The Linguistic Structure of Discourse

Livia Polanyi

November 1995

Abstract

In order to provide a principled foundation for the study of discourse, in this paper we propose answers to three basic questions: What are the atomic units of discourse? What kind of structures can be built from the elementary units? How do we interpret the resulting structures semantically? Inferences and the correct interpretation of deixis and anaphors in discourse depend upon both structural and semantic accessibility relations. Structurally, we argue, discourse is context free and accessibility is determined by the coordination and subordination relations specified by the model of discourse presented here. Semantically, accessibility is controlled by relations among a number of modal contexts (interaction, speech event, genre unit, modality, polarity, and point of view) which determine the discourse world relative to which each primitive discourse unit is interpreted. To demonstrate the validity of our approach, the linguistic discourse model developed here is applied to a problem concerning the distribution of a discourse particle in Mocho and to various problems of discourse interpretation.

1. Introduction. Over the past twenty years, discourse level phenomena have received ever increasing attention from sociolinguists and linguistic anthropologists,

[0] I would like to thank Dorit Ben-Shalom, András Kornai, Laura Martin and Remko Scha for their substantive contributions to this paper. Of course, I alone am responsible for any errors, oversights, or conceptual inadequacies which remain.

functional linguists and structural linguists, as well as from text linguists and literary theorists. But among generative grammarians, even those who occasionally use examples which consist of two sentences instead of only one, there appears to be widespread agreement that discourse is too complex, too messy, too ill defined to be treated in a rigorous manner. In this paper we challenge this view. Recognizing that, for linguists, the concerns of discourse analysis are fundamentally similar to the concerns of grammarians working in other domains, we propose a theory of discourse which provides adequate answers to three basic questions: What are the atomic units of discourse? What kind of structures can be built from the elementary units? How do we interpret the resulting structures semantically? Before proceeding to address these questions directly, let us briefly consider design criteria for a theory of discourse which would be able to answer them satisfactorily.

As in any other scientific enterprise, the theory should not rely upon ill-defined explanatory notions, be circular in the definition of fundamental units and concepts, or use ad hoc methods in the analytic process. Also, a discourse model must be applicable both to naturally occuring disfluent, disorganized, or barely interpretable input as well as to carefully constructed written prose which is easy to read and to understand. Finally, a model which would aspire to psychological reality must strictly limit lookahead to take into account the limitations of short term memory, and cannot assume that the entire discourse is available for study. In this paper we present a highly mechanistic and replicable methodology for discourse analysis built upon an explicit linguistic theory of discourse structure, the LINGUISTIC DISCOURSE MODEL (LDM) which is designed to meet these criteria and brings structural/generative methodology to bear on discourse to allow discourse to be integrated into current theories of phonology, syntax, and semantics

In Section 2 we address the question of units. We introduce two elementary structures, the DISCOURSE CONSTITUENT UNIT and the DISCOURSE OPERATOR and compare them to a number of elementary discourse units which have been proposed in the literature such as the (finite) CLAUSE, the SENTENCE, various PROSODIC UNITS, the TURN AT TALK, the RHETORICAL/COHERENCE RELATION, the DISCOURSE PARAGRAPH, and the DISCOURSE SEGMENT. In Section 3 we consider the structures

which can be built from these units and argue that the key to their proper semantic interpretation lies in the sequential, incremental nature of discourse. To answer the second question, we discuss all major construction types, and present an analytic method which assigns a structural description to a discourse as it unfolds in time and dynamically updates semantics with the information encoded in each gesture parsed, reflecting the fact that as a discourse progresses, semiotic gestures are processed and assigned an interpretation in temporal order one at a time.[1] To demonstrate how the model can provide the analytic machinery for discovering and characterizing the discourse distribution of lower level linguistic structures, our analytic methods are applied to a discourse particle in a narrative text in Mocho, a Mayan language.[2]

Section 4 deals with the question of interpretation. First we develop further the notion of DISCOURSE WORLD and present a partial ordering of the discourse CONTEXTS that play a central role in discourse interpretation. Then we illustrate the power of this approach to address complex and vexing issues for discourse interpretation such as the nature of the 'narrative' in which the progression of narrative time obtains. In Section 5, we revisit the design criteria with which we opened this discussion. After a summary of the discourse model developed throughout the paper, we discuss the LDM in relation to a number of areas of language research including theoretical linguistics, sociolinguistics and psycholinguistics. We end the paper with a brief discussion of directions for future research.

2. ELEMENTARY UNITS OF DISCOURSE STRUCTURE. We take as the goal of discourse analysis explaining how speakers almost without error interpret personal, temporal, and spatial deixis, recover the objects of anaphoric mention, and produce responses which demonstrate that they know what's going on in the talk despite disturbances to orderly discourse development. This is an ambitious task, and one that is far too complex to accomplish without a great deal of future research. This paper

[1] We do not wait for the end of a novel, scientific paper or Doctor/Patient interaction in which we are involved to begin to assess the relationships among the utterances. In fact, interactions are constructed from the reactions of interlocutors to one another's actions.

[2] Section 3.5 is built on collaborative work with Laura Martin, who has generously allowed this material to be included in the paper.

takes the first step toward an adequate theory of discourse semantics by investigating the domain of applicability for discourse rules and processes and by confronting the issue of discourse segmentation.

Because our aim is to characterize discourse interpretation despite surface disruption, we must have a theory of discourse which does not require clear, complete, 'coherent' text as input and yet is able to assign a proper semantic interpretation to each individual utterance. In order to process discourse correctly, it is crucial that only the meanings of utterances which participate in building a representation of the same abstract semantic/pragmatic world be integrated in one semantic model. Propositional information about one situation should not be integrated with propositional information about a second situation merely because the same sort of content is being discussed. For example, consider (1):

(1)

a. *John came to the door*

b. *and left the groceries*

c. *Stop that, you kids!*

d. *Leave those cookies alone!*

e. *I put them away.*

f. *then he left.*

(1) is not 'coherent'; no semantic relationship obtains between the sentences dealing with *John* and *the groceries*, and those which deal with *the kids*. It is, however, perfectly interpretable. Although the same physical being is uttering all of the above sentences, the utterer of *a-b* and *e-f* is a socially constructed being in one INTERACTION with one set of interactants with whom she has established both a previous discourse history and an interactive microsynchrony (Goffman 1979, Goodwin 1981, Kendon 1990, Goodwin and Duranti 1992) while the utterer of *c-d* is in another interaction to which the addressee(s) of *a-b* and *e-f* function as overhearers or UNRATIFIED PARTICIPANTS (Goffman 1977). The shift from one interactional context to another is signaled by a number of diverse linguistic and behavioral phenomena discussed by Gumperz (1982, 1992) under the general term CONTEXTUALIZATION CUES. With-

out arguing for any particular formal semantic machinery for imperatives, we will represent the basic semantic facts about (1) in the following general form ('e' stands for events, 'i' for imperatives):

(2)

speaker:narrator *addressee:X*
John came to the door (e_1 at t_1)
John left the groceries (e_2 at t_2)
I put the groceries away (e_3 at t_3)
John left (e_4 at t_4)

speaker:mother *addressee:kid*
Stop that you kids (i_1 at t_1)
Leave those cookies alone (i_2 at t_2)

Here t_1-t_4 are instants in some past time discourse world **A.** (1c) and (1d) take place in some other world **B.** A correct interpretation of these utterances depends crucially on performing this segmentation and representation operation properly. Hearing this discourse, no one would mistakenly assume that *them* in (1e) should refer to *the children* nor, more importantly, to *the cookies* (see Hanks (1992) for a discussion from an anthropological linguistic perspective of deixis and indexicality in relation to context).

On the basis of the above example and many others which make similar arguments for high level discourse units and modal contexts, we argue that the interpretation of spatial, temporal, object, and event reference is systematically related to the context or set of contexts relative to which an utterance is produced. Therefore we assume that the elementary structure of discourse formation contains both propositional and indexical information. The necessity to index each utterance relative to the correct social and modal contexts will lead us to reject proposals which might posit the sentence or clause as the unit of input to the discourse grammar and to propose two new elementary linguistic units: the propositional DISCOURSE CONSTITUENT UNIT (dcu) and the non-propositional DISCOURSE OPERATOR.

2.1. THE DISCOURSE CONSTITUENT UNIT. We define the dcu as a contextually indexed representation of information conveyed by a semiotic gesture, asserting a single state of affairs or partial state of affairs in a discourse world. Each dcu, whether

linguistically or paralinguistically encoded,[3] expresses an event or in general a state of affairs in some spatio-temporal location, involving some set of (defined or as of yet undefined) participants (Davidson 1967). The event will be either positive or negative, generic or specific. Two dcus which give information about events in the same discourse world will necessarily present information from the same points of view, empathy status, and modality, and relate to the identical genre-defined and socially constructed interactional frames.

In the remainder of this section, we will present arguments for the dcu in terms of arguments against a number of other units which have been proposed as atomic discourse structures. The discussion will also serve as a review of much of the theoretical discourse analytic literature and serve to relate our enterprise to other contemporary approaches.

2.2. THE SENTENCE. In earlier work, in the interest of a perfect fit with sentential syntax, we adopted the sentence as a first approximation as the unit of discourse formation, assuming that sentential syntax and semantics would account for sentence-internal phenomena and pass sentence level information as input to the discourse component. However, in the present formulation of the LDM, we have rejected the sentence as a surface reflex for the dcu because sentences often express more than one state of affairs with interpretations in different discourse worlds. Sentential syntactic structures such as appositives, nonrestrictive relative clauses, detached participial clauses (Thompson 1983) or coordinated clauses can introduce events and entities which may receive an interpretation in a discourse world other than that of a main clause. For example, the appositive *the young millionaire,* in *My eighty-year old uncle Paul, the young millionaire, died penniless,* is interpreted in a discourse world other than the interpretation context of the proposition encoded in the matrix clause.

Subordinate clauses may present similar problems. These constituents which are often prosodically distinct must be treated as separate dcus because they may estab-

[3] Paralinguistic signaling includes the use of deictic hand gestures, ad hoc head nods, eye movements, facial expressions etc. Gestural languages such as ASL and other signed languages encode dcus in linguistic signs realized non-verbally.

lish a new set of spatio-temporal and modal contexts used by subsequent clauses, sentences, and even whole stories for spatio-temporal and entity anchoring as in the *Surgery* text given in (3) below.

(3) *Unfortunately, my sister couldn't eat the salad when she came to dinner last week because she had abdominal surgery a month ago. She got clapped into the hospital, of course, after she told them she had indigestion, and so they took out her gall bladder immediately. Anyway, she had laser surgery and she's healing incredibly fast but no raw vegetables for six weeks post op.*

The temporal index for **hospitalization** and subsequent **laser surgery** was set by the interpretation time of the subordinate clause in the first sentence: *because she had abdominal surgery a month ago.* We infer that *sister* will be able to eat *salad* again two weeks after utterance time, a conclusion based on the information that the surgery was a *month ago* encoded in the subordinate clause.

Let us now consider the temporal interpretation of when **indigestion** and **surgery** occurred and when **vegetable eating** will resume in the text given in (4), which is identical to *Surgery* except that the subordinate clause, *because she had abdominal surgery a month ago* has been omitted:

(4) *Unfortunately, my sister couldn't eat the salad when she came to dinner last week. She got clapped into the hospital, of course, after she told them she had indigestion and so they took out her gall bladder immediately. Anyway, she had laser surgery and she's healing incredibly fast but no raw vegetables for six weeks post op.*

In (4) we assume that the **indigestion** was cotemporaneous with the **dinner** and that the **surgery** occurred subsequent to it. The post op clock in this case has five more weeks to run. In contrast to (3) in which the state of **being unable to eat salad at dinner** is inferentially related to the effects of surgery, in (4) it serves as trigger for the surgery.

Thus while subordinate clauses sententially are dependent units, both syntactically and semantically tied to the matrix, as discourse level structures, they must be viewed as units independent of, though related to, the matrix; able to express a state of affairs in a discourse world different from their syntactically dominant

partners and able to initialize contexts relevant for the semantic interpretation of subsequent linguistic structures. Given the criterion that a dcu expresses one state of affairs in one discourse world, a compound or complex sentence will not be a single dcu and the sentence must be rejected as the elementary unit of discourse formation. In order to deal with the complexities of intra-sentential relations which adopting such a position entails, the resources of sentential syntax must be available at the discourse level of the grammar to combine the syntactic structures of the dcus encoded in matrix and subordinate clauses into well formed sentences. By this reasoning, there must be close collaboration between sentential grammar and discourse grammar – we will return to this matter in 3.4.

2.3. THE CLAUSE. The preceding discussion leaves open the possibility that the dcu is represented in discourse surface structure as a (finite) clause. Indeed, a single clause typically corresponds to a single dcu (and is typically expressed under a single prosodic contour) and in much previous work in discourse analysis the clause has been taken as the elementary segmentation unit of discourse (see for example Givon 1983, Payne 1987, Haiman and Thompson 1988, and many sociolinguistic analyses of spoken or written texts). However, the mapping between surface structure clauses and the expression of one event in a discourse world is not always one-to-one. For example, consider the infinitival complements in the following sentences taken from Sag and Pollard 1991:63:

(5)

a. *Kim wants to go*

b. *Lee is eager to serve on the committee*

c. *Lee promised (Sandy) to behave*

d. *Diana persuaded Lou to attend the meeting*

e. *Rene appealed to Jean to vote for the amendment*

For discourse structural and interpretation purposes we place these sentences in one category because each one refers to two distinct events in two spatio-temporal frames. In support of this position, consider (6):

(6) *I picked Jane up at the airport. On the way home, we stopped at the local diner.*

Jane and I wanted to eat something wonderful. In a busy bistro, we feasted on crisp baguettes, ripe brie, a bottle of Chablis. So we bought doughnuts, of course.

Two distinct interpretational frames of reference are evoked by this text: the realis discourse world of the trip to the local restaurant with Jane, and the irrealis world of the bistro evoked through the act of wanting. In the world of desire there is lovely food to eat in interesting surroundings, questions about the age of the wine could be asked appropriately (Seuren 1985). In the realis world in which the **driving** and **wanting** takes place, there are only doughnuts. How cross-world identity can be maintained between *we* in the irrealis world and *Jane and I* in the realis world will be discussed in 4.1.

2.4. PROSODIC UNITS. Not all sentences involving infinitival complements involve more than one discourse world, however. Some classes of infinitival complement sentences map into one dcu. Verbs like *try* involve only one state of affairs. In *John tried to bake the cookies,* for instance, there is only one event and therefore can be only one discourse world. Because there are two classes of infinitival complements which have importantly different discourse semantics, but identical prosodic structure, we reject the various proposals using prosodic units as the basic building blocks of discourse.

In most cases, segmentation boundaries for intonation contour and breath group fall together, and the unit isolated by prosodic and pausal segmentation criteria corresponds to an individual propositional unit. However, there is sufficient lack of fit between prosodic and semantic structure as pointed out by Chafe (1989:29), to reject a prosodic base for the dcu. As discourse semanticists, our aim is to describe structural and interpretative discourse behavior rather than the fit between the physiological necessity to breathe and the sound/meaning forms speakers make. In this respect, our agenda differs from that of researchers in ethnopoetics and speech understanding for whom understanding the fit between breathing and speaking is a key issue. For modeling speech, careful attention to the breath unit and its more abstract reflex, the intonational phrase or contour, is clearly both practically necessary and methodologically appropriate (for different proposals for prosodic units

see Liberman and Prince 1977, Liberman 1978, Chafe, 1980, 1988, Pierrehumbert 1980, Gee and Grosjean 1983, Selkirk 1984, Ladd 1986, Nespor and Vogel 1986, Hirschberg and Litman 1987). Similarly, for those concerned primarily with verse and poetics, prosodic units based on periodic segmentation are important primitive units, related ultimately to the need of the performer to breathe and to engage listeners aesthetically and physiologically by powerful rhythmic structures and the tension between metrical and structural syntactic units (Hymes 1981, 1987, Tedlock 1983, Woodbury 1985, 1987).

We do not deny that an important relationship exists between prosodic and discourse structures – see, for example, the prosodic reflexes of interruption, return and continuation phenomena explored in Hirschberg and Pierrehumbert 1986. Pauses, and even the perception of pausing, undoubtably segment speaking. Therefore, the pause is a sufficient reason to force a dcu break. But neither pausing nor prosodic contour shifts are necessary conditions for segmentation into elementary dcus. We will return to consider segmentation issues in more detail in Section 2.7 below.

2.5. Other proposed discourse level units. The dcu is a new unit. In defending our choice to introduce a new elementary structure, we began by arguing that units more widely known in linguistics: the sentence, the clause, and the various prosodic units do not yield sufficiently detailed segmentation for the purpose of discourse interpretation. Here we will argue that units specifically introduced for discourse analytic purposes in other frameworks are also inadequate for discourse semantic purposes. We begin by discussing the discourse paragraph, and then argue against adopting the discourse segment, a unit favored in computational treatments of discourse, the turn at talk, an interactional unit, and the rhetorical/coherence relation, a textual unit.

2.5.1. The discourse paragraph. Longacre 1976a:267 proposes that the syntactic unit hierarchy, traditionally viewed as consisting of the morpheme, the phrase, the clause and the sentence be reconfigured in accordance with tagmemic theory to consist of 'eight levels: morpheme, stem, word, phrase, clause, sentence, para-

graph, discourse' [emphasis in original]. According to tagmemic theory, 'the primary elements of construction [of a unit of a given level of structure] are elements from the next lower level of construction' and 'we will therefore expect discourse to be composed primarily of paragraphs, paragraphs to be composed primarily of sentences, sentences to be composed primarily of clauses, clauses primarily of phrases, etc.' (pp. 261-262). The sentence, whether a simple sentence consisting of only one clause or a complex sentence consisting of more than one clause, is the input structure to the paragraph, the constituent structure of discourse. 'The paragraph is the developmental unit of discourse [. . .] It's in general a looser and larger package than the sentence' (p. 276).

The elegance of this proposal is greatly reduced, however, when the levels are discussed in detail. 'Apparently, discourse structure is an obscure tangle of paragraphs, chapters and episodes – with a lot of overlap among them' (Longacre 1976b:4). Of particular interest to our enterprise is the difficulty distinguishing between discourse and paragraph and paragraph and sentence. Without a clear distinction among the levels, or a clear definition of the paragraph, we have no choice but to reject this construct as a fundamental discourse constituent while emphasizing that many of Longacre's binary paragraphs and proposals for discourse relations (summarized in Longacre 1976b:11-20) are not dissimilar to notions which we use here.

2.5.2. Turns at talk. The conversational turn, or exchange of right to the floor, is taken to be the basic unit of the social organization of interactively constructed spoken discourse by many researchers interested in language use, particularly those associated with the Conversation Analysis movement within ethnomethodology. However, for the purpose of semantic interpretation the turn is seldom a relevant structure. Turns often comprise a large stretch of discourse with many shifts in the relevant semantic dimensions necessary for semantic interpretation. Replacing the notion of turn with transition relevance place (TRP), to refer to a moment within a stretch of talk where speaker change may occur (Sacks, Schegloff and Jefferson 1974, Levinson 1983, Fox 1987) does not help, because no existing characterization of the TRP is sufficiently explicit to provide the structures neces-

sary for a semantic account. Nonetheless, we should like to emphasize that conversational analysis provides many important insights into discourse structure which we have retained. Specifically, we use complex constituents such as the ADJACENCY PAIR which refers to interactional structures such as greetings, farewells, compliment/response sequences, etc.[4]

2.5.3. THE DISCOURSE SEGMENT. The arguments against adopting the discourse segment widely used in computational work on discourse are relatively easy to make since this kind of segment has never been defined except in terms of the speaker's INTENTIONS, PLANS, or ATTENTION, none of which have yet received full definition. Although there is a great deal of agreement about many intention and attention segmentation points once people are given an informal tutorial on what these might be (Passaneau and Litman 1993), as Woodbury (1983), Dubois (1987), Chafe (1988), and others have pointed out, clausal boundaries, breath points (or pauses), and propositional or 'idea' unit boundaries tend to cooccur. These cooccurrence points, we would argue, are where, most typically, one would impute a shift in attention or intention to a speaker even if one did not quite know what attention and/or intention shifts might be. However, the interesting cases for purposes of investigating the basic unit or units of discourse structure are precisely those in which clausal, breath, and propositional boundaries are not identical.

Our aim is to provide a correct semantic interpretation of every linguistic expression. Therefore, we reject the attentional and intentional approach because segments based on such cognitive states often persist over stretches of discourse such as *Surgery* discussed earlier or *Apartment* discussed below which are each composed of many clauses and include several shifts in spatio-temporal, modal, polarity, and empathic anchoring and yet the intention of the speaker, to tell a particular story, does not change throughout the telling. One might argue that the attentional focus does

[4] In the analysis of conversational anaphora, Fox 1987 uses the adjacency pair as the primary unit. While this often leads to elegant solutions, we agree with Levinson 1981:107 that conversation does not consist solely, or even mainly, of adjacency constructions. In addition, adjacency pairs need not necessarily coincide with actual TRPs – for example, a rhetorical question/answer pair is most appropriately constructed within one turn.

change, but without an explicit discussion of precisely on what basis one assumes an attention shift, we must conclude that these notions, as they have been presented so far in the literature, while suggestive, fail to provide the information necessary to assign the meaning of each expression to the appropriate semantic representation.

2.5.4. RHETORICAL AND COHERENCE RELATIONS. The aim of rhetorical or coherence relation theory is to characterize the relationships holding between propositional elements in adjacent sentences of coherent written discourse. Often a sentence relates to the sentences immediately preceding it in one of a number of ways: continuing a list or narrative, giving more information about some aspect of the previous sentence, or completing some sort of logical or rhetorical structure initiated previously – supplying a THEN for an IF in the preceding sentence, for example. These relationships have been noted and classified by several scholars under the various headings of BINARY PARAGRAPH STRUCTURES (Longacre, 1976a), RHETORICAL STRUCTURES (Fox, 1987, Mann and Thompson 1988, Hovy 1991, Moore and Paris 1993) or COHERENCE RELATIONS (Hobbs 1979, Reichman 1985). Other work on coherence follows the tradition begun in Halliday and Hasan 1976 which explicitly rejects the notion of structure above the sentence. These several coherence frameworks have so far been unable to characterize COHERENCE adequately and even for easily interpretable texts which present a single subject matter in a conventional manner, the core theoretical construct remains murky and ill-defined, more or less in the category of 'I can't define it but I know it when I see it.'[5] As van Dijk 1977:33 puts it: 'the notion of coherence is not well defined [...] Intuitively, coherence is a semantic property of discourses, based on the interpretation of each individual sentence relative to other sentences.' While we agree with the intuition, we don't find it to be sufficiently strong to serve as the foundation of the theory of discourse.

Our view is that attempts to deal with discourse coherence without a developed theory of discourse structure and interpretation are premature. In addition, theories

[5] COHESION, a related notion, presents similar definitional difficulties. See for example, the discussion of TEXTURE in Hasan 1977:228 where she writes 'Texture is the technical term used to refer to the fact that the lexicogrammatical units representing a text hang together – that there exists linguistic cohesion within the passage [...] Texture is what makes two sentence of a text cohere.'

of coherence relations fail to provide any useful account of discourse which is less coherent, or even incoherent. Faced with discourse in which rhetorical relations do not usefully capture what is going on, as is the case in (1), these theories are reduced to silence. Given our semantic agenda, then, we can not adopt rhetorical or coherence relations as the basic unit of discourse structure. But unlike Grosz and Sidner 1986, we believe that the information about these relationships must be expressed explicitly in a general discourse theory. As we shall see in Section 3, such relations can be stated as complex dcu formation rules.

2.6. DISCOURSE OPERATORS. In addition to semantic structures which express states of affairs about a discourse world, utterances may also involve nonpropositional elements which communicate information helpful to interpreting a discourse properly by making explicit the nature of the links among pieces of information. These DISCOURSE OPERATORS modify discourse constituents and may have scope over long stretches of discourse (Gülich 1970, Reichman 1985, Schiffrin 1987). For example, in the text given in (1) above, *you kids* makes explicit the addressees of the two utterances *Stop that!* and *Leave those cookies alone!*

In the *Surgery* text in (3) above, there are several non-propositional structures which nonetheless carry important discourse information: *but* which logically relates two propositions, and *unfortunately* and *of course*, which express the attitude of the (modeled) speaker towards material in the story. Although some metacommunicative propositional utterances such as *As we were saying before we were interrupted* may function as operators as well as having propositional content, most lexical items functioning as operators such as English *yes, uh, ok, but, because, well, so, if, then, therefore, hello, goodbye, now, or, what, why, and, anyway,* and any proper name used as a vocative do not assert information about states of affairs but give information about the state of the discourse and the relation of discourse entities and discourse representations to one another. As we shall see in Section 4, the interpretation of dcus, though primarily determined by their propositional content, is greatly affected by the contexts in which they occur. Discourse operators, while themselves lacking in propositional content, often make explicit the shift in the indexes of the

content bearing dcus.

2.7. Segmentation and discourse surface structure Based on the arguments against the various linguistic structures we conclude that neither syntactic nor prosodic units necessarily correspond to the dcu, which is a semantically motivated structure expressing one state of affairs in a discourse world. Though there is no completely reliable surface structure reflex for the dcu, both syntactic and prosodic information guide the process of discourse segmentation. As a practical matter, clause boundaries, sentence boundaries, and prosodic contours are all useful in locating single propositional structures. There are several reasons for this. First, a single state of affairs is normally expressed by a single syntactic clause. Second, multiple clauses generally express more than one state of affairs.[6] Third, the standard clause normally encoding a single proposition typically falls under a single prosodic contour. However, if a clause remains incomplete but ends in a noticeable pause, or part of an utterance is elliptical or in other senses fragmentary, the propositional content of that utterance will nonetheless be represented by an independent dcu structure. The methodological justification for treating pauses as absolute discourse unit boundaries comes from the fact that in the real time analysis of interactively constructed spoken language one cannot halt the analysis waiting for the completion of a syntactic or semantic unit given that the expected completion might never happen. Therefore, in order to allow a principled handling of syntactically incomplete utterances, hesitations, false starts, and other key performance factors, the minimal dcu need not encode a complete surface constituent.[7]

Discourse segmentation is determined by semantic criteria and guided by syntax and intonation. Discourse operators also force segmentation breaks on semantic grounds. For example the sentence *I went downtown but Mary stayed home* (discussed in Longacre 1976a:261) is analyzed under the LDM as a three unit structure consisting of two dcus *(I went downtown* and *Mary stayed home)* and a discourse

[6] The most important counterexample is the case of repetition – see Polanyi 1978, Tannen 1989, Walker 1993.

[7] Should the incomplete syntactic structure be resumed after a pause, the rules of sentential syntax will construct a more complete structure from the fragments.

operator *but,* a logical connective with asserts the relationship which obtains between two states of affairs. The sentence *Actually, I slept* expresses a single state of affairs, **speaker slept** which obtains in one world **A** and an attitudinal evaluation of that state of affairs uttered from the point of view of a speaker situated in discourse world **B**. World **B** therefore provides a vantage point from which she can survey and reach conclusions about the state of affairs in **A**. This sentence thus maps into two discourse-level tokens, the discourse operator *actually* and the dcu *I slept.* Though operators are helpful for segmentation, it should be emphasized that the dcu breaks can be computed entirely without operators and it is never necessary to depend on words such as *well, so, anyway,* or phrases such as *in the first place* to determine where a discourse segmentation point is located.

3. Complex discourse units and discourse parsing. In the previous section we proposed to segment discourse maximally: a new dcu is started whenever phonological (i.e. pausal or prosodic) criteria indicate a break, whenever sentential syntactic criteria indicate a clause break (except for a lexically limited set of matrix verbs governing infinitival clauses), and whenever sentential semantics requires a change in any of the contexts (spatial, temporal, modal, etc.) that index the discourse worlds where the events (and in general, states of affairs) are interpreted. This segmentation methodology provides dcus even if discourse operators, the discourse markers of Schiffrin 1987, are entirely absent from the text, and even if the discourse is fragmentary or incoherent. But for a comprehensive theory of discourse syntax and semantics the segmentation of text into atomic constituent units in a uniform and principled fashion is just the first step that must be followed by analysis/generation of higher level constituents.

In this section, we discuss the rules specifying the syntax and semantics of well formed discourse structures recursively built from elementary dcus and develop a typology of higher level constituents. As we shall see, discourse operators are peripheral to this undertaking: the central data structure will be the discourse parse tree (DPT) which has propositional dcus at the leaves. Operators, if present in the text, are treated as clitics attached to the propositional hosts. We will distinguish

three basic types of higher structures: COORDINATION, SUBORDINATION, and BINARY constructions. Nonterminal nodes of the DPT will always be one of these, and will be labeled by *C, S,* or *B* accordingly. We also recognize higher level discourse structures such as GENRE-DEFINED CONSTITUENTS and SPEECH EVENTS. There can be no doubt that stories and other structured discourse genres (such as arguments and negotiations) have a characteristic constituent structure in which expected types of information are presented in a conventionally agreed upon manner, and so do speech events (Hymes 1972) such as doctor patient interactions, formal lectures, business meetings, church services or blind dates etc. for which the participants know when they are in one phase of the activity or in another and behave accordingly.[8] After describing the elementary coordination, subordination, and binary structures in 3.1-3.3, we will discuss the parsing/formation of such higher level constituents in 3.4. The technique is demonstrated in 3.5 on the example of a text from Mocho, a Mayan language.

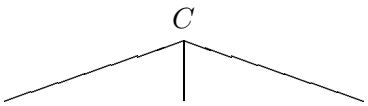
3.1. COORDINATION STRUCTURES. Adding a next item to a list, giving a next episode of a story, beginning a new topic in a conversation when discussion of a previous topic has been concluded, or going on to a next expected activity in a speech event such as a church service can all be analyzed as continuing the development of an ongoing discourse activity. In the DPT such continuing activities are depicted as a sequence of coordinated constituents i.e. as a nonterminal *C* node immediately dominating arbitrarily many constituents that share a single type. LISTS, TOPIC CHAINS, and NARRATIVES, are common sequential structures. Consider the simple discourse fragment given in (7):

[8] Children as young as six years old in traditional schools do not ask what the homework will be for tomorrow when the teacher is in the middle of explaining the new material for the day; they know that the homework phase of the class will occur after the explanation, near the end of the time allotted for that subject. They know, too, that in a story about what happened on the playground, it is necessary to state where the action took place and to introduce the characters before beginning a detailed temporally ordered recitation of the events which took place. Similarly, they know that after the teacher asks a question and a student answers, the teacher will ordinarily evaluate the student's answer. See Sinclair and Coulthard 1975, 1992 for a discussion of classroom discourse.

(7) *I like to read Sci-Fi. I like to ski and I like to sleep late.*

The structure of (7) can be characterized by the tree given in (8):

(8)



I like to read SF *I like to ski* *I like to sleep late*

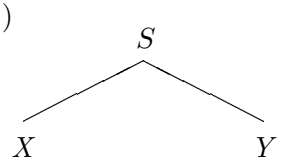
While static characterizations of DPTs can be expressed by ordinary context-free rewrite rules such as $C \rightarrow XX^{+}$, in order to capture the dynamic aspect of DPT construction we will utilize more powerful rules of tree-manipulation, namely ADJUNCTION RULES. We will use both daughter-adjunction (for coordination) and Chomsky-adjunction (for subordination and binary constructions) but only in the special case when the adjoined tree has a single node corresponding to an elementary dcu. We further restrict the power of the adjunction operation by requiring that only nodes dominating the rightmost terminal of the host tree can serve as the site of the adjunction. The significance of this condition of RIGHT OPENNESS will be discussed in detail in 3.4 – for coordination the condition simply means that leaves are added one after the other.

What kind of coordination are we dealing with? In (8), the first dcu, *I like to read Sci-Fi* could be an item on lists of many types such as 'What I like to do', 'What I do on Tuesdays', 'What I like to read', 'What people in my family like to do' and so on (see Barsalou 1985 for a discussion from a cognitive psychology perspective of the problem of constructing ad hoc lists). When the second dcu is encountered, and the information in the proposition **speaker likes to ski** is compared with the information in the proposition **speaker likes to read sci-fi**, using world knowledge we can infer that what is being communicated is a list of items of **fun things the speaker likes to do**. This higher level, more general information, referred to as the COMMON GROUND is used in the DPT as further specification of the C node label (for details of how the computation is done on the lower level dcus to create the specification on the higher level dcu see Polanyi 1985, Prüst 1992, Moens et. al. 1994, Keher 1994, Prüst at al. in press). When the third dcu, *I like to sleep late,* is

encountered, it is compared in form and meaning to *I like to ski*, the immediately preceding clause in the text. A computation of the common ground between dcu 2 and dcu 3 nets the same higher level common ground **fun things the speaker likes to do** as was computed to obtain between the first two dcus. This means that all three dcus are specific instances of the same general list and can be accommodated under the same higher level node.

3.2. SUBORDINATION STRUCTURES. Discourse activities which interrupt the completion of other ongoing activities are treated in a structurally uniform manner. ELABORATIONS on a point just made, DIGRESSIONS to discuss something else, ASIDES, APPOSITIVES, sections of DIRECT DISCOURSE, or true INTERRUPTIONS are all treated as subordinated to activities which continue the development of an ongoing unit, be it a story, a proposal for a course of action, a lecture, or a move in speech event. We also recognize SENTENTIAL SUBORDINATION which obtains between a matrix clause and its subordinated clause,[9] appositive, or parenthetical element. The general constraint in discourse subordination requiring the subordinated element to be to the right of its matrix in the linear ordering of the text (and thus in the discourse parse structure which is strictly bound to text order) is relaxed in sentential subordination, where the normal order of embedding can be reversed. In the general case, the subordinated constituent will be encoded as the right daughter Y in an elementary tree such as (9):

(9)



Notice that the superordinate constituent X does not dominate Y – the fact that the relation between the two is one of subordination is expressed by the label of the mother node. As with coordination, it will be convenient to extend the skeletal node labels given by (9) so as to encode the semantic interpretation of the various

[9] Matthiessen and Thompson (1983) build on Halliday's notions of rankshifting (Halliday 1967) and treat subordinated clauses as discourse-embedded.

units. But unlike in coordination, where the interpretation of the mother node is computed by conjoining the interpretations of the daughters, in structures such as (9) the interpretation of the mother node is the same as the interpretation of the left daughter, and the right daughter has no impact whatsoever. This is particularly clear for the case of interruptions, in which no semantic relationship obtains between the sister dcus.

Perhaps less intuitively, the same structural description (9) is appropriate for elaborations, where the right daughter gives more information about some aspect of the predication or entity encoded in the left daughter, as in the following example (10):

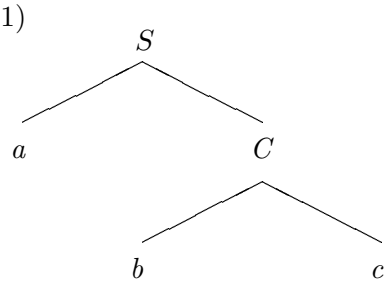
(10)

a. *I like to do fun things on vacation*

b. *I like to read Sci-Fi.*

Should the discourse continue, *I like to ski*, this new dcu would be coordinated to Y *I like to read Sci-Fi* under a newly created *C* node interpreted as **Fun things I like to do** as in (11):

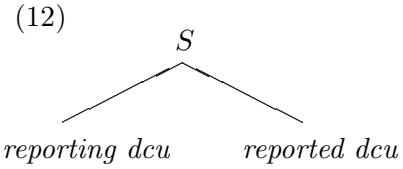
(11)



Reported speech and thought are common in stories, arguments and other forms of discourse. What is spoken or thought by the character is interpreted relative to an interaction in a story discourse world among characters in that discourse world. The narrator, in asserting a reporting clause such as *I said* or *Suzie thought* which typically is an event on the mainline of the narrative, communicates directly to the story recipients in a discourse world which includes narrator and recipient as participants but which excludes the characters in the story. Because the reporting

clauses are events on the story mainline and the reported speech or thought interrupts the development of the narrated world by interposing an interaction among other participants, we subordinate reported speech and thought to the dcu of the reporting narrative as shown in (12):

(12)



Several systemic and other functional grammarians have considered the question of whether the clauses in reported speech should be treated as embedded to the reporting clause (see Matthiessen and Thompson 1983), and it should be noted that in addition to the default construction given in (12), many languages (including English) have a reverse construction, with a distinctive intonation pattern, in which the reported dcu comes first. In our model, this can be treated as sentential subordination or as a case of binary constructions, to which we now turn.

3.3. Binary Structures. Binary structures construct a dcu out of two dcus commonly joined by an explicit or implicit relation. Semantically, binary relations are very complex. Binary relations hold between two constituents related logically (e.g. if/then, then/if, or, therefore), rhetorically, (e.g. sum up), or interactionally (e.g. question/answer, warrant/response, error/repair).[10] The discourse parsing of (13)

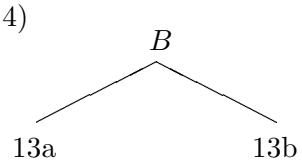
(13)

a. *If John goes to the store*
b. *he'll buy tomatoes*
c. *Otherwise, we'll just have lettuce in the salad.*

begins with setting up an intrasentential binary node dominating both dcus in the first sentence:

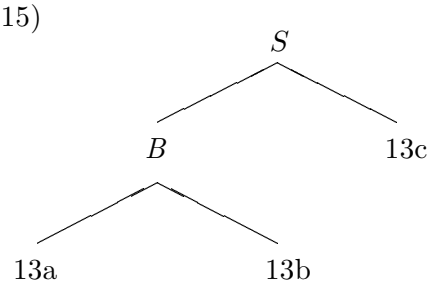
[10] Longacre (1976) refers to logical and rhetorical structures as binary paragraphs.

(14)



When (13c) becomes available, it is subordinated to the B node, since at this point it is a digression:

(15)



At this time it is not known how many types of binary relations (and thus how many binary node types) need to be distinguished, though there is no reason to believe that the number of binary discourse structures commonly found in a language and which should be stipulated in a grammar would greatly exceed the number of complement types that sentential syntax requires us to differentiate. It is to be expected that different languages may have quite different binary relations (Longacre 1976a).

One binary structure deserving special mention is REPAIR, which differs from other discourse relations because, instead of an instruction to semantics to create a new representation or update an existing one, the repair node calls for the removal of information previously added to a representation. Because of repairs, discourse which is syntactically monotonic, is semantically non-monotonic.

3.4. PARSING AND GENERATION. From the preceding discussion of the major construction types it is clear that at a high level of abstraction all DPTs can be described by a simple context-free grammar. Elementary DPTs will either have a C mother node and two or more daughters, or they have S or B mother node and exact-

ly two daughters.[11] But this is no more than the syntactic skeleton of the grammar. As soon as we annotate the nodes by the semantic interpretation of the constituents matters become much more complex. The main difference between DPTs and the trees familiar from sentential syntax is that in DPTs we allow attachment only at the right edge: discourse POPS which resume an interrupted constituent will always close off the interrupting (elaborating or otherwise subordinated) constituents and make it impossible to attach (coordinate or subordinate) any subsequent dcus to them. It is this property of the DPT that we refer to as being RIGHT OPEN.

It should be emphasized that, together with other computational discourse analysts, by stipulating restrictions on dcu attachment we are making a very strong claim about the structure of discourse. The openness of the right edge makes the DPT in this respect equivalent to the INTENTION STACK mechanisms proposed by Grosz and Sidner (1986) and the RIGHT FRONTIER of Webber (1988), as opposed to Reichman's (1985) CONTEXT SPACES, and Johnson-Laird's (1983) MENTAL MODELS which always remain open and available for incrementation. This restriction permits predictions to be made about the encoding forms of incoming propositions. Any attempt to add propositions to a closed unit will be accompanied by an intonational repair or initiation signal and will receive a syntactic and phonological encoding as a new rather than a resumed unit (see Grosz and Sidner 1986, Polanyi 1986, 1988, Hirschberg and Pierrehumbert 1986, Hirschberg and Litman 1987, Webber 1988). The open right edge offers a simple formal mechanism for the analyst to keep track of what is happening at any given moment in a discourse. Ongoing activities that have been interrupted and are expected by the participants to be resumed are all encoded by nodes on the right edge.

In sentence parsing, the real time operation of the model (Yngve 1961, Church 1980, Marcus 1980) is rarely a central issue, since all incoming elementary units (words or morphemes) can generally be assumed to fit into short-term memory

[11] If we use different types of parentheses to encode the type of the mother node, () for coordination, [ ] for subordination, and {} for binary relations, we can describe the language of well-formed DPTs with a single rule $N \rightarrow (NN^+)|[NN]|\{NN\}|t$, where N is the only nonterminal (the start symbol) and t is the terminal denoting an elementary dcu.

and, if needed, can be recalled verbatim to explore alternative parses. In discourse parsing, verbatim recall obviously does not extend to more than a few elementary units, possibly only to one or two. Since the antecedents of anaphors are sometimes recovered across a gap of 15-17 clauses (Givón 1983a:36n6, 1983b), we need some more persistent structure, such as the ACTIVE and the CONTROLLING SEGMENTS of Reichman, 1981, Fox 1987, to store the objects, events, or states of affairs still available for anaphoric mention. In our model, this is the right edge of the DPT. But a performance model must operate in real time or with minimal lookahead, and this is where we depart from the classic model of Yngve, which attributes procedural reality to context-free rewrite rules. At the procedural level, we use the more powerful operation of adjunction at the right edge.

Though the precise details of the division of labor between the sentential and the discourse parser greatly depend on the theory of sentential syntax adopted, we emphasize that the LDM does not put a heavy burden on sentential syntax, since it only requires the recognition of clause level sentential units and the ordinary (intrasentential) subordination and coordination relations. In particular, we do not assume that discourse is structurally organized as a series of conjoined sentences or that some trivial extension of the sentential model (such as the addition of a rule generating strings of sentences) will enable discourse analysis to take a free ride on the parsing mechanism that must be in place for sentential syntax. On the contrary, by treating pauses as absolute dcu boundaries and thus enforcing a single breath group as the outermost limit of dcu size, our model takes responsibility for a broad range of phenomena (such as long lists of phrases) that needlessly stretch the sentential model beyond its normal limitations. In fact, the task of tokenizing a stream of utterances into dcus is relatively trivial compared to the central task of sentential syntax and semantics, namely the analysis and interpretation of dcu-internal structure. The discourse module must have access to sentence-level syntactic structure and semantic interpretation, and we assume that at the interface these are provided by the sentential module in dcu-size chunks (possibly with limited lookahead to accommodate cases when multiple dcus form a single sentential syntactic unit under a single intonation contour).

Suppose a DPT has already been built over the first k dcus $d_1, \ldots, d_k$. When the sentential component provides a new dcu d_{k+1}, we first determine the relationship of this incoming unit to the immediately preceding dcu d_k. If this is an elaboration relationship, then we attach d_{k+1} as the right sister to d_k at a newly created S node, and label this node with the structural and semantic characteristics of d_k. Otherwise, we continue up the open right edge of the DPT, looking for semantic or syntactic matches. When a match is made, we adjoin the newly parsed dcu as a terminal under a higher level existing or newly created nonterminal node. If no match is made, we adjoin d_{k+1} as the right-sister of a newly created S node at the bottom of the DPT assuming that the new dcu is interrupting all on-going discourse activities.

Understanding the history of the discourse is greatly facilitated by adopting the strictly left to right, dcu by dcu attachment procedure outlined above. As a new dcu is processed, it will either be seen to continue or to interrupt an ongoing discourse activity, or to initiate a new discourse activity. Its relationship to the previous discourse will be reflected in how it finds an attachment point on the tree's right edge. A dcu which continues an ongoing discourse activity will be added as a right sister to the dcu currently on the right edge either at an existing or at a newly inserted C node, unless the relation between the new dcu and its sister is one of logical or adjacency pairing, in which case the node label reflects the binary nature of the relationship. While C nodes remain open, and ready to accept new daughters which bear the same relationship to the mother as existing daughters, a B node will not absorb new daughters.

A dcu which interrupts an ongoing discourse activity is adjoined to the DPT as subordinated to the dcu representing the interrupted activity. Interruptions are of two semantically very different types: those which interrupt the flow of an ongoing activity to interject some semantically related material (elaborations) and those which interject semantically unrelated material (true interruptions). The two types are treated similarly: the interrupting dcu is embedded relative to the material interrupted at an S node that has the semantics of the interrupted material. Because this node inherits its semantics from the left daughter, the interrupted material

remains available on the right edge of the tree and thus the discourse activity encoded by the mother remains available for further continuation. A dcu which continues a previously interrupted discourse activity is attached as the right sister of the resumed activity at a C or B node which dominates them both. The new dcu and the node dominating it are thus available for continuation by a next dcu.

A dcu which initiates an entirely new discourse activity will be added to the DPT as a daughter of a high level mother which may be created especially to close off the old discourse activities and begin the new. In this case, the new node is inserted above the highest existing node in the tree and the new daughter becomes the new right sister of the previous discourse, rendering the entire previously existing tree inaccessible. Less dramatically, a dcu which initiates a discourse activity is often the first utterance of a new constituent, such as a new MOVE or EPISODE in an ongoing higher level unit such as a SPEECH EVENT or STORY. Both stories and speech events (linguistically realized socially meaningful activities, see Hymes 1972) are internally organized and while the full details of this organization are complex, the highest level of organization is essentially sequential. For example, in a doctor-patient interaction first there is a greeting, followed by a statement of complaint, an examination, discussion of the findings, suggestions for follow up, and finally, leave taking. If a dcu (such as the doctor's summary of the findings) begins a new move, the previous moves become structurally unaccessible. Interruptions and other real world exigencies do not cause the analysis to fail since they are embedded to the matrix speech event at the moment of occurrence and the speech event is resumed after the digression is ended. Needless to say, there can be attachment ambiguities, but the problem of finding higher level discourse units does not appear to be any more complex than in the sentential case, and since our grammar is context-free, the same techniques of ambiguity resolution are applicable.

3.5. THE DISCOURSE DISTRIBUTION OF A MOCHO PARTICLE. Explaining the distribution of discourse particles is a classic linguistic problem. In this section, we will present an LDM analysis of the distribution of *la,* a phrase initial particle in narrative discourse in the Mayan language Mocho. Because in our model discourse

unit boundaries emerge from the tree construction process described above, without reference to discourse markers, our analysis avoids the methodological circularity (pointed out in Woodbury 1987) of using discourse markers to determine discourse unit boundaries while simultaneously accounting for the distribution of markers in terms of initiation, completion, or resumption of discourse units.

Mocho is a member of the Mayan language family (Greater Kanjobalan subgroup) spoken by approximately one hundred speakers in southeastern Chiapas, Mexico (Kaufman 1967, Martin 1989). Compared to many other Mayan languages, Mocho has an elaborate set of syncategorematic morphemes which do not carry inflection. While the function and distribution of some particles, such as temporal particles *poxo* 'again' and *ni* 'already' conditioned primarily by phrasal or clausal factors are fairly well understood, the behavior of other particles, such as phrase initial *la* has resisted adequate characterization solely in terms of their sentential functioning. As we shall show below *la* is not only a sentential coordinator, but it also joins constituents at the narrative main line.

In order to account for the distribution of discourse *la* we shall first look at the known intra- and inter-sentential coordination functions of this particle. *la* is easily identified as a coordinator at the phrase level where it can serve as a coordinator for nominal, verbal, and sentential constructions as shown in (16-18) below:

(16) Ha'n qa.xangab.oo' pachiih bi', kam.oom ii.wah la kam.oom ii.nu'ul Be
that(one) 1pposs.sandal.1pex before part die.resu 1sposs.father and die.resu 1sposs.brother id
'Those were our sandals before – my late father('s) and my late brother Beto('s.)'

(17) ch.ibeel ngabaal aki la x.k'u7uul ha' aki chaki-a
3sposs.under rain punct.3abs.go-down and 3poss.belly water punct.3abs.go-down down-part
'Under the rain he came and inside the river he came (down)'

(18) ik'i ma' ch.aak'.sa ti beha' la k'iibi eel x.waa'ol
pun.3abs.pass-by perhaps 3erg.wet.caus prep 3ravine and pun.3abs.grow out 3poss.length
'Perhaps he passed through the ravine to wet them and their length grew.'

However, it often makes no sense to coordinate a *la*-initial sentence with the immediately preceding sentence (see for example (19.50) below). But if not coordination, what is the function of this particle? How can its distribution be stated, let alone explained? To investigate this matter, we turn to an episode from a story[12] in which a neighbor discovers that raw leather sandals will stretch when they are wet and shrink if they are heated. In the text given in (19) below, numbered lines correspond to elementary dcus and discourse operators as determined by the semantic, syntactic, and prosodic criteria described in Section 2.5 above. Pausing is indicated by *x* marks in parantheses, each *x* approximately .25 seconds. Capitalization, puctuation, and quotation marks are added for legibility – they are not part of the analytic process. Phrase initial *la* is shown in ALL CAPS, other discourse operators are italicized.

(19)

1	LA	And
2	chaawi ti witz (xxx)	he arrived at the mountain,
3	"muuqu iixangab	"Look at my sandal,
4	*don Luwisito* (xx)	don Luisito,
5	*ma' to*	Gee!
6	chk'iibi eel xwaa'ol te' qaxangab" (x)	the length grows on these our sandals"
7	chuta kamoom iiwah (xxx)	he said to my late father.
8	"*ah*	"Ah
9	indehe xmaña che we' (xxx)	thus is the way of that one,

[12] EPISODES are constituents of the STORY genre. In addition to an optional ABSTRACT or introductory constituent and an optional concluding constituent or CODA, a story will consist of one or more episodes in which the action of the story takes place. Episodes are normally differentiated from one another by a marked change in temporal, spatial or participant structure – see our discussion of indexes in 4.1.

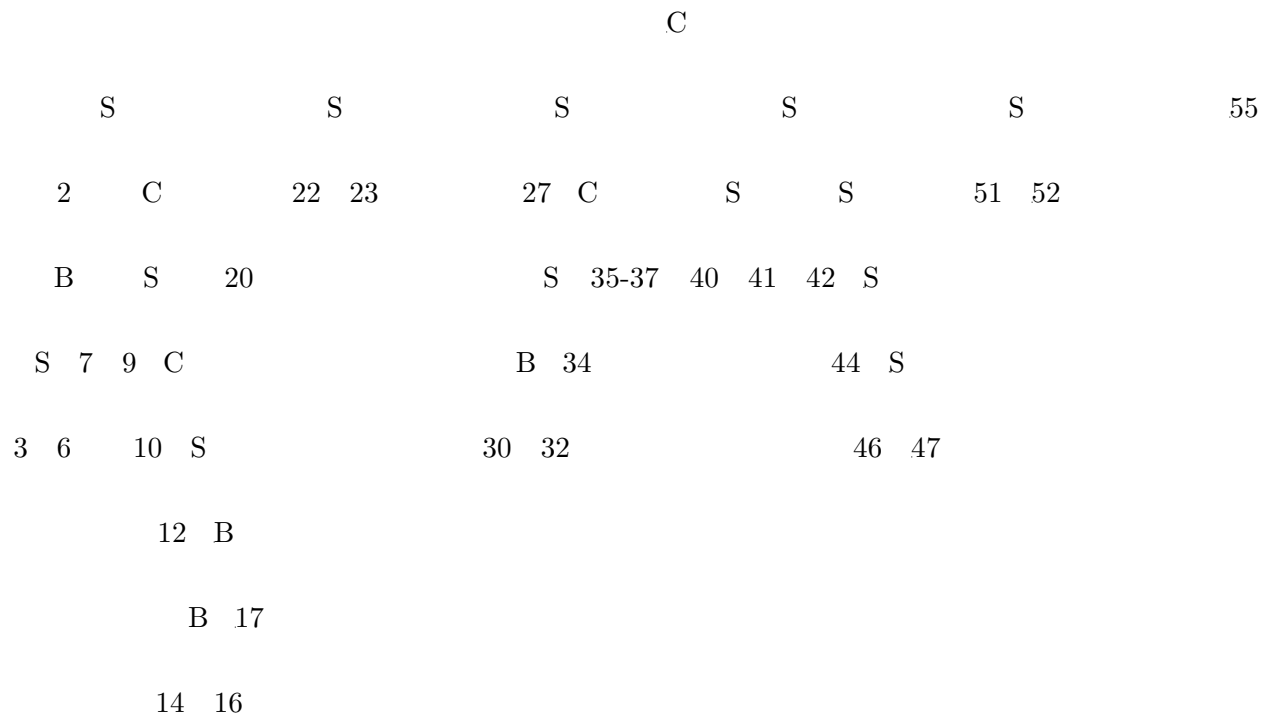
10	*ya che we'*	because that one
11	*komo*	given that
12	ti tz'u'um (xxx)	it's raw leather,
13	LA	and
14	heel ch'aak'	if it gets wet,
15	-*he* (xx)	at that time
16	tu'hbi che we'	that one stinks
17	heel muu kaatsa'	if you don't dry it.
18	-*a*" (xxx)	right there and then"
19	"*ah*	"Ah
20	mehor kiitsa'a'" (xxx)	I better dry it"
21	LA	and
22	tsa'a' (x)	he dried it
23	kene' (x)	leaving it
24	ti tz'a'ik (xxxx)	in the sun.
25	LA	And
26	*komo*	given that
27	xk'upu eel naalang (xx)	he cut a piece off
28	-*he* (xx)	at that time
29	"*komo*	"Given that
30	wila maaq ooki we' iixangab	I saw my sandals grew
31	*tzaani ni* (x)	now already
32	iik'uula emparehar iixangab (x)	I evened up my sandal,
33	*tzaani*	now
34	galan ni kene'" (xxx)	it's good."
35	"*ah*	"Ah
36	galan	good
37	*bi*'" (xxx)	good."
38	LA	And
39	*qeetoo*	as for us,

40	naaba chqamuuquoo	we were just looking at him,
41	moocho' cho' chqalaoo' cheet (xx)	we weren't speaking to him.
42	"moocho' cho' kaawalaqe' cheet (xx)	"Don't you tell him anything,
43	*ya che we'*	because
44	muu ni tz'aqwi ti chooq-	it won't fit his foot.
45	-*a* (xxx)	right (there)
46	iti tzapan xwaa chooki (x)	Little bitty it's gonna get.
47	kaawilaqe' to ban (xxxx)	You'll see.
48	-*a*" (xxxx)	right then and there."
49	LA	and
50	*qeetoo*	as for us,
51	moocho' qamikoonoo (xxx)	we didn't say anything.
52	naaba chqamuuquoo' (xx)	We just kept looking at him.
53	LA	And
54	*cheet*	as for him,
55	xi' aahliino (xxx)	he went off to work.

The DPT for this episode is constructed left to right using rules for discourse coordination (narrative list), discourse subordination (elaboration, direct discourse), and certain binary constructions (IF/THEN, THEN/IF and THEREFORE)[13] as well as rules of sentential syntax which describe complex (coordinated and subordinated) intrasentential constructions.

(20)

[13] Variations of essentially the same binary rules are to be found in taxonomies of discourse relations such as those given in Hobbs 1972, Longacre 1976, Matthiessen and Thompson 1988, though most of these authors do not use the formalism of context-free rewrite rules to express the regularities they discuss.



Inspecting the DPT in (20) we noticed that event or durative/descriptive narrational dcus alternate with sections of (embedded) direct discourse. As examination of a larger corpus of Mocho narratives carried out in the LDM framework has revealed, direct discourse interludes are mandatory constituents in episodes of Mocho stories. In the genre rules given in (21) we use e for event dcus, d for direct discourse, and r for reporting dcus (which follow the dcus reported, and are optional in Mocho).

(21)

EPISODE $\rightarrow e$ BLOCK^{+} e

BLOCK $\rightarrow$ REPORT^{+} e^{+}

REPORT $\rightarrow d^{+}$ (r)

According to these rules, episodes always begin and end in event clauses, and have one or more BLOCKS in the middle. A block contains one or more REPORTS of direct discourse, followed by one or more event clauses. Finally, a report is composed of one or more direct discourse dcus, optionally followed by a reporting dcu. We emphasize here that the genre rules in (21) are not used as input to the discourse parsing

process: we built the DPT in (20) by the general algorithm described in 3.4 and abstracted the genre rules from inspection of several parse trees.

Let us return to the DPT in (20) and ask where the *la*s are attached. The first and the last instances (19.1 and 19.53) are proclitic to the initial and final event dcus (19.2 and 19.55) and because these dcus are peripheral to the whole structure, we are justified in treating them by special rule. Of the remaining five instances, two (19.13 and 19.25) precede event dcus followed with *-he*, and both of these cases can be seen as preposed (binary) clauses in a logical (then/if) relation. Thus we are left with three instances (19.25, 19.38, and 19.49), each of which signal return from embedded direct discourse to the main event line. The generalization emerging from this analysis is that *la* will not occur before any event or durative descriptive dcu, but only before the first *e* in a sequence which attaches to the DPT at a node higher than the attachment point of the immediate preceding constituent, and only after direct discourse (aside from the peripheral and the *la-he* cases).

To sum up, then, Mocho dcu-initial *la* in narrative will occur on main clauses (events and states) which pop to higher levels of discourse structure, as an obligatory mark of popping from reported speech (unless that pop is marked by an attributive verb of speaking). Thus, while *la* can be characterized as a connective, it is more properly seen as a pop marker. This in itself is not surprising – many discourse particles such as English *so*, *anyway,* and *and then* function as discourse pop operators. However, it would be unsatisfactory to say that there is a *la* of coordination and a *la* of temporal succession, just as with English *and*, for two reasons. First, this does not solve the analytical problem of describing the distribution of Mocho *la*, merely 'reduces' it to the (actually more complex) problem of describing the distribution of English *and*. Second, it misses the linguistically significant generalization, true for Mocho *la*, but false for English *and*, that following structurally required reported speech, *la* conjoins the saying event to a subsequent action.

It should be emphasized that we have arrived at a precise description of the environment in which this particle occurs in narrative discourse by mechanically deriving a DPT without *la,* and using this tree as our frame of reference in expressing the distribution of *la* in discourse theoretic terms. No such noncircular analytic

procedure is offered by any other existing method of discourse analysis. Using the DPT as a heuristic tool, we can also discern the relevant genre units and gradually write a genre grammar. Once such a grammar is available we can even go back and restate the generalization in terms of the genre units:

(22)

EPISODE $\rightarrow$ *la e* BLOCK^{+} *la e*

BLOCK $\rightarrow$ REPORT^{+} *la* e^{+}

REPORT $\rightarrow$ d^{+} (r)

$*r$ *la*[14]

Genre-specific regularities are central to many analytic traditions from ethnopoetics to computational linguistics, and we do not belittle their importance. But applying the genre rules mechanically will rarely lead to a satisfactory parse. The reason for this is that many sentential or higher constructions, such as the *la-he* discussed above, are really independent of the genre, and will block the parse according to (21-22). Extending the genre grammar by special rules to cover all such cases is both practically unfeasible (it cannot be realistically hoped that all such constructions can be enumerated in advance) and theoretically unjustifiable, since the overwhelming majority of such constructions is not specific to any genre. In our model the analysis requires no genre-specific knowledge. Rather, the analysis is based on rules expressing general (and aside from some language-specific binary rules, universal) discourse-level grammatical competence, and therefore takes the input at face value, as if it were a generic piece of discourse not subject to any genre constraints. While the genre constraints are seldom available to the analyst in advance, the universal rules of the grammar are always there.

4. Discourse interpretation At the beginning of this paper we started out with three questions: What are the atomic units of discourse? What kind of struc-

[14] This rule is a surface filter expressing the complementary distribution of *la* and verbs of speaking. It can be eliminated in favor of distinguishing reports with or without a final reporting clause, at the price of introducing two versions of the block rule, one referring to reports with, and the other to reports without, r.

tures can be built from the elementary units? How do we interpret the resulting structures semantically? Having introduced the dcu and discourse operator and having described how complex discourse structures can be built, we are now ready to address the problem of discourse interpretation. In 4.1 we introduce the representations central to the semantic component of the model, DISCOURSE WORLDS, and in 4.2 we address the issue of what kind of insights we can gain into the structure of language with this model.

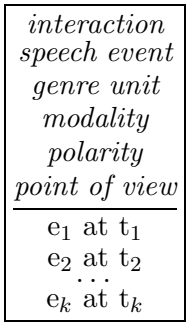
4.1. DISCOURSE WORLDS. Contemporary semantic theory has a great deal to say about isolated propositions, and we believe that a model-theoretic component along the lines of Montague (1973) or Groenendijk and Stokhof (1991) is indispensable for elucidating the meaning of natural language utterances. Yet we find it necessary to use a richer notion of semantic representation both for individual dcus and for larger structures. In addition to the propositional CONTENT of a dcu we will also talk of its CONTEXT and use the formal mechanism of INDEXATION to express the fundamental dependency of propositional content on context.

To some extent, the importance of context has already been recognized in sentential semantics, especially for lexical items such as indexicals where interpretation clearly depends on the identity and location of the speaker (Kaplan 1992). There is an attempt in situation semantics (Barwise and Perry 1983) to incorporate spatio-temporal and polarity indexes, and a growing recognition in the formal semantics community that modality plays a very similar role (Roberts 1987, Farkas 1994). The contribution of the present paper to this enterprise is a systematic, and we believe exhaustive, survey of contextual categories for which an indexical treatment is appropriate. We will argue that the range of contextual categories is considerably broader than generally assumed, and present a hierarchy (partial ordering) of contexts INTERACTION>SPEECH EVENT>GENRE UNIT>MODALITY>POLARITY>POINT OF VIEW.

To the extent that DISCOURSE WORLDS play a more central role in our semantics than the standard notion of possible worlds, our theory is an extension of Discourse Representation Theory (Kamp 1981) and File Change Semantics (Heim 1982), and

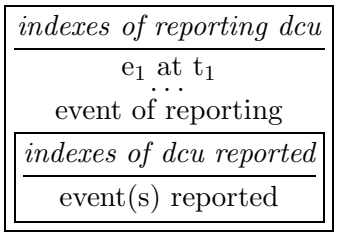
the graphical similarity between our representations in (2-3) above and those used in DRT/FCS is intentional. But the top half of the representation, which in these theories is used for keeping track of variables, will in our discourse worlds be used to keep track of contextual indexes. The change in notation reflects a shift in emphasis. While the central concern of DRT was pronominal reference and the equations between variables that implement coreferentiality, the central concern of the LDM is the setting and resetting of contexts. In the simplest case, we depict a discourse world as in (23):

(23)



Conceptually our discourse worlds are quite similar to Fauconnier's (1985) MENTAL SPACES, Webber's (1986:397) DISCOURSE MODELS, Kameyama's (1992) DISCOURSE SITUATIONS, and Dinsmore's (1993) SPACES, but unlike these authors, we wish to make no cognitive claim about these structures. We treat discourse worlds as purely technical devices of semantics, no more mentally real than variables or generalized quantifiers. For our purposes, discourse worlds are simply intermediate representations between natural language expressions and model structures, much as in DRT. Since discourse worlds are constructed by the discourse, truth relative to discourse worlds is automatic. Truth relative to model structures depends on the embeddability of discourse worlds in possible worlds and will not play a central role in our discussion (for more recent formal semantic work in this direction, see Prüst 1992, Asher 1993). Rather, we talk about the embedding of one discourse world in another, as in the case of reported speech depicted in (24) below:

(24)



In general, discourse worlds can be recursively (but not cyclically) embedded in one another, though, aside from highly crafted literary works such as the Pancatantra, the level of embedding is generally not very deep. Also, they can be related to one another by logical and other relations, which we will indicate by arrows running between the related structures. How logical inferences are drawn on the basis of such relations is a matter too complex and digressive to discuss here (for recent work in this direction see Lascarides and Asher 1991, 1993, Asher 1993, Dinsmore 1993, Farkas 1994) but a brief remark about being the 'same' individual in different discourse world is in order. In the standard model theoretic treatment of cross-world identity (Kripke 1971) the relation of being the same is necessarily symmetric, while in the *want* examples in (5) and (6) the role of the realis and the irrealis individual is asymmetric. Therefore, the use of an explicit COUNTERPART relation (Lewis 1968) seems more appropriate for these cases. Let us also note that discourse worlds are endowed with their own timeline, so that a series of events sharing the same contexts are interpreted relative to a single discourse world. As we shall see shortly, this temporal structure allows for a richer set of representational possibilities than permitted by classical temporal modal logic (for a survey, see van Benthem 1983).

Over the years, many proposals have been put forward to characterize the information status of discourse entities as being in some sense sense NEW in the discourse or OLD (Firbas 1962, Halliday 1967, Dahl 1969, Gundel 1974, Chafe 1976, Prince 1981, 1992, Givon 1983a, Grosz et al. 1983, Sidner 1986, Webber 1991, Valduví 1992). While we do not question the basic intuition that the distribution of sentential syntactic constructions can be conditioned by the relative novelty in the discourse of the information to be conveyed, we see most of these proposals, with the exception of Sidner 1986, as technically inadequate because they consider all

discourse to be a string of clauses with flat (coordinated) structure.

By failing to account for the structure in which the appearance of a reference to an entity can count as new, attempts to characterize discourse constraints on sentential syntax may not make correct predictions in some cases. A particularly striking example of such a failure is discussed in Prince 1988, who concludes that the basic generalization about the Yiddish expletive *es* and postposed subject construction, namely that postposed subjects of *es*-sentences may not represent entities already evoked in the discourse, can only be maintained

> with the unsurprising caveat that discourses have internal structure and may themselves include sub-discourses in each of which some discourse entity may be new (Prince 1991:189).

Here we outlined the basic contextual machinery necessary for structurally describing subdiscourses and thereby delimiting the domain of rules sensitive to information status. To summarize the indexes briefly, at the top level we find the INTERACTION, which is determined by a set of potential participants in a spatiotemporal context. In actual interactions, these slots are filled by actual individuals, locations, and times. Next comes the SPEECH EVENT, where instead of the physical identity of the person their social identity is considered: this slot is filled by socially constructed roles such as JUDGE or WITNESS. Speech events may be composed of smaller constituents called MOVES. While in interactive discourse a great deal of the explanatory burden falls on the interaction and speech event indexes, in stories, arguments, and other genres these stay constant and a greater role is played by GENRE CONSTITUENTS such as invocations, episodes, etc.

The hierarchical ordering of indexes means that any change in the higher index forces a change in the lower index by default. For example, interactional indexes are higher than speech event indexes, because one physical person can smoothly assume a series of socially constructed roles, but the role cannot be assumed by a succession of persons without explicitely establishing the fact that they now fill the role.[15] Similarly, within a single genre constituent there can be several smooth changes of

[15] The default can be conventionally suspended in the theater and other situations where 'true' physical identity is of little relevance.

modality, polarity, and point of view indexes, but a shift in genre constituent always resets the world relative to which modal, polarity, and point of view distinctions are interpreted. Since the necessity of indexical treatment of modality (Roberts 1986, Farkas 1994) and polarity (van Dijk 1977, Barwise and Perry 1983) has been extensively argued in the linguistic literature, in 4.2 we turn to the less well know POINT OF VIEW indexes.

4.2. EXTRACTION AND INSTANTIATION Point of view indexes comprise several conceptually related but grammatically distinct notions such as EMPATHY (Kuno 1977, Kameyama 1985), KNOWLEDGE STATE (Barwise 1989:ch9), and EVIDENTIALITY (Chafe and Nichols 1986). Here we will concentrate on point of view indexes that have a clearly detectable impact on the temporal interpretation of propositions in narrative discourse, a subject extensively investigated since the 1960s. Most researchers subscribe to one or another version of the SEQUENTIALITY CONSTRAINT (SC) that temporal reference in narrative discourse is advanced by telic event clauses which encode non-iterative, non-habitual and temporally bounded propositions in syntactically main clauses (Labov and Waletsky 1967, Vendler 1967, Hopper and Thompson 1980, Kamp and Rohrer 1982, Hinrichs 1981, 1986, Dowty 1986). The SC predicts that clauses $c_1, c_2, \ldots$ encoding event propositions $e_1, e_2, \ldots$ are interpreted to hold true at discrete time instants $t_1, t_2, \ldots$ so that if c_i (immediately) preceded c_j, t_i must (immediately) precede t_j.

Some exceptions to the SC have long been recognized. Event clauses occurring in FLASHBACKS/FLASHAHEADS or MEANWHILE EPISODES are known be interpreted at temporal reference points which precede/follow or may coincide with the last event on the temporal main line. Recently, attention has been focussed on the temporal interpretation of clauses such as *Bill pushed him,* which in the discourse *John fell. Bill pushed him.* may be used to provide explanatory, motivating circumstances for a temporally subsequent event described in a textually preceding clause (Moens 1987, Moens and Steedman 1988, Lascarides and Oberlander 1991, Asher and Lascarides 1993, Kehler 1994). In (25.3-6) we illustrate yet another problematic case for the SC: the GOVERNED MAIN CLAUSE (GMC).

(25) *1. John went to the kitchen. 2. He got the ice cream out of the freezer 3. He opened the door 4. found the unopened container way in the back 5. and grabbed it. 6. Then he jerked it out with one powerful motion. 7. He got himself a bowl from the cupboard 8. served out a big helping 9. and went to the living room to watch TV.*

GMCs are well formed event clauses which are related to the events of the main narrative but do not advance its time line. Semantically, GMCs are elaborations, a special case of our subordination rule (9). In narrative GMC constructions, the granularity of vision of the events in question differs in the two discourse worlds: the governed dcus (25.3-6) provide a more complex, finer grained breakdown of the event in the structurally superior dcu (25.2). The common factor in flashbacks, explanatory/evidential clauses, and the GMC is the shift in point of view, and in order to protect the SC against such counterexamples we must treat it as a regularity that holds only for clauses with the same point of view indexes. To see how this idea can be technically implemented, we will examine a narrative example adapted from a story reported by Linde 1980 in which the narrator recounts how she and her husband decided to take an apartment they had been considering renting:

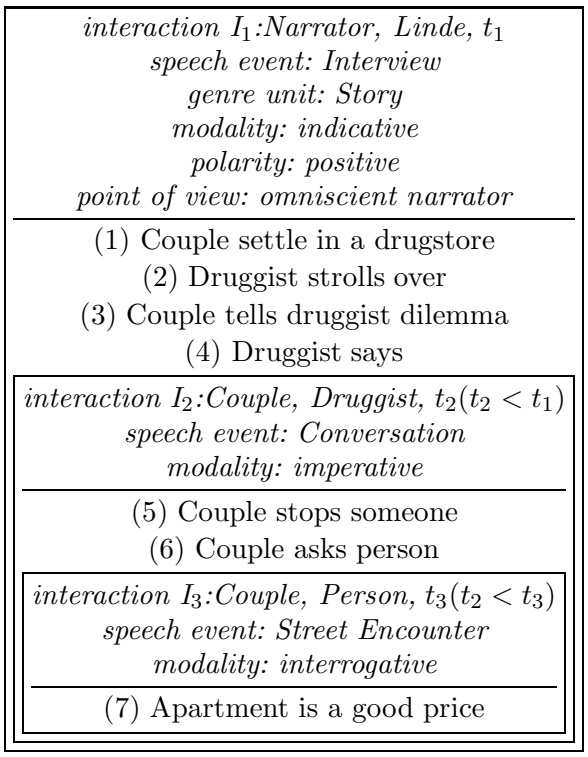
(26) *1. We settled down for a cup of coffee in a drugstore 2. The druggist strolled over to us 3. and we told him our dilemma 4. He told us 5. 'Stop someone on the street, 6. Ask him 7. if the apartment is a good price.' 8. So 9. that's 10. what we did 11. We went to the street corner 12. We stopped a married couple 13. We asked them 14. if we should take the apartment 15. They said 16. 'yes' 17. So we went back 18. and rented it.*

In this short excerpt there are three references to **stopping someone on the street and asking if the apartment is a good price**, encoded in clauses 5-7, 9-10, and 12-14. Nonetheless, we interpret the text as asserting that in the realis world of the story as narrated **stopping and asking** took place only once. Clearly, 12-14 are GMCs expanding on 9-10. Had the discourse terminated at 10, we would have still known that people were stopped and asked about the apartment, exactly as the druggist had suggested. The information that in the realis world of the narrative mainline the advice of the druggist had been followed is made available by two

important discourse semantic processes: DISCOURSE EXTRACTION (which we will refer to as D-EXTRACTION to differentiate it from its sentential counterpart) and DISCOURSE INSTANTIATION (D-INSTANTIATION).

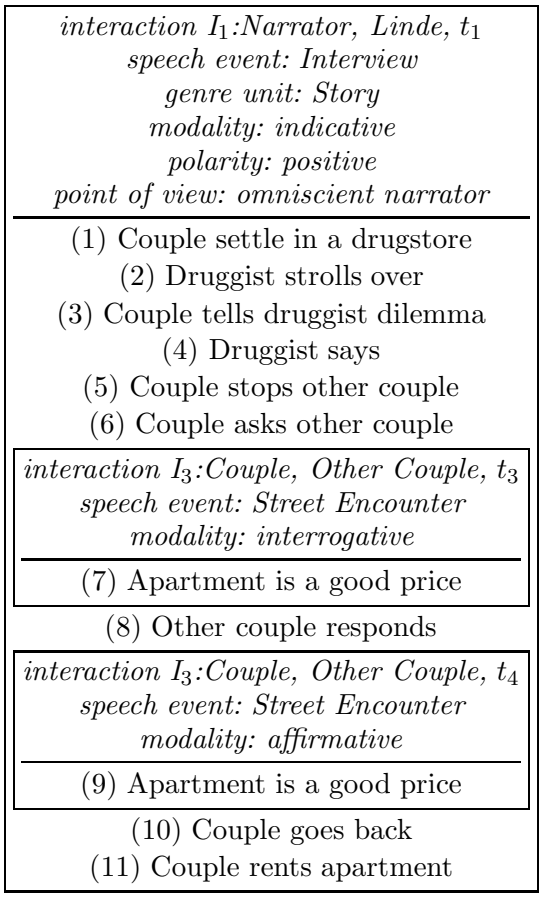
D-EXTRACTION identifies an object in some world **A** which is structurally accessible at a node on the right edge of the DPT, and D-INSTANTIATION establishes a counterpart of the d-extracted object in a dcu interpreted in **World B.** In 9, the anaphoric *that* of the matrix *that's* triggers d-extraction of the entire set of semantic objects in world **A** that make up the interpretation of the C node dominating dcus 5-7, which is still accessible. These objects are instantiated in the realis discourse world of the mainline narrative by the verb of *what we did*. Before these steps, the semantic representation of (26.1-7) is as follows:

(27)



After the extraction steps, the mainline is extended with the events of I_2, now realis, while the material embedded under **druggist says** is now closed off:

(28)



As (28) shows, reported dialog is always embedded relative to the verb of speaking, and always invokes a separate discourse world. It requires a rather complex technical machinery with a number of default assumptions to conclude that the same interaction I_3 (but different speakers) is being reported in both instances. Semanticists will no doubt also note the problems with event anaphors, the delicate issues of maintaining identity between the druggist's *you* and the narrator's *we*, and in keeping this couple distinct from the *married couple* of the text. By the same token, syntacticians could cite the serious unresolved problems with constructions like *apartment is good price* and phonologists could point to the difficulties of segmental phonology we would encounter in attempting to transcribe Linde's original recording.

Because such complexities abound, many linguists think that an attempt to analyze discourse is premature. What we attempted to show with this example is that

the linguistic study of discourse need not be delayed until all other problems of linguistics are resolved. A more exact statement of the d-extraction and d-instantiation rules must await another occasion. Yet the basic insight, that discourse has its own grammar, and that the rules of this grammar are just as amenable to linguistic study as the rules of other domains, need no longer be treated as an article of faith, because the LDM provides a detailed theory of how the structural description and structural change of these rules can be stated.

5. Summary and conclusions. Let us first summarize how the model presented in these pages answers the basic questions we started out with. In the LDM, the basic unit of discourse formation is the discourse constituent unit or dcu, a semantically motivated semiotic structure that expresses a single event or state of affairs in some discourse world. When linguistically encoded, the elementary dcu will typically have a syntactic realization as a single clause, and phonological realization as a single phonological phrase. Elementary dcus are always indexed for context including physical and socal situation of utterance (real or modeled), genre unit, modality, polarity, and point of view. We also recognize discourse operators as elementary structures. These are non-propositional in nature and give information about the relationship (both structural and semantic) obtaining among the dcus as well as pragmatic information linking the discourse to its utterance situation.

Three basic kinds of structures can be built from the elementary dcus or from larger units: *C*oordination (creating lists of various types), *S*ubordination (interruptions, elaborations), and *B*inary structures (adjacency pairs, logical relations, rhetorical structures). Complex discourse is formed by the recursive sequencing and embedding of elementary dcus to dcus. The LDM can be formulated as a limited lookahead parser which accepts a stream of elementary dcus as its input and builds a structural description of the discourse and a semantic representation in lockstep. The structural description of any discourse is an open right Discourse Parse Tree (DPT) composed of S, C, and B nonterminal nodes and elementary dcus as terminal nodes. Each incoming elementary dcu is attached to the developing DPT as the right daughter of an appropriate existing or newly created node along the

right edge of the DPT. Once a node is no longer on the right edge it is no longer a candidate for new dcu attachment.

As a new dcu is attached to the DPT the discourse semantics is also updated, either by adding a new event to an existing world, or by creating a new world and adding the event to it. Worlds can be subordinated (embedded), coordinated, or stand in various logical relations to one another. Though there is a rule-to-rule correspondence (each syntactic rule is coupled with a semantic rule), the semantic structure need not mirror the syntactic structure fully. For example, syntactic subordination of unrelated material, as in interruptions and digressions, does not give rise to semantic subordination, but rather to parallel (coordinated) discourse worlds. In this paper we adopted a generalized version of DRT with Davidsonian events and Lewisian counterparts, but at the level of discourse syntax nothing hinges on these particular choices. For the correct interpretation of deixis, anaphors, and in general for drawing the right inferences we emphasized the importance of finding the relevant structural units, and argued that these are often determined by contexts such as the INTERACTION, SPEECH EVENT, GENRE UNIT, MODALITY, POLARITY, and POINT OF VIEW which can change several times in a single discourse.

We have dealt with discourse as an autonomous linguistic module while, almost paradoxically, insisting that the physical and social identity of the speaker are of crucial importance in discourse interpretation. Where, one might wonder, is the speaker in this theory of discourse? The answer is that accounting for the social concerns, motivations and actions of the speaker along with the cognitive processing apparatus brought into play during discourse production and reception lay well beyond the scope of the present paper. Nonetheless, the model of discourse presented here is potentially of use to the working sociolinguist concerned with the analysis and manipulation of complex interactive data, and the psycholinguist interested in understanding the nature of linguistic competence and performance.

For the sociolinguist, we offer analytic machinery which can handle incomplete utterances, hesitations, repairs, interruptions and changes in social roles and identities (for a survey of contemporary work in the interactive and cultural dimensions of language use see Duranti and Goodwin 1992). The indexing and segmentation

requirements allow the sociolinguist to track what is going on in the discourse in a more consistent manner than has been possible previously. In addition, the definition of the dcu permits the form of encoding of propositional (or operator) material to be non-linguistic. Deictic points, grimaces, or the actions of a machine may all be integrated into the discourse history. The structures of specific instances of a socially recognized speech event can be compared with one another and far more robust descriptions of the sequences of expected actions can be produced.

To the psycholinguist we offer the opportunity to formulate testable hypotheses about discourse processing and to investigate the relationship between discourse structure, sentence form, and memory limitations in terms of an integrated framework. Although we make no specific cognitive claims and pointedly avoid using psychologically appealing terms such as MENTAL MODEL, SALIENCE, or ATTENTION, we nonetheless provide a semantic representation in terms of which one can inquire into the mental model any given speaker might build, the differential salience accorded by a speaker to the entities and events in that model, and the degree of attention entities command (see Levelt 1989 for an overview of much relevant work in this area).

At the beginning of the paper we considered some design criteria for a theory of discourse. Certain criteria, such as avoiding undefined primitives or circularity, are common to all theories, and required special mention only because they are often neglected in theories of discourse. Others, such as the ability to parse any input with limited lookahead, are specific to the generative enterprise and might well be rejected by those who do not view psychological reality as an important goal.[16] But there is one criterion, that of requiring the model to cover all kinds of naturally occurring input, which does not neatly fall into either of these categories. From a general perspective, an empirical field of study can hardly ignore the actual data, and interactive discourse is notoriously disorganized and 'incoherent'. But from the perspective of the syntactician, who routinely deals with such

[16] Highly crafted literary and scientific works which require several passes to fully understand are still constrained by the general rules of discourse formation, but these constraints do not fully determine the higher structure.

data by taking ideally GRAMMATICAL sentences to be the central object of inquiry and analyzes UNGRAMMATICAL sentences only in the light of the grammatical ones, extending this methodology from sentences to larger units as in TEXT LINGUISTICS (for a bibliographic survey see Lohmann 1988, 1989) appears almost inevitable. For example, van Dijk (1977:40) uses (29) and (30) to exemplify the difference between grammatical and ungrammatical TEXTS:

(29) *Mary was badly wounded in a car crash yesterday. Mary lies in a hospital.*

(30) **We will have guests for dinner. Calderón was a great Spanish writer.*

From the perspective developed in this paper it is obvious that there are many ways to construe (30) as grammatical. First, (30.1) and (30.2) can be parts of different interactions, just as in our example (1). Even if addressed to the same listener, (30.1) can be an aside after which (30.2) returns to the main line or the other way around, (30.1) can be part of a larger ongoing discourse interrupted by (30.2). In fact, it does not require a great deal of ingenuity to create a context (such as the planning stage of a conference of literary theorists) where the speaker of (30) can continue with (31):

(31) *They will have a lot to say about him.*

Since *they* refers to the *guests,* and *him* refers to *Calderón,* it is clear that (30) and (31) together form a highly coherent text. Unlike the grammaticality of sentences, which can be (almost by definition) assessed in isolation, the grammaticality of a discourse fragment, unless structurally unaccessible, is always subject to revision. Consider the following example:

(32) *1. We will have guests for dinner. 2. Calderón was a great Spanish writer. 3. And so was Cervantes. 4. They both had a lot to say 5. about the table manners of Spanish nobility. 6. But when you see 7. Don Juan or Donna Maria addressing the host, 8. you'll still 9. be totally amazed.*

At first, all we know about the relationship of (32.1) and (32.2) is that it is not one of coordination.[17] Therefore we subordinate (32.2) to (32.1) as semantically

[17] Unless it is a metalinguistic coordination of some random sentences being read aloud, a possibility mentioned by van Dijk. Such metalinguistic usage would be accompanied by a characteristic

unrelated. Next, a coordination node C is inserted between this S and the terminal node corresponding to (32.2), and this C will now also dominate (32.3). By the time the listener gets the first hint of how (32.1) and (32.2) will eventually be related, the path between these two terminals goes through four nonterminal nodes.

In sentential syntax the dependencies that can hold between main verb and other units are limited, because the list of grammatical relations is finite. But in discourse syntax the variety of structural relations between two adjacent dcus is unlimited, because between two sister dcus originally under the same S node, arbitrarily many new nonterminal nodes can be added by subsequent dcus. The notion of grammaticality can always be viewed relative to a parse tree. In sentential syntax, this would be a distinction without a difference, since the number of possible parse trees over a sentence (string of words) is fixed once and for all, and 'grammatical' can be used without loss to mean 'grammatical relative to at least one parse tree'. But as our example (32) shows, in discourse it is quite possible to restore a seemingly ungrammatical string of dcus to full grammaticality by subsequent dcus, which makes the absolute notion of grammaticality or coherence uninteresting. This is not to say that all strings of dcus can always be restored to full grammaticality by some clever combination of continuing dcus – as we have argued, structurally closed constituent can no longer be affected. Rather, the point that we wished to underscore with this example was simply that acontextual notions of grammaticality or coherence are not the appropriate tools for the analysis of discourse.

In this paper we have provided a better set of tools for systematic investigation of discourse level linguistic phenomena. We have made explicit the nature of our atomic units, the rules for combining them into more complex structures, and the framework in which both simple and constructed units may be interpreted. Linguists, especially the more formally minded, are often held back from the study of discourse by the belief, strongly felt though seldom clearly articulated, that discourse itself is simply an unstructured soup of sentences. Our goal has been to demonstrate that this belief is false: a theoretically well founded characterization of the domain of

list intonation pattern.

rule applicability and the distribution of linguistic structures in discourse is both possible and necessary.

6. References

Asher, Nicholas. 1993. *Reference to abstract objects in discourse*. Dordrecht: Kluwer.

Asher, Nicholas and Alex Lascarides 1994. Intentions and information in discourse. Proceedings of the 1994 Annual Meeting of the ACL.

Barsalou, Lawrence.1981. *Determinants of graded structure in categories*. Ph.D. Dissetation. Stanford CA: Department of Psychology Stanford University.

Barwise, Jon. *The situation in logic*. Stanford, CA: CSLI.

—— and John Perry. 1983. Situations and Attitudes. Cambridge,MA. MIT Press.

van Bentham, Johan. 1991. *The logic of time*. Dordrecht: Reidel.

Birner, Betty. 1994. Information Status and Word Order: An analysis of English inversion. *Lg.* 70:2 pp 233-259.

Chafe, Wallace. 1976. Givennes, contrastiveness, definiteness, subjects, topics and points of view. In Charles Li ed. *Subject and topic*. New York: Academic Press.

—— 1980. The deployment of consciousness in the production of a narrative in Wallace Chafe ed. *The Pear stories: Cognitive, cultural, and linguistic aspects of narrative production* Norwood NJ: Ablex publishing Co. 9:50.

—— 1988. *Linking Intonation Units in Spoken English.* In John Haiman and Sandra A. Thompson (eds.) Clause Combining in Grammar and Discourse. Amsterdam: John Benjamins Publishing Company. 1-29.

Chafe, Wallace and Johanna Nichols (eds). 1986. *Evidentiality : the linguistic coding of epistemology.* Ablex, Norwood, N.J.

Church, Kenneth. 1980. *On memory limitations in natural language processing* Technical Report 245. Laboratory for Computer Science, MIT.

Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Pescher ed. *The logical form of action sentences*. Pittsburgh: University of Pittsburgh Press, 81-95.

Dahl, Osten. 1969. *Topic and comment* Stockholm: Almqvist and Wiksell.

van Dijk, Teun A. 1972. *Some aspects of text grammar.*The Hague: Mouton.

—— 1977. Text and Context. Cambridge: Cambridge University Press.

Dinsmore, John. 1991. *Partitioned Representations.*Dordrecht: Kluwer. Chafe, Wallace ed. 1980. *Pear stories.* Norwood, NJ:Ablex Publishing Co.

Doron, Edit. 1994. *Reader for "Seminar on Point of View"* Linguistics Institute, Columbus OH.

Dowty, David 1986. *The effects of aspectual class on temporal structure of discourse: semantics or pragmatics?* Linguistics and Philosophy 9(1): 37-61

Dubois, John. 1987. *The discourse basis of ergativity. Lg.* 63:4. 805-855.

Dressler, Wolfgang U. 1978. *Current trends in textlinguistics.* Berlin: Walter de Gruyter.

Duranti, Alessandro and Charles Goodwin. 1992. *Rethinking context: language as an interactive phenomenon.* NY: Cambridge University Press.

Farkas, Donka. In prep. *Modal anchoring and NP scope.***1992/1994?**

Fauconnier, Giles. 1985. *Mental Spaces.* Cambridge MA: MIT Press.

Firbas, Jan. 1966. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1:267-280.

Fox, Barbara A. 1987. *Discourse structure and anaphora: written and conversational English.* Cambridge: Cambridge University Press.

Gee, James and F. Grosjean 1983. *Performance Structures:A Psycholinguistic and Linguistics Appraisal* in Cognitive Psychology 11:58-81.

Givon, Talmy. 1983a. Topic continuity in discourse: an introduction. in Talmy Givon ed. *Topic continuity in discourse:a quantitative cross-language study.* Philadelphia: John Benjamins. 1-41.

—— 1983b. Topic continuity in discourse: the functional domain of switch reference, in John Haiman and Pamela Munro (ed.) *Switch reference and universal grammar.* Amsterdam: John Benjamins Publishing Co. 51-82.

—— 1983c. Topic dcontinuity and word-order placement in Ute. in Talmy Givon ed. *Topic continuity in discourse:a quantitative cross-language study.* Philadelphia: John Benjamins. 141-213.

Goffman, Erving. 1979. *Footing.* Semiotica 25 1-29

Goodwin, Charles. *Conversational organization: interaction between speakers and hearers.* NY: Academic Press.

—— and Alessandro Duranti, 1992. Rethinking context: an introduction. In Duranti and Goodwin. 1-42.

Groenendijk, Jeroen and Martin and Stokhof. 1991. Dynamic Predicate Logic. *Linguistics and philosophy.* 14(1)39-100.

Grover, Claire, Chris Brew, Suresh Manandhar, and Marc Moens. 1994. Priority union and generalization in discourse grammar. In *Proceedings of the 1994 ACL*

Grosz, Barbara. 1977. *The representation and use of focus in dialogue understanding.* Unpublished Ph.D. Dissertation. Berkeley, CA: Department of Computer Science. University of California and Berkeley.

——, Aravind Joshi and Scott Weinstein. 1983. Providing a unified account of definite noun phrases in discourse.*Journal of the Association of Computational Linguistics* 21:44-50.

—— 1986. *Attention, Intention and the Structure of Discourse* Computational Linguistics: 175-204.

Gülich, Elisabeth 1970. Makrosyntax der Gliederungssignale im Gesprochnen Französisch. Munich. Wilhelm Fink Verlag.

Gumperz, John. 1982. *Discourse Markers.* NY: Cambridge university press.

—— 1992. Contextualization and understanding. In Duranti and Goodwin. 229-252.

Gundel, Jeanette. 1974. *The role of topic anc omment in linguistic theory.* Ph.D. Dissertation. Austin: University of Texas, Departmkent of Linguistics.

Halliday, Michael A.K. 1967. Notes on transitivity and theme, part 2. *Jornal of linguistics* 3.

—— 1977.Text as semantic choice in social contexts. in Teun A, van Dijk and János Petőfi eds. *Grammars and description.* Berlin: de Gruyter.

—— and Ruquaiya Hasan 1972. *Cohesion in English.* Longman.

Hanks, William. 1992. The indexical ground of deictic reference. In Alessandro Duranti and Charles Goodwin (eds) Cambridge: Cambridge University Press. 43-76.

Hasan, Ruquaiya. 1977.Text in the systemic-functional model. in Wolfgang Dressler, ed. *Current trends in text linguistics* Berlin: de Gruyter. 228:246.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases.* MIT dissertation. Department of Linguistics.

Heydrich, Wolfgang, Fritz Neubauer, János S. Petőfi and Emel Sozer. 1989. *Connectivity and Coherence: Analysis of text and discourse.* Berlin: de Gruyter.

Hinrichs, Erhard. 1985. *A compositional semantics for aktionsarten and NP reference in English.* Unpublished Ph.D. Dissertation. Columbus: Department of Linguistics. Ohio State University.

—— 1986. *Temporal anaphora in discourse of English.* Linguistics and Philosophy 9(1): 63-82

Hirschberg, Julia and Diane Litman. 1987. *Now let's talk about Now: Identifying Cue Phrases Intonationally* in Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, 163-171.

Hirschberg, Julia and Janet Pierrehumbert 1986: *The intonational structuring of discourse.* ACL 24:136-144

Hobbs, Jerry 1979. *Coherence and Coreference.* Cognitive Science, 3(1), 67-82.

—— 1985. On the Coherence and Structure of Discourse. Center for the Study of Language and Information Technical Report 37, Stanford, CA

Hovy, Edward. 1991. Approaches to the planning of coherent text. In Cecile Paris, ed. *Natural language generation in artificial intelligence and computational linguistics.* Boston: Kluwer. 83-102.

Hymes, Dell S. 1972. *Models of the Interaction of Language and Social Setting.* Journal of Social Issues 23(3).

—— 1981. *"In vain I tried to tell you": Essays in Native American ethnopoetics.* Philadelphia: University of Pennsylvania Press.

—— 1987. *Tonkawa Poetics: John Rush Buffalo's 'Coyote and Eagle's Daughter'* In: Sherzer and Woodbury (eds)

Johnson-Laird, Philip. 1986. Mental Models. Cambridge MA: Harvard University Press.

Kameyama, Megumi. 1985. Zero Anaphora: The Case of Japanese. Unpublished Stanford University Ph. D. Dissertation.

—— 1992. Discourse understanding and world knowledge.*Journal of information processing* 15(3) 377-385.

Kamp, Hans A. 1981. *A Theory of Truth and Semantic Representation.*In Jeroen Groenendijk, Theo Janssen and Martin and Stokhof, eds. Formal Methods in the Study of Language, Part 1. Amsterdam: Mathematisch Centrum.

—— and Christian Rohrer. 1983. *Tense in texts*, in Roland Posner and George Meggele eds. *Meaning, use and interpretation.* Berlin: de Gruyter.

Kaplan, David. 1977. That. In Joseph Almog, JohnPery and Howard Wettstein eds. *Themes from Kaplan.* Oxford: University Press

Kaufman, Terrence S. 1967. *Preliminary Mocho vocabulary. Working Paper Number 5.* Berkeley: University of California Laboratory for Language.

Kehler, Andrew. 1994. Common topics and coherent situations: Interpreting ellipsis in the context of discourse inferencing, In *Proceedings of the 32nd Annual Meeing of the Association of Computational Linguistics.* 50-58.

Kendon, Adam 1981. *Conducting interaction: patterns of behavior in focused encounters.* Cambridge: Cambridge University Press.

Kripke, Saul. 1971. 'Identity and necessity', in M. Munitz, (ed.), *Identity and Individuation*, New York University Press, New York, pp. 135-164.

Kuno, Susumo. Generative Discouse analysis in America. in Wolfgang Dressler, ed. *Current Trends in Textlinguistics.* Berlin: de Gruyter. 275:294.

Labov William and Joshua Waletzky. 1967. Narrative Analysis. in June Helms ed.*Essays on the visual and verbal arts.* Seattle: University of Washington Press.

Ladd, D. Robert. 1986. *Intonational Phrasing: The Case for recursive prosodic structure* in Colin J. Ewen and John M. Anderson, eds. *Phonology Yearbook 3.* Cambridge: Cambridge University Press. pp. 311-340.

Lascarides, Alex and Nicholas Asher. 1993. Temporal interpretation, discourse relations, and common sense entailment. *Linguistics and Philosophy* 16(5):437-493.

Lascarides, Alex and Jon Oberlander.1992.Temporal coherence and defeasible reasoning. *Theoretical Linguistics.*

Levison, Stephen C. 1981. Some pre-observations on the modelling of dialogue. *Discourse Processes* 4:93-116.

—— 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, David. 1968. Counterpart theory and quantified modal logic, *Journal of Philosophy* 65:113-126.

Liberman, Mark V. 1975. The intonational system of Englih. MIT. Department of Linguistics Dissertation.

—— and Alan Prince. 1977. *On Stress and Linguistic Rhythm* in Linguistics Inquiry 6 pp. 533-554.

Lohmann, Patricia. 1988. Connectedness of texts: A bibliographical survey (Part I). in János S. Petőfi *Text and discourse constitution. Empirical aspects, theoretical approaches.* 478-501.

—— 1989. Connectedness of texts: A bibliographical survey (Part II). In Heydrich et al. 383-400.

Longacre, Robert E. 1976a. *An Anatomy of Speech Notions*. Lisse, The Netherlands: The Peter de Ridder Press.

—— 1976b. *Introduction to Volume 1* in Robert B. Longacre and Francis Woods eds. Discourse Grammar: Studies in Indigenous Languages of Columbia, Panama and Ecuador Part 1. Dallas TX: Summer Institute of Linguistics. pp. 1-20.

Mann, William C. and Sandra A. Thompson. 1987. *Relational Propositions in Discourse*. Discourse Processes 9 (1): 57-90.

Marcus, Mitch. 1980.*A theory of syntactic recognition for natural language* Cambridge MA:

Martin, Laura 1989. *The interdependence of language and culture in the bear story in Mocho and Spanish*. Anthropological Linguistics 29 (4) 533-48 (Winter 1987).

Matthiesson, Christian and Sandra Thompson. 1988. The structure of discourse and "subordination". in Haiman and Thompson. 275-330.

Moens, Marc. 1987. *Tense, aspect and temporal reference*. Unpublished Ph. D. Dissertation. Edinburgh.

Montague, Richard. 1973. The proper treatment of quantification in natural language. in Hintikka et al. (eds.) *Approaches to natural language*. Dordrecht: D.

Reidel.

Moore, Johanna and Cecile Paris. 1993. Planning texts for advisory dialogues: Capturing intentions and rhetorical information. In *Computational Linguistics* 9 (4) 651-694.

Nespor, Marina and Irene Vogel 1986. Prosodic Phonology. Dordrecht: Foris.

Olsvanger, I. ed. 1947. *Royte pomerantsen*. NY: Schocken Books.

Panchatantra. Freiburgh: B. z. Indiology. Bd 2.

Passonneau, Rebecca J. and Diane Litman. 1993. Intention based segmentation. *Proceedings of the 31st Meeting of the Associarion of Computational Linguistics*.

Payne, Doris. 1987. Information structuring in Papago narrative discourse. *Lg.* 63 (4) 783-804.

Pierrehumbert, Janet. 1980. *The phonetics and phonology of English intonation*. MIT: Ph.D. thesis, Department of Linguitics.

Polanyi, Livia. 1978. False starts can be true. Proceedings of BLS 4 Berkeley, CA.

—— 1985. A theory of discourse structure and discourse coherence, In *Proceedings of the 21st Annual Meeting of the Chicago Linguistics Society*. Chicago: Universaity of Chicago, Dept. of Linguistics,

—— 1987. *Keeping it all Straight: Interpreting Narrative Time in Real Discourse.* WCCFL 6: 229-245.

—— 1988. *A formal model of discourse structure*, Journal of Pragmatics 12: 601-638.

—— and Remko Scha. 1984. A syntactic approach to discourse semantics. Poceedings COLING. Stanford CA.

Prince, Ellen. 1978. *A Comparison of wh-clefts and it-clefts in discourse* Lg. 54. pp. 883-906.

—— 1981. *Toward a taxonomy of given/new information* in Peter Cole (ed.) Syntax and Semantics 9: Pragmatics. NY: Academic Press. pp. 245-59.

—— 1992. The ZPG letter:sbjects, definiteness and information status. Discourse description: diverse linguistic analyses of a fund raising text. ed. by William Mann and Sandra Thompson. 295-325. Philadelphia: John Benjamins.

Prust, Hub. 1992. *On Discourse Structuring, VP Anaphora and Gapping*. Unpublished PhD.dissertation. Alfa-Informatika Dept. University of Amsterdam.

Prust, Hub, Remko Scha and Martin v.d. Berg. In Press. *Discourse Structure and verb phrase anaphora* in Linguistics and Philosophy.

Reichman, Rachel. 1985. Getting computers to talk like me and you. MIT Press.

Roberts, Craigie. 1987. Modal Subordination. Unpublished Ph.D. Dissertation, University of Massachusetts, Amherst.

Sacks, Harvey, Emmanuel Schegloff and Gail Jefferson. 1974. *A simplest systematics for the organization of turntaking in conversation.* Lg. 50. pp. 696-735.

Sag, Ivan and Carl Pollard. 1991. *An integrated theory of complement control* in Lg. 67:1. pp. 63-113.

Scha, Remko and Livia Polanyi. 1988. *An Augmented Context Free Grammar for Discourse* in Proceedings COLING88. Budapest.

Seuren, Pieter. 1985. Presupposition. in Armin von Stechow and Deiter Wundelich (eds.) *Handbuch der Semantik/Handbook of semantics.* Frankfurt: Athenaum.

Schiffrin, Deborah. 1987. Discourse markers. Cambridge: Cambridge University Press.

Selkirk, Elizabeth 1984. Phonology and Syntax. Cambridge: MIT Press

Sidner, Candace. 1979. What the Speaker Means: The Recognitionof Speakers' Plans in Discourse. Technical Report MIT AI-Lab.

—— 1986. Focusing in the comprehension of definite anaphora. In Barbara Grosz, Karen Sparck-Jones and Bonnie Webber eds. *Readings in natural language processing.* Los Altos, CA: Morgan Kaufman Publishers. 363-394.

Sinclair, John and Malcolm Coulthard. 1992. Towards an analysis of discourse. in Malcolm Coulthard *Advances in spoken discourse analysis.* NY: Routledge.

Tannen, Deborah. 1989. Talking voices: repetition, dialogue, and imagery in conversational discourse. Cambridge: Cambridge University Press.

Tedlock, Dennis. 1977. *Towards an Oral Poetics* New Literary History 8. 507-519.

—— 1983. The spoken word and the work of interpretation. Philadelphia: University of Pennsylvania Press.

Thompson, Sandra. 1983. Grammar and Discourse: The English detached participial clause in Flora Klein-Andreu ed. *Discourse perspectives on syntax.* NY: Academic Press. pp 43-65.

—— 1987.'Subordination and narrative event structure'. in Russell Tomlin,ed. *Coherence and Grounding in discourse.*Philadelphia, John Benjamins. 435:454.

Vallduví, Enric. 1990. *The informational component.* Ph.D. Dissertation. Department of Linguistics. University of Pennsylvania.

Walker, Marilyn A. 1993. *Informational redundancy and resource bounds in dialogue IRCS Report No.: 93-45.* Philadelphia: Institute for Research in Cognitive Science, University of Pennsylvania.

Webber, Bonnie Lynn. 1986. So what can we talk about now? in Barbara Grosz, Karen Sparck-Jones and Bonnie Webber eds. *Readings in natural language processing.* Los Altos, CA: Morgan Kaufman Publishers. 395-414.

Webber, Bonnie Lynn. 1988. discourse deixis and discourse processing. Technical Report MS-CIS-86-74. Linc Lab 42. Department of Computer and Information Science. University of Pennsylvania.

Woodbury, Anthony C. 1985. *Functions of rhetorical structure: a study of Central Alaskan Yupik Eskimo.* Language in Society 14:150-93.

—— 1987. *Rhetorical structure in a Central Alaskan Yupik Eskimo traditional narrative.* Native American discourse: Poetics and rhetoric, Joel Sherzer and Anthony C. Woodbury, eds., 176-239. Cambridge: Cambridge University Press.

Yngve, Victor. 1961. The depth hypothesis. *Lg.* 61:283-305.