An Augmented Context Free Grammar for Discourse

Remko SCHA and Livia POLANYI
BBN Laboratories
10 Moulton Street
Cambridge, MA 02238

Abstract

This paper presents an augmented context free grammar which describes important features of the surface structure and the semantics of discourse in a formal way, integrating new as well as previously existing insights into a unified framework. The structures covered include lists, narratives, subordinating and coordinating rhetorical relations, topic chains and interruptions. The paper discusses the problem of parsing discourse, and compares different grammatical formalisms which could be used for describing discourse structure.

1. Introduction

Though a wealth of insights on the structure and meaning of discourse has been gathered by researchers in linguistics, psychology, ethnomethodology and artificial intelligence, these insights have not been integrated into formal grammars which display the breadth, depth and precision of formal treatments of sentential syntax and semantics. In the present paper we make a step towards a formal, integrated description of the surface structure and semantic interpretation of discourse. We introduce a formalism which uses augmented context free rules for specifying discourse grammars, and demonstrate its viability by developing a set of syntactic/semantic rules which covers a number of important discourse phenomena. We discuss the issue of parsing discourse in a semi-deterministic left-to-right fashion, and relate the grammar presented here to the strategies outlined in [24] [25] for building up a structural description of an unfolding discourse. Finally, we compare the formalism used here to some possible alternatives.

2. Discourse Structure and Discourse Semantics

The semantic interpretation of the utterances in a discourse has been shown to depend on the structural relations obtaining among the segments of that discourse. [8] [17] In developing a grammar for the surface structure of discourse, it is our aim to account for the semantically relevant aspects of its structure. Two phenomena of fundamental importance for semantic interpretation depend crucially on discourse structure: *context dependence* and *rhetorical structure*.

Context Dependence of Utterance Meanings

Context dependence of utterance meanings is a pervasive phenomenon in language. When an utterance is analysed in isolation, its meaning is underdetermined in many ways as an effect of the following phenomena:

- **Indexicality.** One rarely makes a statement or asks a question about the universe at large: every utterance presupposes an implicit temporal, spatial and topical framework that constrains the scopes of the meanings of the constituents of the utterance. Every sentence must therefore be evaluated with respect to a frame of reference which is usually left implicit in the sentence itself.

- **Anaphora.** Many utterances contain pronouns and definite descriptions which constitute overt references to the previous discourse. An analysis of the structure of the previous discourse is necessary to resolve such anaphoric references correctly.

- **Implicit arguments.** Many natural language words are semantically unsaturated, needing externally provided arguments in order to be interpreted felicitously. Nevertheless, such constructions can be used in sentences which do not mention their arguments explicitly, if these arguments can be inherited from previous parts of the discourse. *("John is taller [than Peter]." "What is the speed [of John's car]?")* (Cf. [4])

The Rhetorical Structure of Discourse

The insight of work focused on the rhetorical structure of discourse is that a speaker engaged in a discourse may perform speech acts whose illocutionary force has scope over complex propositions which are built up out of individual sentence meanings *and the rhetorical relations between them*. (See, for example, [13], [19], and [20].) These rhetorical relations may be overtly expressed, or they may have to be abducted on the basis of the sentence meanings. Paradigmatic examples of such complex propositions are *"A caused B"*, *"A was caused by B"*, *"A provides evidence for B"*, etc. -- where *A* and *B* may stand for propositions expressed by individual sentences, or may themselves be complex propositions expressed by discourse segments. Because of the

latter case, a correct constituent analysis of the discourse is necessary to establish the arguments of the rhetorical relations. On the other hand, the rhetorical relations themselves constitute an important structure-building component of discourse.

In the approach to discourse semantics outlined in this paper, every sentence is initially interpreted in a local, context-independent fashion. This results in a meaning representation which will usually contain free variables, standing for the discourse-dependent elements in the utterance meaning. When a sentence is integrated into the ongoing discourse, these variables are bound to values picked up from the context into which the sentence is inserted. The next section describes this in more detail.

3. An Augmented Context Free Grammar for Discourse.

Discourses have a hierarchical structure. They are built up recursively out of units of various kinds which can occur as constituents of each other. To account for this, a discourse grammar must be able to assign a tree structure to a discourse. We call this tree structure the *discourse parse tree*. To describe in a formal way how a discourse parse tree is constructed out of constituent sentences, we use a context free grammar whose non-terminal symbols are augmented with attribute/value pairs. (Distinct non-terminal categories have distinct sets of attributes.) Context free rules describe how the constituent segments of a discourse (which we call *discourse constituent units* or *dcu's*) are built up out of their subconstituents. The values of the attributes on a non-terminal represent the relevant structural and semantic properties of the dcu generated by that non-terminal. Every attribute has a fixed set of possible value-expressions. The value-expressions may be of different kinds: they may be atomic, they may themselves be sets of attribute/value pairs, or they may be logical expressions.

Value-expressions often store *partial* information; therefore, they may contain free variables. A value-expression stands for the set of its *ground instances*. (A value-expression without variables thus stands for a singleton set.) When an attribute has a value-expression which stands for the empty set, the complex category symbol that contains this value-expression fails to label a possible dcu.

The context-free rules enforce agreement and upwards-inheritance of the relevant properties of different constituents through the fact that different occurrences of the same variable take on identical values. To put this in more precise terms, we define the meaning of an augmented context free rule as follows. If A, B, C, Y, Z stand for complex category symbols (including attribute/value-expression pairs) and $[dcu1]_Y$ and $[dcu2]_Z$ are legitimate dcu's, the rule "A => B C" legitimizes the dcu $\sigma([[dcu1]_Y [dcu2]_Z]_A)$ iff the substitution σ is the *most general unifier*[1] of the terms <B,C> and <Y,Z>, and $\sigma(A)$ is a legitimate complex category symbol, not containing empty attribute-value expressions.

3.1. Attribute Values and Semantic Interpretation.

The propagation of attribute values between dcu's plays an important part in establishing the semantic interpretation of the utterances in a discourse. We now list some of these attributes, and discuss their role in semantic interpretation and discourse processing.

The **Semantics** attribute records a logical formula representing the meaning of the discourse constituent unit that it is associated with. Anaphoric elements which have not been resolved inside the dcu are represented by free variables. There are two mechanisms for anaphor resolution: (1) unification with value-expressions of attributes of other dcu's, and (2) explicit search processes involving the *Discourse Referents Set* of accessible dcu's.

The **Discourse Referents Set** records the entities introduced in the discourse unit that it is associated with. These, plus the entities in the Discourse Referents Sets of the embedding units (dominating nodes in the tree) are the entities which are available for anaphoric reference in an utterance which extends or expands that discourse unit. Every *Discourse Referent* is a pair consisting of (1) the linguistic expression that introduced it, and (2) the semantic representation that the system attached to that expression. Discourse Referent Sets are accessed in different ways by a number of algorithms, which resolve the meanings of

---

[1] The concepts we use here are essentially the ones that were developed for term unification in first-order logic theorem proving [30]. We assume some straightforward generalization of these concepts which deals with the fact that our "terms" have a structure which, though not interestingly different, is somewhat richer.

different types of context-dependent expressions such as definite descriptions, anaphoric pronouns, demonstrative pronouns, the words *"one"* and *"ones"*, and implicit arguments of function nouns, comparatives, etc.

The **Reference-Time** records the time-interval which is to serve as the temporal index for extensions and expansions of the discourse unit under consideration. The Reference-Time can be reset by the occurrence of explicit temporal adverbs. *Narratives* constitute a specific type of discourse units in which time plays a special role; they are marked by the occurrence of sentences with non-durative aspect (*event*-sentences). In a narrative, the reference-time is reset whenever an event occurs. [15] [12]

The **Spatial Index** and the **Modal Index** play a role in the semantic interpretation of the utterances which is similar to the role of the Reference-Time: they specify a value for components of the index of evaluation of the semantic formula. In English, they can only be updated by the occurrence of spatial and modal adverbials or other similarly explicit information.

In the next section we describe the rules of a discourse grammar. They use some other attributes besides the ones just mentioned. These can be most easily explained as we introduce the rules that use them.

3.2. Grammar Rules

The grammar we present here, though necessarily limited and schematic in many ways, covers a wide range of phenomena that are usually discussed separately. An adequate account of discourse structure, however, depends on their integration into a consistent mechanism. The semantic phenomena that we pay attention to include anaphor resolution, scope of modal indices, movement of reference time, and rhetorical relations. The grammar consists of rules which describe how to build up various kinds of structurally different discourse constituent units. We distinguish the following kinds of dcu's:

- **Subordinations.** These are binary structures in which the first element remains accessible; we view them as units in which all or most of the structurally relevant features are inherited from the left constituent. (In discourse, unlike in the sentence, the subordinating element is always to the left of the subordinated one.) In *semantic subordinations* there is a semantic relation between the two constituents; this is the case in *rhetorical subordinations* and in *topic-dominant chains*. *Interruptions*, on the other hand, though structurally analogous, are semantically very different: in this case, there is no semantic connection whatsoever between the two constituents.

- **Binary Coordinations.** These are binary structures in which the second element has equal status to the first, thus making the first one inaccessible. Under this category we include *rhetorical coordinations* (the counterparts of the rhetorical subordinations), and *adjacency pairs* which are concerned with the *interactional* dimension of the discourse (we include *question/answer pairs* and *request/response pairs*).

- **N-ary Coordinations.** These are flat structures which can contain arbitrarily many elements, of which, at any time, only the most recent one is accessible. We include *lists*, *monotonic lists*, and *narratives*. To generate n-ary coordinations by means of a context free grammar, we must assign them a recursive structure: we build them up by means of binary rules which extend them to the right.

This classification is not necessarily complete, and should be the topic of more extensive discussion. But it does cover the most important structures, and brings some order into the grammar rules that we present below.

Notation: Category symbols have the form "cat $[\phi_1{:}\alpha_1, ... \phi_n{:}\alpha_n]$", where "cat" is the basic non-terminal symbol of the context-free grammar, $\phi_1, ..., \phi_n$ are the attributes, and $\alpha_1, ... \alpha_n$ are expressions which stand for the value-sets of these attributes. Variables are indicated by *italicized* strings. If in a phrase structure rule an attribute has as its value-expression a variable that does not occur anywhere else in the rule, we will, for the sake of readability, leave the attribute out of our specification of the rule.

The grammar does not have a particular start symbol; dcu's of all kinds are recognized as well-formed "discourses".[2] *Sentences* are assumed as elementary dcu's. The following rules specify the internal structure of some important kinds of complex discourse constituent units. We discuss further details about our approach and about the grammar formalism as we introduce these rules.

---

[2]Social constraints on discourses as units of interaction (concerning greetings, farewells, and proper ways of embedding a discourse in a social situation), are dealt with by separate sets of rules, describing "Interactions" and "Speech Events". See [23] for discussion.

Lists:

list [drs: $d1 \cup d2$, schema:s, sem: $s(x)$ & $s(y)$]
=> dcu1 [drs:$d1$, sem: $s(x)$]
    dcu2 [drs:$d2$, sem: $s(y)$]

The intuitive idea behind this rule is that if two adjacent dcu's can be analyzed as having semantically parallel structures, they can be conjoined in a *list* structure. Note that there is no constraint on the category of the dcu's: the main category symbols in the right hand side of the rule are variables.

The formal criterion for the application of the rule is the existence of a λ-formula s such that the semantic content of each of the two dcu's has the form $s(u)$.[3] We require that the semantic representation of the dcu already has the structure $s(u)$, or can be put in that form by means of a very limited repertoire of logical equivalence transformations. (This repertoire needs to be specified in detail. It will probably at least include λ-abstraction.) This restriction excludes undesirable trivial values of s such as (λ u: if u = SEM1 then SEM1 else SEM2). The types of the λ-variables of s are required to be elements in a predefined hierarchy of "natural kinds" which classifies the meanings of lexical entries. The value of s, representing the common denominator between the meanings of the two constituents, is stored as the value of the "schema" attribute on the list-dcu.

Note that this rule often creates ambiguity. To take a simple example: if *"John likes Mary"* is followed by *"Peter likes Mary"*, one can abstract out several different schema's, even without considering different choices for the types of λ-variables. λ P: P(MARY), or λ R,x: $R(x$, MARY), or λ x,y: LIKE(x,y), or λ x: LIKE(x, MARY). The last of these is the preferred one, because it is most specific. Rather than constraining the grammar rule to apply only in the most specific way to every case, we assume that a preference for the most specific rule-application is applied as a heuristic principle during the parsing process.

The "sem" attribute stores the semantics of a dcu. The operation of the "&"-operator which is used here to build up the semantics is more general than logical conjunction: its arguments can have different speakers and have different illocutionary force operators. The semantic representation built up by "&" stores the propositional content of individual utterances in a richly indexed data structure. We shall not attempt to specify this data structure in detail in the present paper. We intend that for the case of purely assertional monologue, the operation of "&" reduces to logical conjunction. (The movement of narrative time, for instance, will therefore not be hidden inside the meaning of "&".)

Note that the notion of "list" that we use here is more general than the one we have used in previous discussions [23]. For instance, this notion of "list" subsumes the notion of a "topic chain".

In the formula stored as the value of the "sem"-attribute of a dcu, typed free variables are used to represent unresolved anaphors, such as pronouns and destressed definite descriptions. The unification process which matches the "sem"-attributes of two dcu's when they are joined to constitute a list may substitute expressions for these variables and thus resolve the anaphoric reference. Thus, some strong reference resolution preferences are explained as following directly from the acknowledgement of parallel structure; in this case, anaphoric elements get resolved without any search through the space of available discourse referents. This accounts for anaphoric reference to *topic* or *attentional focus* [32], as well as anaphoric reference to the corresponding element in a semantically parallel structure.

To accumulate the appropriate candidates for further anaphor resolution processes, the discourse referents (values of the "drs" attribute) of both constituent dcu's are synthesized into the discourse referent set of the list-dcu.

To extend a list-dcu to become a list with one more element, there is the following additional rule:

list [schema:s, drs: $drs1 \cup drs2$, sem: p & $s(y)$]
=> list [schema:s, drs:$drs1$, sem:p]
    dcu [drs:$drs2$, sem: $s(y)$]

A list can be extended by another dcu if this dcu instantiates the structure described by the schema-attribute of the list.

Monotonic Lists:

m-list [schema:s, drs: $drs1 \cup drs2$,
      direction: (if x<y then incr
            else if x>y then decr else fail),
      last:y, sem: $s(x)$ & $s(y)$]
=> dcu [drs: $drs1$, sem: $s(x)$]
    dcu [drs:$drs2$, sem: $s(y)$]

---

[3]If the λ-function s takes more than one argument, we view it as working on n-tuples; the values of x and y are n-tuples in this case.

If in a list the various arguments of the "schema"-function are elements of a linearly ordered domain, the list may be a *monotonic list*. To make it possible to ascertain whether a next dcu can be added to a monotonic list, such lists carry the value of the most recent argument ("last"-attribute), and their "direction" ("increasing" or "decreasing"). ("fail" is a special constant which is not allowed as an attribute value-expression.)

m-list [schema:*s*, drs: *drs1* ∪ *drs2*,
    direction: (if *x<y* ∧ *p*=incr **then** incr
                    else if *x>y* ∧ *p*=decr **then** decr **else** fail),
    last:*y*,   sem: *q* & *s(y)*]
=> m-list [schema:*s*, drs:*drs1*, incr: *p*, last:*x*, sem:*q*]
    dcu [drs:*drs2*, sem: *s(y)* ∧ *y* ∈ *d*]

A monotonic list can be extended by another dcu which instantiates the structure described by its schema-attribute, provided that the increasing or decreasing ordering is maintained. (Monotonic lists are discussed in [24].)

Narratives:

"Narratives" are used to express that a series of states of affairs obtain at successive points along a timeline (*reference times* [29]). How a next main clause of a narrative will interact with the previously established reference time of the narrative, depends on its *aspect*: *durative* clauses behave differently from *non-durative* ones [35].[4] We therefore need two different rules for extending a narrative. We first give the rule for duratives.

narrative [drs: *drs1* ∪ *drs2*, reference-time:*rt*,
    tense: (if *t1* ∈ *xt* then *t1* else fail) , sem: *p* & *[s]*_{*rt*}]
=> narrative [drs:*drs1*, reference-time:*rt*, tense: *t1*, sem:*p*]
    dcu [drs: *drs2*, reference-time:*rt*, x-tenses: *xt*,
        aspect: durative, sem: *s*]

The reference-time attribute indicates the time-interval where the progression of narrative time has arrived, i.e., an interval after the last event in the narrative so far. A durative dcu which extends a narrative is evaluated at the reference-time of that narrative.

The value of the "tense" attribute on a narrative dcu marks the temporally distinct modes of narrative that a particular language allows. For English this includes the distinction between past, present, pluperfect, and future narratives. On a sentence (or more complex dcu) which is to be integrated into a narrative, the relevant attribute is not "tense" (which stores the tense of the dcu), but "x-tenses", which stores the tenses that may be externally imposed on the dcu. A sentence with PRESENT as its tense can be used in a context where the tense is set to either PRESENT, PAST, or FUTURE. Similarly, a PAST tense sentence is compatible with either a PAST or a PLUPERFECT framework. A sentence with the PLUPERFECT tense, however, is only compatible with a PLUPERFECT timeframe. When a dcu extends a narrative, the "tense" of the narrative must be an element of the "x-tenses" of the dcu.

narrative [drs: *drs1* ∪ *drs2*, reference-time:*v*,
    tense: (if *t1* ∈ *xt* then *t1* else fail),
    sem: *p* & *[s]*_{*u*} & *t* <_{*i*} *u* <_{*i*} *v*]
=> narrative [drs:*drs1*, reference-time:*t*, tense: *t1*, sem:*p*]
    dcu [drs: *drs2*, reference-time:*u*,
        aspect:event, x-tenses: *xt*, sem: *s*]

An event-dcu which extends a narrative is evaluated at an interval *u* following the then current reference-time *t* of the narrative. The new reference-time of the extended narrative is established to be another interval *v* after *u*. (Notation: *a* <_{*i*} *b* means "*a* immediately precedes *b*".) The effect of this is that a durative dcu without explicit time-adverbials is interpreted at an interval immediately after the last event, which will be closed off by the next event. (Validity beyond this interval, in both directions, can often be inferred from this by invoking plausible assumptions about the universe of discourse.) Another effect is, that successive events are always separated by a time-gap (though nothing is said about the size of this time-gap).

The above rules only define how to extend a narrative that is already underway. The rules for beginning a narrative are similar have been omitted for reasons of space.

Rhetorical subordinations:[5]

dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index: *i*, sem: *a* & *R(a,[b]_i)*]
=> dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index: *i*, sem: *a*]
    dcu2 [sem: λ *x: R(x,b)*]
    (pop-marker)

This rule parses semantic subordinations which involve an explicitly indicated subordinating rhetorical relation *R* ("for instance," "because"). The meaning of this relation is assumed to be incorporated in the semantics of the subordinated dcu; this dcu therefore has as a value of its "sem"-attribute a λ-function which expects a propositional argument.[6] The attributes and values of a subordination are inherited from the subordinating constituent.

The subordinated constituent is optionally followed by a pop-marker (e.g. "so", "anyway"). All clue-words (push-markers, pop-markers, interruption-markers) are treated as independent units, separate from the sentences that they precede or follow.

In the formulation of the semantic subordination rule we have assumed an attribute called "index", containing reference time as well as spatial and modal index.The rule shows how the subordinated discourse constituent unit is semantically contextualized by the subordinating one.

dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index:*i*, sem: *a* & *R(a,[b]_i)*]
=> dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index:*i*, sem: *a*]
    (push-marker)
    dcu2 [sem: *b*]
    (pop-marker)

This rule parses semantic subordinations for which the rhetorical relation involved is not overtly marked. The variable *R* ranges over all subordinating rhetorical relations.Since its value is not stated explicitly, it must be abducted on the basis of plausibility considerations regarding the resulting semantics. The subordinated constituent is optionally preceded by a push-marker (e.g. "like"), and optionally followed by a pop-marker.

For subordinations we need a more elaborate treatment of semantics than the one assumed in this paper. We need to distinguish between the total accumulated meaning of a discourse constituent unit and its "core meaning", which is considered in computations regarding semantic relations with other dcu's. It is a characteristic property of subordination dcu's that they allow for interpretations in which the core meaning is identical to the core meaning of the subordinating constituent, without any contribution from the subordinated constituent. To represent this, we would need to assume at least two different "sem" attributes, or a more complicated structure for the value of the "sem" attribute.

Rhetorical coordinations:

dcu [ϕ₁: mscg(α₁, β₁), ..., ϕ_n: mscg(α_n, β_n), sem: *a* & *b* & *R(a,b)*]
=> dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, sem: *a*]
    dcu2 [ϕ₁: β₁, ..., ϕ_n: β_n, sem: λ *x: R(x,b)*]

A This rule parses semantic coordinations which involve an explicitly indicated binary coordinating rhetorical relation *R* ("therefore," "thus," "accordingly"). <Ref. Mann, Talmy> As in the subordination case described before, the meaning of the relation is incorporated in the semantics of the clause in which it occurs, which therefore denotes a predicate on propositions.

The function "mscg" computes the "most specific common generalization" of its arguments in the hierarchy of value-expressions of the relevant attribute. (When there is no proper hierarchy defined on the value-expressions of an attribute, mscg degenerates into a function which yields the value of its arguments when the two arguments are equal and which yields a new free variable when they are not.)

dcu [ϕ₁: mscg(α₁, β₁), ..., ϕ_n: mscg(α_n, β_n), sem: *a* & *b* & *R(a,b)*]
=> dcu [ϕ₁: α₁, ..., ϕ_n: α_n, sem: *a*]
    dcu [ϕ₁: β₁, ..., ϕ_n: β_n, sem: *b*]

This rule parses binary semantic coordinations which are not overtly marked as such. Therefore, the semantics of the second dcu is a proposition rather than a predicate on propositions. The variable *R* ranges over all binary coordinating rhetorical relations. As in the corresponding subordination case, the value of *R* must be computed by abduction, magic, or a similar A.I. technique.

Topic-dominant chaining:

dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index: *i*, sem: *s1(y)(x)* & *[s2(y)]_i*]
=> dcu1 [ϕ₁: α₁, ..., ϕ_n: α_n, index: *i*, sem: *s1(y)(x)*]
    dcu2 [ϕ₁: β₁, ..., ϕ_n: β_n, sem: *s2(y)*]
    (pop-marker)

Topic-dominant chainings are subordination structures. In these structures, the subordinated dcu gives information about a constituent of the predicate in the semantics of the subordinating dcu. The above rule requires that there exists an element *y* such that the predicate of the left dcu and the semantics of the right dcu can both be formulated as expressions with the structure f(y) -- where, as before, only a limited

---

[4]Formulated in terms of Vendler's [34] classification, durative dcu's describe *accomplishments* or *achievements*, while non-durative ones describe *states* or *activities*.

[5]To enhance the legibility of the rules, we will from now on leave out the description of the upwards propagation of discourse referents. It occurs uniformly in the same way as in the rules given before.

---

[6]Note that most other rules implicitly assume that they operate on dcu's with a "sem"-value which is a proposition. In the current formulation, they therefore do not operate correctly on explicitly subordinated clauses. The appropriate refinements are not difficult to imagine, but go beyond the limited scope of the present paper.

repertoire of logical transformations can be used to achieve this formulation, starting from formulas which correspond directly to the surface structure of the dcu's. These limitations are to be defined in such a way that the possible values of y correspond to the constituents eligible for *dominance* in [6], or the *forward-looking centers* of [9].

The heuristics of the parsing process prefers applying rules for constructing *list*-structures to the rule for topic-dominant chaining. (Cf. [3])

Adjacency Pairs:

QA [sem: *a(b)*]
=>   *dcu1* [mood: interrogative, sem: *b*]
     dcu2 [sem: *a(b)*]
     (pop-marker)

This rule parses question/answer pairs. The semantics of a yes/no question is assumed to be a proposition; the semantics of a wh-question is assumed to be a set-denoting expression (cf. [31]). The semantics of an answer is a predication on the question-semantics.

RR [sem: *a(b)*]
=>   *dcu1*[mood: request, speaker:*p1*, addressee:*p2*, sem:*b*]
     dcu2 [speaker:*p2*, addressee:*p1*, sem:*a(b)*]
     (pop-marker)

This rule parses request/response pairs. Semantically, these are very similar to question/answer pairs. We have chosen to exclude "rhetorical requests" by requiring the speaker/addressee relation to flip between request and response.

Interruptions:

dcu1[α]
=>   *dcu1*[α]
     (interruption-marker)
     dcu2 [β]
     (pop-marker)

This rule allows for semantically unrelated interruptions of an ongoing discourse (cf. [26] [10]) Interruptions may be introduced by specific markers such as "Oh!".

4. Discourse Parsing

We consider the development of a formal grammar of discourse structure, such as the one sketched in the previous section, to be the first step towards a formal account of the process of discourse *parsing*. We now briefly review some of the issues that would be involved in such an account.

The most important issue in discourse parsing is the necessity of semi-determinism. Spontaneous dialogue involves unpremeditated turn-taking and interruption. In order for this to be possible, there must regularly be points in an interaction at which the interpretation of utterances so far is mutually established as independent of the discourse which is to follow. Moreover, an unanticipated next utterance can operate in a mutually understood way on the *structure* of the discourse so far (for instance, by abandoning a digression to pop to a previously interrupted dcu). Therefore, at points where such a move is allowed, there must also be mutual agreement on the structure of the discourse so far.

The granularity of this "unpredictability without misunderstanding" seems to be the clause or sentence level. We therefore postulate an incremental left-to-right parsing process at this level of granularity, which operates in essentially deterministic mode. In [24] we gave an informal description of such a parsing process, which processes every incoming sentence incrementally, extending an existing discourse tree to the right by node insertion. An important assumption of the parsing process as described there, is that at any point it only uses information on the right edge of the existing discourse tree. This means that interlocutors just need to be aware of the stack of information which corresponds to the labels on the right edge of the tree, rather than the complete details of the discourse that went before.

Inspection of the grammar rules in the previous section suggests that this grammar is compatible with the parsing strategy outlined in [24]: relevant information is always propagated up to the right edge of the tree, and dcu interpretations get propagated up without being influenced by the nature of the intervening nodes.

5. Formalisms for Discourse Grammar.

The augmented context free grammar developed above should be taken as a demonstration of the possibility and utility of formal grammars for decribing structural and semantic phenomena in natural language discourse. We expect that work of this kind, especially if carried out on a larger scale, will constitute a more fruitful path to new insights than approaches which are oriented towards essayistic description or unprincipled implementation.

We should not make exaggerated claims concerning the formalism we have used here. Much more work is needed before it will be clear what kind of formal framework has the best fit with the phenomena. But it is probably useful to articulate our thoughts on how the augmented context free grammar formalism compares to other formalisms that could have been used in this work.

The formalism we have used is a context free grammar augmented with attributes, which propagates feature values through term unification on value-expressions containing variables. Similar formalisms have been used for sentence-level syntax. The closest is the ACFG formalism used for the grammar of BBN's *Spoken Language System* [11] [2], which mainly differs in assuming a more limited syntax of attribute-value expressions. Definite Clause Grammars [21], if decoupled from their commitment to Prolog programming, are also very similar.

Generalized Phrase Structure Grammar [7], as well as related theories such as HPSG and LFG, share many aspects of our approach: an emphasis on context free surface structure, and the use of a unification process to enforce the desired agreement and inheritance behaviour on the values of attributes. However, these frameworks use unification on graphs rather than logical term-expressions. This probably creates additional expressive power, but it goes at the cost of ease of implementation and of conceptual clarity.

It is a major advantage of logical term unification that there is an obvious and simple semantics for it: any object which may contain variables, be it an attribute value, a syntactic tree, or a context free rule, can be viewed as an abbreviation for the set of its ground instances; applying the most general unifier to a set of terms yields a term standing for the intersection of the sets of their ground instances. Compared to the conceptual and computational simplicity of logical term unification, the graph unification formalisms used in modern linguistic frameworks are rather cumbersome. There do not seem to be good linguistic reasons for preferring graph unification.

An interesting perspective on the kind of grammar we have used follows from the realization that it can be viewed as a particular instance of an *attribute grammar* as defined by Knuth [16]: one in which the values of the attributes on a node are always synthesized from the values on the nodes of its immediate constituents. This raises the question whether one might want to formulate this directional dependency explicitly in an attribute grammar notation. More interestingly, it raises the question whether there are phenomena that could be more elegantly described as *inheritance* from a top node to its constituents, rather than the other way around. We expect such phenomena to occur if we would want to integrate more global constraints, related to people's tasks, goals and plans, into our description of discourse.

Finally, we want to reflect on the Augmented Transition Network formulation of discourse structure that we used in previous, more informal papers [23]. ATNs do have some properties which are attractive for discourse. A looping arc which updates a register constitutes a powerful device that doesn't have a direct equivalent in other formalisms. In the current grammar we emulate the effect of such an arc by a recursive binary structure: lists and narratives are built up by repeated extension to the right. Intuitively, one sees lists and narratives as flat structures, and we have described them in those terms in previous papers [24]. The power of ATNs thus makes it possible to account more directly for the structures that seem plausible.

The framework used in the present paper is conceptually simple, and more limited than any of its alternatives. Nevertheless, it seems powerful enough to describe the phenomena we encountered in developing the grammar presented here. Subsequent research will have to answer the question whether it is ultimately powerful enough to describe the full range of discourse phenomena in a felicitous way.

In the present discussion we have ignored what in previous work we have called the level of *speech event structure*, which is concerned with the structure of discourse as a social activity. [22] [23] By the same token, we have left unaddressed the fact that discourse structures may reflect the tasks, goals and plans of the discourse participants, as they would be construed in A.I. based approaches to discourse analysis. Observations reported in [36] show (1) that speech event structure and linguistic discourse structure may be at odds with each other, and (2) that the linguistic discourse structure has the most direct semantic relevance in such cases. Though speech event structure and task structure have in fact considerable semantic relevance and must ultimately be factored in, we want to hold off on dealing with the complexities involved in this issue.

The grammar presented in this paper provides a formal characterization of Discourse Parse Trees by means of the bottom-up rules used in its construction. The problem of mating these bottom-up rules with necessary high-level, top-down rules involving phenomena occurring in the task domain or interaction is essentially the problem of plan-recognition. In an operational system, the functionality of a linguistically based discourse parser would thus be very much enhanced by an efficient Plan Recognizer of the type envisioned in "plan based" and "intention based" pragmatic discourse models [1], [14], [18], [33]. On the other hand, incorporating a discourse grammar would improve the

functionality of existing plan based models, which lack explicit mechanisms for relating sentential syntax and semantics to pragmatic plan structures.

[1]   Allen, James.
        Recognizing Intentions from Natural Language Utterances.
        In Brady, M., and Berwick, R. (editors), *Computational Models of Discourse*. MIT Press, Cambridge, MA, 1983.

[2]   Ayuso, D., Y. Chow, A. Haas, R. Ingria, S. Roucos, R. Scha, D. Stallard.
        Integration of Speech and Natural Language.
        Technical Report 6813, BBN Laboratories, Cambridge, MA, April, 1988.

[3]   Brennan, Susan E.; Friedman, Marilyn W.; Pollard, Carl.
        A Centering Approach to Pronouns.
        In *25th Annual Meeting of the Association for Computational Linguistics*, pages 155-162. Stanford University, Stanford, CA, July, 1987.

[4]   de Bruin, Jos and Remko Scha.
        The Interpretation of Relational Nouns.
        In *Proceedings of the 26th Annual Meeting of the ACL*. SUNY, Buffalo,NY, June, 1988.

[5]   Erteschik-Shir, N. and Lappin, S.
        Dominance and the Functional Explanation of Island Phenomena.
        Theoretical Linguistics, 6:1 :41-86, 1979.

[6]   Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag.
        Generalized Phrase Structure Grammar.
        Harvard University Press , Cambridge, Mass., 1985.

[7]   Grosz, Barbara [Deutsch].
        The Structure of Task Oriented Dialogs.
        In *IEEE Symposium on Speech Recognition: Contributed Papers*, pages 250-253. Carnegie Mellon University Computer Science Dept., Pittsburgh, PA, 1974.

[8]   Grosz, B.J., Joshi, A.K., Weinstein, S.
        Providing a Unified Account of Definite Noun Phrases in Discourse.
        In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, pages 44-50. Association for Computational Linguistics, Cambridge, MA, June, 1983.

[9]   Grosz, B. and Sidner, C.
        Discourse Structure and the Proper Treatment of Interruptions.
        In *Proceedings of the 9th International Joint Conference on Aritificial Intelligence, 1985*, pages 832-839. IJCAI 1985, Los Angeles, CA, August 18-23, 1985.

[10]   Haas, Andrew.
        Parallel Parsing for Unification Grammar.
        In *Proceedings of the 10th IJCAI*. Milan, Italy, August, 1987.

[11]   Hinrichs, E.
        Temporal Anaphora in Discourse of English.
        Linguistics and Philosophy 9 (1):63-82, 1986.

[12]   Hobbs, Jerry R.
        A Computational Approach to Discourse Analysis.
        Technical Report, SRI International, December, 1976.

[13]   Hobbs, J. and Evans, D.
        Conversation as planned behavior.
        Cognitive Science 4(4):349-377, 1980.

[14]   Kamp, H.
        Events, Instants and Temporal Reference.
        In U. Egli and A. van Stechow (editors), *Semantics from a Multiple Point of View*, pages 376-471. de Gruyter, Berlin, 1979.

[15]   Knuth, Donald E.
        Semantics of Context-Free Languages.
        Mathematical Systems Theory 2(2):127-145, 1968.

[16]   Linde, C.
        Focus of Attention and the Choice of Pronouns in Discourse.
        In T. Givon (editor), *Syntax and Semantics, Vol. 12 of Discourse and Syntax*, pages 337-354. Academic Press, Inc., New York, New York, 1979.

[17]   Litman, Diane.
        Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues.
        PhD thesis, University of Rochester, 1985.

[18]   Longacre, R.E.
        An Anatomy of Speech Notions.
        the Peter de Ridder Press, Lisse, 1976.

[19]   W.C. Mann and S.A. Thompson.
        Relational Propositions in Discourse.
        Technical Report RR-83-115, Information Sciences Institute, Marina del Rey, CA, November, 1983.

[20]   Pereira, Fernando C.N. and David H.D. Warren.
        Definite clause grammars for language analysis - a survey of the formalism and a comparison with augmented transition networks.
        Artificial Intelligence 13:231-278, 1980.

[21]   Polanyi, Livia and Scha, Remko.
        The Syntax of Discourse.
        TEXT 3:3 :271-290, 1983.

[22]   Polanyi, L. and Scha, R.
        A Syntactic Approach to Discourse Semantics.
        In *Proceedings of the International Conference on Computational Linguistics*, pages 413-419. Stanford University, Stanford, CA, 1984.

[23]   Polanyi, Livia.
        The Linguistic Discourse Model: Towards A Formal Theory of Discourse Structure.
        Technical Report 6409, BBN Labs, Cambridge:MA, November, 1986.

[24]   Polanyi, L.
        A Formal Model of the Structure of Discourse.
        Journal of Pragmatics 2/3, 1988.

[25]   Polanyi, L.
        A Theory of Discourse Structure and Discourse Coherence.
        In *21st Regional Meeting of the Chicago Linguistic Society*, pages 306-322. Chicago Linguistic Society, University of Chicago, April, 1985.

[26]   Reichenbach, H.
        Elements of Symbolic Logic.
        London:Macmillan, 1947.

[27]   Robinson, J.A.
        A Machine-Oriented Logic Based on the Resolution Principle.
        Journal of the ACM 12(1), January, 1965.

[28]   Scha, R.J.H.
        Logical Foundations for Question Answering.
        Technical Report, Eindhoven: Philips Research Labs, M.S. 12.331., 1983.

[29]   Sidner, C. L.
        Focusing in the Comprehension of Definite Anaphora.
        In Michael Brady and Robert C. Berwick (editors), *Computational Models of Discourse*, chapter 5, pages 267-330. MIT Press, Cambridge, MA, 1983.

[30]   Sidner, C.L.
        What the Speaker Means: The Recognition of Speakers' Plans in Discourse .
        International Journal of Computers and Mathematics, Special Issue in Computational Linguistics 9(1):71-82, 1983.

[31]   Vendler, Z.
        Linguistics and Philosophy.
        Cornell University Press, Ithica, NY, 1967.

[32]   Verkuyl, H.
        On the Compositional Nature of the Aspects.
        D. Reidel, Dordrecht, 1972.

[33]   de Witte, Liesbeth.
        Interacciones Habladas en una Zapateria Madrileña. Un Estudio sobre la Estructura Sintáctica y Semántica del Discurso.
        Unpublished Master's Thesis, Spanish Department, University of Amsterdam.
        1987