the other and reducing the search space, depending on the kind of discourse (a task-oriented one, a conversation etc.). Communicative recognition is hypothesized to be simple, using the knowledge provided by the analysis of surface phenomena and task plan recognition.

Report No. 5338

DISCOURSE AND PROBLEM SOLVING

Diane Litman

July 1983

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



Prepared for:

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

ARPA Order No. 3414

Contract No. N00014-77-C-0378

Effective Date of Contract:
1 September 1977

Contract Expiration Date:
30 September 1984

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by ONR under Contract No. N00014-77-C-0378. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

This research was also supported by the National Science Foundation grant number IST-8210564 and ONR contract number N00014-80-C-0197 to the University of Rochester.

TABLE OF CONTENTS

	Page
1 Introduction	1
2 Background	4
3 Proposals	23
4 Summary	40
5 Acknowledgements	42
6 References	43
7 Appendix	45

CHAPTER 1

Introduction

Natural language research has increasingly emphasized the incorporation of discourse knowledge into understanding and generation systems. Consider a computer system capable of participating in an extended discourse such as Dialogue 1.¹

- (1) User: Could you mount a magtape for me?
(2) It's T376.
(3) No ring please.
(4) Can you do it in five minutes?
- (5) System/ We are not allowed to mount that magtape.
(6) Operator: You will have to talk to the head operator about it.
- (7) User: How about tape T241?
- (8) System: No.
- (9) User: Go ahead.
- (10) System: I am not exactly sure of the reason but we were given a list of users we are not supposed to mount magtapes for and <user> is on it.
- (11) User: I thought we could do it at night.
(12) Is there any time period on that list?
- (13) System: No.
- (14) User: OK.
- (15) System: You might check with Jane.
(16) Perhaps there is supposed to be a time limit, and it was forgotten.
- (17) User: Yes.
(18) I'll do that.

Dialogue 1

¹This is a slightly cleaned up terminal transcript. We thank Bill Mann for providing it.

There are several discourse issues raised by this dialogue, issues which are only now beginning to be addressed. For example, such a system must be able to understand and generate multi-sentential utterances such as (1)-(4), (5)-(6), (11)-(12)... Note that the user could conceivably have said

(1') Could you mount tape T376 for me with no ring please?

in place of utterances (1)-(3). Why did the user produce the multi-sentential utterance? How does the system recognize that utterances (2) and (3) are basically continuations of utterance (1)?

The system must also be able to participate in more than a single question/answer exchange, that is, it must be able to partake in an extended dialogue. Thus, it should be capable of using the information provided by the previous utterances. For example, utterance (7) would be difficult (if not impossible) to understand without the discourse context of utterances (1)-(6). Similarly, the discourse context preceding utterance (7) prohibits the generation of 'Hello, how are you?' instead of (7).

To address such issues, this report proposes a plan-based natural language system that incorporates knowledge of both plan and discourse structure of task-oriented dialogues. An initial representation of communicative (discourse) actions is discussed, in particular how to incorporate knowledge of legal moves as action effects rather than grammars. The subtle differences implicit in various surface realizations are also examined, as well as the structure of these communicative actions in actual dialogues. It is suggested that both local and global discourse structures are necessary (although analysis of the latter has been emphasized here). It is also suggested that planning models need to be extended to include two agent plan execution. Finally, a model of the goal recognition process is presented. Communicative and task knowledge work in parallel, one source dynamically taking control over the other and reducing the search space, depending on the kind of discourse (a task-

oriented one, a conversation...). Communicative recognition is hypothesized to be simple, using the knowledge provided by the analysis of surface phenomena and task plan recognition.

In particular, this paper is organized as follows. Chapter 2 reviews the literature of discourse understanding and generation. The differences between generation and understanding are not a major focus of this paper. Instead, the capabilities provided by incorporating a discourse model into a system (in other words, why a discourse component is necessary) are discussed. The plan-based approach to language adopted here will also be reviewed. Chapter 3 will suggest ways to actually represent, incorporate and use such a component in a natural language understanding (or generation) system. Several examples will be given. Finally, Section 4 presents conclusions and likely future directions.

CHAPTER 2

Background

To look at contemporary studies of language in the cognitive sciences is to see an often times bewildering (and, in terms of sheer quantity, overwhelming) array of linguistic phenomena, data, formalisms, standards, and purposes. [lev79b]

1. Discourse

Researchers in artificial intelligence and related fields have begun to study discourse in order to improve explanations underlying surface linguistic phenomena occurring in natural language. Furthermore, although many researchers have also investigated the understanding and generation of multi-sentential 'utterances' (e.g. text, paragraphs, dialogue, conversation, and stories), they have only recently considered the role of discourse knowledge. This section will review the work in these areas.

1.1. Wilensky

Wilensky [wil78] claimed that to understand stories one must reason about the situations referenced in terms of the intentions (e.g. goals and plans) of the characters. In particular he implemented PAM (Plan Applier Mechanism), a computer program which understood stories by such types of reasoning. Goal-based stories were categorized by the knowledge and inference rules needed to understand them; intensional knowledge was characterized as the kinds of goals which existed, how goals were fulfilled, and how goals interacted. Such knowledge was then applied to make inferences using algorithms for detecting and processing the characterized situations. Furthermore, the processing was both top-down and bottom-up. Figure 2-1 presents a story that PAM understands, as well as PAM's processing loop. For example, to process the first sentence PAM first transforms it into the conceptual dependency representa-

STORY:

John wanted money.
He got a gun and walked into a liquor store.
He told the owner he wanted some money.
The owner gave John the money and John left.

PROCESSING LOOP (informally):

1. sentence -> conceptual dependency
2. explanation found and tested
3. explanation added to story representation
4. goto 1

Figure 2-1. PAM [wil78]

tion. PAM then infers that John has a goal of getting money since it is instrumental to another goal, John likes money, or he might need it in the future. Finally, these inferences are added to the representation. PAM illustrates understanding by answering questions and expressing the story from different points of view.

Wilensky thus extended the work on scripts [sch77] (and frames [min75]), enabling the understanding of novel stories. That is, he still exploited context and the ability to draw inferences; top-down processing provided predictions, coherence, and efficiency (controlled search). However by adding bottom up processing and intentional knowledge the connection among sentences in novel as well as stereotypical situations could now be inferred. Unfortunately, although Wilensky's approach illustrated the importance of cognitive modeling in language understanding, the importance of linguistic phenomena (discussed below) was ignored. Furthermore, it is not clear whether the approach generalizes to genres other than stories, let alone non-goal-based stories.

1.2. Grosz

Grosz [gro77] incorporated the idea of focus of attention into a dialogue understanding system, where focus refers to the effect of linguistic and situational contextual influences.

Without such selective consideration the knowledge necessary for understanding in even simple domains becomes overwhelming. She claimed there were two type of focus of attention, global and immediate. Global focus represented the influence of context (i.e. the preceding utterances of the discourse as well as the situation) and was shown to be useful for the resolution of definite noun phrases. A computational representation of this highlighting was achieved by segmenting the knowledge base into hierarchically structured focus spaces, corresponding to the discourse structure at a given point in the dialogue. Moreover, Grosz observed that task-oriented dialogues subdivided into units just as the tasks subdivided into subtasks. Thus, such a representation could be computationally updated; the task structure could be used as a guide for discourse structure shifts.

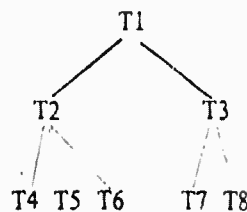


Figure 2-2. A Simple Tree Task Structure [gro77]

Figure 2-2 illustrates dialogue pops.

When task T6 is completed, there is a return to the focus of T2 and possibly directly to T1. Objects that participate only in T4 or T5 are not in focus. Similarly, objects in T2 or T4-T6 cannot be directly referenced from T7 or T8. When T8 is completed, there may be a 'pop' up to T3 or T1. [gro77]

Figure 2-3 illustrates the hierarchical dialogue segmentation (corresponding to the task hierarchy), since the pieces of dialogue between the underlined pronoun and its referent correspond to subtasks.

E: Good morning. I would like for you to reassemble the *compressor*.

...

E: I suggest you begin by attaching the pump to the platform.

... (other subtasks)

E: Good. All that remains then is to attach the belt housing cover to the belt housing frame.

...

A: All right the belt housing cover is on and tightened down.

(30 minutes and 60 utterances after beginning)

E: Fine. Now let's see if *it* works.

Figure 2-3. Pronoun Use Reflecting Dialogue Structure
[gro77]

Immediate focus represented the influence of the linguistic form of an utterance on the succeeding utterance. Although Grosz's work was primarily on global focus, she did show that immediate focus was useful for understanding ellipsis. Sidner [sid83] studied immediate focus in depth; she showed that it was useful for understanding definite noun phrases, pronominalization, *this*, and *that*.

Grosz then was concerned with linguistic phenomena as well as intentional knowledge, i.e. with discourse as well as plan structure. Her use of focus was an alternative to recency explanations of pronoun resolution [win72]. Furthermore, the highlighting was used to constrain the search for the referents of definite noun phrases. However, since her focus updating techniques were based on the correspondence between task and discourse structure her theory needed to be generalized to non-task-oriented dialogues.

1.3. Sidner

Sidner [sid83] presented a computational theory of definite anaphora¹ interpretation using (immediate) focus; focus was the particular discourse element the speaker centered on. She formalized her theory by developing algorithms for finding and moving focus. The first step involved focus recognition, that is choosing an expected focus from the first sentence using syntactic constructions and grammatical relations. For example, the cleft sentence 'It was John who ate the bread' clearly marks John as the focus. (The examples are from [sid83]). The next step was interpretation, using the focus to interpret anaphors in the next sentence. Finally, the focus was confirmed, maintained, or moved; if the anaphora and inference mechanisms yielded contradictions, the expected focus would be rejected. Focus movement was analogous to initial focus recognition.

Consider the following two sentence pairs [sid83].

Last week there were some nice strawberries in the refrigerator.
They came from our food co-op and were very fresh.

Cathy wants to have a big party at her house. She cleaned it up.

In the first excerpt 'some strawberries' is recognized as the expected focus, since it is the subject of a there-insertion sentence. It is set as the value of the current focus and 'last week,' 'refrigerator,' and the verb phrase as alternates. Because 'They' in the second sentence co-specifies with some strawberries, the current focus is confirmed and remains. The cycle then repeats. However, in the second excerpt 'it' is used to reject the current focus of 'big party' in favor of the alternate 'her house' (i.e. focus moves).

Sidner viewed anaphora interpretation as using an already existing specification (cognitive element) of a noun phrase to find the specification of an anaphor, rather than the antecedent. Usually the focus provided a co-specifier or generator for the specification. This allowed her to handle such previously problematic cases as 'My neighbor has a *monster Harley 1200*. *They*

¹Anaphora is the use of words or phrases to point back in the discourse context.

are really huge but gas-efficient bikes.' Furthermore her theory took syntactic, semantic, and inferential knowledge into account. Sidner's model also illustrated the use of focus for controlling inference. Focus predictions were confirmed or rejected based on the presence of inferred contradictions. Finally, unlike many models Sidner's was tractable. Although the model accounted for many surface phenomena there were cases where it fell apart. 'Popping' (as discussed above) violated her proposed rules and indicated the need for discourse structure. Similarly, parallelism or similarity of structure also caused violation of her rules.

1.4. Reichman

Reichman [rei81] pursued the idea that spontaneous dialogues are highly rule governed rather than unstructured. In particular, she performed a structural analysis of discourse and developed the context space theory (presented in [rei78] and [rei79] for informal and technical conversations, respectively) to explain the results. This theory partitions utterances into hierarchical context spaces, characterized by slots (like case frames) and related to one another by conversational moves (communicative goals) such as support, interrupt, and challenge. Much effort was spent characterizing the moves in terms of their preconditions (discourse context which must be present for their appropriate performance), effects on the discourse structure (context space shifts and status reassignment, expectations) and modes of fulfillment. Based on this, Reichman then formalized an abstract process model for well-formed discourse generation and interpretation.

The context space theory delineates a single abstract structure underlying all discourse forms - expository text, argumentative text, narrative text - and based on such structure characterization it is able to specify a single set of 'maxim-abiding,' 'well-formedness' rules applicable to, and governing all discourse forms. [rei81]

The discourse model is written as an Augmented Transition Network (ATN) [bat78]; rules of effective communicative govern use of clue words (e.g. 'incidentally,' 'by the way') and choice of reference. Highlighted portions of the conversation can thus be tracked. Finally, aspects of

the model were shown to complement various theories in cognitive processing.

Figure 2-4 presents the beginning of one of the discourses analyzed by Reichman. The following are examples of the types of analyses given by the abstract process model. The discourse is a debate. Lines 1-4 are an authority support for lines 6-7. An effect of line 8 is the expectation that R will provide further support or evidence on her next turn, since line 8 is a demand for such.

Reichman's context space theory thus provides an abstract, hierarchical discourse structure as well as formalized well-formedness rules applicable to many (or as she claims all) discourse genres. Furthermore, the work is an attempt to formalize Grice's maxims (as well as her earlier work), from which the ATN abstract process model of maxim-abiding discourse was designed. Finally, her model accounted for several linguistic phenomena (pronominalization, nonpronominalization, and clue words) and complemented current theories in cognitive processing (segmentation, selective attention, frame of reference processing, expectations, and cues). Despite the above, Reichman's model could benefit from even further formalization and

-
- R: 1. Except however, John and I just saw this two hour TV
2. show
- M: 3. Uh hum,
- R: 4. where they showed - it was an excellent French TV
5. documentary - and they showed that, in fact, the
6. aggressive nature of the child is not really that
7. much influenced by his environment.
- M: 8. How did they show that?
- R: 9. They showed that by filming kids in kindergarten,
- ...

Figure 2-4. Typical Excerpt [rei81]

connection with surface phenomena; this is true, for example, with respect to the recognition of communicative goals from surface text. Lastly, although Reichman distinguished communicative goals from speaker intent, the interactions and relationships between them were not of primary concern.

1.5. McKeown

McKeown [mck82] demonstrated that both discourse structure and focus constraints are useful for the computer generation of text. McKeown was primarily concerned with what to say and how to organize it effectively, rather than with the transformation into English. The main contribution of her system (called TEXT) was the pairing of rhetorical techniques with discourse purpose, for example the selection of the analogy or identification schemas when replying to a request for a definition. These techniques were represented as (recursive) schemas to reflect the belief that people have preconceived ideas about discourse structure. Discourse purpose was modeled by which database question was to be answered. Each predicate in the schema had associated semantics expressed in terms of the knowledge representation; the schemas were thus filled in by using the semantics to match the knowledge base. Figure 2-5 presents an example of a schema as well as a text that illustrates it. The schema can be read as a rule in a grammar (i.e. Constituency-Schema $\rightarrow$ Constituency Cause-effect*/Attributive*/...). McKeown's system also incorporated global and immediate focus (described above with respect to understanding), which provided what she called relevancy criteria and discourse coherency, respectively.

McKeown thus demonstrated that text could be effectively produced by using communicative strategies instead of tracing the knowledge base. Depending on the question or the focus, the same information in the knowledge base could be described in various ways. Furthermore, the knowledge base didn't need to be designed with text production in mind (as done in [swa81]). The system also illustrated possible interactions between syntax and semantics.

CONSTITUENCY SCHEMA

Constituency
 Cause-effect*/Attributive*/
 { Depth-identification/Depth-attributive
 {Particular-illustration/evidence}
 {Comparison;analogy} }+
 {Amplification/Explanation/Attributive/Analogy}

(notation: {}:optional; /:alternatives; +:may appear 1-n times; *:0-n times)

EXAMPLE

Steam and electric torpedoes. 1) Modern torpedoes are of 2 general types. 2) Steam-propelled models have speeds of 27 to 45 knots and ranges of 4000 to 25,000 yds. (4,367 - 27,350 meters). 3) The electric powered models are similar 4) but do not leave the telltale wake created by the exhaust of a steam torpedo.

CLASSIFICATION OF EXAMPLE USING ABOVE SCHEMA

1. Constituency
2. Depth-identification (attributive)
3. Comparison
4. Depth-identification (attributive)

Figure 2-5. TEXT system [mck82]

For example the strategies determine the final content, yet the available relevant knowledge can help determine the strategy (i.e. the structure chosen). Finally, focus was extended to deal with generation issues.

However there are also several weaknesses of the TEXT system. Although McKown emphasizes that schemas are not grammars of text, they are in effect used as such. That is she acknowledges that there are instantiations of rhetorical techniques not captured by her schemas, but her system does not deal with these. If used for understanding her schemas would then have the same problems as scripts. Furthermore much of the system, for example the semantics of the predicates, basically manipulates the knowledge base (in a sense code which is unlikely to generalize). Since classifications such as found in Figure 2-5 are extremely

subjective it is also unclear if they are even correct. Finally there are omissions which McKeown discusses in some detail, among them a user model, inferencing, discourse context, shifting of focus, and variation of detail.

1.6. Related research

There are numerous approaches relevant to discourse analysis outside artificial intelligence, often covered by the term textlinguistics. The Appendix presents annotations of a small, representative sample. In textlinguistics, written and spoken texts are viewed as the minimal free unit of language (although discourse is a looser term which is also used). Textlinguistics is presented much more narrowly in [1]² however; texts are communicative occurrences meeting seven standards of textuality.

Functional sentence perspective theories describe the sentence from the point of view of its (potential) use in a message (framed in a text or situation).³ [19] Functional syntax [13] is a trend in generative grammar which recognizes that many phenomena previously regarded as syntactic are controlled by non-syntactic factors. In other words, the problems of generative syntax are viewed within the framework of discourse analysis. Kay [12] also argues for functional considerations of grammar. Systemic-functional models [8,9] derive from 'the two notions most fundamental to the text-ness of text: texture and structure' [9] The possession of texture distinguishes a passage with linguistic cohesion (a text) from a random string of sentences. Structure is used to characterize complete texts of a genre. It is controlled contextually; a text is thus viewed as a social event which primarily unfolds linguistically.

Semiotics [18,20] is the study of sign system or codes. Textlinguistics is sometimes considered a subset of semiotics since the latter considers both verbal and non-verbal communication as texts. Methodological properties which characterize a formal integrated text theory of language are presented in [20].

²These numbers correspond to those in the Appendix.

Related work can also be found in the social sciences. For example, psychologists have become concerned with the cognitive processing of discourse.

[The] interpretation of sentences is a function of the verbal and non-verbal context in which a sentence is uttered, and ... the conceptual knowledge structure of our memory not only depends on the interpretation of isolated sentences, but also on the understanding and processing of whole discourse. [24]

In sociology there are the branches of sociolinguistics and ethnomethodology [5, 10, 11, 14, 25].

Sociolinguistics, particularly the field of discourse analysis, has developed methods for collecting, transcribing and analyzing spoken data, and has shown that it is possible to discover regular structure in such spontaneous text ... Related work in ethnomethodology and conversational analysis ... shows why a relation can be presumed to exist between the structures described by the analyst and those which the participants of a conversation themselves use. [5]

Discourse is viewed as a social process, occurring in contexts which influence what actually occurs. Structural regularities are both formal conversational analyses as well as tools used to achieve social regularity.

Finally, other characterizations can be made that are orthogonal to those given above. For example, there are the text grammar [20, 22, 24] and other structural approaches [5, 10, 14, 23], as well as those concerned with characterizing various genres [6, 9, 15, 24].

1.7. Summary

Various studies of cohesive surface phenomena were based on focus and the structure of the discourse. These were in contrast to the earlier and simpler accounts such as the use of recency criteria for pronoun resolution [win72]. The following are examples of alternative artificial intelligence discourse approaches and the phenomena they explain:

focus spaces - definite noun phrases and ellipsis [gro77]

immediate focus - definite anaphora, pronouns, 'this' and 'that' [sid83]

context spaces - clue words, pronouns, 'that', tenses [rei78, rei79, rei81].

As can be seen, the use of focus has been well studied. These issues will not be pursued now;

the results of this earlier work will instead be used to guide and confirm the work reported in this paper.

Early domain-oriented approaches to the organization of multiple sentences were exemplified by frames [min75], scripts [sch77] and plans [wil78]. Cohesion, in particular implicit sentence connections, was of primary concern. Recent approaches were primarily discourse oriented. For example McKeown's work was concerned with both syntactic and semantic approaches to discourse phenomena; she was not concerned with domain (as opposed to communicative) goals. Reichman's work was also concerned with discourse rather than domain intentions although her approach was primarily structural. Although Grosz [gro77] noted the existence of both plan and discourse structures in her expert/apprentice task domain they were nearly equivalent. She used the task domain to determine the discourse structure which was then used to understand surface phenomena. It will be argued in this paper that these contrasting approaches need to be merged. As will be shown a planning model of language appears to be an appropriate framework for a merger of discourse and domain, structural and semantic approaches.

Figure 2.6 is an attempt to approximately categorize researchers with respect to the problems investigated and proposed solutions. The rows represent the phenomena of interest; the columns show the primacy of structure (syntax) versus content (semantics).

2. Language as Planned Action

Several approaches to language have developed the view that acts of communication can be planned, just as physical acts like stacking blocks. More recently this has been extended to the view that language satisfies goals of the participants along various dimensions.

	CONTENT		STRUCTURAL	
COHESION (sentence connections, surface phenomena)	Schank and Abelson		Grosz	
	Wilensky	McKeown	Sidner	
			Reichman	
DISCOURSE STRUCTURE	Grosz	McKeown	Reichman	

Figure 2-6. Summary

2.1. Allen, Cohen, and Perrault

These works [all83,coh79] developed plan-based approaches to speech act recognition and generation. For the purposes of this section Allen's work [1981] is illustrative. Allen's basic claim was that helpful behavior appears when the hearer recognizes and acts on an obstacle in the speaker's plan. A plan-based model of language as cooperative behavior developed which supported this claim. Utterances were viewed as (goal-oriented) speech acts which were executed to modify the hearer's beliefs or goals; the hearer would infer the speaker's plans and detect any obstacles. Figure 2-7 is an example. The obstacles thus detected are that the user needs to know both departure time and location.

Allen's model accounted for helpful responses (providing more information than requested) as well as for responses to indirect speech acts and sentence fragments. These phenomena had been problematic for previous approaches. Furthermore, the hearer used his/her model of the speaker as a context for constraining the inference process. In ARGOT [all82a,all82b] Allen has extended his research by including explicit knowledge regarding discourse structure, improving the representational formalisms used, and investigating the relation to syntactic processing. However, Allen has not extended his model to account for extended dialogue and it could benefit from even further connection with surface phenomena.

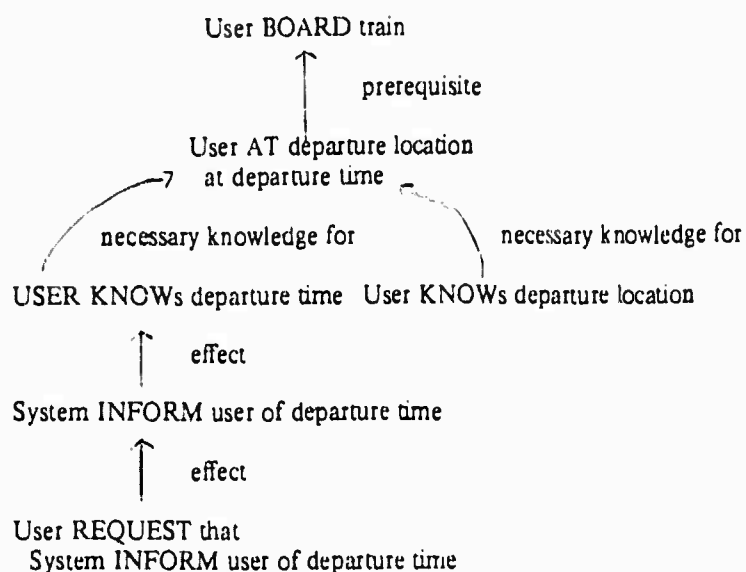


Figure 2-7. Simple Plan Recognized from 'When does the Montreal train leave?' [all82b]

2.2. Levy

Levy [lev79a] investigated a mind based as opposed to text based framework for discourse. That is, he was concerned with the study of mental representations in relation to the process of communication, rather than the study of the text as object. In particular, he developed an initial formulation of communicative goals and strategies within a larger model of language as planning.

Some of these goals (called IDEATIONAL goals) are concerned directly with the communication of these ideas or propositions; some (called TEXTUAL goals) are concerned with the weaving of these ideas into a coherent text; and still others (called INTERPERSONAL goals) deal with presentation of self in relation to the hearer, with matters of status and attitude. [lev79a]³

Communicative goals and strategies thus derive meaning from one's mental activity.

³Levy borrows these terms from [Halliday, M.A.K. (1970) 'Language Structure and Language Function,' in J. Lyons ed., *New Horizons in Linguistics*, Penguin, New York].

Furthermore, since they are satisfied by language they also provide cohesion and connect discourse and syntax (e.g. goals and expressions are connected by strategies). For example, words and phrases represent ideas and concepts, conjunctions interrelate utterances and intonation reflects attitudes. The speaker thus encodes (and the hearer reconstructs) thought processes as well as ideas within an utterance. Figure 2-8 shows a partial description of Levy's 'Refer' strategy (written like a computer program). This strategy represents the mental process used by the speaker to linguistically realize the 'Refer' communicative goal.

Although preliminary, Levy's ideas have been further pursued here and by others. Levy [lev79b] however was primarily concerned with explicating the view of the text held in the cognitive sciences. In particular he perceived text as a designed, communicative artifact, produced by the speaker and reconstructed by the hearer. Text could be treated as either an object or an activity; concept systems (content, activity, and object) served as filters through which text was viewed. Comprehension was a process of convergence on the architecture of the text that could be seen through the various filters.

Refer(object)

Formulate a description of object

If there is more than one description, then

 If this is due to a memory retrieval problem, then

 Express-incompletely-retrieved-description(candidate descriptions)

 else, If this is because a choice has not yet been made, then

 Express-unresolved-choice(candidate descriptions)

Figure 2-8. 'Refer' Strategy [lev79a]

2.3. Appelt

Appelt's [app81] work primarily addressed the interaction between planning and language generation. A planner called KAMP (Knowledge And Modalities Planner) was developed which handled intensional concepts, two agents, and both physical and linguistic actions. The interaction between the production of surface forms from underlying representations and the planning of speech acts was of concern, with emphasis placed on the satisfaction of multiple goals. For example, an utterance could simultaneously inform, request, change focus, and reflect social views.

[An] utterance like, '*Tighten the screw with the long philips screwdriver.*' can realize several illocutionary acts, like a REQUEST to tighten the screw and an INFORM that the tool for tightening the screw is the long philips screwdriver. Given that the speaker knows that the hearer doesn't know that a particular screwdriver is a philips screwdriver, the utterance could in that case also serve to inform the hearer that the long screwdriver is a philips screwdriver. This is contrasted with the case where 'long' is used to distinguish long versus short. [app81]

If formulated as an indirect speech act goals such as politeness would also have been satisfied.

KAMP did not generate extended discourse in any general sense; multiple sentences were only generated when KAMP could not satisfy its goals in one sentence. Coherence thus resulted due to the underlying plan. There were no abstract discourse actions and focus was used only to facilitate reference. Appelt himself points out the need to integrate the results of McKeown [mck82] and Reichman [rei78]. However, KAMP did formalize the incorporation of some of the multiple perspectives advocated by Levy [lev79a] (and also by Grosz [gro79]).

3. ARGOT

Many proposals in this work developed from analysis of ARGOT [all82a,all82b], a plan-based system which claimed that at least two levels of goal analysis were needed to partake in extended discourse. Postulated were the *task level*, which includes goals such as mounting tapes, reading files, etc., and the *communicative level*, which includes goals such as introducing a topic, clarifying or elaborating on a previous utterance, modifying the current topic, etc.

These corresponded to Levy's ideational and textual goals, respectively. Splitting the analysis of intention into the communicative and task levels brought about the problem of identifying and relating the high-level goals of the plans at each level. The high-level goals at the task level were dependent on the domain. The high-level communicative goals reflected the structure of English dialogue and were useful as input to the task level reasoner. In other words, these goals specified some operation (e.g., introduce goal, specify parameter) that indicated how the task level plan was to be manipulated. Hobbs and Agar [hob81] have also begun to explore the relationship between plans at these two levels.

The initial high-level communicative goals were based on the work of Mann, Moore and Levin [man77]. In their model, conversations were analyzed in terms of the ways in which language was used to achieve goals in the task domain. For example, bidding is a communicative action which introduces a task goal for adoption by the hearer. However, not all communicative actions are possible at any given time. For instance, at the start of a dialogue one usually either bids a goal or gets the other agent's attention (a summons), but does not end the dialogue. To capture this knowledge a context-free grammar with these communicative acts as terminals was incorporated, along the lines of Horrigan [hor77]. The grammar indicated what acts were expected at any particular time for both participants. Finally, given a communicative level, plans at this level must be recognized. Neither Mann et al. [man77] nor Reichman [rei78] described in detail the process of recognizing the communicative goals from actual utterances. The recognition algorithm found in Allen [all83], which found an inference path connecting the observed linguistic action(s) to an expected goal in the task level context, was adopted. The algorithm used both the parser's representation of the utterance and the set of possible communicative acts predicted by the grammar as clues when recognizing the actual communicative goal.

The following analysis of utterance (1), "Could you mount a magtape for me?" was taken from [all82a]. The communicative acts expected by the dialogue grammar are

Report No. 5338

user BID-GOAL to system, and
user SUMMON system.

Taking the utterance literally, a linguistic level uses both syntactic and semantic analysis to identify the linguistic actions (speech acts) performed by the speaker. For utterance (1) the speech act is (in an informal notation)

user REQUEST that
system INFORM user if system can mount a tape,

which is sent to the communicative level. The plan recognition algorithm produces BID-GOAL.
acts for two possible goals:

system INFORM user if system can mount a tape (literal interpretation)
system MOUNT a tape (indirect interpretation).

The indirect interpretation is favored, illustrating how goal plausibility depends on what the dialogue participants know and believe. Most people know that operators can mount tapes, so the indirect interpretation is preferred. However, if the user did not know this, the literal interpretation would also have been recognized (i.e., the system might generate 'yes' before attempting to mount the tape). It is important to remember here that the plan was recognized starting from the literal interpretation of the utterance. The indirect interpretation falls out of the plan analysis. Thus, the linguistic level only needs to produce a literal analysis. The recognized BID-GOAL to mount a tape is sent to the task reasoner, which recognizes and accepts a task level plan of mounting the tape. Of course, since the task level reasoner is a general plan recognizer as well, it may well have inferred beyond the immediate effect of the specific communicative action. For example, it might infer that the user has a higher-level goal of reading a file. The BID-GOAL is also made known to the dialogue grammar, to enable the correct production of expected user and system communicative acts. At the task level the goal can then be expanded by the task reasoner and the resultant plan inspected for obstacles. Assuming the

user says nothing further, there is an obstacle in the task plan, for the system does not know which tape to mount. This generates a system goal to identify the tape parameter, which is sent to the communicative goal reasoner. A speech act (or acts) is planned that will lead to accomplishing the goal and which obeys the constraints on well-formed discourse. This would be sent to the linguistic level which would generate a response such as 'which tape?' In Dialogue 1, however, the user utters (2), which will be recognized as a SPECIFY-PARAMETER action at the communicative level. Thus, among the expected communicative acts after the first utterance are

system ASK-PARAMETER of user, and
user SPECIFY-PARAMETER to system.

ARGOT's use of multiple goals for the understanding of extended discourse went further than previous approaches. There were more concrete proposals regarding the communicative level and its relationship with the task level. Furthermore, the data (Dialogue 1) was in some ways a superset of Levy's (a single person response to a question). The weaknesses of ARGOT form the basis of the remainder of this proposal.

CHAPTER 3

Proposals

Consider again Dialogue 1, particularly the initial utterances. Now imagine the following scenario, typical of current planning approaches to language [app81,coh79]. The user's goal is to read a file stored on tape T376. A plan such as Plan1 is thus generated by the user, shown informally below.

CANDO(op,MOUNT())		
HASGOAL(op,MOUNT())		INCORE(file)
<u>BIDTOPIC(usr,op,MOUNT())</u> ----- <u>MOUNT(op,T376)</u> ----- <u>READ(usr,file)</u>		
HASGOAL(op,MOUNT()) ¹	INCORE(file)	HASREAD(usr,file)
MOUNTED(T376)		
Plan1		

The boxes contain actions augmented with the relevant preconditions and postconditions, above and below respectively. Note that the plan contains both physical actions (READ,MOUNT) and linguistic actions (BIDTOPIC). Informally, a BIDTOPIC in this domain is an attempt by one agent to get another agent to carry out some plan.

Most systems stop with the completion of plan generation. However, it seems that to explain Dialogue 1 plan execution and repair² must be considered as well. By producing

¹More accurately the effect of the BIDTOPIC is something like BELIEVES(op,WANTS(usr,HASGOAL(op,MOUNT()))). As will be seen, the original effect results from assuming cooperative behavior. (Although one still could imagine the operator checking preconditions before accepting the goal).

²This work will assume that there exist planners for execution. It is felt that differences resulting from this integrated approach to planning will not be crucial to what is presented here.

utterances (1)-(4) of Dialogue 1, the user successfully executes the BIDTOPIC. The operator now has a goal of mounting tape T376 and enters the planning cycle of plan generation, execution, and repair. However, s/he discovers that s/he is not allowed to mount the tape (a precondition violation) and aborts.

Why does the operator communicate this failure (utterance (5))? Nothing in the failure of the task plan to mount the tape requires (5); if the mounting was self-initiated it is unlikely that the operator would have spoken. It thus appears that the user's BIDTOPIC is partly responsible for the operator's production of utterance (5) (along with the operator's plan to cooperate with the user). Just as actions have k-effects (knowledge-state) and p-effects (physical-state) [app81], communicative actions also have (mutually believed) d-effects (discourse-state effects). For example, the operator's communicative goal to acknowledge is a d-effect of the user's BIDTOPIC. D-effects are a subset of k-effects which describe conventional, domain independent conversational moves.

What are these communicative acts - why do they occur, how are they produced and understood, how are they structured in a discourse and how do they interact with other types of goals? To address these questions this section proposes a plan-based natural language system incorporating both domain (task) and communicative analysis. Two agent plan execution, a simple approach to discourse, and parallel task and communicative interaction are of particular concern.

4. Communicative Goals and Discourse Structure

4.1. Communicative Goals

As implied above, communicative goals and actions come about in at least two ways. They can be part of a plan to achieve some task goal (as is the BIDTOPIC of Plan1), or they can be generated by the d-effects of communicative actions. D-effects capture the legal (and

illegal) moves of the discourse structure. For example, the production of utterance (5) as described above satisfies the d-effect of the BIDTOPIC. Previous approaches to the origin of communicative goals have been based on grammars or schemas [rei81], [mck82]. A plan-based approach provides much more flexibility since all possible discourse structures need not be identified in advance. Although Argot has a 'grammar' which produces discourse expectations, recognition of discourse structure can occur despite expectation violation. Such flexibility will be useful in miscommunication recovery. The approach taken here is probably not isomorphic to Argot's (practically if not theoretically). Grammars tend to imply that a single communicative goal is satisfied by each utterance. With d-effects, however, an utterance can satisfy multiple, interacting goals and expectations.

How are communicative acts formulated as plan operators? Figures 3-1 and 3-2 present loose formulations of two communicative acts, BIDTOPIC and ACKNEG (acknowledge negatively); BIDTOPIC is illustrated by utterances (1)-(4) and ACKNEG by utterance (5). In these figures, the term capacity informally refers to an ability, action, or plan available to an agent (for example mounting tapes or killing jobs in our domain), and *vr* (a term in KL-ONE [inc79]) is a value restriction, a description of potential role fillers. MUTUALLY-BELIEVED loosely means that the dialogue participants know what the speaker intended, each knows the other knows, and so on. Consider Figure 3-1. The REQUEST of the body can be linguistically realized in various ways. The bottom of the figure lists (in order of directness) several plausible surface constructions, along with some observations. Much more than politeness is obviously involved here. Figure 3-2 is a similar description of ACKNEG although slightly less detailed. It should be noted that such formulations should ultimately be domain independent.

There are still questions to be answered before this formulation will be adequate. For example, utterance (4) is part of the BIDTOPIC but seems to generate another operator communicative goal (loosely, to answer yes or no to utterance (4)); do REQUEST's also have d-effects? Another question is what is the nature of the ACKNOWLEDGE goal? Does it

Precondition: BELIEVE(speaker,CANDO(hearer,capacity))

Body: REQUEST(speaker,hearer,DO(hearer,capacity))

Effects: MUTUALLY-BELIEVED(HASGOAL(hearer,DO(hearer,capacity))
 (d-effect) MUTUALLY-BELIEVED(HASGOAL(hearer,ACKNOWLEDGE))

<i>Surface Form</i>	<i>Example</i>	<i>Observations</i>
Imperative	Mount tapel.	Presupposes that operator will mount tape.
Indirect request	Could/Can you mount tapel?	Operator's option to refuse task is explicit.
Inform	I want you to mount tapel.	Unclear if speaker expects hearer to accept task.
	I want to be able to read file foo.	Speaker doesn't know or care how the goal is achieved but knows there is a way (explicit goal, implicit method, as opposed to explicit method and implicit goals above).

Figure 3-1. BIDTOPIC(speaker,hearer,capacity)

PLAN DESCRIPTION (informally)	EXAMPLE
Report task failure.	I can't mount the tape.
Report precondition failure.	We're not allowed to mount that magtape.
Report possible alternate.	(I can't mount tape1 but) I can mount tape2.
Report desired goal state already true (modulo parameters).	A: Add a roleset named nickname to person, with number...and vr text. B: There appears to be an error in the display ...there is a roleset named nickname on person, and it already has a vr.

OBSERVATIONS: When explanation or helpful behavior is explicit, the report of failure can be implicit (subsumed).

Figure 3-2. ACKNEG

acknowledge the acceptance of the bid, task completion, or both? Since agents are assumed cooperative, this work assumes that BIDTOPICS are always successful. That is, the hearer will accept the bid (try to achieve the goal), although s/he might not be able to actually achieve it. Thus, only task completion (and not task acceptance) needs to be acknowledged.³ However, return to the scenario; as a result of utterance (4) the operator wants to mount the tape. Either the operator will (1) successfully complete the planning cycle of generation and execution, (2) unsuccessfully complete the cycle (what actually happens), or (3) interrupt the cycle to get more information. With respect to the discourse, these would be signaled by the communicative acts of ACKPOS, ACKNEG and CLARIFY, respectively. All three of these actions are legal discourse moves; should all be accounted for as d-effects (a disjunction) of the bidtopic? One answer is no; since the user has generated what s/he believes to be a correct plan, s/he only expects a goal of ACKPOS. However, because of the failed mounting a precondition of

³One could also imagine cases where acknowledgment suggests understanding rather than acceptance or completion of a request.

ACKPOS isn't met; ACKNEG arises during debugging of the ACKPOS plan. This is rather odd since the operator probably tries to mount the tape and discovers s/he cannot, rather than tries to answer affirmatively and discovers s/he cannot. Another view is that ACKPOS just has a stronger level of expectation. After all, the user isn't confused when an ACKNEG occurs.

It should also be noted that actions don't necessarily need to be formulated as enumerations of possible bodies. A more general model [all81] describes an action in terms of the conditions under which it could be said to have occurred. A mixture of the two approaches seems practical since the latter involves solving very hard problems. For example, a general formulation of explanation would determine what constitutes an answer to 'why' for any topic; this problem is clearly unsolvable.

4.2. Discourse Structure

Figure 3-3 shows part of Dialogue 1, annotated with the hypothesized discourse structure and important surface phenomena. On the left are the communicative acts; the embedding of acts beneath other acts reflects their proposed structure.⁴ For example, ACKNEG/EXPLANATION is embedded relative to BIDTOPIC. Processes in the operator's head 'demand' this response to BIDTOPIC. On the right the immediate focus [sid83] is noted. The surface phenomena reflecting the foci are highlighted. An embedded act inherits the foci of the outer act. Utterance (5) was less likely to have been 'we are not allowed to mount this magtape,' since the tape is no longer in focus.

How is the structure (the embedding of the acts) recognized? Currently, each communicative act is viewed approximately as a context space [rei78, rei79, rei81] or a focus space [gro77]. Thus, the mechanisms developed by Reichman and Grosz for recognizing shifts of those spaces can be used here as well. Reichman's clue words can be associated with certain acts and

⁴Utterances (2)-(4) are still part of the BIDTOPIC. 'Local structure' within communicative acts will be considered at the end of this section.

embeddings. For example, sequential types of clue words, e.g. 'now,' indicate the same level of embedding. Another example is utterance (9), an explicit request for another communicative act.⁵ Or using the task structure to guide the shifts in the discourse structure [gro77], each explicitly mentioned plan (i.e. a BIDTOPIC) or plan step indicates a shift to a new communicative act. The embedding of the communicative acts corresponds to the plan structure. Utterances (1)-(4) are an example; a new communicative act is indicated since a new goal is bid (although the embedding behavior isn't seen since the goal is the 'root' of the plan structure).

As will be seen again in the next section, the recognition of communicative structure is proposed to be simple, based on surface phenomena and task structure recognition. Using the terminology of ARGOT [all82a], parameters of actions like BIDTOPIC (i.e. what is bid) are basically recognized by the task recognizer since communicative plan recognition would repeat most of this work and the task recognizer must interpret the plan anyway. It is the communicative dimension however which knows that BIDTOPIC must be acknowledged.

Of course the above discussion has left issues unanswered. In particular there are various ways to formulate discourse structure and embedding. For example, the d-effects could include the embedding information and thus generate a discourse structure. This view then raises the following questions. If the discourse structure effects are implicit in the communicative acts, the formulation of this encoding needs to be determined. For example, the d-effects of BIDTOPIC must note that an ACKNEG is embedded relative to the bidtopic since it is a response to the bidtopic. Another difficulty is as follows. After utterance (5) the d-effects predict either a helpful response at the same level or beginning a new BIDTOPIC at a less embedded level. (That is, a bid requires an acknowledgement and optionally a helpful response. Once these discourse requirements have been fulfilled a change of topic (BIDTOPIC) can occur, which since not required by the previous acts is not embedded). However, it seems

⁵Although, this perhaps should be considered another communicative act.

BIDTOPIC	1. Could you mount a magtape for me? <i>FOCUS=magtape</i>
	2. <i>It's T376.</i>
	3. No ring please.
	4. Can you do <i>it</i> in five minutes? <i>FOCUS=mounting</i>
ACKNEG/EXPLAIN	5. We are not allowed to mount <i>that</i> magtape.
HELPFUL RESPONSE	6. You will have to talk to the head operator about <i>it</i> .
BIDTOPIC	7. How about tape T241? <i>FOCUS=tape</i>
ACKNEG	8. No.
	9. Go ahead.
EXPLAIN	10. I am not exactly sure of the reason but we were given a list of users we are not supposed to mount magtapes for and linker is on <i>it</i> . <i>FOCUS=list</i>

Figure 3-3. Example Discourse Structure

that the latter is more like a jump than a pop back to the outermost level (or space). A pop would be appropriate if the user wanted to say something about the already existing BIDTOPIC (see [gro77] and Reichman for discussions of popping). In our case there is a new BIDTOPIC rather than the same one. Furthermore, one could imagine anaphora still referring to previous entities, for example 'Could you do it with tape T241?' A similar example is illustrated in the excerpt below, taken from a scenario for interacting with a KL-ONE layout and graphic editing system [inc79].

BIDTOPIC P: Show me the generic concept called 'employee.'

ACKPOS S: OK.

Suppose P then said 'Make it bigger,' a BIDTOPIC at the same level as the first BIDTOPIC but with anaphora referring to an entity in the 'closed' BIDTOPIC exchange. In such cases the communicative act structure would not coincide with the discourse structure indicated by the

focus information. This suggests that the stack metaphor is inadequate.

A different approach would be to separate the discourse structure from the d-effects and only use embedding to refer to the decomposition of acts (with respect to levels of abstraction). For example, consider an informal communicative analysis of utterances (1) - (5), as shown in Figure 3-4. The discourse structure is represented by the tree structure. Satisfaction of d-effects accounts for the generation of actions pictured horizontally, while decomposition accounts for generation of embedded actions pictured vertically. This view raises questions such as does popping (pictorially, moving up a level) override unfulfilled d-effects at the previous level? Utterance (4) is also a bit problematic and should be examined further. Although part of the BIDTOPIC, it seems to be an embedded information request as well. After (4) there are six expectations:⁶ ACKPOS, ACKNEG/EXPLAIN and CLARIFY from both the BIDTOPIC and the information request. However, utterance (5) only explicitly responds to the BIDTOPIC. What happened to the answer to (4)? Should one assume that satisfaction of an embedded expectation is subsumed⁷ by satisfaction of an outer expectation? How is popping involved? Utterance (5) thus illustrates that the problem of subsumption needs to be

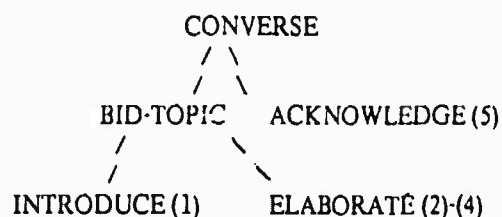


Figure 3-4. Communicative Analysis of Utterances (1)-(5)

⁶These expectations were presented in the discussion of questions regarding Figures 3-1 and 3-2.

⁷An action A1 subsumes another action A2 if A1 and A2 are part of the same plan and action A1, in addition to producing the effects for which it was planned (i.e., the principal effects) also produces the effects for which action A2 was intended.' [app81]

dealt with in understanding as well as generation. Although agreement with surface phenomena will ultimately determine the most fruitful approach, the second view is more analogous with work found in the planning literature.

It should be emphasized that the correspondence between communicative goals and utterances is not one-to-one; neither is the correspondence between speech acts and utterances [app81]. There can be both multiple utterances per goal (utterances (1)-(4) of Dialogue 1) and multiple goals per utterance ('Change the number and vr as indicated and display please' [sid82]). It is not clear that what ties together the utterances within a communicative act is what ties together the communicative acts (described above). In fact, a distinction between local and global discourse structure, or local and global coherence [gro82], might be necessary. This is somewhat analogous to the distinction between immediate and global focus [gro77]. On the other hand, global structure is perhaps recursive (just as McKeown's schemas are recursive [mck82]).

Although this work has concentrated on global structure, some thought has been given to local structure. Several initial observations about local structure follow. The structure of one communicative act seems to lack the embedding and interactiveness (between sentences) of global structure. Perhaps utterance (2) is generated because the user realized the reference of utterance (1) was inadequate [lev79a]. Such repair seems particularly important in spoken, as opposed to written, dialogues. Or maybe there exists a strategy that says if the hearer does not know that the referent exists, generate an indefinite reference followed by an identification; this perhaps extends McKeown's model [mck82] to include a user model. Resource limitations might also play a role. With respect to understanding, perhaps utterances are processed as one goal until discourse clues (anaphora, ellipsis, focus, clue words) indicate a shift or conversely, until an utterance no longer contains any connecting devices. (ARGOT ignores these issues, assuming one communicative goal per utterance. This assumption, however, might be true locally.)

5. Task Goals and Structure

The task dimension proposed here differs from the task level of ARGOT in a fundamental way. Mounting and reading tapes, typical task goals in ARGOT, have now become parameters of goals like generating, executing, debugging, and aborting plans. In other words plan structures (in the sense of Grosz [gro77]) are now parameters of the task goals. These new task actions are similar to the communicative actions of ARGOT.

6. Communicative and Task Goal Interaction

It is obvious (hopefully) that the task and communicative dimensions are different. However, in certain domains the structure at each is nearly equivalent (e.g. [gro77]). Why then should a task/communicative distinction be made explicit? Some parts of an utterance are purely communicative, for example 'Go ahead' of Dialogue 1. On the other hand, as shown earlier (Plan1) a plan can contain steps which involve no communication. Furthermore, though the same information (loosely, the utterance) is used to determine both structures, the actions each dimension takes as a result are different. For example, at the task dimension the user infers that the plan has failed or succeeded and replans or aborts; the planner's response can involve much more than just communication. At the discourse dimension however the user infers things like focus and legal moves, i.e. how utterances fit into the existing discourse context. Finally, indirect speech acts are often responded to literally as well as extraliterally; a purely intentional analysis would not account for both. Also, since the communicative dimension is domain independent it is thus more general. Figure 3-5 shows the structures, actions and subsidiary processes postulated for the two dimensions.

Given these two explicit dimensions one can imagine at least four different strategies of interaction (control structures): (1) task analysis followed by communicative analysis (2) communicative analysis followed by task analysis (3) cascaded analysis (in either direction) [inc79] or (4) parallel (but communicating) analysis. For example, consider strategy (1) and utterance

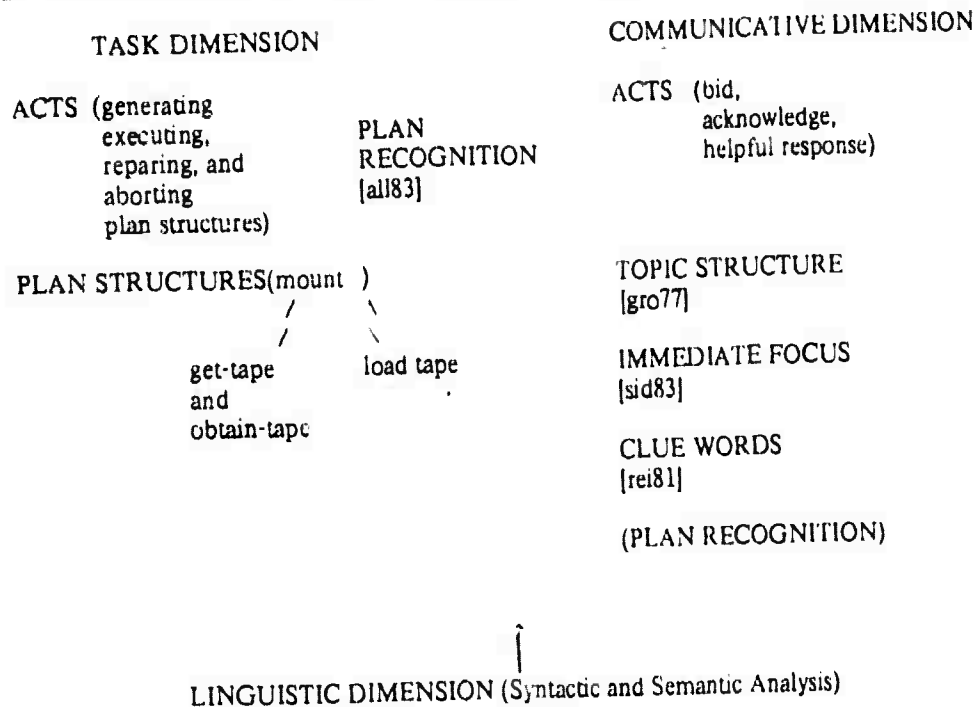


Figure 3-5. System Overview

(6). The user first asks what was the intention behind the utterance and recognizes that the operator is explaining how to debug the plan's precondition failure. The communicative dimension then uses this analysis along with the surface phenomena to determine that a helpful response has occurred. Using strategy (2) the user would first use surface phenomena (and perhaps a recognition procedure) to determine that a helpful response has occurred; the task dimension would then use this information to determine that the operator is suggesting a way to debug the user's plan. Cascaded analysis would do a little at one dimension of analysis, send the results to the other dimension for confirmation, then continue the cycle.

The last strategy is most promising, since it seems to encompass the others but by combining them overcome their limitations as well. For example, suppose the discourse structure

is violated as when a question is ignored; control strategy (2) would likely run into more difficulty than control strategy (1). However, one can easily imagine cases where the reverse is true, as when the user suddenly changes task plan. At the discourse dimension a clue word would likely be used to signal an unexpected shift of topic; however, it is unlikely that there would be such an explicit hint for the task dimension analysis. (But again, if the clue word were missing as well, the reverse would be true.) Strategy (4) allows the appropriate dimension of initial analysis to be determined dynamically; the results can then be used to reduce the search of the floundering dimension. Strategy (4) also seems the most amenable to the addition of other dimensions of analysis, such as the social dimension described earlier.

7. Examples

Consider a simplified analysis of Dialogue 1 using such a strategy. Both the task and communicative dimensions start off with certain expectations, execute-plan and bidtopic respectively. After utterance (1) is spoken the literal speech act identified by the linguistic dimension,

user REQUEST that
system INFORM user if system can mount a tape,

is analyzed along both dimensions. Plan recognition at the task dimension produces the goal of execute-plan with its associated plan structure (p.1) of

system MOUNT a tape (indirect interpretation)

(or even possibly Plan1). Thus the task analysis concludes execute-plan(user,p.1) and expects plan repair or abortion, as well as new plan execution, next. Furthermore, each aspect of the task analysis is immediately made available for communicative analysis.

Simultaneously the immediate focus is determined to be magtape, and after the plan structure is made available the global focus is determined to be mount plan; as in [gro77] the topic structure is thus in terms of the plan structure. Meanwhile, the clue word expert notes

the lack of any relevant phenomena. Finally, communicative analysis concludes $BIDTOPIC(user, system, p.1)$ and the d-effects are noted as expectations. Any communicative effects of the literal speech act (here to answer 'yes' or 'no') must also be dealt with. It was similarly suggested earlier that requests might also have d-effects.

It is also interesting to note what would happen if utterance (1) began with 'please.' Communicative analysis knows 'please' suggests that the intended speech act is a request and makes this information available for task analysis, as shown in Figure 3-6. Since the d-effects are only needed for communicative analysis the formulation of a given action (for example, the REQUEST) could be different along each dimension.

Since analysis of utterances (2)-(4) involves local structure only a few thoughts will be presented at this time. $BIDTOPIC$ could be formulated as an introduction followed by optional elaborations. Communicatively then, each utterance can be viewed as a further elaboration of the first. It should also be noted that communicative actions like elaboration and clarification seem to be universal d-effects, e.g. appropriate anywhere. Viewed along the task dimension, the utterances indicate modifications of the plan structure to be executed. For example, utterance

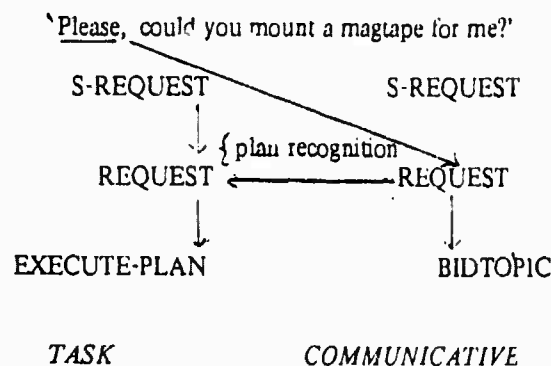


Figure 3-6. Analysis using 'please'

Figure 3-7 presents an analysis of utterance (6), illustrating somewhat different behavior. Each dimension begins with certain expectations determined by the preceding context. It should also be noted that due to the cooperative behavior assumption the system's goal acceptance was assumed. If this was not the case the acceptance would need to be acknowledged. At the task dimension, utterance (5) had indicated that the system's plan (and thus the user's) had failed; the user thus expects the system to either repair or abort the plan (the next stages of the planning cycle). At the communicative dimension (5) satisfied the acknowledge expectation; the d-effects of ACKNEG (helpful response or topic shift) are thus the new expected legal discourse moves. Each dimension then begins its processing. At the communicative dimension the anaphoric 'it' of utterance (6) supports the helpful response; since the focus has not shifted, popping' to the level of embedding of the BIDTOPIC is unlikely. This information is communicated to the task dimension as supporting the repair expectation. The task dimension then limits its recognition attempts to repair and meets with success.

TASK ANALYSIS		COMMUNICATIVE ANALYSIS	
repair	<---->	helpful response	TIME
abort	(expectations)	bidtopic	
execute-plan			
.		.	
.		.	
repair strengthened	<----	anaphora supports	
(search space cut)		helpful response	
.			
.			
repair confirmed	---->	helpful response confirmed	

Figure 3-7. User's Recognition of Utterance (6)

Note that the communicative plan recognizer is hypothesized to be syntactically based, using surface phenomena whenever possible. An alternative approach would be to do full-scale recognition like the task recognizer. This seems to be less desirable since the communicative recognizer would be redoing a lot of the work of the task recognizer, at least using the formulation of communicative acts given above. Of course there will be cases when such analysis is necessary, since no syntactic clues exist. Consider

A: 'Can you go to the movies tonight?'
B: 'I have to study.'

To recognize B's utterance as a refusal, reasoning at the task dimension is necessary (i.e. reasoning that studying is incompatible with going to the movies). It might be reasonable, however, if there are fewer plans possible at the communicative dimension than task dimension or when the task and communicative dimensions are not so tightly coupled. For example after a yes/no question yes, no, or clarification are the only reasonable communicative expectations,

Another unresolved issue involves the communication between the two dimensions. When should communication occur and in what type of language? Should messages be in some common language or should each dimension need to know about the internals of the other?

8. Comparisons

Before concluding it would be useful to compare this approach to previous work. With the exception of Argot, computational discourse structures are either not hooked up with task structure [mck82] [rei81] or collapsed [gro77]. The structural similarity noted by Grosz is reflected in the redundancy of the plan recognizers. With respect to Reichman, since her task level plans are in a sense communicative (teaching [rei79], debating [rei81]), they seem to become part of her discourse level. The objects and plans at the task dimension can in fact be considered the parameters of the actions at the discourse dimension. For example, the plan to

mount tape T376 is an argument of the BIDTOPIC in Plan1. While Argot's levels interact, strategy (3) is useful (at least for the examples); the communicative level indicates manipulations at the task level. Although the example above is like this other examples aren't, e.g. when the discourse structure is violated. Finally, a bit more needs to be said on the need for simple communicative plan recognition. Most models of discourse coherence involve high-level semantic relations, for example amplification [mck82] and illustration [rei78]. Recognition and generation of such relations tends to be left unspecified, i.e. done by magic (and often humans can't even agree); McKeown is able to use them due to her extremely restricted domain. It is thus suggested that although perhaps descriptively nice, they are computationally intractable. It should also be noted that the recognition proposed above uses surface phenomena to determine discourse structure as well as vice-versa.

CHAPTER 4

Summary

The beginnings of a plan-based natural language system that incorporates both communicative and task analysis has been presented (although it should be noted that more dimensions will ultimately be needed). Figure 3-5 shows that identification of the information and processes necessary along these two dimensions is of primary concern. The linguistic dimension will be simulated. With respect to communicative analysis an initial representation of communicative actions was discussed, in particular how to incorporate knowledge of legal moves as action effects rather than grammars. The subtle differences implicit in various surface realizations were also examined, as well as the structure of these communicative actions in actual dialogues. It was suggested that both local and global discourse structures are necessary (although analysis of the latter has been emphasized here). Thus, a syntactic approach to the identification, formulation, and implemented recognition of communicative actions and structures is a major goal of this work. Analysis and incorporation of the results mentioned in the Appendix also needs to be undertaken. Immediate focus [sid83] will be assumed (simulated). To be considered successful, the final model will need to subsume or present an alternative to previous work; furthermore, it will need to include local structure.

With respect to task analysis, it seems that to deal with dialogue one must include knowledge of the complete planning cycle (generation, execution, and repair) rather than just plan generation. For example, in Dialogue 1 the system's plan generation begins in the middle of the user's plan execution. The major question then is what exactly is this level trying to recognize, two-agent planning cycles or domain plans (e.g. mount)? If the former, the necessary model will likely be simulated since it is a thesis of its own. Furthermore, if the recogni-

tion procedure turns out to be exactly Allen's [all83] this will also be simulated

Finally, the two dimensions of analysis will be integrated in a final implementation, envisioned as follows. Communicative and task knowledge work in parallel, one source dynamically taking control over the other and reducing the search space, depending on the kind of discourse (a task-oriented one, a conversation...). Communicative recognition is hypothesized to be simple, using the knowledge provided by the analysis of surface phenomena and task plan recognition. As before, such a system will need to subsume previous work to be considered successful. By incorporating a (domain independent) communicative dimension which interacts with a domain (here task) dimension, we hope to be able to participate in more complex discourses (like Dialogue 1) than in the past.

To reiterate, further determination and clarification of the information and types of processing present along each dimension is of utmost importance. Moreover, many issues remain for possible examination. Although utterance (7) is a BIDTOPIC it is realized very differently than utterance (1). How is this a result of the discourse context? Or perhaps utterance (7) is better viewed as a topic modification rather than a topic introduction. Also, the relationship between the model proposed here and the context space model [rei78] and the ATN formulation [rei81] needs to be determined. Finally there are numerous orthogonal issues which are obviously beyond the scope of such research. For example, determining the effect of intonation (and other distinctions between spoken and written language) as well as the effect of continual output (as opposed to waiting until an utterance has been completely analyzed) is much too ambitious.

CHAPTER 5

Acknowledgements

Much of this proposal grew out of my work with the Knowledge Representation for Natural Language Understanding Group of Bolt Beranek and Newman Inc. I would like to thank Candy Sidner for both the many, many hours spent discussing these ideas and her helpful comments on an earlier draft of this paper. I would also like to thank Marc Vilain and Brad Goodman for their interest and comments.

James Allen was the source of many of the ideas developed above and has continued pushing them with me at Rochester. Finally, thanks to James, Pat Hayes, Henry Kautz and Emil Rainero for their comments on recent drafts of this paper.

CHAPTER 6

References

- [all81]
J. F. Allen, What's Necessary to Hide?: Modeling Action Verbs, *Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics*, 1981, 77-81.
- [all82a]
J. F. Allen, A. M. Frisch and D. J. Litman, ARGOT: The Rochester Dialogue System, *AAAI*, 1982.
- [all82b]
J. F. Allen, ARGOT: A System Overview, Technical Report 101, University of Rochester, April 1982.
- [all83]
J. F. Allen, Recognizing Intentions from Natural Language Utterances, in *Computational Models of Discourse*, M. Brady (ed.), MIT Press, Cambridge, MA, 1983.
- [app81]
D. E. Appelt, *Planning Natural Language Utterances to Satisfy Multiple Goals*, PhD Thesis, Stanford University, 1981.
- [bat78]
M. Bates, The Theory and Practice of Augmented Transition Network Grammars, in *Natural Language Communication with Computers*, (publisher?), 1978.
- [coh79]
P. R. Cohen and C. R. Perrault, Elements of a Plan-Based Theory of Speech Acts, *Cognitive Science* 3, (1979), .
- [gro77]
B. J. Grosz, The Representation and Use of Focus in Dialogue Understanding, Technical Note 151, SRI, July 1977.
- [gro79]
B. J. Grosz, Utterance and Objective: Issues in Natural Language Communication, *IJCAI*, 1979.
- [gro82]
B. J. Grosz, Focusing, Coherence, and Referring Expressions, BBN Seminar, Summer 1982.
- [hob81]
J. R. Hobbs and M. Agar, Text Plans and World Plans in Natural Discourse, *IJCAI*, 1981.
- [hor77]
M. K. Horrigan, Modelling Simple Dialogs, Master's Thesis, Technical Report Number 108, University of Toronto, May 1977.
- [inc79]
B. B. N. Inc., Research in Natural Language Understanding, BBN Report No. 4274 (Annual Report), September 1978 - August 1979.

- [lev79a]
D. Levy, Communicative Goals and Strategies: Between Discourse and Syntax, in *Syntax and Semantics*, vol. 12, T. Givon (ed.), Academic Press, New York, 1979, 183-212.
- [lev79b]
D. M. Levy, *The Architecture of the Text*, PhD Thesis, Stanford University, December 1979.
- [man77]
W. C. Mann, J. A. Moore and J. A. Levin, A Comprehension Model for Human Dialogue, *IJCAI*, 1977.
- [mck82]
K. R. McKeown, *Generating Natural Language Text in Response to Questions about Database Structure*, PhD Thesis, University of Pennsylvania, 1982.
- [min75]
M. Minsky, A Framework for Representing Knowledge, in *Psychology of Computer Vision*, Patrick Henry Winston (ed.), McGraw-Hill, New York, 1975.
- [rei78]
R. Reichman, Conversational Coherency, *Cognitive Science* 2, (1978), .
- [rei79]
R. Reichman, Conversational Coherency in Technical Conversations, Working Papers 43, Institut dalle molle d'etudes semantiques et cognitives, Universite de Geneve, 1979.
- [rei81]
R. Reichman, Plain Speaking: A Theory and Grammar of Spontaneous Discourse, Bolt, Beranek and Newman Report No. 4681, 1981.
- [sch77]
R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals, and Understanding*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.
- [sid82]
C. L. Sidner, Protocols of Users Manipulating Visually Presented Information with Natural Language, Bolt Beranek and Newman Report 5128, September 1982.
- [sid83]
C. L. Sidner, Focusing in the Comprehension of Definite Anaphora, in *Computational Models of Discourse*, M. Brady (ed.), MIT Press, Cambridge, MA, 1983.
- [swa81]
W. R. Swartout, Explaining and Justifying Expert Consulting Programs, *IJCAI*, 1981.
- [wil78]
R. Wilensky, Understanding Goal-Based Stories, Research Report #140, PhD Thesis, Yale University, September 1978.
- [win72]
T. Winograd, *Understanding Natural Language*, Academic Press, London, 1972.

CHAPTER 7

Appendix

1. R. Beaugrande and W. Dressler, *Introduction to Text Linguistics*, (publisher?), 1981.

Views the text by how it functions in human interaction. In particular, text is viewed as a communicative occurrence which meets seven standards of textuality - cohesion, coherence, intentionality, acceptability, informativity, situationality, intertextuality.
2. L. H. Carlson, *Dialogue Games: An Approach to Discourse Analysis*, PhD Thesis, MIT, 1982.
3. W. L. Chafe, Givenness, Contrastiveness, Definiteness, Subjects and Topics, in *Subject and Topic*, Charles Li (ed.), Academic Press, New York, NY, 1976.

Discusses the ways in which a speaker accommodates his/her speech to temporary states of the addressee's mind. In particular, nouns have packaging statuses - how the content is transmitted rather than the content itself (case status) - dependent on the hearer's cognitive state. Discusses 6 packaging phenomena - those in the title and empathy. Good.
4. N. E. Enkvist, Stylistics and Text Linguistics, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Discusses stylistics, in particular how textlinguistics aids stylistics as well as vice-versa. Example - the study of style markers which require description in text-linguistic terms. Reviews relevant work and examples.
5. J. A. Goguen, J. L. Weiner and C. Linde, Reasoning and Natural Explanation, Miscellaneous Report, 1982.

Presents a precise and computationally effective model of the structure of (naturally occurring) human explanation (viewed as social process). Explanations are represented by trees whose internal nodes correspond to types of justification; production is represented by a sequence of transformations. Focus is represented by pointers, and shifts by ptr. movement. The ordering and embedding of explanations are considered. Discusses implications for a.i.
6. J. E. Grimes, Context Structure Patterns, *Nobel Symposium on Text Processing*, 1980.

A framework for text typology is presented, based on the notion of context spaces and structure - a referential core description (who, when, where, theme) and pointers indicating development or subordination. Vague.

7. J. E. Grimes, Narrative Studies in Oral Text, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977.

Discusses 3 recurring themes which he believes are 3 partially independent sub-systems of language - content (what to say), cohesion (how to relate to what has gone before), staging (perspective to stage what to say).

8. M. A. K. Halliday and R. Hasan, *Cohesion in English*, Longman, London, 1976.
9. R. Hasan, Text in the Systemic-Functional Model, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977.

Views text as a linguistic entity, not a super sentence. Concerned with the 2 notions most fundamental to textness - texture (what makes a string of sentences become a text) and structure (allows incomplete vs. complete texts, genres) as well as how structure/genre is controlled contextually. Texts realize (not constitute) genre structure. Good.

10. G. Jefferson and J. Schenkein, Some Sequential Negotiations in Conversation, in *Studies in the Organization of Conversational Interaction*, Jim Schenkein (ed.), (publisher?), 1979.

Extends intuitive observations on conversational data into reflections of underlying structural phenomena. In particular, views passes of moves as sequential expansions of unexpanded versions of projected action sequences. Furthermore, the passes are sensitive to the projected action sequence possibilities (i.e. the position before an acknowledgment inherits efforts to continue, so it might be negotiated). Finally, certain positions may be considered as pairs rather than individually. Somewhat interesting but poorly written.

11. G. Jefferson, Sequential Aspects of Storytelling in Conversation, in *Studies in the Organization of Conversational Interaction*, J. Schenkein (ed.), (publisher?), 1979.

Conversation is occupied by activities relevant to the telling of a story, where the story itself occupies a portion of the fragment. Stories emerge from turn-by-turn talk, they are locally occasioned by it (often predictable) and upon completion, stories re-engage negotiable turn by turn talk (are sequentially implicative).

12. M. Kay, Unification Grammar, (*Conference?*), Summer 1982.

Argues for reintroduction of functional considerations of grammar, since no fundamental inconsistency between this and structural/generative considerations and has potential for contributing to a more revealing account of discourse phenomena than by either alone. The grammar produces outputs in response to specific functional inputs which the linguistic component then unifies with part of the grammar. Different syntactic forms are not arbitrary (and just charted) but reflect meaningful choices of a speaker. Unification grammar relates sentences to both their logical form and (orthogonal) function; the only specifically syntactic devices are concerned with linear ordering. Not particularly relevant.

13. S. Kuno, Generative Discourse Analysis in America, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Presents a new trend in generative grammar known as functional syntax, which views the problems of generative syntax within the framework of discourse analysis. I.E. all factors which control linguistic phenomena (rather than those statable within the syntactic component of TG) are considered. Presents accounts of phenomena such as pronominalization and empathy. Nice.

14. C. Linde, The Organization of Discourse, in *The English Language in its Social and Historical Context*, Shopen, Zwicky and Griffen (ed.), (publisher?), .

Presents support for the existence of discourse units. Then using the units of joke, narrative, and apartment description, investigates the following: What are the boundaries and internal structure? How are the syntax and focus of the surface phenomena affected? How do beliefs and attitudes (social factors) influence? Various principles of coherence are explored as well - temporal ordering, trees, social norms. Interesting stuff which would seem to benefit from the rigor of c.s.

15. R. Longacre and S. Levinsohn, Field Analysis of Discourse, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Presents classification of discourse genre (by parameters) with associated schemas of deep structure, as well as lists of cohesive devices. A methodology for systematically displaying material (how and what) is also discussed. Not a must-read (except for list of surface phenomena).

16. M. Marslen-Wilson, E. Levy and L. K. Tyler, Producing Interpretable Discourse: The Establishment and Maintenance of Reference, in *Speech, Place, and Action*, R.J. Jarvella and W. Klein (ed.), (publisher?), 1982.

Embedded in a comprehension system with on-line interpretation and cooperation of knowledge sources. Presents a distributional analysis of reference, illustrating dependence on both the narrative function and informational context; the use of names, descriptions, pronouns, gestures... must be explained w/r/t discourse history and cognitive functions underlying referential use. A pragmatic inference view of reference resolution is argued for.

17. D. Metzger, Parsing Task-Oriented Dialogue Interactions, University of Bielefeld.

18. W. Noth, The Semiotic Framework of Textlinguistics, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Explicates semiotics (study of sign systems or codes, verbal and non-verbal communication as texts) and its relationship to textlinguistics. Presents example analyses within various semiotic frameworks of textlinguistics. Bizarre.

19. Z. Palkova and B. Palek, Functional Sentence Perspective and Textlinguistics, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

FSP is a description of the sentence from the point of view of its (potential) use in a message. The paper discusses how the phenomena described by FSP can help a text-grammar approach, as well as the resulting methodological requirements. Basically introduces FSP "rules" (more like tendencies) for basic concepts.

20. J. S. Petofi, A Formal Semiotic Text Theory as an Integrated Theory of Natural Language, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Overviews methodological properties of an integrated formal theory - what is taken from past work, other reasons for considering certain aspects as important, then presents a particular theory.

21. L. Polanyi, Literary Complexity in Everyday Storytelling, in *Spoken and Written Language*, vol. IX, Deborah Tannen (ed.), Ablex Publishing Corporation, 1982.

Shows that oral stories demonstrate the same complexities as found in literary language- point of view, identity of reference, and multiplicity of meaning. Thus, these features do not define literariness. Some aspects discussed bring to mind the work of Reichman however - shifters (deictics, pronouns...).

22. H. Rieser, On the Development of Text Grammar, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Surveys the approaches taken to text grammar and discourse analysis - pregenerative, generative (interpretative, vs. sentence, semantics, logic), and montage. Presupposes a lot of knowledge for understanding this.

23. H. Sacks, E. A. Schegloff and G. Jefferson, A Simplest Systematics for the Organization of Turn-Taking for Conversation, *Language* 50, 4, Part 1 (December 1974), .

Presents a simplest systematics (components and rules), grossly observable phenomena (facts), and how the system accounts for the facts, of an independent turn-taking system for conversational speech-exchange systems. The system is locally and interactively managed, has general abstractness/local particularization. Syntax conceived in terms of its relevance to turn-taking.

24. T. A. Dijk and W. Kintsch, Cognitive Psychology and Discourse: Recalling and Summarizing Stories, in *Current Trends in Textlinguistics*, Wolfgang U. Dressler (ed.), Walter de Gruyter, Berlin, 1977 .

Assumes discourse processing (understanding, organization, retrieval) is a function of the structures assigned to the discourse during input. A theoretical framework is presented (a theory of discourse, a theory of discourse structure processing, and more general theory for complex cognitive information processing). Also discussed are psychological hypotheses and supporting evidence, in particular: macro-structures are stored in memory and used as cues; (narrative) schema needed to comprehend; macro-structures constructed in comprehension. Good paper.

Official Distribution List

Contract N00014-77-C-0378

	<u>Copies</u>
Defense Documentation Center Cameron Station Alexandria, VA 22314	12
Office of Naval Research Information Systems Program Code 137 Arlington, VA 22217	2
Office of Naval Research Code 200 Arlington, VA 22217	1
Office of Naval Research Code 455 Arlington, VA 22217	1
Office of Naval Research Code 458 Arlington, VA 22217	1
Office of Naval Research Branch Office, Boston 495 Summer Street Boston, MA 02210	1
Office of Naval Research Branch Office, Chicago 536 South Clark Street Chicago, IL 60605	1
Office of Naval Research Branch Office, Pasadena 1030 East Green Street Pasadena, CA 91106	1
Naval Research Laboratory Technical Information Division Code 2627 Washington, D.C. 20380	6

cont'd.

Naval Ocean Systems Center Advanced Software Technology Division Code 5200 San Diego, CA 92152	1
Dr. A. L. Slafkosky Scientific Advisor Commandant of the Marine Corps (Code RD-1) Washington, D.C. 20380	1
Mr. E. H. Gleissner Naval Ship Research & Development Ctr. Computation & Mathematics Dept. Bethesda, MD 20084	1
Capt. Grace M. Hopper, JNR Naval Data Automation Command Code 00H Washington Navy Yard Washington, D.C. 20374	1
Mr. Paul M. Robinson, Jr. NAVDAC 33 Washington Navy Yard Washington, D.C. 20374	1
Advanced Research Projects Agency Information Processing Techniques 1400 Wilson Boulevard Arlington, VA 22209	1
Capt. Richard L. Martin, USN 507 Breezy Point Crescent Norfolk, VA 23511	1
Director, National Security Agency Attn: R54, Mr. Page Fort G.G. Meade, MD 20755	1
Director, National Security Agency Attn: R54, Mr. Glick Fort G.G. Meade, MD 20755	1
Major James R. Kreer Chief, Information Sciences Dept. of the Air Force Air Force Office of Scientific Research European Office of Aerospace Research & Development Box 14 FPO New York 09510	1

cont'd.

Mr. Fred M. Griffiee
Technical Advisor C3 Division
Marine Corps Development
& Education Command
Quantico, VA 22134

1