

# Links Between Situation and Language Use: Towards a more Coherent Interaction

*Michael Elhadad*

Department of Computer Science  
Columbia University  
New York, NY 10027  
Elhadad@cs.columbia.edu

CUCS-289-87

26 February 1988

## Abstract

This paper is a survey of the techniques used to take into account the influence of situation in Natural Language (NL) systems. A model of the influence of situation on language use will provide a coherent interaction between a program and its users. We propose a classification in three types of the characteristics of situation according to the way they affect language use : linguistic, semantic and pragmatic, and we study the different techniques developed for each of them. We also contrast these techniques to linguistic theories dealing with the issue.

## Table of Contents

<b>1. Introduction: The Need to Take Situation into Account in a Natural Language System</b>	<b>1</b>
<b>2. The Concept of Situation: a Classification</b>	<b>2</b>
2.1. The Concept of Situation: a First Classification	2
2.2. Problems with the Classification	3
<b>3. Linguistic Situation: Text Organization and Textual Cohesion</b>	<b>4</b>
3.1. Tools for Cohering a Text: Global Models - Rhetorical schemas	4
3.1.1. Schemas and Discourse Organization in TEXT	5
3.1.2. Text Organization in RST	5
3.1.3. Problems with the Schemas	6
3.2. Local Models of Cohesion	7
3.2.1. Linguistic Cohesive Devices	8
3.2.2. The Theme/Rheme Articulation	9
3.2.3. Focus of Attention	10
3.2.4. Grosz's model of Focus Space	11
3.2.5. Sidner's Analysis of Immediate Focus	12
3.2.6. McKeown's Study of Focusing for Generation	12
3.3. Summary	14
<b>4. Links with the Extra-linguistic Situation</b>	<b>15</b>
4.1. Ms. Malaprop: the Need for Inference	16
4.2. From Scripts and Plans to MOPS: the Importance of Memory	17
4.2.1. Scripts	17
4.2.2. MOPs and Dynamic Memory	18
4.2.3. Why MOPs are Representative: the Principle of Distinction	19
4.3. Memory and Communication: a Critique of Current Knowledge Representation Techniques Applied to NLP	19
4.3.1. MOPs and the Intra-personal Study of Language	20
4.3.2. Distinction and the Interpersonal Dimension of Language	20
4.3.3. Links between Pragmatic and Semantic Situation	21
4.4. Situation Semantics	21
4.5. Summary	24
<b>5. Pragmatic Aspects of the Interaction: Influence of the Situation of Utterance on Language Use</b>	<b>25</b>
5.1. The Planning Theory of Communication	25
5.1.1. The Speech-Act Theory	25
5.1.2. SAs in AI: Allen and Perrault's Planning Theory	26
5.1.3. Developments of the Planning Theory of Communication	27
5.2. Grice's Theory of Meaning - Conversational Implicatures	29
5.2.1. Using Implicatures in AI Systems	32
5.3. What Must Be Integrated in a Representation of the Pragmatic Situation ?	33
5.4. Summary	36
<b>6. The Notion of Situation from a Theoretical Point of View</b>	<b>36</b>
6.1. The Concept of Situation in Linguistics: Who Cares ?	36
6.2. The Role of Situation in Theories	37
6.3. Situation in Theories: Definitions	39
6.4. The Distinction Between Semantic and Interactive Analysis	41
<b>7. Towards a Model of a Coherent Interaction</b>	<b>41</b>
<b>ACKNOWLEDGEMENTS</b>	<b>42</b>

<b>List of Figures</b>		
<b>Figure 3-1:</b>	<b>Example of a Rhetorical schemas defined in Mann's RST</b>	<b>6</b>
<b>Figure 3-2:</b>	<b>Lexico-grammatical Cohesive Devices from [Hasan 84, p.185]</b>	<b>8</b>
<b>Figure 3-3:</b>	<b>Focus Used as a Constraint on Proposition Ordering in the TEXT System</b>	<b>13</b>
<b>Figure 5-1:</b>	<b>Allen &amp; Perrault's Model of SAs</b>	<b>27</b>
<b>Figure 5-2:</b>	<b>Distinction between Speaker's and Hearer's Beliefs in [Pollack 86]</b>	<b>28</b>
<b>Figure 5-3:</b>	<b>The Principle of Co-operation in (Grice-67)</b>	<b>31</b>
<b>Figure 5-4:</b>	<b>Different Types of Implicatures, in [Levinson 83]</b>	<b>31</b>
<b>Figure 5-5:</b>	<b>Using Scalar Implicatures</b>	<b>33</b>
<b>Figure 6-1:</b>	<b>Linguistic Determinism according to [Fishman 72]</b>	<b>38</b>
<b>Figure 7-1:</b>	<b>Links between situation and linguistic phenomena</b>	<b>43</b>

## 1. Introduction: The Need to Take Situation into Account in a Natural Language System

The way situation affects language use is an important issue in linguistic research. Very little work in natural language processing (NLP) has been aimed at a general model of situation influence on language use. However, existing programs have met the issue and have taken situation into account in some ways in their description of natural language. A model of the influence of situation on language use will provide a coherent interaction between a program and its users. This paper is a survey of the techniques which have been developed to make natural language systems situation sensitive and to make interaction with them more coherent.

“Situation” refers to the set of all conditions affecting language use in an interaction. We consider it as a more general term than “context,” that will be restricted here to denote the linguistic surroundings of an utterance (also called co-text). The following examples illustrate aspects of situation influence on language use.

- (1) He wrote this program. He is not very smart.
- (2) He wrote this program, but he is not very smart.

What is the meaning of the proposition “*He wrote this program*”? AI researchers at first would attempt to describe this meaning as:

- (3) program(P) & write(X, P).

But then, in (1), the writing of the program in question seems to be the proof of its author’s limited intellectual faculties, whereas in (2), the same proposition could lead to the conclusion that the program is a complex piece of code, and that being able to compose such a work is a respectable achievement. Thus the representation (3) does not fully explain the “meaning” of the proposition since in (1) and (2), the juxtaposition of the two propositions generates additional meaning.

Sometimes, additional meaning comes from other sources than the textual surroundings of a proposition:

- (4) I’ll go to school tomorrow with Ron.
- (5) He wanted to come to the party, but he lost the keys to his car.
- (6) He then discovered the spaceship. The machine was a roaring, impressive, symbol of the Luxemburgese technology.

To interpret (4), one needs to know when it was said and who was speaking in order to correctly instantiate the values of “*I*” and “*Tomorrow*.” In addition, a complete interpretation requires knowing what school the locutor attends, and knowing his/her acquaintances to determine what “*school*” refers to and who “*Ron*” is.

Similarly, to interpret (5), world knowledge is necessary to understand that a car is generally used to move from place to place and that a car needs keys to be operated. In (6), world knowledge is again necessary in order to understand that “*machine*” refers to “*spaceship*” and that a spaceship is a complex technological machine.

The interpretation of these sentences shows that the meaning of language depends on the physical situation of its use. To interpret pronouns, proper names or deictics one needs to refer to the physical circumstances surrounding an act of utterance and to use additional domain knowledge.

Knowledge about who is speaking to whom and for what purpose is also relevant to understanding sentences.

such as:

- (6) Come here !
- (7) Could you please come ?
- (8) You will go to Paris tomorrow.

(6) and (7) roughly convey the same information and roughly achieve the same effect. However, (6) would be much more imprudent than (7) from an employee to his/her boss. The relationships between the locutors of (6) and (7) determine two completely different linguistic realizations of the same act. Similarly, knowing the relationships between the speaker of (8) and its hearer would help determine whether (8) is a promise, a threat or an assertion.

In this paper, I survey how the problem of situation and its influence on language use has been addressed in Artificial Intelligence (AI) research. In the next section, I propose an *a priori* distinction between three types of situations : linguistic (the textual surroundings), semantic (the physical circumstances), and pragmatic (the relations between the locutors). In the three following sections, I review typical works, techniques and concepts used in AI to deal with each type of situation. Finally, section 6 presents concepts developed in theoretical linguistics that highlight some of the limitations of current AI approaches. Several notions proposed by ethnolinguists and sociolinguists are very useful for studying certain aspects of situation in a disciplined manner.

The major thesis defended throughout the paper is that, in order to be able to deal with the full complexity of meaning, Natural Language Processing (NLP) must address the whole process of communication and study language use within practical human interactions, as opposed to an abstract study at the sentence level. In particular, certain phenomena which seem hard to explain with a purely linguistic analysis, become easier to capture if more attention is paid to the situation. We put particular emphasis on one aspect of situation that is often neglected - the pragmatic situation - and claim that it could explain a number of hard problems in NLP if it were more extensively used.

## 2. The Concept of Situation: a Classification

In this section, a first approach to the concept of situation as currently used in AI research is presented. An *a priori* distinction between three kinds of situations is proposed, based on Morris' definition of syntactic, semantic and pragmatic levels. These three perspectives are used here as classifying categories applied to the notion of situation, and allow one to discern between the multiple and heterogeneous parameters involved in the study and delimitation of this notion.

### 2.1. The Concept of Situation: a First Classification

Morris' distinction of syntax, semantics and pragmatics has been one of the single most important methodological contributions to the way linguistics is currently shaped [Morris 71]. These three levels of analysis correspond to distinct complementary perspectives on the same complex object - the language use in context. The same general approach can be applied to help clarify the notion of situation.

- **Syntax:** answers the question "How are words combined - what are the constraints on this

combination?" The key word at this level is **grammaticality**. From this perspective, situation can be described as **the linguistic context**. The management of linguistic context has as a goal to achieve **textual cohesion**. It relies on the study of how basic linguistic units (propositions or sentences) are related to each other, and cohere to realize a structure like a text.

- **Semantics**: answers the question "How are utterances related to the world?" The key words in this perspective are **reference or denotation**. For a logician's point of view, it is the study of what is **true or false**. We shall say that the **extra-linguistic situation** is constituted by all the elements that can affect the truth value of a sentence, all that is referred to by words.
- **Pragmatics**: answers the question "Who is talking to whom, to which aim, and in which manner?" This perspective focuses primarily on the usage of language and its function. The key word is now **appropriateness** similar to the notion of grammaticality for well-formed sentences. Research here is primarily concerned with the link between language and its users. This approach applied to the notion of situation defines the **social or interpersonal context**. We shall refer to it as **the pragmatic situation**: its study is concerned with the analysis of **speaker/hearer characteristics and intentions**, the influence of **social factors** on the language used, and the **linguistic devices** allowing speakers to make their contributions as appropriate as possible.

## 2.2. Problems with the Classification

The three levels of classification presented correspond to different perspectives, motivations and goals in the study of language. But these three levels of analysis must also be combined in a global theory. Recent arguments between linguists have shown that such an integration is not an easy problem: is there a sequential order between the three levels? Is it possible to study one level without any link to the others? Is it possible to build a theoretical machine composed of three isolated components activated successively ?

The distinction between syntax and semantics is a well known discussion inside AI. Purely syntactic parsers such as the one used in LUNAR [Woods and Kaplan 72], produce a parse tree to be analyzed by a separate semantic analyzer. In contrast, in semantic grammars the parsing process consists in the unification of an actual text with semantic patterns without reference to the syntactic structure (LIFER [Hendrix 77]). Integrated parsers (IPP e.g. [Lebowitz 80]) or the use of cascaded ATNs ([Woods 80]) where both syntactic and semantic analysis are achieved together merge features of both approaches. This evolution indicates that no clear boundary can practically be installed between syntactic and semantic processing.

Similarly, the separation of semantics and pragmatics is questionable. For example, the truth value or the denotation of sentences containing words called indexicals ('*here, now, you, I ...*'), is determined by the pragmatic situation. Other more difficult problems arise in trying to define more precisely the boundary between these two levels when studying verbs of attitude such as '*believe, want, hate*' [Barwise & Perry 83]. Linguists like [Anscombe & Ducrot 83] argue that the semantic description of words like '*but, almost, already*' necessarily involves reference to the pragmatic situation, speaker's intentions and beliefs about the hearer, pointing the need for *integrated pragmatics*. Jameson's program, IMP, [Jameson 83] also provides a "semantico-pragmatic" description of words like '*unfortunately, by the way*' based on a user description and his/her expectations.

At our level of analysis, the same general problem of organization faces this tentative classification of different aspects of the notion of situation. There is no reason to believe that the linguistic context can be studied or modeled in isolation from the semantic or pragmatic situation. It serves here only as a useful *a priori* distinction, without any

claim of theoretical validity. But we keep in mind that the interaction between the different levels is a problem, and try to inspect how the different techniques developed to deal with a certain level could be interfaced to other techniques dealing with other aspects of situation.

### 3. Linguistic Situation: Text Organization and Textual Cohesion

The notion of textual cohesion provides an explanatory framework for numerous empirical observations; the basic idea is that a text is a complex whole whose properties are more than those of a collection of isolated units. For example, the problem of 'organizing' a text is familiar to writers; it seems empirically verified that a strongly structured text is easier to read and understand [Calfey & Curley 84]. Isolated linguistic units such as propositions or sentences often do not have a complete meaning; part of their meaning is determined by their textual surrounding, their linguistic context.

This section presents work done in AI related to the notion of textual cohesion that is concerned with the task of handling the linguistic context. Most of this presentation will be devoted to describe the linguistic theoretical framework of this AI research, as it appears to have strongly influenced this field.

Two approaches exist to the study of text coherence: global and local. Top-down global analyses of coherence focus on the text and try to determine its structure. Such methods find their roots in the venerable field of rhetoric, which is the object of a current resurgence of interest. A second approach consists of locating the features of more basic linguistic units that relate them to other units. Such bottom-up (local) approaches are mainly founded on the work of the functionalist schools of linguistics. These two approaches are obviously not exclusive since rhetoric has to account for how the overall structure of a text is recognized by readers/hearers and how a given rhetorical decision imposes constraints upon the writer/speaker to organize low-level units. Similarly, those who don't follow a theory of rhetoric agree that language is a resource to realize higher-level 'functions,' and that low-level cohesive devices are tools to build such a higher level meaning.

#### 3.1. Tools for Cohering a Text: Global Models - Rhetorical schemas

In AI, theories of text organization have mostly been developed in relation to research in text generation. In trying to generate text segments beyond the sentence limit, researchers have discovered that grammar rules alone are not sufficient to express what it means to be a "good," easy to read text. It appears that text-generation programs require a specification of some way to organize their output, and that if this organization is not realized according to certain rules humans are used to, the text generated turns out to be awkward, inappropriate, even misleading. As pointed out in [Mann 84]: "people commonly recognize that well-written text is organized and that it succeeds partly by exhibiting its organization to the reader."

Because of these common expectations from readers, a text needs to follow certain rules of organization in order to be well understood. First approaches in that direction have tried to describe "text-grammars" [Van Dijk 77] or story-grammars [Rumelhart 75]. But the nature of the rules involved in the process of text-organization is seemingly different than the nature of grammatical rules: nothing can prove that a given organization is rigidly necessary under certain circumstances. On the contrary, as noticed by McKeown, "many experienced and talented

writers purposely **break** norms in order to achieve a striking literary effect'' [McKeown 85]<sup>1</sup>.

Thus, organizational patterns can be considered as resources available to the communicant to achieve certain communicative functions instead of normative rules constraining his/her potential expressive ability. This is the option followed by McKeown in her implementation of the TEXT system [McKeown 85], and by Mann, as the basis of his Rhetorical Structure Theory [Mann 84] (see also [Hobbs 78]; [Cohen 83]).

### 3.1.1. Schemas and Discourse Organization in TEXT

The TEXT system [McKeown 85] is an interface to a data-base which is capable of generating paragraph-long answers to user queries. The range of questions supported by the system include queries for description of entities contained in the data-base, comparison between different entities, and explanation of the structure of the data-base. In order to handle these different types of questions, McKeown has developed a set of "rhetorical schemas." These schemas are used to guide the text-generation system in determining the content and the form of an answer. Different schemas exist related to the different types of questions. The choice of a schema thus appears as one of the linguistic devices available to the system to convey a certain kind of information.

Previous work in linguistics has studied the different organizations a text can follow [Shepherd 26, Levin 78]. The work of McKeown has consisted in making "high-level" advice on how to compose a text operational for a computer program. In the domain studied, the schemas developed for the TEXT system were:

- Identification - requests for definitions
- Attributive - requests about available information
- Constituency - requests for definitions - requests for available information
- Compare and contrast - requests about the differences between objects

We note here that each schema has a well defined function, a communicative purpose, and corresponds to a certain category of request.

The building units of the schemas are called *rhetorical predicates*. A rhetorical predicate correspond to a basic communicative function (like "illustration" or "comparison"), and is realized by a single linguistic unit (proposition or sentence). Schemas indicate how rhetorical predicates can be composed together to form a paragraph.

### 3.1.2. Text Organization in RST

In the same direction, Mann [Mann 84] is currently working on a general theory of text organization, which he calls *Rhetorical Structure Theory* (RST). His aim is to build a descriptive theory that reliably accounts for the diversity of naturally occurring texts. Mann also uses *schemas* to describe the overall organization of a text. Each schema is formed of *spans*. A schema is constituted of *nuclear* and *satellite* spans. *Satellites* are linked to their

---

<sup>1</sup>It is important to note that this assertion does not imply that there exists one norm, literature being defined as a departure from this neutral (hypothetical) form. Rather, it means that text organization is governed by multiple conventions. The same may be true at the sentence level





When one knows the overall structure of a text segment, it is possible to identify the rhetorical functions of its constituents, but there is little said as of now, on how to detect the status of a unit from purely linguistic clues. This limitation is best explained by the orientation of this research; the future of the RST for example, is to be a "constructive theory" for use in text-generation. If it were also intended to be a theory of text interpretation, then formal rigorous methods would have been necessary to determine the rhetorical function of linguistic segments. But for text generation, there is no necessity to design a theory with any kind of "recognition" device or interpretive mechanism. This limitation is also related to the second difficulty with these rhetoric theories.

It is necessary to view the study of rhetoric as part of a more general theory of action since overall, rhetoric is part of a theory of communication and action. This interpersonal aspect of rhetoric is mis-perceived, since by its historical tradition, Rhetoric has become mostly normative, and has propagated the illusion that there exists one "good way" of writing, one "good style." This tradition has often hidden that a text is as much a "piece of communication" as are speech or body gestures. A text is produced in certain conditions and addressed to a certain public. Therefore, it is most probable that the rhetorical aspects of a text are determined by the conditions of its production. If rhetorical schemas are communicative tools, their form is likely shaped by their use. And this use is essentially interpersonal.

All these remarks participate in the same argument: (1) the descriptive categories of a rhetorical theory cannot be only "semantic." (2) There is no "internal logic" to text organization. Rhetorical predicates and relations are not just textual. They are part of an interpersonal action - namely persuasion, argumentation, information etc. Thus, they have to acquire that interpersonal dimension. Paris' work is an interesting confirmation of this objection [Paris 85]. She has demonstrated that the structure of descriptions are different given the level of expertise of the audience. When describing technical objects to experts vs. naive persons, people not only vary the level of precision of their descriptions, but overall, they change the structure of their descriptions - **they use a different rhetorical schema.** Such results clearly confirm that there is no such thing as "one logical way of describing," but several, sharply different methods determined by the interpersonal conditions in which the descriptions were produced.

In summary, the highly intuitive explanatory power of schemas and the good results obtained in integrating such high-level structures in text-generation programs confirm the relevance of a kind of rhetorical theory in the realm of AI. Further developments of the theory, however, clearly indicate that the level of macro-textual cohesion is only a partial aspect of the general issue of coherence; particularly, the categories used in the description of a schema (*i.e.* for the example schema "inform," "elaboration," etc.) are not to be taken for granted, but rather should be related to a theory of interpersonal relations.

### 3.2. Local Models of Cohesion

Another approach in studying textual cohesion is more inductive in nature. Its aim is to account for "what can follow immediately at any given point" [Goldberg 83]. It consists in analysing how language as a system is used to reflect (or *realize* in the functionalist terminology) unity of a text. It is thus more of a bottom-up method, and focuses on local aspects of the textual cohesion, the lexico-grammatical devices, or the semantic articulation of the language available to make a text a "coherent" whole.

This section first presents the linguistic background of the study of local models of cohesion. This includes the study of lexico-syntactic devices of cohesion as defined by [Halliday & Hasan 76], and leads to the analysis of the semantic articulation of text variously called *theme/rheme*, *given/new*, or *topic/comment*. This notion is compared to the psychological notion of *focus of attention* as it is widely used in AI research. A survey of significant works incorporating focusing mechanisms in discourse processing closes the section.

### 3.2.1. Linguistic Cohesive Devices

The functionalist linguistic tradition has developed the important notion that a text is the result of a series of choices or decisions. Each level of decision is said to be "realized" in the next lower level - the process ending in the choice of lexical items. In this view, each known function of a text determines one of its linguistic aspects - syntactic form, lexical item, etc. One of the important functions posited by the theory and described in [Halliday & Hasan 76] is the so-called *textual function*; its effect is to cohere a linguistic segment into a text, "to provide texture and organize discourse as relevant to the situation" (Kress introduction to [Halliday 76]). According to the general orientations of the theory, language as a system provides speakers with a set of devices allowing the realization of that function.

Linguistic Devices Creating Cohesion	
<ul style="list-style-type: none"> <li>• <i>Reference</i> <ol style="list-style-type: none"> <li>1. pronominal</li> <li>2. definite article</li> <li>3. demonstrative</li> <li>4. comparative</li> </ol> </li> <li>• <i>Lexical</i> <ol style="list-style-type: none"> <li>1. reiteration               <ol style="list-style-type: none"> <li>a. repetition</li> <li>b. synonymy</li> <li>c. superordinate</li> <li>d. general word</li> </ol> </li> <li>2. collocation</li> </ol> </li> </ul>	<ul style="list-style-type: none"> <li>• <i>Substitution &amp; Ellipsis</i> <ol style="list-style-type: none"> <li>1. nominal</li> <li>2. verbal</li> <li>3. clausal</li> </ol> </li> <li>• <i>Conjunction</i> <ol style="list-style-type: none"> <li>1. cohesive conjunctive               <ol style="list-style-type: none"> <li>a. additive</li> <li>b. adversative</li> <li>c. temporal</li> <li>d. causal</li> </ol> </li> <li>2. continuative</li> </ol> </li> </ul>

**Figure 3-2:** Lexico-grammatical Cohesive Devices from [Hasan 84, p.185]

Figure 3-2 gives examples of the linguistic cohesive devices that are available in English. Hasan gives the significance of these elements: "Each of the devices entered in [figure 3-2] is a potential cohesive device; it represents a resource for the creation of cohesion. Cohesion is actually created through the use of a member of any one of these categories, a semantic bond is created between this member and some other element in the textual environment" [Hasan 84]. These devices create essentially *nonstructural relations* as opposed to the schemas reviewed in section 3.1.

The notion of cohesion is not widely used in AI. For text understanding, problems exist with the definition of

lexical cohesion which could produce the most interesting results: collocation is so wide a concept that it is not operational<sup>2</sup>. [Hasan 84] proposes alternative more workable concepts, but the notion of lexical cohesion remains difficult to handle for text understanding. In contrast, the notion of cohesion could be very helpful in text production (see [Granville 83]).

### 3.2.2. The Theme/Rheme Articulation

A direct analysis of the semantic articulation of a text leads to the important distinction of **theme/rheme**, **topic/comment** or **given/new**. There is no real agreement on a precise definition of these terms and each has a slightly different meaning. For this presentation, I will restrict myself to the notion of **theme/rheme** as understood by the Prague School [Firbas 66, Sgall 74, Sgall, Hajicova & Benesova 73].

According to the Prague School theory, the **theme** of a sentence is its "contextually bound part;" that is to say, the set of items already made explicit by the context of the sentence. The theory states that the **theme** of the sentence is always distinguished from the "contextually non-bound part," that is, its **rheme** [Sgall 74]. Among the semantic cases of a sentence, the **theme** is the one that can be considered as answering a potential preceding question. For example, the sentence:

(1-a) Jack was here yesterday

can be considered as an answer to the question:

(1-b) When was Jack here?

in this case, the **theme** of the sentence is "yesterday"<sup>3</sup>.

The problem associated with this definition is that it does not provide formal means to recognize the **theme** of a sentence. If intonation were available, it would give a precise indication on the location of the **theme** (almost always associated with the stress). But in written text one has to rely on other kinds of information. The order of the words in the sentence is at the basis of the **theme/rheme** strategy as defined by the Prague school. They state that sentences generally start with the **theme** (known or recoverable from the context), and then introduce the **rheme**, that is, information not derivable from previous discourse or not assumed to be shared by the audience. Thus for example, in:

(2) Jack was in Paris yesterday

the **rheme** is more probably "yesterday," whereas in:

(3) Yesterday, Jack was in Paris

the **rheme** might rather be "in Paris."

In an active sentence the order of constituents is generally "subject verb complement." One way to signal the

---

<sup>2</sup>roughly, there is a collocation between two words if their meaning "have something in common," as between "sea" and "boat," "sailor" and "fish" etc...

<sup>3</sup>"Here" can also be an answer to "Where was Jack?" this points out that a sentence in isolation does not have a determined theme.

theme/rheme articulation in the surface form is to change this unmarked order by using a syntactic construction like passivation, there-insertion or it-extraposition. This marked order allows one to put the theme towards the beginning of the sentence [Halliday 76].

Research is currently under way to devise formal means to identify the theme in a sentence. An important factor is that the theme is related to the context, and is not totally determined by syntactic form. For instance, example (3) can also be considered as an answer to:

(4) Who was here yesterday?

and then the theme of the sentence would be "Jack." Contextual information can even override the syntactic clues: [Ducrot & Todorov 72] give an example where the indication given by the it-extraposition rule is contradicted by the context. In "it is Jack who came first" the theme indicated by the it-extraposition rule is "Jack." But in a context of sentences describing Jack's activities, the preceding question is more likely "what did Jack do?" than "who came?" (for example in "It is Jack who came first, it is Jack who spoke, and it is he again who concluded.")

The opposition theme/rheme only appears in the context of a segment of several sentences. Thus the linguistic context is necessary to determine the theme/rheme of its parts; furthermore, rules can be devised indicating how the activated items of a segment (the rheme) may be connected to other items to build the new theme of subsequent units. The themes and rhemes of the parts (propositions) are combined to predict and determine the theme of larger segments<sup>4</sup>.

But behind this difficulty of formalization, it remains clear that the theme/rheme distinction has an effect on the linguistic surface structure of sentences considered alone and on their combination into texts.

### 3.2.3. Focus of Attention

In the view presented so far, the structure theme/rheme is a property of the sentence. It is a purely linguistic phenomenon. Work in AI has been using more generally the notion of focus of attention to study a set of phenomena roughly equivalent to the one accounted for by the notion of theme. The notion of focus is derived from the field of psychology; it refers to a conscient activity of a speaker or listener engaged in a conversation. As its name indicates, the focus is the information on which a speaker or a listener is currently centering his/her attention, the set of semantic items somehow "activated" in the consciousness or memory of a locutor. When speaking, locutors will generally present the information under focus as the theme of the sentences they produce.

Several computational models have been proposed based on the notion of focus. Earlier models were concerned with interpretation of texts. In this perspective, Grosz's distinction between immediate and global focus is particularly important [Grosz 77]. Sidner built on this distinction, and introduced the notion of thematic development between immediate foci in a computational model [Sidner 79]; she devised algorithms to determine when to maintain or to shift immediate focus. McKeown later used these principles in the framework of her text-generation research [McKeown 85], using the notion of focus development in association with her rhetorical

---

<sup>4</sup>This aspect of semantic composition is studied *per se* by the structural semantics of [Greimas 66], and leads to the notion of isotopy. Here the clues to predict the isotopy of a text are more centered on the words used (called lexemes) and their semantic decomposition into semes. This approach does not use any kind of syntactic information.

schemas (cf section 3.1).

### 3.2.4. Grosz's model of Focus Space

Grosz used the notion of focus of attention to help in dealing with the interpretation of definite descriptions. She founded her theory on an extensive analysis of transcripts of conversations between two people, an expert and a novice, cooperating to complete a task (assembling a toy water pump). This analysis has demonstrated the importance of correctly interpreting definite descriptions as in the following example:

- (1) Do you see the two screws?
- (2) Which screws, I see 4 of them.
- (3) The ones on the left of the wheel.

The problem of determining what is referred to by definite noun phrases like "two screws" is difficult even for the locutors involved in the discussion as demonstrated by this kind of clarification dialogue. Grosz hypothesized that people are actually focusing on a set of entities when trying to achieve a given task and that the way they use language reflects this focusing. This led to the definition of the **global focus** as the set of entities currently under focus and thus susceptible of being referred to by a definite NP. The particular entity being spoken about in the current utterance is called the **immediate focus**. Focusing can be described in this model as the operation of narrowing a set of entities to one of its subsets, such as moving from one focus space to another, more restricted one. Furthermore, certain linguistic clues are regularly used to signal a shift of focus, indicating the linguistic relevance of such a concept. Using that model of focus, Grosz discovered that in the corpus of conversations she studied the structure of the dialogues follows the structure of the task people are currently working on; that is, the order of appearance of foci in the linguistic surface can be predicted by an analysis of the task, or the structure of entities being discussed. For example, when the topic of the current discussion is to mount a wheel on a car, the focusing mechanism will highlight the operation of choosing an adequate screw, and putting it on the wheel, etc. which are the subtasks necessary to achieve the task under current focus.

Grosz implemented a computer system incorporating focusing mechanisms to help understanding task-oriented dialogues. She chose to represent the task domain as a semantic network, and the focus as a partition of that network [Hendrix 75 on network partitioning]. Focusing on a node of the network actually brings its child nodes under implicit focus. This mechanism elegantly models the phenomenon studied. In this work, however, the link between the focusing mechanism and the linguistic surface has not been extensively studied. The emphasis has been rather put on the **global organization** of the focus over a whole dialogue.

The main result of this approach, besides the distinction between global and immediate focus, has been to state that there exists a mapping between the semantic situation and the linguistic context (in other words, that the structure of the dialogue follows the structure of the domain) in the task domain. The consequence of this assumption is to give a primordial importance to the semantic situation, and to remove any specificity to the linguistic context. This assumption has also been held by [Litman & Allen 84] in another domain.

### 3.2.5. Sidner's Analysis of Immediate Focus

Developing the focus model introduced by Grosz, Sidner worked on immediate focus. In her research she investigated the link between focusing and surface linguistic phenomena - in particular the use of pronouns [Sidner 79]. Her theory is that tracking the evolution of immediate focus over a sequence of propositions can help disambiguate the use of an anaphorical expression: past immediate foci are plausible candidates for the role of referent, and therefore, keeping a trace of the focus evolution can reduce the search for a valid antecedent. For example, in the following sequence of utterances, the focus shifts from the person being described to her eyes and back:

- (1) Jane and Janet are beautiful
- (2) They have such beautiful eyes
- (3) They are black and deep
- (4) They seem to always be asleep

Keeping track of the structure of this evolution enables correct interpretation of the pronouns involved: the referent of "they" in (4) is checked against the "eyes" first, which are currently in focus, and then against "Jane and Janet," which was previously under focus.

Sidner specified algorithms to track immediate focus over several sentences. The algorithms determined whether the immediate focus of a sentence shifts to a new focus, or is maintained. In the proposed model, the linguistic context (for a sentence) is represented by the three following data-structures:

1. a list of potential foci
2. the current focus
3. a stack of past foci

Given these three data structures, the algorithm decides whether the focus remains the same, shifts back to a previous focus, or shifts to a new focus (from the list of potential foci). This algorithm intertwines tracking of focus and interpretation of focus.

The model proposed is descriptive, in that it does not help predicting a shift of focus, nor explaining why a speaker actually chose one option rather than another (keep the same focus rather than shift for example but see [Grosz, Joshi & Weinstein 83] on centering). Such an explanatory framework would be necessary to extend the model, or to merge the information given on that particular function with other linguistic features of discourse, such as the use of personal pronouns, titles and different form of address. But as it is, the model provides valuable linguistic information, grounded on a plausible psychological theory and modeled by a solid computational method.

### 3.2.6. McKeown's Study of Focusing for Generation

McKeown used the notion of immediate focus in her program TEXT (see section 3.1 for a general presentation). The use of focus in the perspective of text generation introduced interesting new methods. Here, immediate focus is used as a local constraint helping to decide what to say next in the context of a rhetorical schema. It is mainly conceived as a constraint on the transition from one proposition to the following. The algorithm used in the TEXT system is such that when the constraints generated by the current rhetorical schema leave the possibility for more than one proposition to be activated and included in the generated text, focus constraints are checked to help narrow the set of possible successors. McKeown proposed a scale, ordering the different options offered to a

speaker at any given point : (1) shift focus to a topic mentioned in previous discourse as a potential focus, (2) maintain focus, (3) return to topic of previous discourse, (4) select the proposition with the greatest number of implicit links with the current potential focus list. Thus, focus plays the role of a local constraint on the discourse, which is used to help order propositions within a text.

#### The Use of Focus Constraints on the Order of Propositions

- |   |                           |
|---|---------------------------|
| (1) John is working at ZXC.                           | Focus=John; Potential=ZXC |
| (2) ZXC is a large California-based software company. | Focus=ZXC                 |
| (3) John is also a very good cook.                    | Focus=John                |
|   |                           |
| (1) John is working at ZXC.                           | Focus=John; Potential=ZXC |
| (3) John is also a very good cook.                    | Focus=John                |
| (2) ZXC is a large California-based software company. | Focus=ZXC                 |

In this example, the first order 1-2-3 is much more coherent than the second 1-3-2 : It is better to shift to potential focus ZXC as soon as it is mentioned than to maintain focus. (John-ZXC-John better than John-John-ZXC).

Figure 3-3: Focus Used as a Constraint on Proposition Ordering in the TEXT System

Figure 3-3 shows how rules (1) and (3) allow the system to choose between the two orderings of propositions. Shifting to potential focus as in (1-2 John-ZXC) and then back to previous focus (2-3 ZXC-John) gives a more coherent paragraph than maintaining focus (1-3) and then introducing a new focus (3-2).

Furthermore, McKeown proposed to use this focus information to determine what syntactic form to use. In her model, propositions are internally represented under a form similar to: predicate[arg1,...,argn], where each argument corresponds to a semantic case of the predicate. A unification grammar [Kay 79] translates this deep structure into a sentence. During this process, the focus information is used to choose, for example, between the passive and active forms: if the focus is on the AGENT case, then the active form is selected; if the focus is on the OBJECT case, then the passive is used. In both situations, the case under focus will appear as the surface subject of the generated sentence. Other syntactic forms are similarly selected using the focus information.

[Danlos 84], arguing that no general rule of this kind can be valid, criticized this approach on two grounds: the use of focus to order the propositions in the text is contradicted when there exist general semantic links (e.g., temporal or causal) between the propositions. She gives the following example:



The 3 propositions to be ordered are:

- (1) Max was married to Marie
- (2) Max cheated on Marie
- (3) Marie committed suicide

Both orders (2)-(3) and (3)-(2) are acceptable in isolation:  
 Max cheated on Marie, she committed suicide.  
 Marie committed suicide, Max cheated on her.

But after (1), only (2)-(3) is acceptable:

- (4) Max was married to Marie.  
 She committed suicide. He cheated on her.
- (5) Max was married to Marie.  
 He cheated on her, she committed suicide.

The second critique was about the use of focus to decide on syntactic form. For example, certain verbs cannot be constructed at the passive voice. Furthermore, the relation between the semantic cases and the surface syntactic roles is not always so direct: for verbs denoting a state, as in "I miss Max" the subject cannot be assimilated to the semantic case of AGENT.

Both these objections indicate the incredible complexity of natural language, and overall the difficulty of isolating one linguistic function from the others. Although this shows that focusing mechanisms are not as general as one might wish, this does not demonstrate their invalidity. In the TEXT system, the rules of focus are only one part of the global process of question answering. It is assumed that the rhetorical schemas and syntactic information could override predictions based on focus analysis alone. The most productive attitude towards language seems to look for the diversity of functions that it can assume, and to try to articulate all of them in a coherent framework. The large body of knowledge and methods already based on the notion of focus establishes that something similar to focusing is prevalent in language use. The problem is to study how this mechanism interacts with other resources available to locutors.

### 3.3. Summary

The main points made in this section are:

- It is useful to distinguish between local and global aspects of text coherence.
- The main tools used in AI to account for global linguistic coherence are schemas. The notion of global focus also accounts for the global coherence of discourse.
- Local cohesion can be studied as a linguistic or as a psychological phenomenon. As a linguistic phenomenon, two notions are useful: (1) cohesive devices and (2) the theme/rheme articulation. The notion of linguistic cohesive devices is not widely used in AI. The notion of focus is the psychological correlate of the theme/rheme articulation. It is widely used in AI both for understanding and generation.

We have also noted some problems associated with current AI approaches to text coherence:

- General rules using focus to describe or plan the structure of discourse, to order propositions in generating a paragraph, or to choose a syntactic form, cannot work in isolation. Information derived from focus mechanisms needs to be often overridden by other aspects of discourse (syntactic constraints, general semantic links between propositions, or pragmatic situation).
- Rhetorical schemas are only "compiled textual routines." There is as now no formal way to recognize a schema. The status of rhetorical predicates (the components of the schemas) is therefore unclear (what

constitute the specificity of a predicate, what effect does it have on the linguistic surface form, how many predicates exist, how would it be possible to define a new predicate, what are the relations between predicates...).

- One way to clarify the status of both predicates and schemas would be to relate them to their interpersonal function in the framework of a communicative analysis. There is no “linguistic necessity” explaining discourse coherence, but constraints to insure communication efficiency.

This last point is the most important conclusion. It does not imply that a theory of textual coherence must also be a theory of interpersonal relations, but rather that the links to such a theory have to be stated explicitly. Attempts to explain discourse structure by factors other than pragmatic (for example the structure of the task in the task domain) are bound to remain *ad-hoc* and limited.

#### 4. Links with the Extra-linguistic Situation

This section surveys of techniques used in AI to investigate the influence of the extra-linguistic situation on language. The focus here is on the “semantic situation.” The question at hand is “what has to be known about the world in order to carry out a conversation or understand a discourse ?” The problem is to determine how this information is to be structured so that it can be used effectively in the process of understanding or generating utterances. In conformity with the general perspective of this paper, we also investigate how the techniques and strategies developed to handle the semantic situation can be articulated with techniques dealing with other aspects of the situation (linguistic and pragmatic).

It is not our purpose in this section to survey the semantic theories often underlying the systems we mention. Rather, we want to illustrate, through examples, how the semantic situation is represented and used in typical AI programs. In the first part, we present, through the example of Charniak’s program Ms. Malaprop, two problems that AI programs dealing with semantic situation must face: choice of a knowledge representation scheme and control of inference. Many solutions have been proposed to these problems (using frames, semantic nets, scripts, logical representations etc.).

In the second part, I present one family of solutions that are particularly representative: the theory of memory organization packets (MOPs). This solution emphasizes the importance of memory in the treatment of semantic situation. We will try to demonstrate why this aspect makes MOPs illustrative of the solutions generally retained to deal with the problem in AI.

The third part is a critique of certain aspects that AI solutions generally share. It argues for an integration of the treatment of semantic situation with pragmatic considerations, within the framework of an interpersonal communicative analysis. Finally, the last part of this section is a presentation of the still young theory of situation semantics which offers some interesting perspectives for the problem of treatment of semantic situation.

#### 4.1. Ms. Malaprop: the Need for Inference

Early programs in AI demonstrated that the interpretation of discourse involved large amounts of world knowledge. Charniak proposed methods to handle this aspect of language comprehension with a program called "Ms. Malaprop" [Charniak 77]. Ms. Malaprop was not supposed to understand English directly, the program was instead supposed to reason from an intermediate semantic representation of utterances. This simplification allowed demonstration of the importance of the knowledge representation used and of the inferential processes necessary. Ms. Malaprop demonstrates the existence of problems more than it solves them; therefore we will focus here more on the definition of these problems than on any particular technique which would help solving them.

Ms. Malaprop was a program intended to answer questions about stories dealing with painting. Its competence was based on an extensive representation of the processes involved in any painting episode: the purpose of painting an object, what paint is, how it sticks to objects, what a brush is, why brushes are generally absorbent etc. When reading a story stating that an actor got a newspaper, Ms. Malaprop inferred from its knowledge of painting episodes that the paper was going to be used to protect the floor from the paint, and not to be read. This example illustrates the theory investigated by the program: understanding a story consists in fitting incoming information with pre-existing frames describing a situation. Further, this process of "matching" new information with pre-existing knowledge requires the making of inferences: to match the information that an actor got some paper with the already known fact that people generally protect the surroundings before painting, one has to infer that paper can protect the floor, that before protecting the floor with something, one has to get this protection from somewhere, etc.

The implementation of this theory is bound to face two serious problems: determining a good scheme to represent pre-existing and incoming information in a general, preferably domain-independent way, and how to control the process of inferencing.

Charniak addressed the first problem by using a large base of "frames"<sup>5</sup>. In his representation, complex events like painting a chair were described as a succession of simpler events such as applying a brush on a surface, involving diverse "actors," "instruments" etc. When reading the story, the system "activates" certain frames that are mentioned in the text: these frames represent the **context of the story** (that is, the semantic situation). One very interesting aspect of this process is that it views understanding as a process of recognizing an *a priori* known situation rather than as trying to build a representation of the situation from the story alone. The fact that the input is already represented in the form of a frame hides a real difficulty of this approach: how to relate a frame to the language. Charniak gives no indication that the gap between linguistic surface ("words") and internal semantic representation (frames) can ever be filled. But the idea that the incoming discourse does not "contain" all the information needed to understand it is of great importance.

The second problem, that of controlling inferences, remains a major one: from given information and rules of inference, there are a tremendous number of possible paths of inference. The problem is that there are no criteria proposed in the theory stating when a conclusion is "interesting" and when it is not. Moreover, all the inferences do

---

<sup>5</sup>A frame is best described as a structured chunk of information. In the terminology introduced by [Minsky 75], a frame consists of several "slots," each slot has a name and a value, and all slots describe a particular aspect of the data represented by the frame as a whole. For example, a frame representing a person would have a slot describing their name, the color of their eyes etc.

not produce correct conclusions. In the example discussed above, it is possible that the painter got a paper just to read it. Actually, this problem is at the heart of the link between language and situation: to understand why the sentence "Jo got some paper" is included in a story about painting, one has to know why certain aspects of a situation are explicitly described in a text, and why certain aspects remain unstated. This problem will be further discussed in the next section.

In summary, the program described by Charniak demonstrates the need for inferential processes within a theory of text comprehension, and proposes a theory of frame-based knowledge as a representation of the semantic situation but does not propose a solution to the problem of control of inference.

## 4.2. From Scripts and Plans to MOPS: the Importance of Memory

### 4.2.1. Scripts

Among the proposals for plausible models for organizing knowledge about situation, *Scripts, Plans, Goals and Understanding* ([Schank and Abelson 77]) has been particularly influential. The main claim defended in this book was that "human text comprehension involves the use of high-level knowledge structures which serve to generate expectations about what is coming next" [Abelson 82, p.viii]. One of the structures they hypothesized was the script. A script is a "prepackaged chain of causal relations between events and states" [Schank and Riesbeck 81, p.6]. The classic example illustrating the notion is that of the "restaurant script:" when entering a restaurant, people correctly interpret what is going on in the establishment, role of the waiter is, the cashier, the patrons, how to get food served, when to pay, etc. The theory explains this kind of anticipation by stating the existence in memory of scripts describing the typical course of events associated with the notion of restaurant. Scripts actually provide the knowledge necessary to find the coherence of stories, even if the story does not contain descriptions of all the events.

The first direct implementation of the theory was the program SAM (Script Applier Mechanism) developed at Yale by Cullingford ([Cullingford 78]). SAM was supposed to read and understand simple stories, and to answer questions about them in a second stage. In a way, SAM addresses the same problems as Ms. Malaprop: representing knowledge about a situation and providing knowledge necessary to perform the right inferences. But the psychological theory behind the program allows a more systematic approach to these issues. In SAM, three types of inferences were discovered to be necessary: causal chain completion, role instantiation and role merging. Causal chain completion denotes the process of understanding that intermediate events occurred when learning that two isolated states in a script are reached (like understanding that when "John left the restaurant" he first ate his meal and paid for it). Role instantiation is the ability to affect a scriptal role to a person mentioned in the text (like deducing that "John" is the waiter of the restaurant), and role merging that of recognizing that multiple references refer to the same role filler. These three types of inferences are explained by the theory, and their effective use demonstrated that something similar to script knowledge may actually be used by humans when processing stories.

But clear limitations of a purely "scriptal" approach to understanding were also demonstrated: SAM as a text understander is not "robust," in that it is not able to deal with unpredicted events. If an event occurs in the story that is not represented in the current script, the program is bound to fail. The SAM approach - relying exclusively on

memory to understand stories - appears rigid and domain-dependent. This aspect demonstrated the need for a more powerful management of memory in two ways: systems need to be designed with a huge memory, or must be able to learn and update their memory; second, "general" principles have to be designed to avoid the domain-dependence. One approach concerning this second point has been to study "planning" as a general principle allowing intelligent systems to deal with non-stereotypical situations [Wilensky 78]. Questions remain whether planning alone may be the only general principle needed to add flexibility to a system. But the main conclusion that can be learned from the SAM experiment is that scriptal knowledge alone is useful but not sufficient to account for understanding, and therefore has to be complemented by other sources of knowledge.

#### 4.2.2. MOPs and Dynamic Memory

The development of SAM showed the need for organizing multiple scripts in memory so that they can be found and used when needed. The problem is deciding when a particular script is relevant to interpret what is happening in a text. The difficulty of searching memory and detecting the relevance of a particular script led Schank to develop a theory of memory, called "dynamic memory" [Schank 82]. This theory is of psychological nature. It aims to explain how high-level structures such as scripts are organized in human memory, in hierarchies representing situations of varying generality. To this end, it posits the existence of memory structures that Schank calls Memory Organization Packets (MOPs). MOPs are developed to store "information about how memory structures are ordinarily linked in frequently occurring combinations" [Schank 82, p 83]. The theory further describes structures called TOPs (for Thematic Organization Packets) which account for the analogical ability of human memory. TOPs store information about the indexation of lower-level structures in chains representing "related" events or scenes.

The theory of dynamic memory is highly attractive in regard to the two problems mentioned about Ms. Malaprop. MOPs handle knowledge representation with a particular stress on generalization. The theory provides two ways to control inference: inferences of a certain type only can be drawn and memory structures can be used to control inference. IPP [Lebowitz 80] was one of the first implementations of the theory. IPP is a computer program designed to read and generalize from a large number of news stories. It has been used in the domain of international terrorism. In IPP, bottom-up inferences of a limited type are made during the parsing of the news story. For example, from the presence of certain types of words called Event-builders (e.g. "killed"), it is possible to activate an "Action Unit" (a memory structure representing an event, e.g. SSHOOT). Other bottom-up inferences are used to deal with NPs (e.g. "terrorist") and fill roles of the memory units activated. These inferences are controlled by higher-level memory structures (MOPs) which generate expectations in a top-down manner. MOPs (e.g. S-DESTRUCTIVE-ATTACK) once instantiated, predict which AUs can occur in the story, and explain their role in a larger framework, as results, causes or goals. The organization of memory and the mixing of bottom-up and top-down inferences allow a more flexible use of these limited inference methods than in previous systems, but the same overall problem of rigidity remains: how to be able to parse unpredicted events or events not fitting in memory. The approach of limiting the kind of possible inferences does not completely solve the problem of controlling inference.

### 4.2.3. Why MOPs are Representative: the Principle of Distinction

Most work on structured knowledge representation techniques (frames, scripts, semantic networks, cf [Levesque & Brachman 83] for an overview) shares with MOPs the use in one form or another the property of inheritance. Philosophers (e.g. [Prieto 75]) have referred to this property under the name of distinction: objects are only significant to the extent that they differ from a related prototype. This property explains that in memory, for example, different situations can be represented in a generalization hierarchy: a situation appears in the hierarchy only if it contains specific features that are not part of a higher level situation. Further, this situation is always meaningful when contrasted to another situation: when comparing a “restaurant situation” to a “school situation,” just saying “school” or “restaurant” is enough. But to contrast a “Burger King” from a “McDonald” situation, more detail is required.

In this perspective, work on knowledge representation is very similar to the classic structuralist approach to linguistics which states that objects are not known by themselves, but only within a structure of similar and different objects [De Saussure 49, Prieto 75]. One extremely important consequence of this principle of distinction is that understanding cannot be considered as translating a message from one code to another, but necessitates the contrasting of two entities. Distinction presupposes a duality. Work on memory leading to methods as memory-based parsing [Riesbeck & Martin 85] can then be described as an application of the distinction principle between text and memory: understanding does not mean translating an input text into an output semantic form, but rather locating input information in an existing network of information, or memory. This theoretical decision is far from trivial: it implies that the link we were looking for between language and situation is replaced by a link between language and memory - that is between language and another semiotic system representing the situation. In the universe of “closed” symbolic computation that might describe current AI research, that mediation between situation and language through a semiotic system is necessary.

MOPs are also representative of two other aspects of knowledge representation schemes in general: they highlight the incompleteness of knowledge bases both in content and structure, and the difficulty of capturing the implicit aspects of language. Trying to increase the size or structural complexity of knowledge bases is not, however, the only way to overcome these limitations. We argue, in the next subsection, that an integration of the treatment of semantic situation with pragmatic situation might bring some insight to this problem.

## 4.3. Memory and Communication: a Critique of Current Knowledge Representation Techniques Applied to NLP

This subsection is a general critique of the use of current knowledge representation (KR) techniques applied to NLP. It argues that in developing current KR techniques, the communicative function of Natural Language has been often overlooked, and that certain difficulties met with approaches dealing with the semantic situation could be reduced, if the interpersonal dimension of language is granted an adequate treatment. Here again I begin with an analysis of the examples of MOPs, and try to generalize to knowledge representation techniques.

#### 4.3.1. MOPs and the Intra-personal Study of Language

One can note about research using MOPs that it is based on a purely intra-personal study of language: the question addressed by the research reviewed in the previous section can be stated as: "what faculties - knowledge, memory, reasoning - must an individual have to understand language as a representation of the world." Although this exclusive perspective is legitimate to study memory, it hides certain aspects of the language/situation link in which we are interested. Indeed, a purely intra-personal theory cannot account directly for the fact that communication is possible through the use of language.

If one tries to apply a theory using the property of distinction in a situation of communication, one must first decide what will be the two terms of the contrasting operation presupposed. One term must be what a locutor wants to express. Deciding that the second term is the memory of the speaker alone - as is described by Schank's theory for example - can lead to serious difficulties: memory is acquired and updated through personal experience. As a consequence, the structure of memory is highly idiosyncratic. But an utterance is directed towards a certain audience which may or may not share the same memory. The need to be understood, which is inherent in the act of speaking, actually introduces a "third" element in our schema of distinction: what is known by the audience and what is not. As a consequence, one may try to say that the difference between what is said and not said relies on what is shared and not shared between the speaker and his/her audience. This is the first intrusion of an interpersonal theory, one which deals with the problem of assessing what is the supposed audience, the knowledge that is shared, and that is not, and how isolated individuals perceive these categories. I argue next that these factors also affect the way language is understood.

#### 4.3.2. Distinction and the Interpersonal Dimension of Language

From a perspective that includes the interpersonal, the property of distinction alone is not sufficient to account for what is said and what remains implicit. Conversations are usually highly redundant. Conversations on the weather, or even descriptions of physical objects are not directly "efficient" in terms of information theory. On the other hand, certain utterances are too brief and concise to convey effectively what they are supposed to mean. Linguistic taboos exist which prevent the use of certain words, even if they could convey most effectively a given meaning. In certain conditions whole topics are proscribed from conversation. For example, [Tannen 84] gives numerous examples of circumlocutions people use in different situations to refer to the act of making love. Both cases illustrate that the principle of distinction, if used alone to decide what to say, provides either too much or not enough material compared to real occurring utterances. Thus, the role the audience plays, that we have identified as a "third element" complementing the text and the memory within the language production process, actually modifies the whole form and substance of utterances: as we have seen in the previous section, the pragmatic situation can influence the structure of discourse as well as its surface manifestations.

As a consequence, the *conditions of production* of a discourse are at least part of its meaning. If we try to close the loop, and consider that all discourse has been produced with a certain view of its audience, then this process of tailoring must have an effect on the understanding of the discourse. Here is the second intrusion of an interpersonal theory within our model of text-understanding: the incoming text given as input to a system has already been mediated through a set of social rules, constraining what has been said and how it has been said. This kind of "bias" must be taken in account to describe the result of the understanding process. [Schank 82, p143] mentions that when

forming his/her memory, a person has to come up from physical to societal generalizations. He gives the following example: a terrorist attack can be considered as a violent attack (this is its "physical" characterization), or as an instance of a negotiated contract (this is one of its "societal" aspects).

The important point is to capture what people directly perceive: Schank's theory seems to be that people first perceive the physical properties of the situation, and then "mediate" them through their memory to activate a physical memory structure, and that in a second stage, societal generalizations enable a more powerful localization and indexation within the memory.

#### 4.3.3. Links between Pragmatic and Semantic Situation

If we now try to apply this theory to a computer program, given written texts, it is important to note that the input is produced with societal generalizations already activated. The input is already mediated by a social system. In the example of the domain studied by IPP, this social system prescribes how to present a terrorist attack and how to speak about it, and proscribes other perspectives on the event. Without the level of awareness provided by an analysis of the production conditions of an incoming text, a program will be somehow brittle. Unable to criticize its input, it will "ingest" it and update its memory in the same way, whether the text comes from the victim of the attack or from its perpetrator.

A text is thus a representation of a situation mediated through two different semiotic systems, memory and the social system surrounding the author and reader of the text. Both systems are intimately intermingled, but several facts indicate the relevance of the social level of mediation on linguistic data. The whole system of modals in English (use of would, must, might, could...), for example, is explained by [Halliday 78] in terms of social rules. The system of address in languages that include one (e.g. French, Japanese) is similarly driven by social considerations. Lexical choice is also heavily conditioned by phenomena of ephemeral trends or social pressure or by the context of a fast evolving micro-culture (the use of the word "paradigm" within the AI community is a good example of such effects).

In summary, the study of the semantic situation has shown that the real situation itself is not accessible directly to a symbolic system, but has to be mediated by another semiotic system. In this framework, memory may appear to be a good way of doing this. Understanding text could be viewed as a process of putting relevant features of the represented situation in their appropriate place in memory, using methods derived for most of them from the principle of distinction. However, trying to extend this theory to account for a complete process of communication indicates that the medium of memory alone may not be sufficient, and a theory of interpersonal relations has to be devised to account for multiple linguistic phenomena.

#### 4.4. Situation Semantics

Among all the theories reviewed in this paper, Situation Semantics is the one occupying the most central place to our topic: it is being developed with the explicit concern of explaining the link between "situations and attitudes" and language. The theory is an ambitious attempt at defining "meaning" and was initiated in 1981 by Barwise and Perry (e.g., [Barwise & Perry 83], [Linguistics and Philosophy 85]). Developing the idea that meaning is not completely conveyed by the expression, but also resides in the situations surrounding the use of language, Barwise



and Perry propose a semantic theory focusing on the external significance of language. Barwise and Perry have developed their theory with a constant interest in complexity and computational issues, it is therefore particularly relevant to AI, even if no implementation is yet available.

Situation Semantics is a complex theory, addressing most of the hard problems encountered by more classical ‘‘logic’’ semantics. We will limit our presentation here to only some general considerations.

Situation Semantics considers transmission of information as the main function of language. In this view, semantics is a theory used to classify events. Meaning is defined in this theory as *a relation between situations*. This ‘‘relation theory’’ of meaning is at the basis of a powerful mathematical formalism. The assumptions behind this theory are that people are attuned to similarities between situations, that they can recognize relationships between these uniformities, and that they can classify situations according to them.

Classification according to similarities is a highly complex cognitive task. It is indeed very similar to the principle of ‘‘distinction’’ mentioned in the previous paragraph. Similarity and contrast are two dual notions: developing the concept of similarity leads to the inclusion of relations as objects of the theory, whereas developing the concept of contrast leads to the definition of the notion of ‘‘opposition.’’

Connections between situations can be of various kinds. A few examples demonstrate the generality of the ‘‘relation theory’’ of meaning:

1. *smoke means fire*: here the relation is ‘‘natural;’’ all situations involving smoke also involve a fire.
2. *kissing means touching*: the relation is here more fine-grained and between two actions culturally defined.
3. *The ringing bell means class is over*: the relation between the two events is here purely conventional. But the cognitive activity necessary to recognize such a relation is exactly the same as is necessary to recognize natural relations. This cognitive activity is the only common point between the three examples given.

It should now be clear that the ‘‘relation theory’’ of meaning is far from being restricted to linguistic meaning. This means that Situation Semantics is part of a more general ‘‘semiotic’’ theory, but actually it also means that communication through language is nothing special. There is no distinction of a special level of meaning attached to language by the virtue of linguistic particularities. Instead, Situation Semantics presents the meaning of a simple declarative sentence, for example, as a relation between the situation of utterance and the situation described by the sentence. The interpretation of a statement is then viewed as the set of all possible situations compatible with the description given (that is, the interpretation is the extension of the relation). Barwise and Perry insist on this point that they call ‘‘the priority of information:’’ information exists independently of the language that can convey it. Therefore, there must exist a way of representing the world independent of language (this is an assumption generally accepted by all AI practitioners).

Barwise and Perry describe a mathematical formalism to represent and describe situations. The mathematical object representing a situation is called an abstract situation in the theory. It also represents the cognitive ability to recognize certain similarities. Abstract Situations are built from space-time locations and situation-types. A

situation-type describes individuals having properties and standing in relation to each other at various spatio-temporal locations. The particular individuals, properties and relations can all be recognized across different situation. Thus, for example, let us consider the two situations described in the following statements:

Molly, the dog, barked yesterday in the garden

Molly, the dog, barks in the kitchen

The following formula can describe the relation between the two situations:

in  $s$ : barks, Molly; yes.

This formula defines a situation-type,  $s$ . It is constituted by a list of the specific predicates characterizing this situation-type. Predicates are represented in the form <predicate>, <args>. They are followed by an indicator which can be either "yes" or "no." "Yes" indicates that the event represented by the predicate must be recognized in a situation for it to be an instance of  $s$  whereas "no" means that the predicate must be recognized as false. Thus, the situation-type  $s$  describes all the situations in which Molly the dog barks. Furthermore, this description is **partial**: it is silent on all the events not described in the formula. We can't know given the description  $s$ , if Molly is breathing or eating or anything else or where and when she was barking.

In this purely relational framework, objects, attributes and relations between them are just tools to classify situations. An object is at most a uniformity across multiple situations. An event is a situation occurring at a given space-time location  $l$ . **Event-types** can furthermore be defined as classes of events all standing in relation to each other. Finally, systematic relations between event-types create a flow of information. In order to extract this information, organisms have to be attuned to these systematic relations. Therefore, the semantic theory also provides a representation of these constraints between event-types, which are what ultimately creates "meaning."

This general theory is applied to linguistic meaning by considering a basic "linguistic situation." A situation describing an individual telling something to someone else:

Discourse Situation: at the location  $l$ ,  
                           speaking  $a$ ; yes  
                           addressing  $a$ ,  $b$ ; yes  
                           saying  $a$ ,  $\alpha$ ; yes;

This situation defines four basic roles involved in all linguistic activities: the speaker ( $a$ ), the addressee ( $b$ ), the expression ( $\alpha$ ), and the discourse space-time location ( $l$ ).

The meaning of the expression  $\alpha$  for declarative sentences is defined as the systematic relation existing between the discourse situation  $DU(\alpha)$  and the described situation  $e$ . This meaning is noted:  $DU[\alpha/e]$  ( $[\alpha]$  is the name of the relation and the situations  $DU$  and  $e$  stand in relation). Building on these premises, Barwise and Perry study several classical problems facing semantic theories, and propose, for example, an attractive treatment of reference, harmoniously unifying classical theories.

Situation Semantics, though highly original, remains very close to a purely "logical" treatment of semantics (especially possible world semantics). Although the authors mention that they "dream" to develop their theory to a point "beyond" the speech-act/truth theoretic approaches to semantics, Situation Semantics in its current shape is

definitely on the truth-theoretic side of the debate.

It is true however that this aspect is not due to the theory itself. Situation Semantics seems to only provide a very powerful tool to reason about meaning in the terms of the relational theory. Since it is so young and general, it does not demonstrate how to practically use this tool. It seems to have potential, but has not yet been fully explored. The logic language CIL proposed in [Mukai & Kurosawa 85] is an interesting approach in this direction: it allows use of Situation Semantics types of objects (complex indeterminates, situation types) in a Prolog-like environment.

As for the main concern of this paper, Situation Semantics provides at this date few insights on how the interpersonal dimension of language might be handled. But, the basis of the linguistic developments of the theory - the definition of the discourse situation - explicitly defines a speaker and an addressee. This means that meaning in Situation Semantics has to be built using these "pragmatic" categories. This also means that the proposed framework includes the essential tools to represent the conditions of production of linguistic expression. Describing pragmatic constraints, of course, remains to be done, and might have implications on the developments of the theory; but Situation Semantics is a theory that at least "leaves the door open" for this direction (see Winograd's critique [Winograd 85] and Barwise and Perry's answer in [Linguistics and Philosophy 85] on this point).

#### 4.5. Summary

In this section, we have been interested in the link between the external world and language use. The main points mentioned are:

- In order to understand or to produce a discourse, one needs to make inferences about the situation described.
- NLP programs are based on a representation of an adequate portion of the external world. Structured knowledge representation techniques are most commonly used for that purpose.
- The theory of Dynamic Memory is representative of several features of AI techniques: it uses the principle of distinction (inheritance) and highlights the incompleteness of memory both in content and in structure. It also puts a particular emphasis on the role of memory in language use.
- Situation Semantics develops the "relation theory of meaning," which states that meaning is a relation between situations. This theory avoids problems associated with earlier semantic theories since it requires only minimal philosophical claims on the correspondence between objects and relations in the world and their representations in language.

These points are related to the problems AI faces when dealing with the external world:

- Control of inference remains a very hard problem. Too many possible paths of inference exist for any piece of information. There is no general technique to decide which ones are worth being pursued and which ones should be abandoned.
- Dealing with incomplete knowledge is still a problem.

We also mentioned that the focus of AI researchers on knowledge representation as a metaphor for memory has caused them to overlook important pragmatic aspects of language. The relation between the world and language use is indeed doubly indirect. Language use necessitates a mediation through memory and through a set of social conventions. Paying proper attention to these social conventions might help in dealing with problems NLP currently

faces, such as determining what needs to be said in a text, and why a text contains certain information and not other information.

## 5. Pragmatic Aspects of the Interaction: Influence of the Situation of Utterance on Language Use

The pragmatic situation could be considered the frame of the interpersonal aspect of communication. The underlying hypothesis justifying the study of the pragmatic situation is that *language is symptomatic of the causes which originated its use*. If language is not only the “expression of thought,” other causes have to be investigated. In this section we survey the techniques and concepts used in AI to deal with the pragmatic situation.

In the first part, we present the Speech Act (SA) theory, and research in AI derived from it. We present first Allen and Perrault’s theory, using a planning technique, and the developments it has undergone. In the second section, we present the theory of implicatures, and AI programs applying some of its aspects. We then try to induce from this survey what needs to be represented in a model of pragmatic situation, and point out limitations of current approaches.

### 5.1. The Planning Theory of Communication

#### 5.1.1. The Speech-Act Theory

The main pragmatic approach used within the framework of Artificial Intelligence is the theory of Speech Acts (SA) [Austin 62, Searle 69]. At the basis of this theory is Austin’s distinction of a special class of English verbs that he calls *performatives*. These verbs have the property that uttering them actually accomplishes the act they denote. For example, someone declaring to his friend (1) “*I promise you I’ll finish this work tomorrow*” is actually promising by saying this sentence.<sup>6</sup> The main conclusion reached through the study of these verbs is that an utterance can be used to accomplish certain kinds of acts.

Austin distinguishes three kinds of acts related to the act of utterance:

- The *locutionary act* is the act of uttering (the physiological activity).
- The *illocutionary act* is the act accomplished immediately by the uttering, like promising in (1).
- The *perlocutionary act* is an act which is performed as an indirect effect of the utterance. For example, saying “*this building is really well protected*” can have as effect to reassure a worrying person.

Searle extended these first insights by positing that any utterance has an illocutionary value - that the act of uttering always has an immediate effect similar to the one produced by performative verbs. Performing such an effect is a speech-act.

SA theory builds on this principle: speaking is a process of performing SAs. Further, Searle has shown that the

---

<sup>6</sup>Actually, deciding to which extent the cause of the performative effect is the verb used itself or the syntactic form (1) “*I <verb> to <action>*” is an important problem in the theory [Berrendoner 81, Benveniste 66].

kind of act accomplished by speaking is very special. To perform a SA, a locutor has to let "the hearer recognize his intention to produce it, and intend to get the hearer to recognize this intention by virtue of the hearer's knowledge of the rules that govern the utterance of the sentence" [Searle 75]. Thus the SA is a kind of institutional act since it has to be recognized as the realization of a cultural, conventional way of performing a special class of acts.<sup>7</sup>

One of the most important aspect of the SA theory is that it is constitutively interpersonal: it is based on the categories of audience (hearer), speaker's intention and effect on the audience. This feature distinguishes it from previous "semantic" theories, and places it in the framework of pragmatics.

### 5.1.2. SAs in AI: Allen and Perrault's Planning Theory

Allen and Perrault were among the first to develop an artificial intelligence program based on the SA theory [Allen & Perrault 80]. Their main contribution has been to integrate the notion of SA and a theory of planning. A theory of planning has been successfully used in Artificial Intelligence for a long time (e.g., STRIPS [Fikes and Nilsson 71]). The basic categories of the theory are the notion of state of the world, goal state and operators. A goal is a desired state of the world. States of the world are represented conceptually as sets of logical propositions. A rational agent manipulates a set of operators and is given the capacity to plan a sequence of operator applications that will lead the world from its initial state to the desired goal. An operator, finally, is the representation of an act; it is characterized by a set of preconditions which determine whether or not the operator can be applied in a given state of the world, and an effect, which is the encoding of the modification of the world caused by the application of the operator. Operators represent the legal transitions from state to state.

The idea of Allen and Perrault has been to represent SAs in this framework. The main assumption underlying this move is that in an adequate representation of the world, SAs can be represented as operators, and *can be used in plans to achieve goals (whether linguistic or not)*. It is important to note here that the theory relies on the description of the pragmatic situation as a set of propositions using the categories of knowledge, beliefs and wants. Once that representation is chosen (and accepted), it is possible to represent a SA as an operator whose effect is to modify the speaker's and hearer's beliefs, knowledge and wants and to build a theory of communication as a planning process of basic acts.

Allen and Perrault's program is aimed to be a model of "helpful behavior and of its linguistic effects." Their model of cooperation relies on the following process:

1. People, as rational agents, form and try to execute plans.
2. This rationality gives them the ability to infer others' plans.
3. When they detect an obstacle in other agents' plan, they try to help them.

Figure 5-1 gives an overview of the formalism Allen and Perrault used in their program. The predicates *WANT*, *BELIEVE* and *KNOW* are particularly important. SAs are a local concept, applicable to the level of a single utterance. Allen and Perrault's Planning Theory considers the SA as the basic units of communication and provides

---

<sup>7</sup>One of the problems of the theory consists in describing this class of acts doable as SAs. Several classifications exist, but there is no real agreement. The problem is further complicated by the existence of indirect SAs [Searle 75].

### Overview of the Allen & Perrault's Formalism

#### 1. Belief - Knowledge - Wants:

Agents are determined by a set of propositions describing their beliefs, knowledge and wants. The notation used is the following:

Agent A believes that P is true:  $AB(P)$

A believes that 'P is true and S believes that P is true:'  
 $AB(P \& SB(P))$

A knows that P is true:  $A\text{KNOW}P \iff P \& AB(P)$

A wants P to be true (A has a goal to achieve P):  $AW(P)$

#### 2. Plan Representation, Construction and Inference

A plan is a path of actions from an initial state of the world to a goal. Actions are parameterized procedures formed of Preconditions and Effects. Plan construction is viewed as a backward chaining from the goal to the initial state.

Planning = set of planning rules + control strategy

Plan inference: inferring the plan of an agent from original beliefs about others' goal AND general knowledge of how plans are (rationally) constructed.

Method: applying planning rules in reverse, reconstruct the plan from observed actions. Leads to a search through "partial inferred plans."

Plan Inference = inference rules + control strategy

#### 3. Speech Acts Representation

Speech Acts are represented as intentional acts. That is they have the following property:

$ACT(A) \implies AW(ACT(A))$

(A does ACT only if A wants to do ACT)

SAs have as parameters: Speaker (S), Hearer (H) and Propositional content (P). Preconditions and Effects are expressed as conditions and modifications of H's and S's beliefs and wants. Their execution is the production of an utterance.

*INFORM*(speaker, hearer, prop)  
 precondition: speaker *KNOW* prop  
 effect: hearer *KNOW* prop  
 body: hearer *BELIEVE* speaker *WANT* hearer *KNOW* prop

Figure 5-1: Allen & Perrault's Model of SAs

a mean of composing several SAs together to achieve a communicative function. The model introduces the important notion of **Plan Inference**: Understanding an utterance is actually discovering the intention which caused its production. Or in other terms, speech is symptomatic of the speaker's plan and intentions.

#### 5.1.3. Developments of the Planning Theory of Communication

This Planning Theory of Communication (PTC) has been developed in several directions by a number of researchers, but the basic framework remains roughly the same.

[Pollack 86] has shown the importance of distinguishing the speaker's and hearer's perspectives on the world in the process of Plan Inference: that is, each agent has his/her own representation of the world which is not

### Distinction between Hearer's and Speaker's Beliefs in Plan Inference

- In a question-answering system dealing with the Operating System domain:

(1) I want to prevent Tom from reading my mail-file. How can I set the permissions on it to faculty-read only?

It is important to detect that the speaker has already formulated a partial plan to solve his/her problem, and that this plan is based on incorrect assumptions on the domain.

- (2) There is no way for you to set the permissions on a file to faculty-read only. But you can move it into a password protected subdirectory; that will prevent Tom from reading it.

To be able to generate such a helpful answer, the system must understand the invalid beliefs of the speaker, even though it does not share them.

- If it fails to understand these beliefs, it will simply be unable to make sense of the question as in:

(3) I want to talk to Kathy, so I need to find out how to stand on my head.

**Figure 5-2:** Distinction between Speaker's and Hearer's Beliefs in [Pollack 86]

necessarily the same as his/her interlocutor, and similarly, each agent has his/her own set of operators and does not assume that the others participants have the same. This distinction allows a framework for studying a user's misconceptions in a domain, or invalid plans as illustrated in figure 5-2.

The other axis of development has been concerned with the distinction between immediate goals and longer term intentions or plans. The difficult part of the plan inference is to recognize a plan when knowing only some of its parts. Allen and Perrault originally used a heuristic search through the different plans embodying a given act and developed several "Partial Plans" concurrently. The main limitation of this approach is that it allows recognition of intentions in a single utterance but without providing any way to relate it to the ongoing conversation.

[Sidner and Israel 81, Carberry 83, McKeown *et al* 85, Litman & Allen 84] all develop the basic PTC, and try to apply it to the problem of understanding or generating whole conversations as opposed to isolated utterances. The problem is that a single utterance can be viewed as a step to achieve several high-level goals pursued during the conversation. Therefore, several utterances in a sequence are not independent, but can all participate to the realization of the same high-level intentions as the different steps of a plan. This observation leads to consideration of actions, goals and plans of different levels of abstractions.

The solution generally involves the structuring of the set of available known plans in a hierarchy of goals - from the more general intention to the basic communication unit. The process of plan inference is then to gradually refine the current goal of the user by relating it to higher level plans.

Several studies using the PTC have focused on the structure of the conversation. First approaches have assumed that in task oriented dialogues, the conversational structure follows the task structure; that is, the utterances are directly related to a step in the task, and therefore, the structure of the conversation is similar to the plan followed to solve the task at hand ([Litman & Allen 84, Grosz 77], see section 3.2.4). This hypothesis is totally

dependent on the framework of the PTC. The isomorphism between the semantic situation and the structure of discourse becomes in the PTC a direct mapping between linguistic acts (SAs) and the tasks of the domain. Although this approach has proven to be useful so far, it must be emphasized that it is only a partial account for the structure and cohesion of discourse. It is an explanation of the semantic coherence only in the task domain. It is however clear that the linguistic cohesive devices mentioned in paragraph 3.2.1 or interpersonal considerations similar to those described in 4.3 are also of primordial importance in accounting for the global coherence of a conversation. Therefore, any theory of discourse coherence has to be open to all these dimensions and allow an extension towards a more extensive theory accounting for more of the functions of language. This does not seem to be the case for the “isomorphism assumption.”

Altogether, the PTC is a recent field of research providing an interesting account for many of the “traditionally” difficult problems of computational linguistics. It helps in determining the focus of a sentence (given the intention of the speakers), is at the basis of a unified account of definite noun phrase usage in generation [Appelt 85]. It is commonly viewed as a working solution to the problem of Indirect Speech-acts, and can serve as the basis of ellipsis understanding [Carberry 83].

Another possible development of the PTC would be to integrate knowledge of short-term plans and goals of agents into long-term user models (see [Rich 79, Paris 85]). Long-term characteristics of users can indeed help determining immediate (short-term) goals and intentions of users. For example, the level of expertise of a user allows one to predict what kind of operations s/he may want to perform with the system. The problem of long-term characteristics is that they are hard to recognize. It is not clear whether long-term user-models may be inferred from discourse using techniques similar to the ones used to infer immediate intentions of utterances.

## 5.2. Grice's Theory of Meaning - Conversational Implicatures

The motivation of the theory of conversational implicatures is to explain occurrences where people understand more in an utterance than has been actually said.

1. *Can you close the door ?*
2. *Do you have the time ?*
3. *What happened to the roast beef ?*
4. *The dog seems happy...*
5. *I entered a house*
6. *The flag is white*
7. *The flag is blue, white and red*

(1) is rarely understood as a question on the ability of a person, but more often as a request to close a door. (2) is almost always understood as a request of information, not as a yes/no question. (4) answered to (3) can perfectly mean that perhaps the dog has eaten the roast beef. In (5), everyone would conclude that the house the locutor entered is not his/her house. If (6) is uttered, then people generally don't agree that (7) is true, even though there is no logical contradiction between (6) and (7).



[Grice 75] proposed to explain this apparent discrepancy between the “literal meaning” of a sentence and what is intended by a derivational process based on non-linguistic principles that he called maxims of conversation which form the co-operative principle. Before detailing what is meant by these terms, it is useful to first describe the theoretical framework in which they are defined.

As we have seen, Situation Semantics does not differentiate in nature between intuitively different occurrences of the term “meaning.” The field of pragmatics in linguistics, in contrast, is concerned with distinctions such as those generally felt between the utterances of “smokes means fire” (a law of nature) and “John means it’s ugly” (the expression of an opinion). A first step in this direction is marked by the definition of meaning-*nn* by Grice [Grice 57]. Grice stated a distinction between natural meaning and non natural meaning (meaning-*nn*). [Levinson 83] gives the following definition of Grice’s meaning-*nn*:

- S* meant-*nn* *z* by uttering *U* if and only if:  
 (i) *S* intended *U* to cause some effect *z* in recipient *H*  
 (ii) *S* intended (i) to be achieved simply by *H* recognizing that intention (i).

As for all definitions based on the concept of “intentionality,” this one causes some difficult philosophical problems, mainly related to the recursive nature of the concept of “intention.”<sup>8</sup> But it also primarily determines the framework of all pragmatic investigation: human intentional communication.

To explain how people can “mean-*nn*” more than “what they say,” it is necessary to highlight the word “recognizing” in the definition. One can assume that people use a set of known - or natural - conventions to derive a new meaning from a “literal” one, in order to explain that people can “recognize” meaning other than what linguistic information alone conveys.

[Grice 75] has hypothesized a set of such conventions that he presents as rational principles guiding the use of language in conversations. He presents them as four maxims following Kant’s categories of Reason. These maxims implement what Grice calls the co-operative principle (see figure 5-3).

The real power of these maxims comes from their versatility. They explain how “implicated” meaning can be derived; the assumption is that when people try to “make sense” of an utterance they take into account the co-operative principle. Thus for example, an application of the maxim of Quantity explains why sentence (6) above seems to contradict (7): if the flag is not only white, then the locutor will say it to be “as informative as possible,” since only “white” is mentioned, the flag must be completely white. Grice calls these kind of conclusions implicatures (as opposed to logical implications).

Grice classifies the several kind of implicatures as shown in figure 5-4. He emphasizes that maxims are not always followed. When saying “war is war” a locutor is apparently violating the maxim of Quantity by providing no information at all. But, hearers still try to understand the utterance as informative: the blatant violation of a maxim plays the role of a marker indicating that another meaning must be derived (here for example, that in exceptional circumstances such as war, certain horrible events do occur). Grice calls this operation “flouting of a

---

<sup>8</sup>an act is intentional only if it is done with the intention to be intentional.

- *The Co-operative Principle*: make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.
- *The Maxim of Quality*: try to make your contribution one that is true, specifically:
  1. do not say what you believe to be false
  2. do not say that for which you lack adequate evidence
- *The Maxim of Quantity*:
  1. Make your contribution as informative as is required for the current purposes of the exchange
  2. Do not make your contribution more informative than is required
- *The Maxim of Relevance*: make your contribution relevant
- *The Maxim of Manner*: be perspicuous, and specifically:
  1. avoid obscurity
  2. avoid ambiguity
  3. be brief
  4. be orderly

Figure 5-3: The Principle of Co-operation in (Grice-67)

#### Different Types of Implicatures

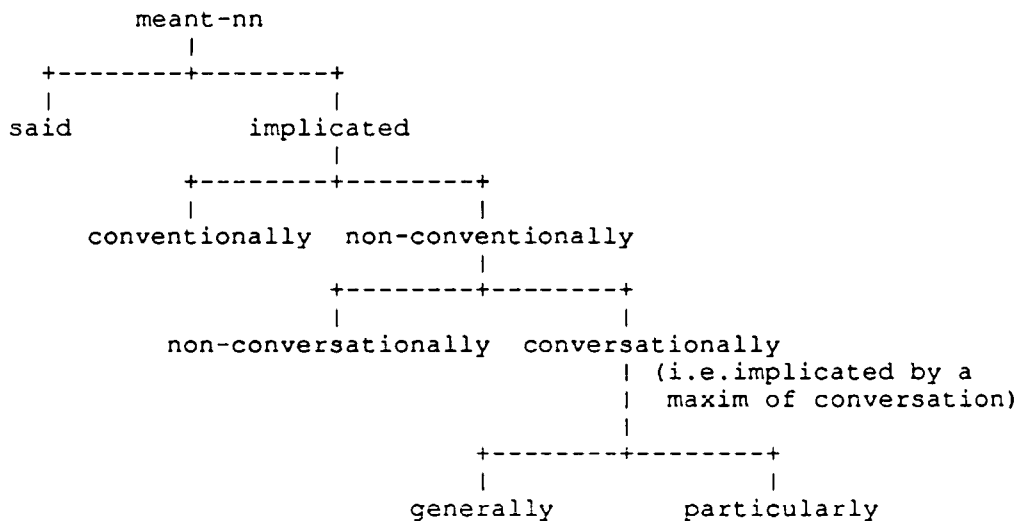


Figure 5-4: Different Types of Implicatures, in [Levinson 83]

maxim.”

Grice further distinguishes between general and particular implicatures. For example, the implicature derived from (5) is independent of any context (it is general) whereas the implicature derived from (4) is only valid in the context of the question (3) (it is particular).

Finally, Grice defines conventional implicatures as implicatures associated with particular words or expressions, and not with non-linguistic principles. For example, according to Grice, the word "but" has the same truth conditional meaning as "and," but adds the implicature that there is some contrast between the two clauses it connects.

The notion of implicature is one of the most powerful introduced by the study of pragmatics. It very clearly demonstrates the need to take into account the pragmatic situation in the operation of meaning derivation, and provides a theory explaining the existence of implicit meaning in language use.

Pragmatics is, however, extremely hard to formalize; the assumption that implicated meaning must be calculable is still an axiom, not a result of the theory: the interpretation of an implicated utterance requires the intervention of various and numerous sources of knowledge. Implicatures seem to be too high-level to be currently implemented as specified by Grice.

### 5.2.1. Using Implicatures in AI Systems

There have been several attempts to embody the theory of implicatures in AI systems. The Co-Op program [Kaplan 82] is a data-base query system able to generate "cooperative" answers to user queries.

- (8) How many students are taking the NLP class this semester ?
- (9) 0.
- (10) There is no NLP class this semester.

In response to (8), answer (10), although indirect, is much more helpful to the user than answer (9) which is literally correct. To provide such a helpful behavior, Kaplan has studied one particular case of implicatures that are called **presumption**. The existence of presumptions explain why sentences like "John is here but I don't believe it" are pragmatically anomalous; the fact that a locutor uses the definite noun phrases "the NLP class this semester" in (8) "implicates" that s/he believes that such a class is given. Kaplan called questions containing such presumptions "loaded questions." He defined the following strategy to deal with loaded questions: if the set presumed by the question is empty in the data base (no record can be found corresponding to its description), then the question may reveal a misconception of the user. The maxim of Quantity states that such misconceptions should be corrected in a cooperative answer.

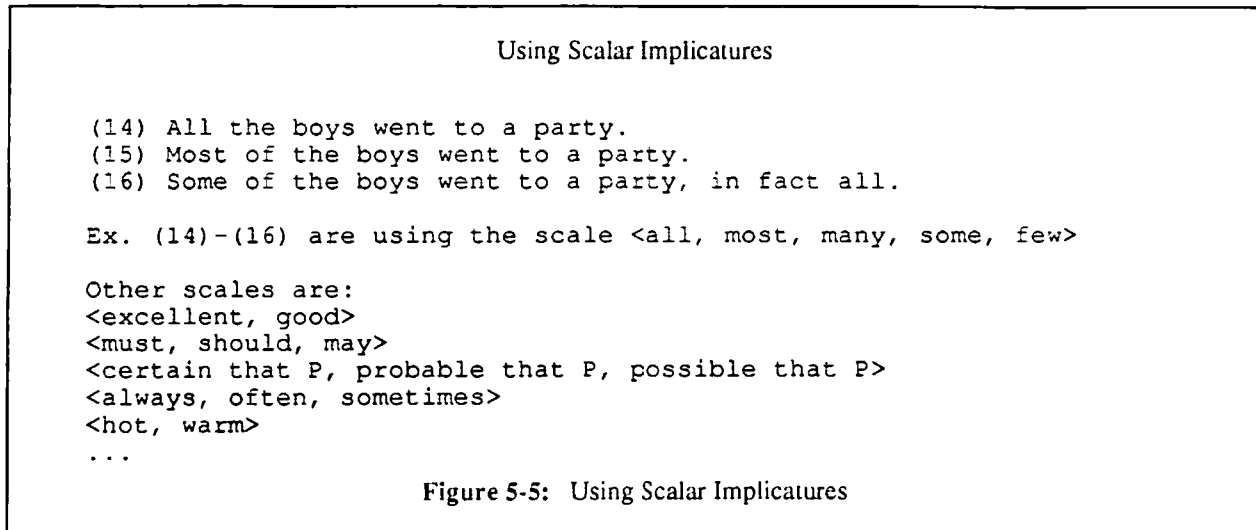
Interestingly, existential presumptions can be detected by a purely linguistic analysis and don't require world knowledge. Other types of presumptions however necessitate some kind of domain knowledge - as in example (11)-(13). In these examples, the status of the implicated meaning of the question is more related to what is called **presupposition** in its more general form (e.g., [Mays 80]).

- (11) List the undergraduates who taught this semester.
- (12) None.
- (13) Undergraduates cannot teach.

To be able to generate the cooperative answer (13) to query (11), one needs to know the rules of universities. This reasoning is much harder than in the case of existential presumptions.

[Hirschberg 83] proposed a treatment of another type of implicature related to the maxim of Quantity: scalar implicatures. Such implicatures are based on the concept of scale: "A linguistic scale consists of a set of linguistic

alternates or contrastive expressions of the same grammatical category, which can be arranged in a linear order by degree of informativeness or semantic strength" [Levinson 83, p.133] (see figure below). This concept gives an operational aspect to the "information" term mentioned in the maxim of Quantity.



Using an element of the scale which is not the highest degree compatible with what is described would be considered as an infraction of the maxim of Quantity (for example, saying (15) in place of (14) in figure 5-5). Hirshberg extends the definition of scale to different kind of semantic ordering: temporal scales, process/stages, prerequisite orderings... Her approach allowed her to deal with yes/no questions and generate cooperative indirect answers.

(17) Have you baked the cake?  
 (18) I just put it in the oven.  
 Based on the scale <mix; put in oven; bake>

Using scales is a very attractive way to operationalize the most productive of Grice's four maxims. But some problems are attached to the definition of the concept of scale: whether the scalar organization of expressions is a property of the lexicon or a part of the information contained in the Knowledge Base, the linguistic markers indicating that a scale is active in a sentence, whether it is possible to organize all the different scales in a usable representation system, how the activation of relevant scales is achieved. Many of these problems are related to problems with scripts mentioned in section 4.2.

### 5.3. What Must Be Integrated in a Representation of the Pragmatic Situation ?

The survey of pragmatically motivated works demonstrates the tremendous importance of the categories of **intention, belief, and mutual belief** within the process itself of communication. It also highlights the relevance of the study of communication as a framework for NLP research.

All the works surveyed in this section have in common their interest in the conditions of production of utterances: "who spoke, why and how" in addition to the classical "what was said" or "is it true" or "is it correct" approaches to language.

Finding answers to these questions necessitates representing the Pragmatic Situation. From what we have seen, some of the components of this representation are:

- The Beliefs of the locutor [Allen & Perrault 80]
- The Beliefs of all the participants [Pollack 86]
- The Intentions of all the participants. Intention being a difficult notion, it seems preferable to use the notions of Plans and Goals.
- The Mutual Beliefs of the participants,<sup>9</sup> [Particularly see [Appelt 85] for an effective use of the notion].

These are the main techniques currently used to represent the pragmatic situation in AI systems. But the underlying theories also show that many other parameters must be part of this representation as well. In addition to the general “conversational maxims” presented by Grice, it seems that several “norms” or guidelines constrain the conduct of conversation: Grice mentions that in addition to the conversational maxims, there exist other kinds of nonconversational maxims that are “aesthetic, social or moral in character.” See [Sadock 78, Kerbrat-orecchioni 86] for a critique of Grice’s classification of implicatures. [Brown & Levinson 87] have demonstrated the importance of what they call politeness in communication. [Bourdieu 82], [Labov 70] hint on the importance of sociological influences on language use. A whole tradition of works on the structure of conversation has demonstrated the existence of social rituals [Goffman 67] that are followed by conversationalists, and highlighted the importance of a “micro-sociology” of the interaction. The list of pragmatically relevant processes<sup>10</sup> having an influence on language use is large and each of these bring with it a set of variables that should be included in the representation of situation (see next section for a more detailed account of these point).

This suggests that building a theory of situation representation is difficult: too many variables need to be defined and structured, and above all, the influence of these variables on language use is extremely varied and not easily specified. It seems therefore a good strategy to try to bypass, in some way, the difficulty of defining situation and to study rational processes of general explanatory scope as the basis of the communicative behavior of individuals. Such processes would be the rational mechanisms explaining how communication works and would be applied to a uniform, well determined set of simple variables (see [Cohen & Levesque 85] for an example of this approach).

The Planning Theory definitely plays this role in current AI research: it represents language use as a uniform set of acts and the pragmatic situation as a uniform set of beliefs and describes the link between them as a simple coherent process.

But this methodological approach faces several serious problems. The technique of planning itself is not clearly defined. Notions of plan and realization are not always correctly related. Winograd expresses some of these uncertainties:

[Orthodox AI Planning] works reasonably well for some limited domains, but falls far short of being a general theory

---

<sup>9</sup>A and B mutually believe that P if A believe that P and B believes that P, A believes that B believe that P and B believes that A believes that P and so on.

<sup>10</sup>Pragmatically relevant means, in this context, that such processes have to do with the authors of utterances and their motivations, not only with the utterances

of intelligent action. It does not work well in the small (how I operate my finger muscles, or where an amoeba slithers), or in the large (how I conduct my life or where my research is headed). Even in the cases of clearly explicit rational planning (e.g., planning a bank robbery), the relation between plan and execution is not easy to capture (what happens when the teller sneezes?). [Winograd & Flores 86]

Actually, this critique presupposes that Planning is used as a **psychologically plausible** model of human behavior. And it correctly points out that it is not plausible at all as a complete cognitive theory. Several arguments support this claim:

- It is not practical to represent everything as beliefs (this would generate huge sets).
- Planning is a uniform mechanism: it does not account for cultural differences in communication behavior (cf [Brown & Levinson 87, p.9] for a discussion of the cross-cultural validity of a rational study of communication).
- It does not account for language evolution (if language is rational, why should it change?)

But, if Planning is not psychologically plausible as the total model, it may remain a valuable **abstract formalism**; After all, if one can find a powerful mechanism to account for the aggregation of lower level acts into patterns of behavior, then does it matter if it is plausible or not? Planning may be viewed as such a mechanism, working on **representations** of acts, and therefore distinct from any actual processes occurring in human behavior. If planning currently does not take into account influences from the linguistic context or social background of the speakers, then it is perfectly possible to integrate these influences into the preconditions and actions of a plan.

This clarification unfortunately demands a reconsideration of the methodological shift avoiding the problem of situation: if planning is not plausible, it is certainly not an underlying (explanatory) mechanism actually playing a role in human communication; therefore, it cannot account for many aspects of human communication if only based on belief variables as it would require additional parameters.

So the problem of defining situation comes at a later stage in this approach: the description of the linguistic acts (operators) available to the Planning mechanism must include heterogeneous parameters provided by an analysis of situation.

Furthermore, implementations of the SA theory based on planning tend to overlook the notions of **legitimacy, commitment and felicity**. These terms have been particularly developed in recent theoretical work on the SA theory (see [Habermas 79], [Winograd & Flores 86] for a presentation focusing on the notion of commitment). These notions derive from an analysis of what is actually achieved when a SA is performed, and what acts are doable using SAs. There is no single class of such acts, but rather, within a given social insitution, a set of conventions (felicity conditions) rule who has the legitimacy to achieve certain acts using SAs, and what kind of commitments this action impose. For example, in a company, only the chairman has the legitimacy to declare a meeting open. Conventions particular to the company, moreover, rule who can speak next and what kind of problems one faces if one does not follow these conventions (only such constraints can give sense to the notion of commitment).

In contrast to such an analysis of the social context of interactions, planning theory focuses on a single hearer/speaker relation, viewed in an empty social background, free of any conventions and commitments. Such a

view, in addition, makes it difficult to relate the SA theory to the theory of implicatures, since implicatures are also based on a set of conventions.

It is therefore clear that planning is not a unifying theory (as claimed in [Cohen & Levesque 85]) and must be complemented by a thorough description of situation - conventions and "network of commitments" (cf [Winograd & Flores 86]) particular to only certain kind of interactions. In the next section, we review in more details work in linguistics and sociology that might help developing a global model of situation.

#### 5.4. Summary

In this section, we have looked at the link between the interpersonal situation and language use. We have seen that research in AI applies the SA theory and tries to use the notion of implicatures. The main technique used is an extension of planning. Planning is commonly viewed at the basis of a theory of communication which is outlined in the research we have surveyed.

It is unclear, however, whether a logic of communication can be developed using the categories of beliefs, goals and planning as exclusive tools that could account for the complexity of human communication if it does not consider communication as a form of social action embedded within larger social structures.

### 6. The Notion of Situation from a Theoretical Point of View

The notion of situation as we have used it so far comes from linguistics. Many of the techniques used in AI derive from linguistic work. It is therefore interesting to try to relocate the notion in the field of linguistics and analyze what aspects of the linguistic study of situation could be usefully brought into AI.

#### 6.1. The Concept of Situation in Linguistics: Who Cares ?

An important precaution to take with a fuzzy notion, such as situation, is to determine what status, if any, has been relegated to it. Indeed, linguistics, in its historical struggle to define its own autonomous object of study, has often rejected the study of situation into a horizon of utopian un-scientific preoccupations and non-operational considerations (cf [Parret 74] for numerous examples of this approach). Many of the well known dichotomies used by linguists have as their only purpose, the rejection of situation out of the field of linguistics, *e.g.*, *parole/langue* (speech/language), *competence/performance*.

But, as some linguists' interests moved from the description of language as a system to the study of language use, the necessity of describing situation became clear. A theory of language use must explain how the same expression can give rise to an infinity of effects when used in different situations. We already mentioned the philosophical work which originated most of the theory of linguistic pragmatics: the speech-acts of Austin and Searle and the conversational principles of Grice both embody a theory of situation.

Other trends in linguistics are also worth mentioning: both ethnolinguistics and sociolinguistics propose interesting concepts for the study of situation. The aim of ethnolinguistics is to study communication within a given culture. The field was "officially" constituted in 1964 with the publication of [Hymes 64] and [Gumperz & Hymes

64]. Considering communication in different cultures allows a more acute perception of the importance of situation: if one has no idea on the meaning of a language, a sense of “what’s happening” is essential to any form of understanding.

Sociolinguistics develops the idea that even within the same culture, variation of the social situation has an influence on language use. [Labov 72] and [Fishman 71] are representative works of this trend. Such studies evaluate, for example, the impact of socio-economic variables on certain aspects of linguistic usage, the particularities of the speech pattern of a certain social class. Sociolinguistics attempts to capture the determinations that social structure imposes on language use. A sociolinguistic description of situation would be mainly composed of indications on the socio-professional status of locutors, and would necessitate background knowledge of the organization of society.

In contrast, [Goffman 64] has pointed out that interactions have a structure of their own and generate constraints on their participants often different from those induced by larger social structures. He has called these “micro” social structures and the constraints they generate interaction rituals. Ethnomethodology<sup>11</sup> is the study of the constraints of rituals on language use - or in the terminology they use, the analysis of social implicit.

Ethnomethodology puts the focus on the actor. It attempts to elucidate how social actors use and are constrained by the social methods (or rituals) that are available in a given culture. In contrast to the “macro” perspective of socio-linguistics, it does not study the source and origins of the social constraints, but how these constraints are practically used and translated in the behaviour of actors. An ethnomethodologist’s description of situation would model a specific ritual and roles and constraints determined by the ritual. Marking works in that framework are presented in [Garfinkel 68]. [Sacks, Schegloff & Jefferson 78, Atkinson & Heritage 84] on analysis of conversation are also related.

It is, however, necessary to organize the very large number of different rituals existing in a given society into a consistent framework. [Fishman 72] defines the notion of domain as a tool describing the organization of multiple “rituals” within society. This notion allows a fruitful articulation of the sociolinguistic and ethnomethodologist approaches. Figure 6-1 gives an idea of the role domains could play in a representation of social situation.

## 6.2. The Role of Situation in Theories

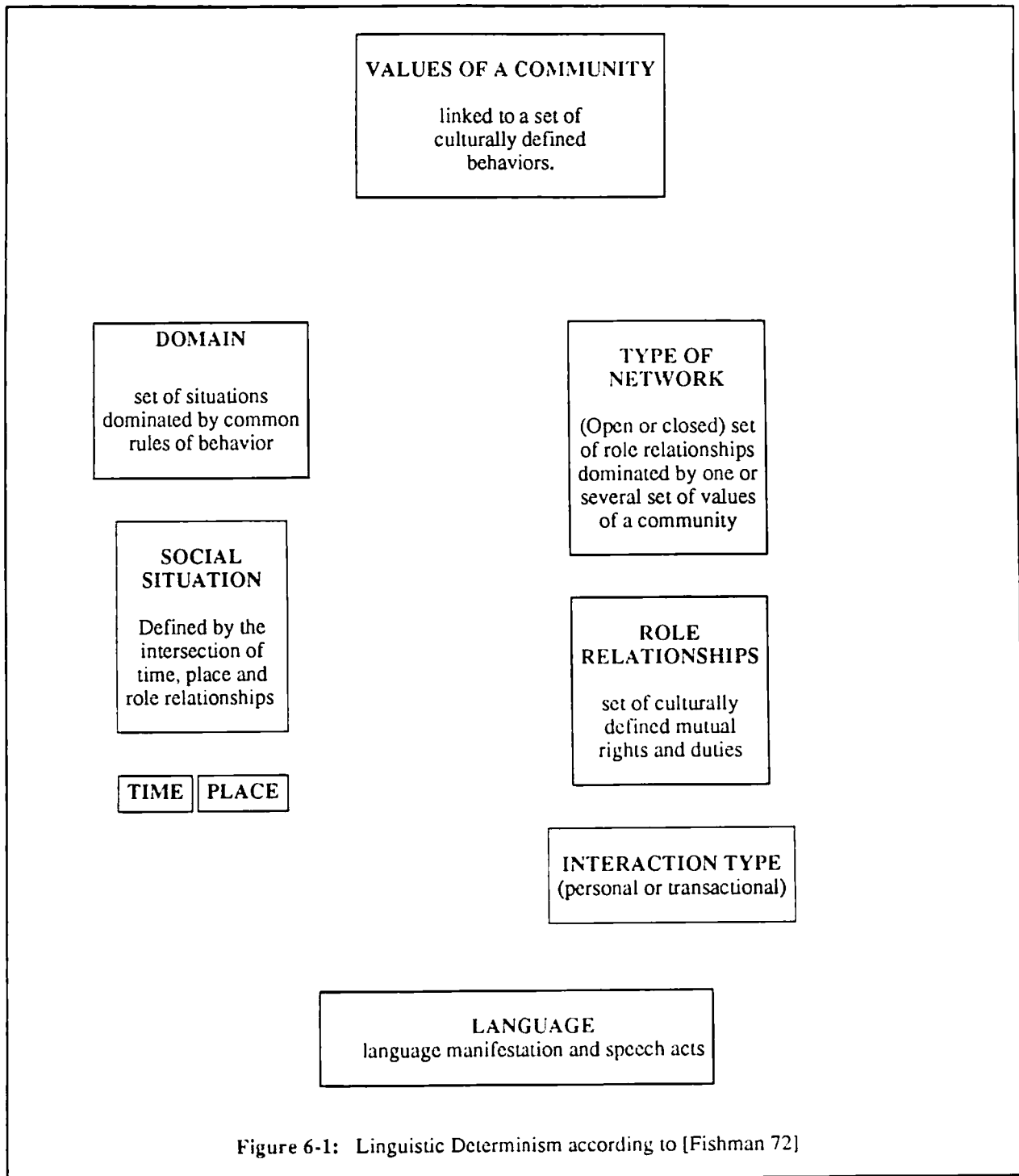
The role that different linguistic trends allot to situation depends on many theoretical decisions. The way situation is defined and used in different theories obviously depends on what role it is assumed to play. I therefore quickly survey, among the different linguistic theories, the most significant perspectives on that role.

For the Chomskian linguistics, situation is an “unfortunate” filter that alters the process of an ideal transparent communication. Situation is mainly a source of noise interfering with pure linguistic data [Chomsky 74]. In contrast, Halliday’s theory tries to find in language the marks of its social function: that is, a description of language

---

<sup>11</sup>The word comes from [Garfinkel 68]. The word methodology here does not refer to the scientific method used by this group of sociologists. It describes the object of the study: the methods used by social actors in their social life.





is a description of how it can be used in certain situations [Halliday 78]. Here, the notion of situation is prevalent at all levels.

Often the resort to situation in language use is mainly viewed as a convenient device that allows a more economical expression (cf [Prieto 66] for example). Saying "this" instead of describing an object within the field of vision is a major economy in many cases. This approach is supported by many psychological results showing the difficulty with communicating in the absence of a common visual reference. But [Benveniste 66] has criticized this theory in an important text proposing an interesting theory of pronoun use: using "I" and "you" instead of names is not a simple economy but has consequences on the behavior of locutors implied in a conversation. By using "you," I recognize that the other locutor is a member of my "discourse universe," and I grant him/her the right to become an "I." As a consequence, the speech acts realized by using pronouns are not realized when using the corresponding names ("I order you to come" vs. "John orders Peter to come.") Therefore, resorting to situation in a discourse often means more than just saving on the cost of expression. These ideas have been developed by [Anscombe & Ducrot 83] who propose a theory centered around the notion of act of utterance.

[Ducrot 72] proposes an original approach: he distinguishes between two components in a theory of semantics. Here, the linguistic component creates a series of semantic instructions that are given to the second component (rhetorical component), which in turn instantiates and executes these instructions in a given situation to produce the meaning of an expression. The originality of this scheme is that no meaning and no structure is manipulated by the linguistic component: only abstract and general "instructions" that indicate how to build the meaning of an expression in a given situation.<sup>12</sup>

Finally, it is important to add a word about the direction of influence between situation and language use. Most of the theories mentioned in this section view situation as a constraining agent, determining certain aspect of language use. In contrast, the theory of speech acts has first indicated how language use could change and affect situation. This mutual influence is quite hard to account for. It seems that the link between situation and language use is best viewed as a dialectic relation as proposed by [Bang & Door 79]. This consideration could have important consequences on both the representation of situation and of speech acts.

### 6.3. Situation in Theories: Definitions

When it comes to describing situation and its structure, unfortunately, linguistic theories are not very precise. Among the definitions proposed, some are global and high-level, and therefore, most often non-operational. Others are a long and detailed list of potentially relevant features, but remain often rather imprecise.

Examples of the first tendency include [Bloomfield 33], [Prieto 66] and [Goffman 64]. Bloomfield's definition of situation is "whatever is in the world of the locutor, including his internal dispositions when he emits his message." Prieto similarly defines situation as the collection of "all the facts known by the locutors" [Prieto 66]. Goffman's definition of social situation is more restricted and functional, but still hardly useful from an AI perspective: "an environment of mutual monitoring possibilities" [Goffman 64].

Lyon's definition is an example of the second tendency. It distinguishes three elements within situation: a

---

<sup>12</sup>The notion of semantic instruction is best viewed as the linguistic correlate (and extension) of the ethnomethodologist notion of "pragmatic instructions" proposed by [Cicourel 74].

spatio-temporal location, a global discursive context (immediate and past discourse), and a discourse universe composed of the beliefs and conventions of locutors directly relevant to the situation. Of course, the problem resides in the imprecision of this universe, and of the kind of relevance that is necessary to define it [Lyons 68].

Ethnolinguists, in contrast, when studying communication have defined situation in a more concrete manner. [Hymes 72] has proposed a model of situation analysis known as the *SPEAKING* model, which stands as an acronym for:

Setting	Key
Participants	Instrumentalities
Ends	Norm
Acts	Genre

This model was intended to provide a "grid" of analysis of communicative events. Some of the dimensions proposed are not independent, but this model is one of the richest available, and has proven to be useful in empirical research.

An example of such analysis is provided in [Bachmann *et al* 81]. The interaction described is a family lunch in a middle-class family. The *setting* is described as the living-room of the family home. The date is precised - a day in July. Under *participants*, we find a description of all the present family members. Their relation within the family is precised. Psychological and sociocultural features are given for each of the participants. For the *ends* of the interaction, Hymes distinguishes between the intentions of each of the participants and the results of the interaction. Here, the ends of the activity is to have something to eat, and at a more implicit level, to maintain family ties. The description of *acts* is more fuzzy. It includes a list of the themes mentioned during the conversation, and the style under which they have been approached. For each of the acts, the *key* indicates the modality used for the communicative function described: an advice is given, whispered or shouted, or presented as a joke, a child cries. The *instrumentalities* feature describes the different media of communication and for each of them, the code used: language, para-linguistic, proxemic... The codes of importance are those playing a role in the interaction: for example, the clothing may be of importance in a social gathering. The *norms* relevant for the analysis are of two kinds: norms of interaction and norms of interpretation. Norms of interaction constrain the way a conversation flows, who can speak first, how interruptions can be handled, the role of silence, etc. The norms of interpretation form a background of presuppositions, and help in expliciting the meaning of messages: for example, when the grand-father announces that his friend had eight children, it is understood as a compliment, in view of the grand-father's known positions on large families. Finally, the *key* is a categorization of the interaction in a network of known cultural activities: is it a story-telling, a commercial transaction, a business meeting, etc.

The problem, from a more strictly linguistic perspective, is that situation is relevant only to the extent that it can influence language use. But since language is always "about" something, there seems to be no limit to this effect of influence. A solution to the problem of the ever extending scope of situation is proposed by Halliday with his definition of registers. Halliday's approach is greatly influenced by Hymes' model and forms in a way its "linguistic application" with some modifications. Considering situational factors as determinant of a text, Halliday defines the categories of *field*, *tenor*, and *mode* which are supposed to "represent the situation in its generative aspects:"

Field refers to the ongoing activity and the particular purposes that the use of language is serving within the context

of that activity; *tenor* refers to the interrelations among the participants (status and role relationships); and *mode* covers roughly Hymes' channel, key and genre. [Halliday 78, p.62].

These definitions may seem still vague, but Halliday takes a resolutely text-oriented approach: each of these factors is relevant to language use because it directly affects one of its aspects: the three components of situation correspond to the three components of Halliday's system; that is, to the system of choices that produce well formed and appropriate discourse.

By and large, it is the ideational component of the system that is activated by the choice of field, the interpersonal by the *tenor*, and the textual by the *mode*. [Halliday 78, p.63].

This approach is particularly attractive in that it clearly represents the link between situation and language (viewed as a system). From an AI perspective, it seems best fitted to a generation system (see [Matthiessen 85] for an application to the Nigel System of some of these ideas).

#### 6.4. The Distinction Between Semantic and Interactive Analysis

I want to stress the richness of the ethnolinguistic tradition and of the concepts it brings to the analysis of communication. Hymes' *SPEAKING* model, the notions he introduced of a functional analysis of communication, and the related linguistic theories indicate that an analysis of a communication event cannot remain at a traditional semantic level but needs to be extended to an interactive analysis.

An interaction must be represented as an activity located within larger domains in the sense defined by [Fishman 72]. It is therefore above all a cultural event governed by cultural norms or rituals in Goffman's sense. Furthermore, it is performed by participants assuming various roles, and standing together in various interpersonal relationships. The interaction has a certain purpose, and when achieving this purpose, language can assume certain conventional functions, and therefore certain aspects of its use are determined. Finally, a communication act is "about something." Only this "aboutness" is accounted for by a traditional semantic analysis. An interactive analysis needs to account for the whole communicative event.

The mastery of all the constraints and determinations existing in an interaction has been called *communicative competence* by Hymes and Gumperz, and, from another perspective by [Habermas 70]. Since NLP has as an aim to model the communicative competence of individuals within certain restricted domains, research must deal with some of the issues addressed by the theories briefly mentioned in this section.

### 7. Towards a Model of a Coherent Interaction

The techniques we have surveyed in this paper give an NLP system the ability to behave in a situation-sensitive way. Such techniques are all parts of a global model of what would be a coherent interaction between a program and its user(s). Studying situation allows one to account for a number of linguistic phenomena. I believe it could also lead to a better functional definition of the role of an NLP system in human/computer communication (see [Winograd & Flores 86, Suchman 87] for insightful reflections on this point).

In this paper, we have distinguished between three perspectives on situation: linguistic, semantic and pragmatic. We have also stressed the difficulty with studying the interaction between the three types of situation presented. It is

important to remember that the textual, semantic and pragmatic situations are not three different objects, but only three different aspects of the same problem. They all contribute to the global coherence of an interaction. For each of these aspects, we have looked at the techniques used in AI to deal with the information it can bring, and described what linguistic phenomena can be accounted for by a study of situation. Figure 7-1 summarizes this survey.

The study of textual situation accounts for both global (text organization) and local (anaphora, lexical choice, syntactic form) aspects of discourse coherence. The techniques used in AI are mainly based on the use of focus and of rhetorical schemas. We have seen that for both techniques problems arise if one tries to use them alone, without paying attention to semantic and pragmatic issues.

Similarly, the study of semantic situation accounts for the ‘‘content’’ of discourse. We have looked at some AI techniques of knowledge representation, based on the principle of distinction, and at Situation Semantics, based on the relation theory of meaning. We have pointed out that, here again, problems arise if one tries to account for content using only information from the semantic situation. Indeed, language being used within a social matrix, and serving interpersonal goals, one must take into account these social aspects to properly understand or produce language.

Finally, a recent trend in AI focuses on the study of the pragmatic situation. New techniques, based on the speech-act theory or on the implicature theory, account for a number of linguistic phenomena. AI techniques implementing the speech-act theory generally use the concept of planning, viewing speech acts as operators that speaker and hearer use to achieve their goals. This approach suffers serious limitations: it views the speaker/hearer in an idealized vacuous social situation, and therefore overlooks the notions of legitimacy, commitment and felicity, which are, in contrast, the primary concern of recent developments of the speech-act theory. We have argued that a more careful study of the global social situation could accord a proper treatment to the notions of legitimacy, commitment and felicity, and eventually reconcile the use of speech-act with the use of implicatures.

The development of a global model of a coherent interaction is definitely out of reach for still some time. But several pieces of such a model already exist or are currently being developed. A study of the ideas developed by sociolinguists such as Fishman or ethnolinguists like Hymes could help in the development of such a global model, and in the articulation of the existing trends. Merging different sources of information indeed becomes more and more necessary as the phenomena studied become more complex. We have tried to indicate that taking into account the pragmatic situation could help in dealing with phenomena thought to be purely linguistic or semantic, *e.g.*, text organization or recovery of implicit information. A global model developed on the lines suggested by Fishman (cf section 6) or expanding Hymes’ SPEAKING model would be useful in this attempt to organize different types of information and techniques.

## ACKNOWLEDGEMENTS

I would like to thank my advisor Kathy McKeown for her invaluable comments on countless previous versions of this paper. This paper would not exist without her support. I would also like to thank Michael Lebowitz and Terry Boulton for their help and judicious comments.

<u>Techniques used in AI:</u>	<u>Linguistic phenomena accounted for:</u>
<u>Textual Situation</u>	
<b>Global:</b> Schemas Global focus	Text Organization
<b>Local:</b> Theme/rheme Cohesive devices Focus	Anaphora Coreference Constraints on lexical choice Word disambiguation Constraints on choice of syntactic form
<u>Semantic Situation</u>	
<b>Knowledge Representation</b> Frames, Scripts, MOPs, Semantic nets ...	Reference Deixis Constraints on lexical choice Word disambiguation Thematic development
<b>Control of inference</b> Incompleteness in content and structure Principle of Distinction	
<b>Situation Semantics</b> Relation theory of meaning	
<u>Pragmatic Situation</u>	
Speech Act models Planning Plan inference Intention inference Misconceptions detection	Performative effects Modals Time aspects Personal pronouns/forms of address Reference Politeness Implicit Ellipsis Determining focus
User modeling	
Implicature	

**Figure 7-1:** Links between situation and linguistic phenomena

## References

- [Abelson 82] Abelson, P.R.  
Preface.  
*Dynamic Memory: A Theory of Reminding and Learning in Computers and People.*  
Cambridge University Press, New York, 1982.
- [Allen & Perrault 80] Allen, J.F. & Perrault, C.R.  
Analysing Intentions in Utterances.  
*Artificial Intelligence*, 15(1):pages 143-178, January, 1980.
- [Anscombre & Ducrot 83] Anscombre, J.C. & Ducrot, O.  
*Philosophie et langage: L'argumentation dans la langue.*  
Pierre Mardaga, Bruxelles, 1983.
- [Appelt 85] Appelt, D. E.  
*Planning Natural Language Utterances.*  
Cambridge University Press, Cambridge, England, 1985.
- [Atkinson & Heritage 84] Atkinson, J.M and Heritage, J.  
*Structures of Social Action: Studies in Conversation Analysis.*  
Cambridge University Press and Maison des Sciences de l'Homme, Cambridge and Paris, 1984.
- [Austin 62] Austin, J.L.  
*How to do Things with Words.*  
Harvard University Press, Cambridge, Mass., 1962.
- [Bachmann *et al* 81] Bachmann, C. *et al.*  
*Langage et communications sociales.*  
Hatier-Credif, Paris, 1981.
- [Bang & Door 79] Bang, J.C. and Door, J.  
Language, Theory, and Conditions for Production.  
*Januara Linguarum 85. Pragmalinguistics - Theory and Practice.*  
Mouton, The Hague, 1979, pages pages 21-50.
- [Barwise & Perry 83] Barwise, J. & Perry, J.  
*Bradford Book: Situations and Attitudes.*  
The MIT Press, Cambridge, Mass., 1983.
- [Benveniste 66] Benveniste, E.  
*Problemes de linguistique generale.*  
PUF, Paris, 1966.
- [Berrendoner 81] Berrendoner, A.  
*Propositions: Elements de Pragmatique linguistique.*  
Les Editions de Minuit, Paris, 1981.
- [Bloomfield 33] Bloomfield, L.  
*Language.*  
Harper and Row, New York, 1933.

- [Bourdieu 82] Bourdieu, P.  
*Ce que parler veut dire.*  
Fayard, Paris, 1982.
- [Brown & Levinson 87]  
Brown, P. & Levinson, S.C.  
*Studies in Interactional Sociolinguistics 4: Politeness : Some universals in language usage.*  
Cambridge University Press, Cambridge, 1987.
- [Calfey & Curley 84]  
Calfey, R.C and Curley, R.  
Structure of Prose in the Content Areas.  
*Understanding Reading Comprehension.*  
International Reading Association, 1984.
- [Carberry 83] Carberry, S.  
Tracking User Goal in an Information Seeking Environment.  
In *Proceedings of AAAI-83*. American Association for Artificial Intelligence, 1983.
- [Charniak 77] Charniak, E.  
Ms. Malaprop, a Language Comprehension Program.  
In *Proceedings of the Fifth IJCAI*. IJCAI, 1977.
- [Chomsky 74] Chomsky, N.  
Discussion with Herman Parret.  
*Janua Linguarum. Discussing Language.*  
Mouton, The Hague, 1974.
- [Cicourel 74] Cicourel, A.V.  
*Cognitive Sociology : Language and Meaning in Social Interaction.*  
Free Press, New York, 1974.
- [Cohen 83] Cohen, R.  
*A Computational Model for the Analysis of Arguments.*  
PhD thesis, University of Toronto, Department of Computer Science, October, 1983.  
Tech Report No. 151.
- [Cohen & Levesque 85]  
Cohen, P.R. and Levesque, H.J.  
Speech Acts and Rationality.  
In *Proceedings of the 23rd Conference of the ACL*, pages pages 49-59. Association for Computational Linguistics, 1985.
- [Cullingford 78] Cullingford, R.  
*Script Application: Computer Understanding of Newspaper Stories.*  
PhD thesis, Yale University, Department of Computer Science, 1978.  
Technical report 116.
- [Danlos 84] Danlos, L.  
*Generation de textes en langage naturel.*  
Masson, Paris, 1984.
- [De Saussure 49] De Saussure, F.  
*Cours de linguistique generale, 4eme edition.*  
Librairie Payot, Paris, 1949.
- [Ducrot 72] Ducrot, O.  
*Collection Savoir: Dire et ne pas dire - Principes de semantique linguistique.*  
Herman, Paris, 1972.



- [Ducrot & Todorov 72] Ducrot, O. and Todorov, T.  
*Points: Dictionnaire encyclopedique des sciences du langage.*  
 Editions du Seuil, Paris, 1972.
- [Fikes and Nilsson 71] Fikes, R. E. and Nilsson, H. J.  
 STRIP: A New Approach to the Application of Theorem Proving to Problem Solving.  
*Artificial Intelligence*, 2:pages 189-205, 1971.
- [Firbas 66] Firbas, J.  
*Travaux Linguistiques de Prague: On Defining the Theme in Functional Sentence Analysis.*  
 University of Alabama Press, 1966.
- [Fishman 71] Fishman, J.A.  
 The sociology of language: an interdisciplinary social science approach to sociolinguistics.  
*Advances in the Sociology of Language.*  
 Mouton, The Hague, 1971.
- [Fishman 72] Fishman, J.A.  
 The link between macro- and micro- sociology in the study of who speaks what to whom and when.  
*Directions in sociolinguistics.*  
 Holt-Rinehart-Winston, New York, 1972.
- [Garfinkel 68] Garfinkel, H.  
 The origin of the term 'ethnomethodology'.  
*Proceedings of the Purdue Symposium on Ethnomethodology.*  
 Purdue University, Purdue, Indiana, 1968.
- [Giglioli 72] Giglioli, P.P.  
*Language and Social Context.*  
 Penguin Books, Harmondsworth, England, 1972.
- [Goffman 64] Goffman, E.  
 The neglected Situation.  
*The Ethnography of Communication, special issue of the American Anthropologist.*  
 American Anthropological Association, Washington D.C., 1964, pages pages 133-136.  
 also in [Giglioli-72].
- [Goffman 67] Goffman, E.  
*Interaction ritual. Essays in face-to-face behavior.*  
 Doubleday, New York, 1967.
- [Goldberg 83] Goldberg, J. A.  
 A Move Toward Describing Conversational Coherence.  
*Series in interpersonal communication. Conversational Coherence, Form, Structure and Strategy.*  
 Sage Publications, Beverly Hills/London/New Delhi, 1983, pages pages 25-45.
- [Granville 83] Granville, R.  
*Cohesion in Computer Text Generation : Lexical Substitution.*  
 Technical Report MIT/LCS/TR-310, MIT, 1983.
- [Greimas 66] Greimas, A.  
*Semantique structurale.*  
 Larousse, Paris, 1966.

- [Greimas 74] Greimas, A.  
*Structural Semantics : an attempt at a method.*  
University of Nebraska Press, Lincoln, 1974.  
English translation.
- [Grice 57] Grice, H. P.  
Meaning.  
*Philosophical Review*, 66(3):pages 377-388, 1957.
- [Grice 75] Grice, H. P.  
Logic and conversation.  
*Syntax and Semantics: Speech Acts, Volume 3.*  
Academic Press, New York, 1975.
- [Grosz 77] Grosz, B. J.  
Representation and use of focus in a system for understanding dialogs.  
In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence.* Morgan Kaufmann, Los Altos, CA, 1977.
- [Grosz, Joshi & Weinstein 83] Grosz, B.; Joshi, A.K. and Weinstein, S.  
Providing a Unified Account of Definite Noun Phrases in Discourse.  
In *21th Annual Meeting of the Association for Computational Linguistics*, pages pages 44-50.  
ACL, Cambridge, Mass., June, 1983.
- [Gumperz & Hymes 64] Gumperz, J.J. and Hymes, D.  
*The Ethnography of Communication, special issue of the American Anthropologist.*  
American Anthropological Association, Washington D.C., 1964.
- [Habermas 70] Habermas, J.  
Towards a Theory of Communicative Competence.  
*Inquiry*, (3):pages 360-75, 1970.
- [Habermas 79] Habermas, J.  
What is Universal Pragmatics?  
*Communication and the Evolution of Society.*  
Beacon Press, Boston, 1979, pages pages 1-68.
- [Halliday 76] Halliday, M.A.K.  
*System and Function in Language.*  
Oxford University Press, London, 1976.
- [Halliday 78] Halliday, M.A.K.  
*Language as Social Semiotics. The social interpretation of language and meaning.*  
Edward Arnold, London, 1978.
- [Halliday & Hasan 76] Halliday, M.A.K and Hasan, R.  
*English Language Series: Cohesion in English.*  
Longman, London, 1976.
- [Hasan 84] Hasan, R.  
Coherence and Cohesive Harmony.  
*Understanding Reading Comprehension.*  
International Reading Association, 1984, pages pages 181-219.
- [Hendrix 75] Hendrix, G.G.  
*Partitioned Networks for the Mathematical Modeling of Natural Language Semantics.*  
Technical Report NL-28, University of Texas Austin, Department of Computer Sciences, 1975.

- [Hendrix 77] Hendrix, G. G.  
Human engineering for applied natural language processing.  
In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*. Morgan Kaufmann, Los Altos, CA, 1977.
- [Hirschberg 83] Hirschberg, H.  
Towards a Redefinition of Yes/No Question.  
In *Proceedings of the 22nd annual meeting of the ACL*. Association for Computational Linguistics, Stanford University, California, 1983.
- [Hobbs 78] Hobbs, J.  
*Coherence and Coreference*.  
Technical Note 168, SRI International, 1978.  
Menlo Park, California.
- [Hymes 64] Hymes, D.H.  
*Language in Culture and Society : a Reader in Linguistics and Anthropology*.  
Harper and Row, New York, 1964.
- [Hymes 72] Hymes, D.H.  
Models of the Interaction of Language and Social Life.  
*Directions in Sociolinguistics : the Ethnography of communication*.  
Holt-Rinehart-Winston", New York, 1972.
- [Jameson 83] Jameson, A.  
Impression Monitoring in Evaluation-Oriented Dialog. The Role of the Listener's Assumed Expectations and Values in the Generation of Informative Statements.  
In *Proceedings of the IJCAI-83*. International Joint Conferences on Artificial Intelligence, Karlsruhe, 1983.
- [Kaplan 82] Kaplan, S. J.  
Cooperative Responses from a Portable Natural Language Query System.  
*Artificial Intelligence*, 2(19)1982.
- [Kay 79] Kay, Martin.  
Functional Grammar.  
In *Proceedings of the 5th meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, 1979.
- [Kerbrat-orecchioni 86] Kerbrat-orecchioni, C.  
*Linguistique: L'implicite*.  
Armand Colin, Paris, 1986.
- [Labov 70] Labov, W.  
The study of Language in its Social Context.  
*Language and Social Context*.  
Penguin Books, Harmondsworth, England, 1970, pages pages 283-307.
- [Labov 72] Labov, W.  
The Logic of Nonstandard English.  
*Language and Social Context*.  
Penguin Books, Harmondsworth, England, 1972, pages pages 179-215.
- [Lebowitz 80] Lebowitz, M.  
*Generalization and memory in an integrated understanding system*.  
Technical Report 186, Yale University Department of Computer Science, New Haven, CT, 1980.  
PhD Thesis.

- [Levesque & Brachman 83] Brachman, R. and Levesque, H.  
*Readings in Knowledge Representation.*  
Morgan Kaufmann Publishers, Los Altos, Ca., 1983.
- [Levin 78] Levin, G.  
*Prose Models (4th edition).*  
Harcourt Brace Jovanovich, New York, 1978.
- [Levinson 83] Levinson, S. C.  
*Pragmatics.*  
Cambridge University Press, Cambridge, England, 1983.
- [Linguistics and Philosophy 85] Linguistics and Philosophy.  
*Special Issue on Situation Semantics.*  
Reidel Publishing Corp., Dordrecht, Holland, 1985.
- [Litman & Allen 84] Litman, D.J. and Allen, J.F.  
A Plan Recognition Model for Clarification Subdialogues.  
In *Coling84*, pages pages 302-311. COLING, Stanford, California, July, 1984.
- [Lyons 68] Lyons, J.  
*Introduction to Theoretical Linguistics.*  
Cambridge University Press, London, 1968.
- [Mann 84] Mann, W.C.  
*Discourse Structure for Text Generation.*  
Technical Report ISI/RR-84-127, Information Sciences Institute, February, 1984.  
4676 Admiralty Way, Marina del Rey, California 90292-6695.
- [Matthiessen 85] Matthiessen, C.  
The Systemic Framework in Text Generation : Nigel.  
*Systemic Perspectives on Discourse.*  
Ablex, Norwood, New Jersey, 1985.
- [Mays 80] Mays, E.  
Correcting Misconceptions about Data Base Structure.  
In *Proceedings 3-CSCSI*. Canadian Society of Computational Studies of Intelligence, Victoria,  
B. C., May, 1980.
- [McKeown 85] McKeown, K. R.  
*Using Discourse Strategies and Focus Constraints to Generate Natural Language Text.*  
Cambridge University Press, Cambridge, England, 1985.
- [McKeown et al 85] McKeown, K.R., Wish, M., and Matthews, K.  
Tailoring Explanations for the User.  
In *Proceedings of the IJCAI*. IJCAI, 1985.
- [Minsky 75] Minsky, M.  
A framework for representing knowledge.  
*The Psychology of Computer Vision.*  
McGraw-Hill, New York, 1975.
- [Morris 71] Morris, C.W.  
*Approaches to Semiotics: Writings in the General Theory of Signs.*  
Mouton, The Hague, 1971.

- [Mukai & Kurosawa 85] Mukai, K. and Kurosawa, H.  
Complex Indeterminates in Prolog and its Application to Discourse Models.  
*New Generation Computing*, (4):pages 441-466, 1985.
- [Paris 85] Paris, C. L.  
Description strategies for naive and expert users.  
In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*.  
Chicago, 1985.
- [Parret 74] Parret, H.  
*Janua Linguarum: Discussing Language*.  
Mouton, The Hague, 1974.
- [Pollack 86] Pollack, M.E.  
A Model of Plan Inference that Distinguishes between the Beliefs of Actors and Observers.  
In *24th Annual Meeting of the Association for Computational Linguistics*, pages pages 207-214.  
ACL, Columbia University, June, 1986.
- [Prieto 66] Prieto, L.J.  
*Principes de Noologie*.  
Mouton, The Hague, 1966.
- [Prieto 75] Prieto, L.J.  
*Le sens commun: Pertinence et pratique*.  
Les editions de Minuit, Paris, 1975.
- [Rich 79] Rich, E.  
User modeling via stereotypes.  
*Cognitive Science*, 3(4):pages 329-354, 1979.
- [Riesbeck & Martin 85] Riesbeck, C.K. and Martin, C.E.  
*Direct Memory Access Parsing*.  
Technical Report YALEU/DCS/RR#354, Yale University, Department of Computer Science,  
January, 1985.
- [Rumelhart 75] Rumelhart, D. E.  
Notes on a schema for stories.  
*Representation and Understanding: Studies in Cognitive Science*.  
Academic Press, New York, 1975.
- [Sacks, Schegloff & Jefferson 78] Sacks, H. Schegloff, E.A. and Jefferson, G.  
A Simplest Systematics for the Organization of Turn Taking in Conversation.  
*Studies in the Organization of Conversational Interaction*.  
Academic Press, New York, 1978.  
also in *Language*, 1974, 50, 696-735.
- [Sadock 78] Sadock, J.M.  
On Testing for Conversational Implicatures.  
*Syntax and Semantics: Pragmatics*.  
Academic Press, New York, 1978.
- [Schank 82] Schank, R. C.  
*Dynamic Memory: A Theory of Reminding and Learning in Computers and People*.  
Cambridge University Press, New York, 1982.

- [Schank and Abelson 77]  
 Schank, R. C. and Abelson, R. P.  
*Scripts, Plans, Goals and Understanding.*  
 Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.
- [Schank and Riesbeck 81]  
 Schank, R. C. and Riesbeck, C. K.  
*Inside Computer Understanding.*  
 Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1981.
- [Searle 69]  
 Searle, J.  
*Speech Acts: An Essay in the Philosophy of Language.*  
 Cambridge University Press, Cambridge, England, 1969.
- [Searle 75]  
 Searle, J. R.  
 Indirect Speech Acts.  
 In P. Cole and J. L. Morgan (editor), *Syntax and Semantics.* Academic Press, New York, 1975.
- [Sgall 74]  
 Sgall, P.  
 Focus and Contextual Boundness.  
*Topic and Comment, Contextual Boundness and Focus.*  
 H. Buske Verlag, Hamburg, 1974.
- [Sgall, Hajicova & Benesova 73]  
 Sgall, P. and Hajicova, E. and Benesova, E.  
*Topics, Focus and Generative Semantics.*  
 Scriptor Verlag, Kronberg/Taunus, Democratic Republic of Germany, 1973.
- [Shepherd 26]  
 Shepherd, H. R.  
*The Fine Art of Writing.*  
 Macmillan Co., New York, N.Y., 1926.
- [Sidner 79]  
 Sidner, C. L.  
*A Computational model of co-reference comprehension in English.*  
 PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1979.
- [Sidner and Israel 81]  
 Sidner, C. and Israel D.  
 Recognizing Intended Meaning and Speakers' Plans.  
 In *Proceedings of the IJCAI.* International Conferences on Artificial Intelligence, August, 1981.
- [Suchman 87]  
 Suchman, L.A.  
*Plans and Situated action: The problem of human-machine communication.*  
 Cambridge University Press, Cambridge, 1987.
- [Tannen 84]  
 Tannen, D.  
*Conversational Style. Analysing Talk among Friends.*  
 Ablex Publishing Corp., Norwood, New Jersey, 1984.
- [Van Dijk 77]  
 Van Dijk, T.A.  
*Text and Context : Explorations in the Semantics and Pragmatics of Discourse.*  
 Longman, New York, 1977.
- [Wilensky 78]  
 Wilensky, R.  
*Understanding goal-based stories.*  
 Technical Report 140, Yale University Department of Computer Science, 1978.

- [Winograd 85] Winograd, T.  
*Moving the Semantic Fulcrum.*  
Technical Report CSLI-85-17, Center for the Study of Language and Information, Stanford University, 1985.  
also in *Linguistics and Philosophy*, 8:1 (1985), 91-104.
- [Winograd & Flores 86]  
Winograd, T. and Flores, F.  
*Language and Being: Understanding Computers and Cognition - A New Foundation for Design.*  
Ablex Publishing Corporation, Norwood, New Jersey, 1986.
- [Woods 80] Woods, W. A.  
Cascaded ATN grammars.  
*American Journal of Computational Linguistics*, 6(1)1980.
- [Woods and Kaplan 72]  
Woods, W. A. and Kaplan, R. M.  
*The lunar sciences natural language information system: Final report.*  
Technical Report BBN Report 2265, Bolt Beranek and Newman, Inc., Cambridge, MA, 1972.