Foundations of Cognitive Science

edited by
Michael I. Posner

"DISCOURSE" GROSZ, B.J., M. POLLACK, AND C. SIDNER

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

ıtification in reading
Psychology: Learning,

ology 1:59–85.

Perception and Psycho-

11 Discourse

Barbara J. Grosz, Martha E. Pollack, and Candace L. Sidner

Research in discourse analysis, discourse structure, and discourse processing considers discourses, rather than isolated sentences, to be the basic units of language. Central to all endeavors is a desire to uncover the mechanisms that make certain sequences of utterances coherent and others not and to understand how context affects the use and understanding of various kinds of linguistic expressions. In this chapter we focus on three collections of discourse problems: those that arise in trying to determine discourse structure, those relating to phrase-level phenomena (for example, pronouns, definite descriptions, ellipsis), and those that concern how the meaning or effect of a discourse or discourse segment is derived. Because more research has been done in the cognitive sciences than could be reported in this chapter, we focus on work that has taken computational or processing questions to be central.

We begin with a survey of early approaches to discourse within artificial intelligence. These early approaches made evident the problems in existing theories, the need to determine the role of nonlinguistic information (including domain knowledge, differences among beliefs of speaker and hearer, reader and writer, and cognitive state) in discourse processing, and the need for computationally tractable theories adequate to support discourse processing. Next we describe approaches to discourse structure, discussing various ideas about the kinds of discourse segments that exist and the kinds of relations that join them. Following previous work (Grosz and Sidner 1986), we distinguish the linguistic structure of the discourse from its intentional structure and its evolving attentional state.

In section 11.3 we examine different kinds of information about discourse structure that is carried explicitly in the linguistic signal itself. In section 11.4 we describe research on the effects of discourse structure and discourse context on the interpretation of various phrase-level phenomena, including definite descriptions, pronouns, quantified noun phrases (for example, *every girl on the block*), event reference, and ellipsis. Section 11.5 describes various formulations of plans and techniques that have been developed for plan recognition. Once we take seriously the idea that language is used by agents to affect their world, it becomes crucial to understand the effects an individual utterance can have (this was the initial purview of speech-act theory), and the ways in which utterances can combine to form larger actions (research on plan recognition in discourse). Linguistic actions can thus be seen to affect discourse structure and discourse meaning.

11.1 Early Approaches to Discourse Processing

The earliest attempts to deal directly with the computational problems of discourse were carried out within the context of building complete computer-based natural-language processing systems in the early 1970s. These systems maintained no model of discourse structure in the sense that it is now understood. However, system creators experimented with techniques for pronoun understanding and a limited form of intention recognition. The LUNAR system (Woods et al. 1972, Woods 1978) provided mechanisms for resolving pronouns, but because each question-and-answer pair was an independent exchange, the discourse model was simple: only a list of previously mentioned entities were maintained for use in pronoun understanding. Charniak (1977) attempted to identify the referents of definite descriptions and pronouns in the process of "understanding" children's stories by encoding domain information in rules of inference and associated triggering programs. The most extensive dialogue system, Winograd's SHRDLU system (Winograd 1971, 1972) participated in a dialogue with a user about constructions of toy blocks. SHRDLU interpreted the user's intentions as programs to perform actions on its part and kept a history of actions it had performed. It could interpret some of the personal pronouns, some definite descriptions, and "the one"-type anaphoric expressions using "word specialists," that is, programs that determined reference based on heuristics of plausibility for the objects most recently mentioned in the previous sentence.

The next generation of natural-language systems, built toward the end of the 1970s, concentrated either on the use of domain knowledge in discourse or on methods of incorporating discourse phenomena other than pronoun understanding. Most systems dramatically limited the domain of the discourse to one small class of actions or events and focused on interactions that conformed to a narrow range of discourse behavior (for example, SAM (Cullingford 1977, 1981) and GUS (Bobrow et al. 1977)). However, Lehnert (1977) experimented with a system to capture the nonliteral meaning of questions posed to a story-understanding program. Grosz and colleagues (Grosz 1977, Walker 1978) developed the Task Dialogue Understanding System, in which distinctions were made among domain knowledge (represented in a task model), discourse information (represented in the global focus and associated algorithms for interpreting references), and intention recognition (treatment of questions about and updates to the task model).

This system was the first to consider the interaction between the (linguistic) structure of a discourse and the interpretation of expressions in the discourse. It did not distinguish with sufficient clarity, however, the different roles of a task model in discourse processing.

Early approaches to discourse processing taught researchers that discourse processing could not be viewed as an agglomeration of procedures for reference understanding, intention recognition, and manipulation of domain knowledge. It requires a means of distinguishing among these processes and providing mechanisms for each as well for their interaction.

11.2 Discourse Structure

The utterances a discourse comprises are not random sequences, but rather have structure much in the way that the words in an individual sentence have syntactic structure. Analysis of a variety of types of discourse has established that discourses divide into discourse segments, and that these segments may bear different kinds of relations to one another. Among the types of discourse that have been analyzed are task-oriented dialogues (see, for example, Grosz 1978a, Mann et al. 1975, Sidner 1982), descriptions of complex objects (Linde 1979), narratives (Polanyi 1985, Schiffrin 1982), arguments (both informal (Reichman-Adar 1984) and formal (Cohen 1984)), negotiations (Linde and Goguen 1978), and explanations (Reichman-Adar 1984).

An understanding of discourse structure is important both for theories of discourse meaning and for language processing. Theories of discourse meaning depend in part on a specification of the basic units of a discourse and the relations that can hold among them. Discourse processing requires an ability to determine to which portions of a discourse an individual utterance relates. Thus the role of discourse structure in discourse processing derives both from its role in delimiting units of discourse meaning and from constraining the units of discourse considered to be relevant to the interpretation of any individual utterance.

Furthermore an account of the interpretation of a variety of linguistic expressions depends on an explanation of the role of some of these expressions in determining discourse structure as well as on an understanding of the effect of discourse structure in constraining the interpretation of others. Section 11.3 discusses the use of cue phrases and intonation to mark discourse structure. Section 11.4 describes research on the interaction between discourse structure and the interpretation of anaphoric expressions.

Although there is general agreement on the fact that discourses are composed of segments, and on the interaction between discourse structure and the interpretation of various classes of linguistic expressions, theories of discourse structure postulate different types of information as central to the computation of discourse structure. Furthermore theories differ in the kinds of relations between segments considered important. Early work on text understanding (vanDijk 1972, Rumelhart 1975) proposed text (or story) grammars analogous to sentence grammars;[1] the earliest work on dialogue (Grosz 1974) argued that task-oriented dialogues had a structure that depended on the structure of the task being done. Subsequent work has taken one of the following approaches: (1) adaptions of the notion of grammar, (2) specification of a small set of rhetorical or textual relations as the basis of discourse structure, (3) investigation of domain-specific or commonsense knowledge as the source of discourse structure, or (4) examination of intentions broadly construed (a generalization of task structure) and relations among them as the foundation of discourse structure.

Segmentation and Structure of Utterances

Research on the possible relationships among the utterances that a discourse comprises has addressed such questions as which utterances group together into a segment, and what relations may hold between these segments (Polanyi 1985, Hobbs 1979, Linde 1979, Mann and Thompson 1986). In previous work (Grosz and Sidner 1986) we have referred to this structuring of the discourse into groups of utterances as the linguistic structure of the discourse. Our terminology reflects the focus placed on the structure of the language behavior itself. In theories that primarily consider linguistic structure, discourse structure is typically represented by a tree with clause-level utterances or groups of them as the nodes.[2]

Theories differ with respect to the information on which they base decisions about the embedding relationships encoded in the tree, and with respect to the kinds of specific embedding relationships they consider possible. Although some differences stem from the types of discourse treated (for example, narratives as opposed to arguments), others reflect disparate views of the goals of discourse processing.

Linde (1979) and Polanyi (1986) propose models derived in part from the sociolinguistic tradition, in which explanations are expressed solely in terms of surface behavior. They are thus concerned with providing an explanation of discourse structure in terms of the surface connections that obtain among utterances. Polanyi claims that the hierarchical structure of discourse "emerges from the structural and semantic relationships obtaining among the linguistic units which speakers use to build up their discourses" (Polanyi 1986, p. 4). In her model a discourse tree is constructed using a set of discourse grammars. The nodes of the tree are clauses (in some cases with their associated semantics). That is, the discourse structure is taken to be a tree structure with linguistic elements as the nodes.

Research in this paradigm typically takes full interactions (for example, extended personal conversations, doctor-patient interactions) as the largest unit of discourse, and clauses, perhaps grouped by coordinating or subordinating relations (for example, lists, expansions), as the smallest unit. Intermediate units of explanation may be socially motivated units (for example, service encounters, medical examinations), or linguistically motivated ones (for example, stories). Frequently labels are applied at these levels (for example, question/answer) that indicate the intentions of the discourse participants, but the theories avoid any reference to intention in their explanations.

Rhetorical Relations

Reichman-Adar (1984) studied informal arguments; the structure she proposed is based on a grammar for arguments that is analogous to story or text gammars (vanDijk 1972, Rumelhart 1975). For example, she claims an argument consists of a claim followed by support, and that support decomposes into an authority citation or a narrative or a proposition. Although the relations she posits between utterances and discourse segments combine semantic, pragmatic, and intentional properties, Reichman's analysis of processing is focused at the linguistic level.

Other research that builds a tree structure with linguistic nodes has focused on the relationships that can hold among meanings of utterances (that is, on semantic relationships). Cohen (1987) has analyzed a variety of arguments and investigated the problem of inferring evidence relations among the propositions expressed in them. She presents a processing model for arguments that uses a combination of cue-phrase information and (an initial set of) methods for inferring evidence relations to determine argument structure.

Much of the research described so far, as well as work by Hobbs (1979), Lehnert (1981), Mann and Thompson (1986), and McKeown (1985), offers an underlying stratum of rhetorical relations as the basis for deriving relationships among utterances and segments. Each such relation requires processing of domain information to determine how a relation could be recognized or produced in discourse processing. For example, Hobbs (1979) defines a set of coherence relations (for example, parallel, enablement, contrast; a complete list may be found in Hobbs 1983) that hold between discourse segments. Determination of the relation that holds between two successive utterances depends in inferences drawn on the basis of domain facts. Lehnert's relations derive from an underlying model of mental states and events. A set of primitive configurations, which she dubs *plot units*, describes the permissible transitions between states or events and thereby directly encodes domain information. These plot units then serve as the relations among utterances of segments. In McKeown's work on generation of discourses, a collection of rhetorical predicates (including, for example, attribution, identification, and comparison) are organized into schemas that abstractly define the set of acceptable discourse types; one or more individual utterances are produced as instantiations of the predicates in a schema.

Intentions and Attention in Discourse Structure

Levy (1979) and Grosz and Sidner (1986) have argued that plan-based or intentional relations are the root of discourse structure and postulate only an embedding relationship in the linguistic structure. The embedding relationships between segments depend in part on certain linguistic features of the segment (such as intonation or cue phrases), and in part on the intentions conveyed by the utterances of the segment. The segment-level intention is not a simple function of the utterance-level intentions, however, but a complex function of utterances, domain facts, utterance-level intentions, and inferences about these.

We argue further (Grosz and Sidner 1986) that discourse structure is a composite of three interrelated structures; in addition to linguistic structure there are a structure of intentions and an attentional state. The intentional structure comprises discourse-segment purposes and relationships between them. Discourse-segment purposes are intentions of the discourse participants that lead in part to the discourse and are, like Gricean utterance-level intentions (1957, 1968), intended to be recognized. Unlike the rhetorical relations discussed previously, discourse-segment purposes are not drawn from any special set of intentions; almost anything that can be intended can be a discourse-segment purpose. However, two relations among intentions are defined as common in many discourses: *domination* and *satisfaction precedence*. These relations represent, respectively, the fact that satisfaction of one intention contributes to the satisfaction of a second, and the fact that one intention needs to be satisfied before another. Recognition of these two relations plays the same role in our theory that recognition of rhetorical relations plays in the theories described previously. Determination of discourse-segment purposes depends in part on recognition of these relationships, in part on domain knowledge, and in part on other features of the discourse context.

Attentional state reflects the focus of attention of discourse participants as the discourse progresses. It is modeled as a (pushdown) stack of focus spaces, one for each segment of the discourse. The focus spaces contain representations of the discourse-segment purpose and the entities referred to in the segment. The stack grows when segments are introduced into a discourse and shrinks as the intentions of the segments are satisfied. Following standard terminology, we say that a new space is *pushed* onto the stack when a new segment is introduced, and that a space is *popped* when its purpose is satisfied. The focus space model of attentional state constrains processing as the discourse unfolds. Those entities and purposes represented in the stack are most salient to the discourse. Constraints may use this fact to stipulate when various linguistic expressions can be used and to help determine when a given discourse-segment purpose may dominate or satisfaction-precede another.

11.3 Linguistic Indicators of Discourse Structure

The moral of this chapter should by now be clear: discourses exhibit structure. Every discourse can be intuitively partitioned into segments, and discourse segments can embed and be embedded in other discourse segments, resulting in a hierarchical structure. Recognizing the structure of a discourse is an essential part of comprehending it.

But how is it that the structure of a discourse is recognized? As we noted in the preceding section, the answer to that question is still widely debated—not surprisingly because it is the central motivating question in research on discourse. Virtually all researchers agree, however, that speakers have available certain powerful linguistic devices for assisting hearers in recognizing the structure of an ongoing discourse. These devices, such as cue phrases, intonational patterns, and gesture, can indicate both where a discourse segment begins and ends, and how it is related to other segments.

Cue Phrases

The most widely studied class of linguistic indicators of discourse structure are the *cue phrases* (sometimes called *clue words*, *discourse markers*, or *discourse particles*). Cue phrases are expressions such as *now*, *in the first place*, and *by the way*, which do not make a direct semantic contribution to an utterance, but instead convey information about the structure of the discourse containing the utterance. Of course because different theories make different claims about what the structure of a discourse amounts to, there are also variations among theories about precisely what structural information is signalled by a cue phrase.

Researchers who view the structure of a discourse as solely a linguistic structure maintain that cue phrases provide a direct indication of how distinct segments are related in the hierarchy (Cohen 1984). Those who argue that the structure of a discourse consists of an underlying stratum of rhetorical relations linking discourse segments view cue phrases as suggesting the particular relation that holds between two discourse segments. For example, Hobbs (1985, p. 31) claims that "'That is' or 'i.e.' suggests *elaboration*, 'similarly' suggests *parallel*, 'for example' suggests *exemplification*, and 'but' suggests *contrast* or *violated expectation*."[3] Similar claims have been made in Rhetorical Structure Theory (Mann and Thompson 1986) and in accounts of discourse parsing (Polanyi and Scha 1984, Reichman-Adar 1984).

A contrasting claim about the role of cue phrases has been made by Grosz and Sidner (1986). As already noted, we have argued that the structure of a discourse actually comprises three interrelated structures. In this view cue phrases can provide information about one or more components of the overall (three-part) structure (Grosz and Sidner 1986, pp. 196–199). For instance, certain cue phrases, such as *that reminds me* and *anyway*, indicate changes to the attentional state—the former indicating a push to a new focus space and the latter a pop to a previously established space. Notice that although the change to attentional state signalled by *that reminds me* is accompanied by the addition of new components to the intentional structure, the cue phrase itself does not specify what this change is. With *anyway* there is no change to the intentional structure at all. In contrast other cue phrases directly indicate changes to the intentional structure. The expression *incidentally* provides information that the speaker is embarking on a digression, and consequently the intentional structure is to be extended with a new intentional hierarchy distinct from the existing one. The expression *for example* indicates that the intention underlying the upcoming discourse segment—the one whose starting boundary is marked with the cue phrase—is dominated in the intentional structure by the intention underlying the preceding discourse segment.

Most researchers, regardless of their stance on what discourse structure actually constitutes, believe that cue phrases are in general neither necessary nor sufficient for determining discourse structure. There are many cases in which it is quite possible to determine the structure of a discourse, or a portion of one, that lacks any cue phrases. Likewise there are many discourses, or portions thereof, containing cue phrases that only suggest the underlying structure or, put another way, provide constraints on the range of possible structures. Ultimately the structure of a discourse depends on the information conveyed by the utterances it comprises and the way in which that information is interconnected. Cue phrases simplify the task of determining those connections. It has been shown that the process of determination of intersentential semantic relationships (Cohen 1984) and plan recognition (Litman and Allen 1988) can be constrained by taking into account cue phrases.

Intonation and Gesture

Intonation is another powerful tool for indicating discourse structure. Studies of spontaneously produced discourses have shown that changes in pause length (Chafe 1979, 1980) and speech rate (Butterworth 1975) correlate with discourse-segment boundaries. Hirschberg, Pierrehumbert, Litman, and Ward have provided evidence of a close correspondence between particular intonational features and specific components of the three-part discourse structure proposed by Grosz and Sidner (Hirschberg et al. 1987, Hirschberg and Pierrehumbert 1986). *Pitch range* can indicate discourse segment boundaries (Silverman (1987) came to the same conclusion); *accent* can provide information about the attentional state, and *tune* can mark intentional structure. In addition phrasing and accent can help distinguish the use of an expression such as *now* as a cue phrase from a use in which the expression contributes directly to sentence meaning (Hirschberg and Litman 1987).

Gesture is another device that is useful, at least in face-to-face conversation, for signalling discourse structure. Studies have shown that gesture tends to be coincident with discourse-segment boundaries (Kendon 1972, Marslen-Wilson et al. 1982) and can provide information about focus of attention and about intentional structure (McNeill 1979, McNeill and Levy 1982).

11.4 Phrase-Level Phenomena

Discourse context affects the interpretation of individual phrases and clauses within a single utterance. The meaning of pronouns and definite descriptions is quite obviously influenced by the context in which they are used; among the most widely studied problems in discourse processing are those concerning the influences of context on the processes of generating and interpreting such phrases. The related problem of appropriately representing quantified noun phrases has also been addressed. We discuss a range of approaches to these problems of noun-phrase reference in the following subsection. Reference may be made to events as well as objects, raising problems discussed in the second subsection. Other phrase-level problems that have received treatment from the perspective of discourse processing are modifier attachment and discourse ellipsis.

Of the three components of discourse structure, attentional state has the greatest effect on problems of phrase-level interpretation. Research on referring expressions has most directly used attentional models, whereas approaches to modifier attachment and discourse ellipsis have only begun to make use of them. Although there is as yet no full account of the range of discourse constraints (nor of perceptual ones) relevant to any of these phrase-level problems, researchers have identified and experimented with a range of representations and algorithms central to them.

Pronouns and Definite Descriptions

The primary concern of computational theories and models of the use of referring expressions has been with determining the entity (or entities) to which a pronoun or definite description refers.[4] Both the problem of specifying the range of possible referents an initial description makes available and the problem of choosing among possibilities have been investigated.

Two approaches have been taken to the problem of identifying the referent of an anaphoric[5] phrase. In one approach (Hobbs 1979) referent identification is subsumed by more general processes of inference. In the other approach questions of how referring expressions interact with attentional state are considered primary (Grosz 1977, Sidner 1981, Reichman-Adar 1984, Grosz and Sidner 1986). Research following the first approach considers only problems of interpreting referring expressions, whereas research following the second is concerned as well with specifying constraints on generation of appropriate referring expressions.

For the second approach the concepts of focus and the process of focusing (Grosz 1977, Grosz 1978b, Grosz 1981, Grosz and Sidner 1986) have played central roles in the treatment of definite descriptions, used both for first reference and as anaphoric noun phrases. In this work focusing is defined as the movement of the focus of attention of the discourse participants as the discourse progresses. Two levels of focusing, global and local, have been identified (Grosz 1977, Grosz and Sidner 1986). Global focusing is modeled by a stack of focus spaces; it affects the use and interpretation of definite descriptions. Focusing at the local level is modeled with centers and centering. The center of a given segment is an element of the attentional state, and at the beginning of each new segment a new center is introduced. Centering affects the use and interpretation of pronouns.

Each individual space on the global focus stack contains representations of the entities that the participants focus on during a corresponding discourse segment, as well as the discourse-segment purpose. The entities currently in focus (that is, in some space on the focus stack) are the primary candidates for referents of definite descriptions; they are also the source of implicitly focused entities (that is, a phrase may refer to an item related to something in a current focus space; for example, *the cover* may be used to refer to the cover of a book when that book is in focus). The set of entities in the global focus also provides constraints on the content of subsequent definite descriptions; for example, a speaker must include sufficient descriptors to distinguish the entity to which he wishes to refer from other entities in focus.

Grosz (1977) devised a set of focusing mechanisms for the interpretation of definite noun phrases in a system that participated in a dialogue about a task.[6] The mechanisms brought entities into focus as the discourse moved to a subtask of the overall task and moved them from global focus when the subtask was completed. Position in global focus was relative to the most immediate subtask; other incomplete subtasks and their associated entities were in global focus, but were less salient than those associated with the most immediate subtask. Procedures for the interpretation of anaphoric definite noun phrases and of implicitly focused first references chose items from among those in global focus. The focusing techniques could predict the anaphoric referent of a definite noun phrase such as *the screw* when the screw in the wheelpuller had been brought into focus. They also limited the search necessary to find the representation of a previously mentioned entity—those entities not in focus were unavailable, and those in focus and associated with the most immediate subtask were searched before less salient ones. For example, *the screw* would be understood as a referent to the screw of the wheelpuller only when the wheelpuller was more salient than some other screw (for example, the drill screw) in the global focus; otherwise the drill screw would be chosen as the referent.

Appelt and Kronfeld (1987) have used the mechanisms for focusing in the generation of first-use referring expressions. In addition Kronfeld (1986) has provided the concepts of functionally relevant descriptions (that is, those whose content is needed to distinguish the referent) and conversationally relevant descriptions (that is, those whose content provides information about the referent's relation to other aspects of the conversation). These concepts redefine Donnellan's (1966) classical distinction between referential and attributive descriptions in terms of the role a phrase plays as a referring tool and as a contributor to the content of discourse.

Immediate focus (Sidner 1979) guides the interpretation and generation of third-person pronouns as well as the anaphoric uses of *this* and *that*. Immediate focusing operates within individual discourse segments; it tracks the entity most relevant at any utterance within the segment, based on features of the preceding context, especially properties of the preceding utterance. Sidner (1979, 1981, 1983) developed a set of algorithms for predicting the choice of local foci in discourse and their movement as the discourse progressed. Her algorithms used the concepts of a discourse focus and an actor focus, that is, entities mentioned in an utterance that were locally in focus because of the syntactic structure and thematic relations of the utterance. A set of rules, using the immediate foci and a set of potential new foci, predicted the intended interpretation of pronouns in subsequent utterances. McKeown (1985) and McDonald (1983) adapted the focusing algorithms to generate pronouns in text.

The centering theory of Grosz, Joshi, and Weinstein (Joshi and Weinstein 1981, Grosz et al. 1983) replaced the notion of immediate foci with that of centers: a backward-looking center (Cb) that corresponds roughly to the discourse focus, and several forward-looking centers (Cf) corresponding to the potential new foci. As in Sidner's theory the centers shift according to the behavior of referring expressions in each utterance of discourse. The theory differs in having no correlate of actor focus and accounting for multiple uses of the same pronoun by allowing entities other than the Cb to be pronominalized so long as the Cb is. The centering theory simplifies Sidner's account of local focusing by eliminating the need for both actor and discourse foci and by providing an explanation of particular instances of pronoun interpretation that are problematic in the immediate focus theory.

Although the focus of attention is crucial in the understanding of anaphora, the form and content of the phrases and utterances that first evoke the entities to which anaphora refer vitally contribute to their understanding as well. Webber (1980, 1983) introduced the notion of phrases "evoking discourse entities." By this she meant that phrases brought into the discourse, or *naturally evoked*, a well-structured collection of representations of the (actual) entities referred to. Discourse entities were available for interpretation of definite anaphora and verb phrase ellipsis. She defined the concept of *discourse entity invoking description* (ID) to formalize her notion and posed a set of specific representations and rules for creating these IDs by successive application of the rules.

Webber's rules operated on complex semantic representations, which included embedded quantifiers and quantification over sets of individuals. The rules produced IDs for each possible interpretation of the noun phrases and verb phrases of an utterance, but Webber did not explore how to determine computationally which interpretation was intended on the basis of subsequent anaphoric descriptions. Sidner (1983), however, sketched an account of how focusing might provide the necessary information. In subsequent work Kamp (1981) and Heim (1982) have proposed similar alternative formalisms; Guenthner and colleagues (1986) report on a system using Kamp's formalism. Two system-building efforts have investigated methods for computing the representations of complex noun phrases as part of the process of determining (a representation of) the meaning of a sentence (Dahl 1987, Pollack and Pereira 1988).

Among the referring expressions least investigated in computational discourse research is the use of deixis.[7] Fillmore (1975) subcategorized four types of deixis: spatial, temporal, social, and discourse. Discourse structure may constrain the latter two types as is exemplified by the phrases *the former*, used to refer to the first of two items in a list (an instance of discourse deixis), and *we* (an instance of social deixis). Sidner (1979) explored instances of discourse deixis, but a general treatment remains to be provided.

Reference to Events

Discourse research has considered two problems caused by references to events: identifying to which event an anaphoric reference refers, and determining the relative temporal ordering of the events described in the discourse.

Anaphoric references typically use pronominal forms either bare (*it, this, that*) as in example 1, or with the proverb do (*do so, do it, do that*), as in example 2. The *do* form and the bare pronoun appear to behave differently. Although some research treats *do* forms (Sidner 1979, Sidner 1981, Robinson 1981), a comprehensive account of all directly anaphoric references to events remains to be given.

(1) John runs every day of the week.
That's his main form of exercise.
that = John's running every day.

structured collec-
ed to. Discourse
ıaphora and verb
ntity invoking de-
of specific repre-
ıve application of

entations, which
r sets of individ-
rpretation of the
Webber did not
terpretation was
criptions. Sidner
ıg might provide
(1981) and Heim
Guenthner and
formalism. Two
r computing the
f the process of
tence (Dahl 1987,

n computational
) subcategorized
ourse. Discourse
emplified by the
ems in a list (an
of social deixis).
t a general treat-

ed by references
ence refers, and
nts described in

s either bare (*it*,
o, do it, do that),
ppear to behave
ner 1979, Sidner
rectly anaphoric

(2) John runs every day of the week.
He *does it* to stay healthy.
does it = act of running every day of the week.

Relative temporal ordering is affected by verb tense and aspect, as well as by adverbial modifiers; it also depends on attentional state changes. In the following discourse (from Webber 1987) the event described in utterance 1, E1, occurs at some point in the past, utterance 2 describes an event E2 that occurs before E1, the event E3 of utterance 3 occurs after E2 and before E1 and is situated in the location described in utterance 2, and the event E4 in utterance 4 occurs after E1.

(1) John went over to Mary's house.
(2) On the way, he stopped by the flower shop for some roses.
(3) He picked out 5 red ones and 3 white ones.
(4) Unfortunately they failed to cheer her up.

Two major proposals for a theory of event reference, that of Webber (1987) and that of Hinrichs (1986) and Partee (1984), each provide an account of the relative order of events. Webber also provides an account for the recognition of the ordering of events based on temporal focus and focus movement, concepts analogous to the immediate-focus treatment (Sidner 1983, Grosz et al. 1983) for definite anaphora.

Modification Relations in Noun Phrases

Modification relations among parts of a complex noun phrase pose two problems for the interpretation or generation of reference. The first concerns the choice of a single underlying (intended) description of the structural and functional relations among the entities described by the complex noun phrase; this single description must be chosen from a number of possible such descriptions. For example, the intended description of the phrase *the pump dispenser* rests on uncovering the functional relationship between the two objects: pump and dispenser. The second problem for interpreting modification relations is identifying the entity referred to (using of course some technique for determining possible relationships). Identification problems can be quite difficult for complex noun phrases with prepositional modifiers used in a context that includes several possible referents for each simple part of the noun phrase. The problem is illustrated by the phrase *the cat in the hat* used in a context in which there are two cats, two hats, but only one cat in a hat (one of *the* hats). Although the referent is clear here, no simple search for either a unique cat or a unique hat in context can serve as the basis of the search for the referent.

Research on nominal compounds has explored frameworks for using linguistic information to predict the possible structure of complex noun phrases (compare Isabelle 1984). Computational approaches to these problems (Finin 1980, Hobbs and Martin 1987) have addressed the joint problems of appropriately representing the domain knowledge needed to determine the intended interpretation and defining processes for searching this knowledge. Although attentional state obviously affects the identification of appropriate interpretation, this issue has not yet been explored computationally.

By contrast recent approaches (Mellish 1982, Haddock 1987) to identifying the referent (which have considered prepositional phrase attachment rather than nominal compounds) have concentrated on the concept of incremental reference evaluation using constraint-satisfaction techniques. These techniques assume that the head noun, noun phrase embedded in the prepositional phrase, and the relation specified by the preposition form a set of constraints on the choice of candidate referents for the whole phrase; these constraints may be applied incrementally as the phrase is constructed to search for the intended referent. These approaches assume some representation of attentional state that delineates a small set of possible referents and serves to limit the search undertaken by incremental constraint satisfaction. A theory that incorporates aspects of incremental evaluation with the search methods proposed for nominal compounds previously discussed has only recently been explored (Pollack and Pereira 1988).

Discourse Ellipsis

Discourse ellipsis is the omission (or *elision*) from an utterance of a syntactically required phrase when the content needed to determine the interpretation of the elided utterance can be recovered from a previous utterance. An adequate treatment of discourse ellipsis requires that a discourse (rather than a sentence) be considered the primary unit of communication. Two kinds of discourse ellipsis have been investigated. In the first type the elided material can be recovered directly from a representation of the meaning of the previous utterance. In the second the intentional structure provides the source of the elided material; in this case the elided content may not appear directly in any prior utterance.

Approaches to the first type (Hendrix et al. 1978, Webber 1983) have considered both noun-phrase and verb-phrase ellipsis. They have viewed ellipsis as similar to discourse anaphora because an ellipsis is also interpreted by locating previously mentioned phrases and their discourse representations as the source of the material to be reconstructed. Unlike discourse anaphora elided phrases need not refer to the same act or entity as the previously mentioned phrase, as shown in the following example:

(1) Fred kissed his mother.
  John did 0 too.
  0 = kissed his own mother or, alternatively, kissed Fred's mother.

Rather the discourse representation of the previously mentioned phrase provides material to construct the representation for the elided phrase.[8]

Approaches to the second type have used the intentional structure as the source of elided material. Elliptical utterances are presumed to be intended to contribute to the overall purpose of a discourse or discourse segment. As long as the discourse environment contains sufficient information, a full utterance need not be used; a phrase or portion of a sentence will contribute just as well. Elliptical fragments such as (1) (from Allen 1979) and (2) (from Carberry 1985) below can be seen to contribute directly to the discourse purpose; no reconstruction is needed of a full sentence in which the fragment is embedded.

(1) (said to information booth clerk at a railroad station) The train to Windsor?

(2) I want to get a degree. CS major. No courses at night.

Allen (1979) demonstrated that such phrases can be directly associated with a speech act which in turn identifies a portion of a speaker's plan to act in a domain of interest. The plan recognition models of Allen and Perrault (1980), Carberry (1985), and Litman (1985) underscored Allen's position and related the fragment to speech acts and the plans of speakers (that is, plans in a domain of action). By viewing action as a root of the explanation for why discourse occurs, these researchers illustrated that some forms of ellipsis could be viewed in terms of the intentional structure of discourse rather than in terms of a set of relations to other utterances.

11.5 Plan Recognition

Contemporary accounts of plan recognition in discourse understanding derive largely from two major theories of the philosopher Grice: his theory of *nonnatural meaning*[9] (Grice 1957, 1968) and his theory of *implicature* (Grice 1975, 1978). In Grice's theory of nonnatural meaning what speakers mean when they use language depends crucially on what they intend. The theory of implicature rests on the observation that much of what is intentionally conveyed during language use is not explicitly expressed; the theory itself shows how speakers can convey more than they explicitly express in their utterances. In both theories speakers are seen as producing utterances with the intention that their hearers recognize the intentions underlying those utterances. The collection of intentions underlying any utterance will thus include, but will not be limited to, an intention to have at least some portion of the collection itself recognized.

Understanding language thus requires determining what intentions speakers have: figuring out what plans they are pursuing in part by making an utterance. Plan recognition in conversation is a feasible task precisely because the speaker intends the hearer to perform it. The speaker cannot achieve her intended effect unless the hearer recognizes the speaker's plan; hence the speaker will include in the utterance what she believes to be sufficient information to make plan recognition possible for the hearer. This characteristic of plan recognition in conversation distinguishes it from the problem of determining the plans of an agent merely by observing the agent's actions without interacting with her. The latter problem, dubbed "keyhole recognition" because of its similarity to the problem of watching an agent through a keyhole and inferring the next action, is in general much more difficult.[10]

In constructing detailed models of the process of plan recognition in discourse, researchers have drawn not only on the work of Grice but also on the related insight of Austin (1962) and of Searle (1969, 1975) that language is used not merely to "say" things but also to "do" things. In other words communicative behavior should be viewed primarily as purposeful action. It follows that to a large extent the same tools that are used to analyze nonlinguistic action can be applied to the analysis of linguistic action—an idea originally developed largely in the work of Cohen and Perrault (1979), Allen (1983), and Perrault and Allen (1980). In particular, AI models of plan recognition in discourse have made use of techniques for representing and reasoning about action originally developed by those interested in the problem of automatic plan formation.[11] Before turning to a discussion of these models, however, it is worthwhile noting the range of discourse phenomena for which plan recognition is important.

Consider a person who walks up to the attendant at a train-station information booth and says, "Do you know when the next train to Detroit leaves?" with the intention that the attendant tell her the departure time of the next train to Detroit.[12] The speaker's communicative plan is to make a request—to be told when the next train to Detroit leaves—and to make this request by uttering her query. If the attendant recognizes this plan, she can cooperatively respond by performing the requested action (assuming of course that she is able to do so and has no reason to prefer not doing so). That is, she may reply by saying "At 12:30" rather than by saying simply "Yes."

Usually communicative plans are performed as part of larger plans, and it is often necessary for a hearer to infer these larger plans as well. In the current example if the attendant recognizes that the speaker's communicative plan is likely to be part of a plan to go to Detroit, she may be able to provide the speaker with additional information that will facilitate her goal. She may, for example, tell her what gate the train departs from or she may critique her plan, telling her that although the next train to Detroit leaves in 15 minutes, it is a local train, and the express that leaves in 45 minutes will get her to her destination sooner.

Plan recognition is thus important to responding appropriately to a speaker's utterances in interactive discourse.[13] It is also important to understanding the coherence of multiclause utterances. Consider a different person, who says to the information-booth attendant, "I'm going to Detroit. Where's gate 7?" It is by recognizing what plan the speaker may have that the attendant can determine why the speaker's utterance is coherent; indeed plan recognition enables the attendant to understand the utterance fully. It is essential to be able to understand the coherence of multiclause utterances in noninteractive as well as in interactive discourse. In fact several studies of plan recognition have focused on story understanding, an essentially noninteractive form of discourse (Bruce 1981, Wilensky 1983).

The Basic Technique

As noted in the previous subsection, the idea that language use should be viewed as purposeful action led to models of discourse that make use of techniques already developed for reasoning about (nonlinguistic) actions. In particular the representation of plans and actions used in most plan-recognition systems is a direct outgrowth of the representation first developed in the STRIPS system (Fikes and Nilsson 1971) and later expanded in the NOAH system (Sacerdoti 1977). Because the STRIPS representation has certain limitations in expressiveness, there has been since the mid-1980s a resurgence in interest in the plan formation process, and a number of alternative representation schemes have been devised.[14] However, the basic plan recognition techniques have been developed for STRIPSlike representations of actions.

In these representations states of the world are modeled as sets of propositions, and actions are modeled with *operators* that map one state of the world to another. Each operator has a *header*, which names the action being represented, a *precondition list*, which describes what propositions must be true for the action to be performed, and an *effect list*, which describes what becomes true as a result of performing the action. A *plan* is analyzed as a sequence of actions and states of the world. Specifically $[\alpha_1, S_1, \ldots, S_{n-1}, \alpha_n]$ is a plan to transform a state S_0 to a state S_n provided that

- all the preconditions of α_1 are true in S_0;
- all the effects of α_n are true in S_n;
- in each intermediate state S_i all the effects of α_i and all of the preconditions of α_{i+1} are true.

Plan formation is the process of finding a plan that transforms the current state of the world into a state in which some set of goals are true. In the simplest case plan formation can be cast as a *graph-search* process, in which the nodes of the graph represent actions or states of the world, and legal connections between nodes are determined by the relations between states and actions specified in the set of operators (see chapter 7 of Nilson 1980).

This simple view can be refined by allowing each operator to include a *body* as well as a precondition and an effect list. The body of an operator may be either a list of subactions whose performance constitutes performance of the action named in the header or a list of subgoals whose achievement constitutes performance of the header action. When operators include bodies, the plan formation process may be *hierarchical:* a sequence of actions and intermediate states that are relatively abstract can first be sought and then elaborated into more and more detailed plans. Hierarchical plan formation is often much more efficient than linear plan formation (Sacerdoti 1977, Stefik 1981, Tate 1984, Wilkins 1984).[15]

In plan recognition one agent observes another agent perform some action or sequence of actions and then attempts to determine what plan the latter agent is carrying out. In discourse the observed actions are the utterances of a speaker;[16] thus we can let *S* (for speaker) denote the actor whose plan is being inferred, and *H* (for hearer) denote the agent who is inferring *S*'s plan. Suppose *H* observes *S* perform some action α in a state S_1 in which all of the preconditions of α are true. Then *H* may conclude that *S*'s purpose in performing α is to bring about one or more of α's effects. Indeed *H* may be able to reason further. If some effect of α, say *e*, is a precondition of some other action β, then *H* may conclude that *S* intends to perform β. The process may again be iterated: *H* may conclude that *S* intends to achieve some effect of β, and so on. Typically this reasoning process terminates when *H* determines that *S* intends to achieve some typical domain goal.

One of the earliest and most influential formulations of plan recognition was by Allen (1979, 1983). Allen provided a set of plan-recognition rules, along with a heuristic strategy for controlling their application. A typical plan-recognition rule is the Action-Effect Rule:

$$\mathrm{Bel}(H,\ \mathrm{Int}(S,\alpha)) \rightarrow \mathrm{Bel}(H,\ \mathrm{Int}(S,e)) \qquad \text{if } e \text{ is an effect of } \alpha.$$

This rule can be glossed as "if the hearer (or more generally the inferring agent) believes that the speaker (or more generally the actor) intends to perform some action α, then the hearer may decide that the speaker intends to make some proposition *e* be true, if *e* is an effect of performing α (in the state of the world in which α will be performed)."[17]

An assumption of intentionality is necessary for the plan-recognition process to begin. That is, given that *H* observes *S* perform some action α, the Action-Effect Rule only applies if *H* concludes that *S* did α intentionally. An assumption of intentionality is not problematic in cases in which the plan recognition is being performed in conversation: if *S* says something to *H*, it is quite reasonable for *H* to suppose that *S*'s utterance was intentional.

Other rules link propositions to actions of which they are preconditions and link subactions to actions containing them. These rules, plus the rule discussed, are used for so-called *forward-chaining:* reasoning from an observed action to effects of that action, and from those effects to further actions that are thereby made possible, and so on. Plan-recognition systems typically also include rules for *backward-chaining*. An inferring agent can also reason about what goals an actor is likely to have and backward from that to reason about what actions would achieve those goals, what the preconditions of those actions are, and so forth. Allen's Effect-Action Rule is one example of a plan inference rule to be used in backward-chaining:

$$\text{Bel}(H, \text{Int}(S,e)) \rightarrow \text{Bel}(H, \text{Int}(S,\alpha)) \qquad \text{if } e \text{ is an effect of } \alpha.$$

In addition to the simple forward- and backward-chaining rules, many plan-recognition systems have additional rules that apply only to information-seeking actions. For example, many systems relate the action of "finding out whether *p*," for some proposition *p*, to the action of "achieving *p*." These rules, along with rules for handling so-called nested plan inference, however, can be viewed as special cases of the simpler rules (Kautz 1985).

It is important to note that the plan-inference rules are not to be construed as logical entailments but rather as descriptions of "likely" but nondeductive inferences that an agent may make in performing plan recognition. Kautz (1988) presents a precise formalization of the process of using these likely inference principles to perform plan recognition, providing model, proof, and algorithmic theories that are founded on McCarthy's theory of circumscription (McCarthy 1980, 1984).

The application of plan-inference rules is typically controlled by a set of heuristics, which are designed both to make the recognition process more efficient and to ensure that more plausible plans are found before less plausible ones. Some of the control heuristics are based on commonsense notions about the nature of plans and apply equally well to the process of inferring the plans of one's conversational partner and inferring the plans of another agent whom one is merely observing. For example, one such heuristic biases a plan-recognition system against consideration of candidate plans that contain actions whose effects are true at the time that the action is to be performed. This heuristic is reasonable whether or not the plan-recognition process is taking place in discourse: in general plans do not contain superfluous actions.

In contrast a number of important control heuristics are only justified for plan recognition in conversation because they are founded on the Gricean notion of *intended recognition:* that speakers intend for their hearers to recognize at least some subset of their intentions. This idea motivates plan-recognition heuristics such as the *forking heuristic*, which asserts that the likelihood that a candidate plan is the one that the speaker actually intends is inversely proportional to the number of alternatives into which it can be expanded (Allen 1983), and the *single-branch assumption*, which asserts that if the speaker believes that more than one plan might be inferred at some stage of the discourse, it is her responsibility to make known the one she intends—hearers need only infer to the point of a potential split (Sidner 1985). Such heuristics mediate against difficult inferences on the assumption that speakers in general attempt to produce utterances that facilitate relatively simple plan recognition.

As just one example of the application of the basic technique for plan recognition in discourse, consider a traveler who says, "When does the next train to Detroit leave?" Following Allen, we can encode this utterance action as

REQUEST(*S*, *H*, INFORM-REF(*S*, *H*, time1)).

That is, *S* has requested *H* to inform her of the reference of some constant time1, which happens to denote the departure time of the next train to Detroit.[18] By applying the Action-Effect Rule to the REQUEST operator

Header: REQUEST(Agent1, Agent2, Action)
Effect: WANT(Agent2, DO(Agent2, Action))

H can conclude that *S* may want *H* to want to inform *S* of the Detroit train's departure time. Further reasoning allows *H* ultimately to infer the entire plan shown in figure 11.1.[19]

The arcs of the plan graph have been labeled with forward-chaining plan-inference rules, but of course some portion of the graph might actually have been found using backward-chaining. A number of additional operators are also necessary for inferring this plan. One is the BOARD operator; construction of the others is left as an exercise for the reader.

Header: BOARD(Agent, Train, Station)
Precondition: AT(Agent, dept-loc(Train), dept-time(Train))
Effect: ON-BOARD(Agent, Train)

Extensions to the Basic Technique

The basic technique for plan recognition has several limitations. Perhaps most striking is the fact that it is restricted to cases in which the inferring agent makes a single observation. When applied to discourse, this means that the basic technique is useful for reasoning only about single-clause utterances. But one of the defining features of discourses is that they are extended in time. Speakers convey information using several clauses and in interactive discourse using several utterances (or "turns").

Incremental recognition techniques have been developed to extend the basic technique for plan recognition to deal with sequences of utterances (Carberry 1988, Sidner 1983, 1985). In incremental recognition the sys-

AT (*S*,detroit)

↑ action-effect

TRAVEL-TO (*S*,detroit,train1)

↑ precondition-action

ON-BOARD (*S*,train1)

↑ action-effect

BOARD (*S*,train1,detroit)

↑ precondition-action

AT (*S*,locl,time1)

↑ action-effect

GOTO (*S*,loc1,time1)

↑ precondition-action

KNOW-REF (*S*,time1)

↑ action-effect

INFORM-REF (*S*,*H*,time1)

↑ precondition-action

WANT (H,INFORM-REF(*S*,*H*,time1))

↑ action-effect

REQUEST (*S*,*H*,INFORM-REF(*S*,*H*,time1))

time1 = dept-time (train 1)

loc1 = dept-loc (train 1)

Figure 11.1 A plan to board a train

tem playing *H*'s role, begins by using the basic technique to infer as much as possible from *S*'s first (discourse-initial) utterance. It may not be possible yet to determine which of several plans *S* is pursuing, and it may not be possible to determine what *S*'s plan is to a level of detail sufficient to respond appropriately. The interim result thus may be a set of candidate partial plans. Once these have been computed, processing is suspended. Upon hearing each subsequent utterance, the system attempts to expand as many of the already constructed partial plans as possible. The expansion process again makes use of the basic technique: plan-inference rules are used to relate some node of an existing subgraph to the action performed by making the utterance and then to reason from that action to other actions. It may prove impossible to expand certain of the existing subgraphs: these are then eliminated from consideration.

Incremental plan recognition can be made more efficient by taking into account the speaker's focus of attention. As we noted in section 11.2, in task-oriented dialogues at least, the structure of the task influences the structure of the discourse. When a speaker is talking about a plan to perform some task, she is not likely to skip around in the presentation, talking first about some subtask, then about another, then returning to the first, and so on. Rather at each point in the discourse, some portion of the speaker's overall plan will be more salient than others, and the speaker's utterances are most likely to concern the salient portion. This idea can be used to constrain the plan-recognition process: *H* can prefer to expand those portions of candidate subplans that are in focus to those that are not (Carberry 1988, Litman and Allen 1989).[20]

It is often necessary in discourse processing to recognize several related plans that a speaker is executing. In particular speakers often interrupt their discussion of some domain plan to execute a discourse plan, for example, inviting clarification or correcting an earlier misconception, with the execution of domain plans. In the following dialogue, adapted from one in Litman and Allen 1989, the passenger engages in a plan to clarify the information he has already received in the midst of carrying out his plan to determine the departure time and location of a particular train:

Passenger: The train to Montreal?

Clerk: Gate seven.

Passenger: Where is it?

Clerk: Down this way. Second one on the left.

Passenger: And what time does it leave?

Clerk: Nine o'clock.

Additional operators that represent discourse plans can be introduced. Litman has proposed the use of a stack of partial plans in plan recognition: representations of discourse plans are to be stacked on top of the representations of the domain plans on which they depend (Litman and Allen 1989).

A further difficulty arises when the observed behavior is the result of two or more interacting domain plans. Wilensky (1983) has proposed handling such cases by providing the inferring agent with a set of *meta-plans* that operate on other domain plans to construct more complex plans: *resolve-conflicts* is one example of such a meta-plan.[21] In noninteractive text there is also a problem in reasoning about the interactions among the separate plans of multiple agents, for example, the characters in a story; this problem has been studied in detail by Bruce and colleagues (Bruce 1981).

Another kind of restriction inherent in the standard model of plan recognition has been pointed out by Pollack (1986, 1988): it assumes that the inferring agent (*H*) and the actor (*S*) have extremely similar beliefs about the domain of action. Consider again the Action-Effect plan-inference rule. It was glossed as "if the *H* believes that the *S* intends to perform some action α, then the *H* may decide that the *S* intends to make some proposition *e* be true, if *e* is an effect of performing α." Note that this rule leaves unstated precisely who it is—*H* or *S*—that believes that *e* is a precondition of α. If we take this to be a belief of *H*, it is not clear that *H* will infer *S*'s plan; on the other hand if we consider it to be a belief of *S*, it is unclear how *H* comes to have direct access to it. In practice there is only a single set of operators relating preconditions and actions in standard models of plan recognition; the belief in question is regarded as being both *H*'s and *S*'s.

In many situations an assumption that the relevant beliefs of *S* are identical to those of *H* results in failure not only of the plan-recognition process but also of the communicative process that plan recognition is meant to support. In particular it precludes the principled generation of appropriate responses to queries that arise from invalid plans. Pollack (1986, 1988) has proposed a model of plan recognition in conversation that distinguishes between the beliefs of *S* and those of *H*. The model rests on an analysis of plans as mental phenomena: "having a plan" is analyzed as having a particular configuration of beliefs and intentions. The plan-recognition process can take advantage of diverse techniques of belief ascription. Judgments that a plan is invalid are associated with particular discrepancies between the beliefs that *H* ascribes to *S*, when *H* believes that *S* has some plan, and *H*'s own beliefs.

Furthermore discourse participation may require an ability to represent the joint plans of multiple agents. Grosz and Sidner (1989) have shown that such plans cannot be defined solely in terms of the private plans of individual agents. They have proposed a representation that extends Pollack's definition of a plan to the plans of two or more agents as they collaborate on a task and communicate in discourse.

One other recent trend in plan recognition in discourse has been to relate more closely speech-act theories with fundamental theories of rational action. Cohen and Levesque (1989) have shown that illocutionary acts, such as requesting or promising, need not be treated as primitives in a model of plan recognition in discourse; instead the proper behavior of the discourse participants can be analyzed in terms of a model of rational principles of belief and intention adoption. Similarly Perrault (1989) has shown how an account of speech acts can be based on a simple theory of belief adoption and action observation. Like Cohen and Levesque, Perrault avoids the need for treating illocutionary acts as primitives in the model. His theory differs in making use of Reiter's (1980) default logic, which allows him to present axioms that are more context independent.

Notes

1. Story grammars as the basis of processing have been widely criticized; Levy's (1979) article on discourse structure and syntax contains a good summary.

2. The tree is taken to be analogous to parse trees for individual sentences. In a previous paper (Grosz and Sidner 1986) we argue that the embedding (or constituency) relationships represented in the tree are derived from the stack behavior of attentional state.

3. Hobbs (1985) argues that not only are cue phrases used by speakers to signal discourse structure, but that attempts to insert them in a complete discourse can prove useful for an analyst attempting to discern its structure.

4. This emphasis differs from that of much of American linguistics, in which the constraints that prevent cospecification of pronouns have been central (see, for example, Reinhart's (1983) and Lasnik's (1976) research on pronominal anaphora).

5. We use *anaphora* and *anaphoric* here to refer to those phrases that refer to entities that have been mentioned previously in the discourse and that cospecify (compare Sidner 1983) with some phrase occurring previously in the discourse.

6. For this system discourse-segment purposes were approximated by tasks and subtasks.

7. *Deixis*, from the Greek for "pointing," is used to refer to demonstrative phrases (for example, *this book* and *that*) and to other linguistic constructs whose interpretation similarly involves pointing at some circumstance of the utterance (for example, *here* and *now*).

8. This view contrasts with that of Halliday and Hasan (1976), who view ellipsis as a form of textual substitution.

9. Nonnatural meanings include the sort of meanings conveyed in language and are to be distinguished from natural meanings, for example, that smoke "means" fire.

10. For examples of systems attempting keyhole recognition, see Fischer et al. 1985, Genesereth 1979, McCue and Lesser 1983, and Schmidt et al. 1978.

11. The work of Austin and of Searle has also been tremendously influential outside of AI, inspiring an entire field of research called *speech-act theory*, with practitioners in each of the major disciplines of cognitive science. Levinson (1983, chapter 5) provides a good overview of work in speech act theory. See also the papers in Cohen, Morgan, and Pollack (1989).

12. Examples from the "train-station" domain have been widely discussed in the literature. They were inspired by a set of dialogues recorded at the Toronto train-station information booth (Horrigan 1977). The particular example described here was first analyzed by Allen (1983).

13. Appropriate response generation may also depend on distinguishing between intended responses and unintended (though still helpful) responses (Sidner 1983, 1985).

14. Georgeff (1987) provides a good overview of these recent developments; Georgeff and Lansky (1986a) have edited a collection of relevant papers.

15. Rosenschein (1981) provides a logical analysis of hierarchical plans; Wilkins (1985)

clarifies a number of important distinctions that should be made among hierarchical planning processes.

16. Of course here, as throughout this chapter, we mean by "speaker" any agent who is producing language, whether that language is spoken or written.

17. Actually a rule such as the Action-Effect Rule is used to infer not just that S intends to make e true, but that S intends to bring about e by doing α—that is, that S's plan includes the subsequence $[\alpha, e]$.

18. Allen's encoding of speech acts derives largely from Cohen's formalization of speech-act theory using AI planning notions (Cohen and Perrault 1979).

19. Most of the operators shown in figure 11.1 should be self-explanatory. KNOW-REF (agent, value) should be understood as "the agent knows the reference of (or the value of) the constant." Arguments to the operators that are capitalized (except S and H) are variables; arguments in lowercase are constants.

20. Linguistic indicators of discourse structure, such as *cue phrases*, can also be used to constrain incremental recognition; see section 11.3 for a discussion of such indicators.

21. Recent work in plan formation has also incorporated the idea of meta-plans such as these; see, for example, Georgeff and Lansky 1986b.

References

Allen, J. F. 1979. *A Plan-Based Approach to Speech Act Recognition*. Technical report 131, University of Toronto, Toronto, Canada.

Allen, J. F. 1983. Recognizing intentions from natural language utterances. In M. Brady and R. C. Berwick, eds. *Computational Models of Discourse*. Cambridge, MA: MIT Press, pp. 107–166.

Allen, J. F., and Perrault, C. R. 1980. Analyzing intention in dialogues. *Artificial Intelligence* 15(3):143–178.

Appelt, D., and Kronfeld, A. 1987. A computational model of referring. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, pp. 640–647.

Austin, J. L. 1962. *How to Do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press.

Bobrow, D., and the PARC Understander Group. 1977. Gus, a frame driven dialog system. *Artificial Intelligence* 8:155–173.

Bruce, B. 1981. Plans and social action. In *Theoretical Issues in Reading Comprehension*. Hillsdale, NJ: Lawrence Erlbaum.

Butterworth, B. 1975. Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research* 4:75–87.

Carberry, M. S. 1985. A pragmatics based approach to understanding intersentential ellipsis. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, Chicago, pp. 188–197.

Carberry, M. S. 1988. Pragmatic modeling: Toward a robust natural language interface. *Computational Intelligence* 3.

Chafe, W. L. 1979. The flow of thought and the flow of language. In T. Givon, ed. *Syntax Semantics 12: Discourse and Syntax*. New York: Academic Press, pp. 159–182.

Chafe, W. L. 1980. The deployment of consciousness in the production of a narrative. In *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex, pp. 9–10.

Charniak, E. 1977. A framed painting: The representation of commonsense knowledge fragment. *Cognitive Science* 1(4).

Cohen, P. R., and Levesque, H. 1989. Rational interaction as the basis for communication. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Cohen, P. R., Morgan, J. L., and Pollack, M. E., eds. 1989. *Intentions in Communication*. Cambridge, MA: MIT Press.

Cohen, P. R., and Perrault, C. R. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science* 3:177–212.

Cohen, R. 1984. A computational theory of the function of clue words in argument understanding. In *Proceedings of COLING84*, Stanford, CA.

Cohen, R. 1987. Analyzing the structure of argumentative discourse. *Computational Linguistics* 13(1–2).

Cullingford, R. E. 1977. Script application: Computer understanding of newspaper stories. Doctoral dissertation, Yale University, New Haven, CT.

Cullingford, R. E. 1981. Sam. In *Inside Computer Understanding: Five Programs Plus Miniatures*. Hillsdale, NJ: Erlbaum, pp. 75–119.

Dahl, D. A. 1987. Nominalizations in *PUNDIT*. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, pp. 131–139.

Donnellan, K. S. 1966. *Reference and Definite Descriptions*. Vol. 75. Reprinted in Steinberg and Jacobovits, eds. *Semantics*. Cambridge, Engl.: Cambridge University Press.

Fikes, R. E., and Nilsson, N. J. 1971. STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2:189–208.

Fillmore, C. J. 1975. *Santa Cruz Lectures on Deixis*. Technical report, Indiana Linguistics Club, Bloomington, IN.

Finin, T. 1980. *The Semantic Interpretation of Compound Nominals*. Ph.D. thesis, University of Illinois, Urbana, IL.

Fischer, G., Lemke, A., and Schwab, T. 1985. Knowledge-based help systems. In *Pro-*

ceedings of the CHI'85 Conference on Human Factors in Computing Systems, ACM SIGCHI, San Francisco, pp. 161–167.

Genesereth, M. R. 1979. The role of plans in automated consultation. In *Proceedings of the 6th International Joint Conference on Artificial Intelligence*, Tokyo, pp. 311–319.

Georgeff, M. P. 1987. Planning. In J. Traub, ed. *Annual Review of Computer Science*. Palo Alto, CA: Annual Reviews.

Georgeff, M. P., and Lansky, A. L. 1986a. *Reasoning about Actions and Plans. Proceedings of the 1986 Workshop*. Los Altos, CA: Morgan Kaufmann.

Georgeff, M. P., and Lansky, A. L. 1986b. Procedural knowledge. *Proceedings of the IEEE, Special Issue on Knowledge Representation*, pp. 1383–1398.

Grice, H. P. 1957. Meaning. *Philosophical Review* 67.

Grice, H. P. 1968. Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of Language*, 4:1–18.

Grice, H. P. 1975. Logic and conversation. In P. Cole and J. L. Morgan, eds. *Syntax and Semantics Vol. 3: Speech Acts*. New York: Academic Press, pp. 41–58.

Grice, H. P. 1978. Further notes on logic and conversation. In P. Cole, ed. *Syntax and Semantics Vol. 9: Pragmatics*. New York: Academic Press, pp. 113–128.

Grosz, B. [Deutsch]. 1974. The structure of task oriented dialogs. In *IEEE Symposium on Speech Recognition: Contributed Papers*, IEEE. Pittsburgh: Carnegie Mellon University, Computer Science Dept.

Grosz, B. J. 1977 (July). The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA.

Grosz, B. J. 1978a. Discourse analysis. In D. Walker, ed. *Understanding Spoken Language*. Ch. IX. New York: Elsevier North-Holland, pp. 235–268.

Grosz, B. J. 1978b. Focusing in dialog. In *Proceedings of Workshop on Theoretical Issues in Natural Language Processing-2*, Urbana, IL.

Grosz, B. J. 1981. Focusing and description in natural language dialogues. In A. K. Joshi, I. Sag, and B. Webber, eds. *Elements of Discourse*. Cambridge, Engl.: Cambridge University Press.

Grosz, B. J., Joshi, A. K., and Weinstein, S. 1983. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, Cambridge, MA.

Grosz, B. J., and Sidner, C. L. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12(3).

Grosz, B. J., and Sidner, C. 1989. Plans for discourse. In Cohen, Morgan, and Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Guenthner, F., Lehmann, H., and Schoenfeld, W. 1986. A theory for the representation of knowledge. *IBM Journal of Research and Development* 30(1):39–56.

Haddock, N. J. 1987. Incremental interpretation and combinatory categorial grammar. In N. Haddock, E. Klein, and G. Morril, eds. *Edinburgh Working Papers in Cognitive Science, Vol. 1: Categorial Grammar, Unificiation Grammar and Parsing*. Edinburgh, Scotland, pp. 71–84.

Halliday, M., and Hasan, R. 1976. *Cohesion in English*. English Language Series, Title No. 9. London: Longman's.

Heim, I. 1982. The semantics of definite and indefinite noun phrases. Ph.D. dissertation University of Massachusetts, Amherst, MA.

Hendrix, G., Sarcedoti, E., Sagalowicz, D., and Slocum, J. 1978. Developing a natural language interface to complex data. *ACM Trans. on Database Sys.* 3(2).

Hinrichs, E. 1986. Temporal anaphora in discourses of English. *Linguistics and Philosophy* 1(9).

Hirschberg, J. B., and Pierrehumbert, J. 1986. The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, New York, pp. 136–144.

Hirschberg, J. B., and Litman, D. J. 1987. Now let's talk about now: Identifying cue phrases intonationally. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, pp. 163–171.

Hirschberg, J. B., Litman, D. J., Pierrehumbert, J., and Ward, G. L. 1987. Intonation and the intentional structure of discourse. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, pp. 636–639.

Hobbs, J. 1979. Coherence and co-reference. *Cognitive Science* 1:67–82.

Hobbs, J. 1983. Why is discourse coherent? In F. Neubauer, ed., *Coherence in Natural Language Texts*. Hamburg: H. Buske Verlag.

Hobbs, J. R. 1985. *On the Coherence and Structure of Discourse*. Technical report CSLI-85-37, Center for the Study of Language and Information, Stanford, CA.

Hobbs, J., and Martin P. 1987. Local pragmatics. In *Proceedings of the 10th International Conference on Artificial Intelligence*, Milan, Italy.

Horrigan, M. K. 1977. *Modelling simple dialogues*. Technical report TR 108, Dept. of Computer Science, University of Toronto, Toronto, Ontario.

Isabelle, P. 1984. Another look at nominal compounds. In *Proceedings of the Tenth International Conference on Computational Linguistics*, Stanford, CA, pp. 509–516.

Joshi, A., and Weinstein, S. 1981. Control of inference: Role of some aspects of discourse structured-centering. In *Proceedings of International Joint Conference on Artificial Intelligence*, Vancouver, BC, pp. 385–387.

Kamp, H. 1981. A theory of truth and semantic representation. In J. Groenendijk, T.

orv for the representation
-56.

ɪry categorial grammar. In
Papers in Cognitive Science.
nburgh, Scotland, pp. 71–

.anguage Series, Title No.

hrases. Ph.D. dissertation

ʼ78. Developing a natural
;. 3(2).

. Linguistics and Philosophy

structuring of discourse.
putational Linguistics, New

out now: Identifying cue
ting of the Association for

:. L. 1987. Intonation and
ternational Joint Conference

67–82.

ed., Coherence in Natural

Technical report CSLI-85-
1, CA.

·s of the 10th International

·t TR 108, Dept. of Com-

eedings of the Tenth Inter-
. 509–516.

ome aspects of discourse
e on Artificial Intelligence,

ı. In J. Groenendijk, T.

Janssen, and M. Stokhof, eds. *Formal Methods in the Study of Language*. MC TRACT 135, Amsterdam, The Netherlands.

Kautz, H. A. 1985. *Toward a Theory of Plan Recognition*. Technical report TR162, University of Rochester, Rochester, NY.

Kautz, H. A. 1988. A circumscriptive theory of plan recognition. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Kendon, A. 1972. Some relationships between body motion and speech. In *Studies in Dyadic Communication*. New York: Pergamon.

Kronfeld, A. 1986 (June). Donnellan's distinction and a computational model of reference. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*. New York: Association for Computational Linguistics.

Lasnik, H. 1976. Remarks on co-reference. *Linguistic Analysis* 2(1):1–22.

Lehnert, W. 1977. A conceptual theory of question answering. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*. Palo Alto, CA: Morgan Kaufman.

Lehnert, W. G. 1981. Plot units and narrative summarization. *Cognitive Science* 5(4):291–331.

Levinson, S. C. 1983. *Pragmatics*. Cambridge, Engl.: Cambridge University Press.

Levy, D. M. 1979. Communicative goals and strategies: Between discourse and syntax. In T. Givon, ed. *Discourse and Syntax*. New York: Academic Press, pp. 183–210.

Linde, C. 1979. Focus of attention and the choice of pronouns in discourse. In T. Govon, ed. *Syntax and Semantics*. New York: Academic Press, pp. 337–354.

Linde, C., and Goguen, J. 1978. Structure of planning discourse. *J. Social Biol. Struct.* 1:219–251.

Litman, D. 1985. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues*. Technical report TR170, University of Rochester, Rochester, NY.

Litman, D. J., and Allen, J. F. 1987. A plan recognition model for subdialogues in conversations. *Cognitive Science* 11:163–200.

Litman, D. J., and Allen J. F. 1989. Discourse processing and commonsense plans. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Mann, W. C., Moore, M. A., Levin, J. A., and Carlisle, J. H. 1975. *Observation Methods for Human Dialogue*. Technical report, Information Services Institute, Marina del Rey, CA.

Mann, W. D., and Thompson, S. A. 1986. Relational propositions in discourse. *Discourse Processes* 9(1):57–90.

Marslen-Wilson, W., Levy, E., and Tyler, L. K. 1982. Producing interpretable discourse: The establishment and maintenance of reference. In *Speech, Place, and Action*. Chichester, Engl.: Wiley.

McCarthy, J. 1980. Circumscription—a form of non-monotonic reasoning. *Artificial Intelligence* 13:27–39.

McCarthy, J. 1984. Applications of circumscription to formalizing common sense knowledge. In *Proceedings of the AAAI Workshop on Nonmonotonic Reasoning*, New Paltz, NY, pp. 295–324.

McCue, D., and Lesser, V. 1983. *Focusing and Constraint Management in Intelligent Interface Design*. Technical report COINS 83-36, University of Massachusetts, Amherst, MA.

McDonald, D. D. 1983. Natural language generation as a computation problem. In *Computational Models of Discourse*, Cambridge, MA: MIT Press.

McKeown, K. 1985. *Text Generation*. New York: Cambridge University Press.

McNeill, D. 1979. *The Conceptual Basis of Language*. Hillsdale, NJ: Erlbaum.

McNeill, D., and Levy, E. 1982. Conceptual representations in language activity and gesture. In *Speech, Place, and Action*. Chichester, Engl.: Wiley.

Mellish, C. S. 1982. *Incremental Evaluation: An Approach to the Semantic Interpretation of Noun Phrases*. Technical Report, University of Sussex Cognitive Studies Programme, Sussex, Engl.

Nilsson, N. J. 1980. *Principles of Artificial Intelligence*. Palo Alto, CA: Tioga.

Partee, B. H. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7:243–286.

Perrault, C. R., and Allen, J. F. 1980. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics* 6:167–182.

Perrault, C. R. 1989. An application of default logic to speech act theory. In P. R. Cohen, J. L. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Polanyi, L. 1986. *The Linguistic Structure of Discourse: Towards a Formal Theory of Discourse Structure*. Technical report 6489, Bolt, Berenek and Neuman, Cambridge, MA.

Polanyi, L., and Scha, R. 1984. A syntactic approach to discourse semantics. In *Proceedings of the 10th International Conference on Computational Linguistics*, Stanford, CA: pp. 413–419.

Polanyi, L. 1985. A theory of discourse structure and discourse coherence. In W. Eilfort, P. Kroerber, and K. Peterson, eds. *Proceedings of the 21st Regional Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press, pp. 306–322.

Pollack, M. E. 1986 (June). A model of plan inference that distinguishes between the beliefs of actors and observers. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, New York, pp. 207–214.

Pollack, M. E. 1988. Plan as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, eds. *Intentions in Communication*. Cambridge, MA: MIT Press.

Pollack, M. E., and Pereira, F. C. N. 1988. An integrated framework for semantic and oning. *Artificial Intel-*

mmon sense knowl-
New Paltz, NY, pp.

in Intelligent Interface
Amherst, MA.

n problem. In *Com-*

v Press.

aum.

guage activity and

nterpretation of Noun
rogramme, Sussex,

oga.

ilosophy 7:243–286.

direct speech acts.

y. In P. R. Cohen,
nbridge, MA: MIT

Theory of Discourse
ge, MA.

tics. In *Proceedings*
CA: pp. 413–419.

nce. In W. Eilfort,
ting of the Chicago

hes between the
of the Association
, New York, pp.

. J. Morgan, and
ress.

or semantic and

pragmatic interpretation. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, Buffalo, NY.

Reichman-Adar, R. 1984 (Mar). Extended person-machine interface. *Artificial Intelligence* 22(2):157–218.

Reinhart, T. 1983. Co-reference and bound anaphora: A restatement of the anaphora question. *Linguistics and Philosophy* 6:47–88.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence* 13:81–132.

Robinson, A. 1981. Determining verb phrase referents in dialogs. *Computational Linguistics* 7(1).

Rosenschein, S. J. 1981. Plan synthesis: a logical perspective. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. IJCAI, Vancouver, BC, pp. 331–337.

Rumelhart, D. E. 1975. Notes on a schema for stories. In D. G. Bobrow and A. Collins, eds. *Representation and Understanding*. New York: Academic Press.

Sacerdoti, E. D. 1977. *A Structure for Plans and Behavior*. New York: American Elsevier.

Schiffrin, D. 1982. *Discourse Markers: Semantic Resource for the Construction of Conversation*. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Schmidt, C. F., Sridharan, N. S., and Goodson, J. L. 1978. The plan recognition problem: An intersection of artificial intelligence and psychology. *Artificial Intelligence* 10:45–83.

Searle, J. R. 1969. *Speech Acts*. Cambridge, Engl.: Cambridge University Press.

Searle, J. R. 1975. Indirect speech acts. In P. Cole and J. L. Morgan, eds. *Syntax and Semantics Vol. 3: Speech Acts*. New York: Academic Press, pp. 59–82.

Sidner, C. L. 1979 (June). *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. thesis. Technical report 537, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Sidner, C. L. 1981. Focusing for interpretation of pronouns. *Computational Linguistics* 7(4):217–231.

Sidner, C. L. 1982. *Protocols of Users Manipulating Visually Presented Information With Natural Language*. Technical report, Bolt, Berenek and Newman, Cambridge, MA.

Sidner, C. L. 1983. What the speaker means: The recognition of speakers' plans in discourse. *International Journal of Computers and Mathematics* 9(1):71–82.

Sidner, C. L. 1985 (Feb). Plan parsing for intended response recognition in discourse. *Computational Intelligence* 1(1):1–10.

Silverman, K. 1987. *Natural Prosody for Synthetic Speech*. Ph.D. thesis, Cambridge University, Cambridge, Engl.

Stefik, M. 1981. Planning with constraints. *Artificial Intelligence* 16:111–140.

Tate, A. 1984. Goalstructure-capturing the intent of plans. In *Proceedings of the Sixth European Conference on Artificial Intelligence*, Pisa, Italy, pp. 273–276.

vanDijk, T. A. 1972. *Some Aspects of Text Grammars*. The Hague: Mouton.

Walker, D. 1978. *Understanding Spoken Language*. New York: Elsevier North-Holland.

Webber, B. L. 1980. *A Computational Approach to Discourse Anaphora*. New York: Garland.

Webber, B. L. 1983. So what can we talk about now. In M. Brady, ed. *Computational Approaches to Discourse*. Cambridge, MA: MIT Press.

Webber, B. L. 1987. The interpretation of tense in discourse. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Stanford, CA, pp. 147–154.

Wilensky, R. 1983. *Planning and Understanding*. Reading, MA: Addison-Wesley.

Wilkins, D. E. 1984. Domain-independent planning: Representation and plan generation. *Artificial Intelligence* 22:269–301.

Wilkins, D. E. 1985. *Hierarchical Planning: Definition and Implementation*. Technical report 370, Artificial Intelligence Center, SRI International, Menlo Park, CA.

Winograd, T. 1971. *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*. Ph.D. thesis. Report no. TR-84, Project MAC, Massachusetts Institute of Technology, Cambridge, MA.

Winograd, T. 1972. *Understanding Natural Language*. New York: Academic Press.

Woods, W. A. 1978. Semantics and quantification in natural language question answering. In M. Yovits, ed. *Advances in Computers*. New York: Academic Press, pp. 1–87.

Woods, W. A., Kaplan, R. M., and Nash-Webber, B. L. 1972 (June). *The Lunar Sciences Natural Language Information System: Final Report*. BBN Report 2378, Bolt, Beranek and Newman, Cambridge, MA.