Discourse Structure and Anaphoric Accessibility

MASSIMO POESIO
University of Essex and University of Edinburgh
Massimo.Poesio@ed.ac.uk

BARBARA DI EUGENIO
The University of Illinois at Chicago
bdieugen@eecs.uic.edu

ABSTRACT. We evaluated the predictions of Grosz and Sidner's theory of anaphoric accessibility using a corpus of tutorial dialogues whose discourse structure was annotated according to Relational Discourse (RDA) Analysis. We found support for Moser and Moore's proposal that only segments with an intentional core should be viewed as introducing new focus spaces; we also found that both embedded cores and embedded contributors should remain open as long as the RDA-segment in which they occur is open, and discuss the implications of this finding for Grosz and Sidner's theory.

1 Introduction

In this paper, we present the results of an empirical study of the relation between discourse structure and anaphoric accessibility. Seminal theories such as (Reichman 1985; Grosz and Sidner 1986; Fox 1987) have been around for about fifteen years. However, only now is it possible to subject them to rigorous empirical testing because of the recent improvements in annotation methodology leading to more reliable annotation techniques, and the resulting increased availability of corpora annotated for discourse structure (Carletta, Isard, Isard, Kowtko, Doherty-Sneddon, and Anderson 1997; Moser, Moore, and Glendening 1996; Marcu 1999).

Much of this recent effort on discourse structure annotation has been based on Rhetorical Structure Theory (RST) (Mann and Thompson 1988) or on similar theories that restrict the range of possible structures. One reason for this is that (each version of) RST specifies a fixed (although fairly large) repertoire of intentions to choose from. While it is still an open question whether it is possible to come up with an exhaustive list of discourse intentions (Grosz and Sidner 1986), in order to achieve reliable annotation it is essential to do so.

As a result, however, the corpora annotated in this way have mainly been used to test new theories of anaphoric accessibility that build directly on RST, such as

Paper presented at the ESSLLI 2001 Workshop on
Information Structure, Discourse Structure and Discourse Semantics
Edited by Ivana Kruijff-Korbayová and Mark Steedman.

the recently proposed Veins Theory (Cristea, Ide, and Romary 1998). Our own aim, at least at this stage, is not to develop a new theory of anaphoric accessibility, but to study the empirical validity of existing theories of anaphoric accessibility, and specifically of the best-known among these, Grosz and Sidner's theory (1986). This used to be a problem because although Grosz and Sidner's theory has originated a coding manual (Nakatani, Grosz, Ahn, and Hirschberg 1995) that has been used at least once (Nakatani 1996), as far as we know there is no sizable corpus coded accordingly. However, recent proposals concerning the mapping between rhetorical structure and intentional structure such as (Moser and Moore 1996b) produced annotated corpora that can be used to investigate the predictions of the theory.

The Sherlock corpus collected at the University of Pittsburgh is a case in point. This corpus is a collection of tutorial dialogues annotated according to Relational Discourse Analysis (RDA), a theory of discourse structure that attempts to merge RST with Grosz and Sidner's theory. One of the features of RDA is to distinguish rhetorical relations into intentional and informational (Moore and Pollack 1992; Moser and Moore 1996b): intentional relations pertain to the effects that the speaker intends his discourse actions to have on the hearer, whereas informational relations pertain to domain relations between the entities being talked about. For example, among RST relations, **evidence** is intentional, whereas **cause** is informational (Moore and Pollack 1992; Moser and Moore 1996b).

In this paper, we report on the results of a study of anaphoric accessibility using this corpus. We first briefly review both Grosz and Sidner's focus-space based theory of the attentional state and RDA. We then discuss how RDA structures can be used to specify opening and closing operations on focus spaces. In the evaluation part of the paper, we present first of all our results concerning the effect on anaphoric accessibility of the more distinctive aspect of RDA, the distinction between intentional and informational relations, and of Moser and Moore's proposal that only the former express DSPs. We then look at a few aspects of the relation between Grosz and Sidner's theory and RDA left open by Moser and Moore, in particular, how embedded segments of different types should be treated. Finally, we analyze the results and evaluate Grosz and Sidner's proposal.

2 Background

2.1 Grosz and Sidner's Theory

According to Grosz and Sidner (G&S), the structure of a discourse is determined by the intentions that the people producing it intend to convey, or DISCOURSE SEGMENT PURPOSES (DSPs). In a coherent discourse, the DSPs are related to form an INTENTIONAL STRUCTURE by either **dominance** relations (in case a DSP is interpreted as contributing to the satisfaction of another intention) or **satisfaction-precedes** relations (when the satisfaction of an intention is a precondition for the satisfaction of a second one).

Anaphoric accessibility of entities in a discourse is modeled by its ATTENTIONAL STRUCTURE, which, according to Grosz and Sidner, is a stack of FOCUS SPACES. G&S propose that when a segment is open, its corresponding focus space, which includes the discourse entities introduced in that segment, is pushed onto the focus stack; when the segment is closed, the focus space is popped, and the discourse entities associated with that focus space are not accessible any more. G&S also argue that the pushing and popping of focus spaces on the stack reflects the intentional structure, in the sense that a new focus space is pushed on the stack whenever the discourse introduces a new DSP subordinate to the present one, and the focus space of the current is popped whenever the associated DSP is satisfied.

This claim about anaphoric accessibility was illustrated in the original paper with a few examples; however, as far as we know, it has not been empirically tested. There are no sizable corpora annotated according to both G&S's treatment of discourse structure and anaphora. More in general, there are no guidelines about how to identify the DSPs in a discourse. Our purpose is therefore twofold: to use RDA to make more specific claims about the DSPs in discourses (of a certain genre and in a given domain), and then to test G&S's claims (within this genre).

2.2 Relational Discourse Analysis (RDA)

Relational Discourse Analysis (RDA) (Moore and Pollack 1992; Moser and Moore 1996b) synthesizes Grosz and Sidner's approach and RST. RDA owes to Grosz and Sidner the idea that discourse is hierarchically structured, and that discourse structure is determined by intentional structure; each RDA-segment originates with an intention of the speaker. But in RDA segments have additional internal structure: each segment consists of one CORE, i.e., that element that most directly expresses the speaker's intention, and any number of CONTRIBUTORS, the remaining constituents in the segment, each of which plays a role in serving the purpose expressed by the core. The notions of core and contributor derive of course from the notions of nucleus and satellite in Rhetorical Structure Theory (RST) (Mann and Thompson 1988), which claims that in each "segment" (text span, for RST) one component should be identified as the 'main' one, and the others as secondary. However, in RST there is a distinction between nucleus and satellite for (almost) all RST relations, whereas in RDA core and contributors are only identified if a segment purpose has been recognized. In this case, each contributor is linked to the core by one intentional relation, and one informational relation. This is unlike RST, in which only a single relation can obtain between nucleus and satellite — see (Moore and Pollack 1992).

In RDA, segment constituents may in turn be other embedded segments, or simpler functional elements: these elements may be either basic UNITS, i.e., descriptions of actions and states, or relational CLUSTERS. Clusters are spans that only involve constituents linked by informational relations; no *core:contributor* structure exists, but they can themselves be embedded.

Unlike G&S's theory and like RST, RDA is based on a fixed number of relations;

1.1 Before troubleshooting inside the test station,
1.2 it is always best to eliminate both the UUT and TP.
2.1 Since the test package is moved frequently,
2.2 it is prone to damage.
3.1 Also, testing the test package is much easier and faster
3.2 than opening up test station drawers.

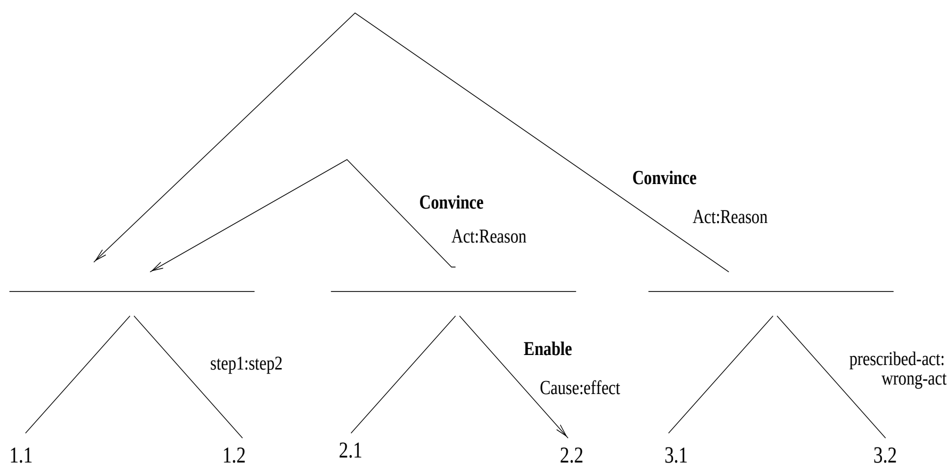

Figure 14.1: A tutorial excerpt and its RDA analysis

in particular, RDA assumes four intentional relations – **convince**, **enable**, **concede**, **joint**–and a larger set of informational relations; this latter set is expected to be domain dependent. In the Sherlock corpus, 23 informational relations are used, of which 13 pertain to causality (they express relations between two actions, or between actions and their conditions or effects) (Moser, Moore, and Glendening 1996).

Figure 14.1 shows a small excerpt from one of the dialogues in the Sherlock corpus, and its corresponding RDA analysis. The text is broken into clauses (UUT is "Unit under test", TP is "test package"). The analysis shows the text to be analyzed as an intentional segment whose core spans 1.1 and 1.2. This segment has two contributors, spanning 2.1 and 2.2, and 3.1 and 3.2 respectively. Graphically, the core is at the end of the arrow whose origin is the contributor; moreover, the link is marked by two relations, intentional (in bold), and informational. In this specific case, the two contributors carry the same intentional and informational relations to the core, but this doesn't need to be the case. The core and the two contributors are further analyzed. The core and the second contributor are analyzed as informational clusters, whereas the first contributor is recognized as having its own intentional structure.[1] Clusters are marked by one informational relation, but

[1]According to the manual used for the annotation (Moser, Moore, and Glendening 1996), an **enable** relation holds "if the contributor [2.1] provides information intended to increase the hearer's

not by intentional relations.

We will come back to the analysis of the text according to G&S in the next section. As regards the analysis of the text according to RST, the structure would presumably be the same, although no double relations would exist, and for every relation one relatum would be considered as the nucleus, the other(s) as its satellites.

3 Anaphoric Accessibility in RDA

In order to use an annotation based on RDA to test Grosz and Sidner's claims about anaphoric accessibility we have to specify a mapping from an RDA structure to focus spaces: which RDA constituents correspond to separate focus spaces, and which focus spaces should be open when a given anaphoric expression is encountered. This mapping is not entirely trivial, because the structure of a discourse according to RDA is much more detailed than the structure that would be assigned to that discourse by Grosz and Sidner. In RDA, each clause is treated as a distinct constituent, whereas in a G&S-style analysis, multiple sentences can be considered a single constituent. Furthermore, G&S make no distinction between cores and contributors, and only allow two relations between intentions, whereas in RDA many types of intentional relations are possible.

Moser and Moore partially specify a mapping between RDA notions and an intentional structure in the G&S sense based on the following principles:

- Every DSP must be associated with a core.
- Constituents of the RDA structure that do not include cores - i.e., clusters (see above) - do not introduce DSPs.

In terms of segments / focus spaces, these principles mean, first of all, that a segment in the G&S sense should always be a segment in the RDA sense (an RDA-SEGMENT): an element with a core and one or more contributors. I.e., no focus space should be pushed on the stack unless a core is recognized. (Moser and Moore leave open the question of whether the reverse should also be true, i.e., whether there should be a 1:1 mapping from G&S-segments to RDA-segments.) Secondly, informational relations do not affect the attentional state, unless associated with intentional relations.

The segment structure that–according to what we have seen so far–we can derive from the RDA analysis in Figure 14.1 should then be as in Figure 14.2. Notice that because informational relations don't give rise to intentional segments, the informational clusters 1.1-1.2 and 3.1-3.2 are not assigned separate segments. In particular, the unembedded core in 1.1-1.2 is not assigned a separate focus space, since it expresses (part of) the DSP associated with the overall RDA-segment.

understanding of the material presented in the core, or to increase the hearer's ability to perform the action presented in the core." (p. 6).

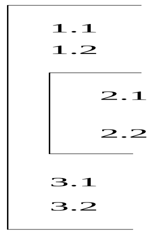

Figure 14.2: A G&S-style analysis for the text in Fig. 14.1

Even these first simple mapping principles already result in different predictions concerning accessibility than one would get from a pure RST analysis. In Fox's analysis, for example (see below), even if the constituent spanning 2.1 and 2.2 were a cluster, not an intentional segment, no antecedent introduced in these propositions would be accessible from 3.1 and 3.2. This is the first claim whose correctness we have to test.

However, the principles proposed by Moser and Moore leave a number of aspects of the mapping open. One question Moser and Moore themselves raise is whether embedded cores, i.e., cores that are themselves RDA-segments (whose possibility they consider, but do not analyze in detail) should be treated as embedded G&S-segments or as part of the same G&S-segment as their embedding RDA-segment. (Examples in which the antecedent of a pronoun is contained in an embedded nucleus are discussed by (Fox), p. 101.) A second question is how should embedded and not embedded contributors be treated: whether all of them should count as separate G&S-segments, or only the embedded ones.

Third, even if (some) contributors push focus spaces, when should these focus spaces be popped? Immediately, or only when the intentional relation is completed? E.g., in the example in Figures 14.1 and 14.2, should segment 2.1-2.2 be popped as soon as we are done processing it, or should it remain on the stack until the whole segment is over, given that it participates in the intentional relation that determines the superordinate segment? Fox's data seem to suggest that antecedents introduced by 'active' non-embedded satellites should be accessible; but even antecedents introduced by active embedded contributors might.

4 The Study

We tested all of these possible ways of using an RDA structure to guide the focus space construction mechanism using the Sherlock corpus. Our results show that the version in which both embedded core and embedded contributors remain on stack has the least number of reference violations.

4.1 Existing Data

The corpus we used is a collection of tutorial dialogues between a student and a tutor, collected within the Sherlock project (Lesgold, Lajoie, Bunzo, and Eggan 1992). The corpus includes seventeen dialogues between individual students and one of 3 expert human tutors, for a total of 313 turns (about 18 turns per dialogue), and 1333 clauses. The student solves an electronic troubleshooting problem interacting with the Sherlock system; then, Sherlock replays the student's solution step by step, schematically criticising each step. As Sherlock replays each step, the students can ask the human tutors for explanations. The student and tutor communicate in written form.

The Sherlock corpus had been previously annotated using RDA to study cue phrases generation (Moser and Moore 1996a; Di Eugenio, Moore, and Paolucci 1997). The research group which proposed RDA discusses the following reliability results (Moser and Moore 1996a). 25% of the corpus was doubly coded, and the κ coefficient of agreement was computed on segmentation in a stepwise fashion.[2] First, agreement at the highest level of segmentation was computed. After computing agreement at level 1, the coders resolved their disagreements, thus determining an agreed upon analysis at level 1. The coders then independently proceed to determine the subsegments at level 2, and so on. The deepest level of segmentation was level 5; the κ values were .90, .86, .83, 1, and 1 respectively (from level 1 to 5).

4.2 Our Methods

We annotated about half of the Sherlock corpus for anaphoric information, using a much simplified version of the annotation scheme developed by the GNOME project (Poesio 2000b). More specifically, we marked each NP in the corpus, specified its NP type (proper name, pronoun, the-np, indefinite NP, etc) and then we marked all 'direct' anaphors between these NPs (i.e., no bridges). This scheme has good results for agreement (Poesio 2000a) and has already been used for studying anaphoric accessibility (Poesio, Cheng, Henschel, Hitzeman, Kibble, and Stevenson 2000). We annotated a total of 1549 NPs, 507 of which were anaphoric.

One problem we had to address was that in the RDA annotation, only tutor turns had been annotated (because the students' questions are very short), but many of the antecedents of anaphoric references were discourse entities introduced in the preceding student turn asking the question. We followed Fox and made the first elements of adjacency pairs part of the accessibility space of anaphoric expressions in the second part. To do this, we enclosed each student turn in a special `student-turn` element, marked the NPs it contained, and made this turn a special focus space accessible from the entities in the tutor turn. We dealt with antecedents contained in 'tied' adjacency pairs (Fox 1987), i.e., in turns further away, by count-

[2] It is unknown to us whether κ was also computed on clusters, and on the specific informational relations used.

ing the anaphoric expressions whose antecedent was unaccessible for these reasons and factoring them out.

A second problem to solve was the fact that a large proportion of the anaphoric expressions whose antecedent is not on the stack are proper names. Because it can be argued that these expressions do not access the stack to find their antecedent, we also counted them separately.

We ran a script over the annotated corpus that simulates focus space construction under the several possible ways of mapping RDA structures into focus spaces that we considered, and attempts to find the antecedent for an anaphor in the focus space stack accessible to the anaphor according to each of these possibilities. The variants considered are:

1. **All:** Push a new focus space on the stack whenever a non-atomic RDA unit (both intentional segments and informational clusters) is encountered, and pop this focus space when the constituent ends. (E.g., in Figure 14.1, push a new focus space for all three constituents of the top segment: 1.1-1.2, 2.1-2.2, and 3.1-3.2.)
2. **Intentional Only / Imm Pop**: Only push a new focus space when an intentional segment is encountered; pop it as soon as the segment is completed. (E.g., in Figure 14.1, only push a new focus space for segment 2.1-2.2, and pop it as soon as that segment is completed. 1.1, 1.2, 3.1, and 3.2 are just added to the top focus space.)
3. **Intentional Only / Delay pop of cores**: Only push a new focus space when an intentional segment is encountered. Pop focus spaces introduced for contributor segments immediately; but only pop the focus space introduced for a core segment when popping the whole segment .
4. **Intentional Only / Partial delay pop of tribs** Like the previous version, but in addition keep a focus space introduced for a contributor on the stack as long as the segment in which it occurs is still on. (E.g., in Figure 14.1, do not pop the focus space for segment 2.1-2.2 before processing 3.1-3.2.)

5 Results

Table 14.1 illustrates the impact on anaphoric accessibility of the distinction between intentional and informational relations by showing the percentage of anaphoric antecedents which are on the stack according to the first two variants of the mapping algorithm. The line indicated as ‘All’ shows the results obtained by treating both informational and intentional relations as introducing new focus spaces: ’OK’ indicates the number of anaphoric antecedents which are accessible, ’NO’ indicates the number of antecedents which are not accessible, ‘Out of AP’ the cases in which the antecedent is not accessible because it’s outside the current Adjacency Pair, and ’PN’ the number of cases in which the antecedent is not accessible but the anaphoric expression is a proper name (which, presumably, can access it denotation through long term memory rather than through the stack). The table shows that separating intentional constituents (which introduce new focus spaces) from informational clusters (that don’t) makes more antecedents accessible; the result is highly significant by the χ^2 Test ($\chi^2 = 29.47, p \leq 0.001$).

	OK	NO	Out of AP	PN
All:	199	74	63	158
Intentional only / Imm Pop:	280	20	63	131

Table 14.1: The informational / intentional distinction and accessibility

	OK	NO
Int. only / Imm pop of emb core and trib	280	20
Int. only / Delay pop of emb cores	287	16
Int. only / Delay pop of emb trib	310	8

Table 14.2: Effect of the differences in popping.

Table 14.2 shows the differences among the different ways of fixing the options left open by Moser and Moore (variants 2-4 above). The first line shows the results obtained if both embedded cores and embedded contributors were to push new focus spaces, popped as soon as these embedded constituents are completed; the second line the results if we keep embedded cores open until the end of the RDA-segment; the third line the results if we treat embedded contributors within an RDA-segment as remaining on the stack until the segment is closed off. In this table we have ignored both cases in which the antecedent is inaccessible but the anaphoric expression is a proper name, and the 63 cases in which the antecedent is inaccessible because it's not in the same adjacency pair (see discussion above). The differences are not so great in this case, but the correlation is still significant ($\chi^2 = 6.09, p \leq 0.05$) and in particular there is a clearly significant difference between the simplest possible treatment of embedded intentional relations, in which they are always popped, and the last model.

6 Discussion

As said above, we studied two separate issues: how best to use ideas from RDA to make G&S's theory more specific, and to evaluate G&S's proposals concerning anaphoric accessibility under this mapping.

6.1 Mapping RDA into Focus Spaces

The first goal of our work is to use RDA structures to gain a more detailed understanding of when focus spaces should be opened and closed. In this respect, our first result is a significantly better characterization of anaphoric accessibility if we assume that new focus spaces are only pushed on the stack when cores are recognized, as opposed to also being pushed when a purely informational structure is observed. This result is especially interesting when compared with Fox's proposal (discussed below). According to Fox, informational relations also affect accessibility. It is also interesting to contrast this result with the proposals of Veins Theory, which also makes no distinction between informational and intentional relations.

The second interesting finding in this respect is that the best results concerning accessibility are obtained when embedded contributors, as well, are only popped when an RDA-segment is closed. This is not something that would be predicted on the basis of either Fox's work, or any obvious interpretation of the notion of RDA-segment; in fact, as we will see shortly, it's not immediately obvious how to account for this result in terms of Grosz and Sidner's theory, either.

6.2 An Evaluation of Grosz and Sidner's Theory

If we assume that contributors stay on the stack until the RDA-segment is completed, only 8 anaphoric antecedents are outside the currently open focus spaces. Of these, 5 are cases of definite descriptions that might be viewed as referring deictically to parts of the circuit, 1 is a cataphoric discourse deixis, one is a temporal deixis (*at this point*), and one a bit unclear. In other words, under the suggested rules for opening and closing focus spaces, virtually all anaphoric antecedents are accessible.

The question then is whether these rules are consistent with the G&S framework: the answer, we think, is that while the simplest treatments of contributors and cores probably are, the most complex ones probably aren't. In fact, we didn't even test what some people might think of as the most natural mapping - make the unembedded core of an RDA segment part of the focus space for that segment, but treat unembedded contributors as introducing subordinate focus spaces - since this would violate one of Moser and Moore's basic hypotheses (that only RDA-segments result in G&S-segments). The approach of leaving unembedded cores and contributors on the stack, but popping all embedded RDA-segments, can probably be still viewed as consistent with the theory;[3] but delaying the popping of embedded cores and contributors on the stack may be considered beyond its limits, whether we view the problem in algorithmic terms or in terms of intentional structure. Leaving these focus spaces open means that the attentional state cannot be properly seen as a stack anymore, but has to be seen as a list, since there is no guarantee that the focus space associated with the core will be the last element on the stack. Figure 14.3 shows an example in which it is necessary to leave the embedded contributor 24.13-24.14 on the stack, in order to solve *the "other" voltage* to *some other voltage*; at the end of the segment, this focus space has to be removed while leaving the focus space for the core on top of the stack. Furthermore, note that in the case of Figure 14.1 we need a 'discontinuous focus space' to make the embedded contributor accessible without eliminating the principle that new entities are always added to the focus space on top of the stack.

Looking at the problem from the point of view of intentional structure, these cases of accessibility can only be handled within the G&S framework by hypothe-

[3]This however means that the intentional relations contained in the segment - e.g., **convince** in Figure 14.1–should become part of the DSP for that segment, which again may or may not be consistent with G&S's ideas.

24.13a Since S52 puts a return (0 VDC) on it's outputs
24.13b when they are active,
24.14 the inactive state must be some other voltage.
24.15 So even though you may not know what the "other" voltage is,
24.16 you can test to ensure that
24.17a the active pins are 0 VDC
24.17b and all the inactive pins are not 0 VDC.

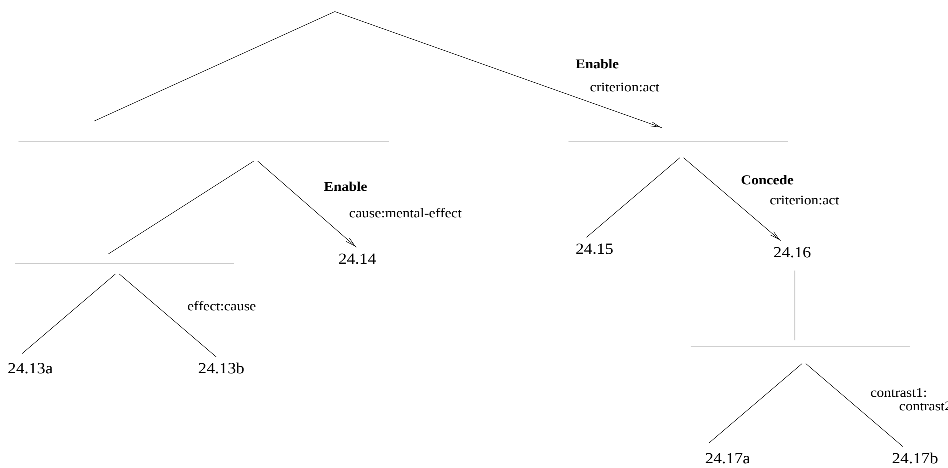

Figure 14.3: A contributor that precedes the core

sizing that the DSPs associated with embedded cores and contributors **satisfaction-precede** the core and the other contributors. More in general, one has to wonder whether Fox's (and before that, Reichman's) suggestion that the material introduced by 'active' propositions (i.e., the other constituents of a rhetorical scheme) remains accessible can really be formulated in terms of Grosz and Sidner's intentional structure.

6.3 Open Issues

Keeping more entities on the stack is obviously going to increase the likelihood that an antecedent will be found on the stack, but we have to be careful not to lose the crucial property of the attentional state, reducing search ambiguity (otherwise, we could just make every discourse entity accessible at all times). In order to properly compare models of anaphoric accessibility it is necessary to compute what we might call their 'perplexity'; unfortunately, in order to properly do this it is necessary to have an anaphoric resolution mechanism . We plan to do this evaluation next.

A major problem with any attempt at looking at the empirical import of discourse structure is that it's very hard to get researchers to agree on what the structure of a given discourse is. Work such as Marcu's has shown that in general we can expect people to roughly agree on the main segments of a discourse, but it is much harder to get agreement on the relative importance of the constituent segments, and even harder to get agreement on the particular relations. At the moment, the only way to achieve a reliable annotation is by extensive training and discussions, as done when constructing the Sherlock corpus. So, we do not expect everybody to agree on the particular analyses proposed here; we would want to point out however that these annotations do have a certain amount of objectivity in that they were not produced by us.

7 Related Work

7.1 Fox

(Fox), although only concerned with references to singular and human antecedents, is perhaps the most extensive study of the effects of discourse structure on anaphora in both spoken and written discourses. Fox uses different methods for analyzing the two genres: she uses concepts from Conversation Analysis, and in particular the notion of Adjacency Pair, for spoken conversations, and RST to analyze written texts. Her main proposal about written texts is as follows:

> A pronoun is used to refer to a person if there is a previous mention of that person in a proposition that is ACTIVE or CONTROLLING; otherwise a full NP is used.

(Where a proposition is ACTIVE if it's part of the same RST scheme as the proposition in which the pronoun occurs; whereas a proposition is CONTROLLING if it's part of a scheme which dominates the scheme in which the pronoun occurs.)

Fox's proposals concerning pronominalization apply less well to references to objects (and even in her corpus there are many references for which the hypothesis above would licence the use of a pronoun are actually realized by a definite NP, which she explains by arguing that the principle above is only one of many interacting principles that determine the realization of a NP); nevertheless, she makes a lot of compelling points about structure. In particular, she makes it very clear that active propositions should be accessible for as long as the scheme is open; and produces several examples showing that material introduced in active embedded nuclei is accessible. Fox didn't find references inside active embedded satellites (but then again none of these is made via a pronoun in our corpus). In addition, our study suggests that proper names behave differently from definite descriptions in that the former are much less sensitive to discourse structure than the latter, so the two classes should not be conflated like Fox does; and not separating informational relations from intentional ones restricts too much the range of accessible antecedents, even if it may be correct as far as pronominalization is concerned.

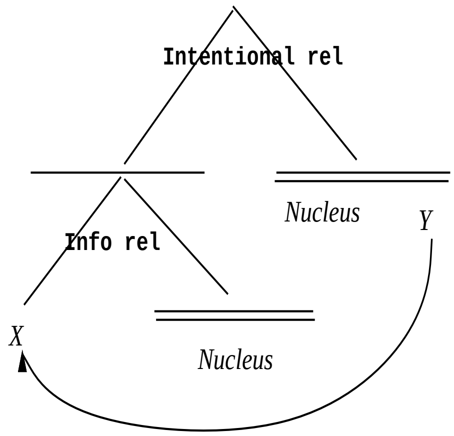

Figure 14.4: A tree distinguishing Veins Theory from the proposals discussed in this paper

7.2 Veins Theory

Veins Theory (VT) (Cristea, Ide, and Romary 1998; Cristea, Ide, Marcu, and Tablan 2000; Ide and Cristea 2000) is a recently proposed theory of the effect of discourse structure on anaphoric accessibility, which relies on RST for its definition of discourse structure, and whose predictions have been tested using an RST annotated corpus of newspaper texts (Cristea, Ide, Marcu, and Tablan 2000). The propositions accessible to an anaphoric expressions are computed by an algorithm that operates directly over an RST tree and involves two steps: a bottom up step in which the 'heads' of each node in the tree are computed (where the head of a non-terminal node is the concatenation of the heads of its nuclear daughters) followed by a top-down computation of the VEIN EXPRESSIONS. The crucial idea of VT is that material introduced in nuclear nodes 'percolates up' veins, where veins are paths in the tree all of whose arcs connect nuclear nodes; the antecedents introduced in any node along the vein are accessible from all the nodes of the subtree which has the top of the vein to which that node belongs as its root. The second idea of the theory is that antecedents introduced in a satellite node to the left of a nucleus remain accessible to all nodes controlled (in Fox's sense) by that nucleus.

In some respects, the proposal presented here can be viewed as a generalization of the proposals of VT: the material introduced in core constituents percolates up in a similar way, but we also allow antecedents introduced by embedded contributors to the *right* of the nucleus to be accessible as long as these contributors are active (in VT, only binary trees are considered).

The one point of contrast between the two theories is that in our proposal, we do not consider all nuclei, but only the nuclei of intentional relations. We have seen that treating informational relations as introducing focus spaces makes a big difference in terms of accessibility; the difference is significant even if we allow these additional relations to remain open as long as the dominating relation is open. The crucial case distinguishing the two theories are trees with the structure in Fig 14.4: in our theory X would be available as an antecedent of Y, whereas in Veins Theory it wouldn't.

Bibliography

Carletta, J., A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A. H. Anderson (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics 23*(1), 13–32.

Cristea, D., N. Ide, D. Marcu, and V. Tablan (2000). Discourse structure and co-reference. In *Proc. of COLING*, pp. 208–214.

Cristea, D., N. Ide, and L. Romary (1998). Veins theory: A model of global discourse cohesion and coherence. In *Proc. of COLING*, pp. 281–285.

Di Eugenio, B., J. D. Moore, and M. Paolucci (1997). Learning features that predict cue usage. In *Proc. of the 35th ACL*, Madrid.

Fox, B. A. (1987). *Discourse Structure and Anaphora*. Cambridge.

Grosz, B. J. and C. L. Sidner (1986). Attention, intention, and the structure of discourse. *Computational Linguistics 12*(3), 175–204.

Ide, N. and D. Cristea (2000, October). A hierarchical account of referential accessibility. In *Proc. of ACL*, Hong Kong.

Lesgold, A., S. Lajoie, M. Bunzo, and G. Eggan (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin and R. Chabay (Eds.), *Computer assisted instruction and intelligent tutoring systems*, pp. 201–238. Hillsdale, NJ: Erlbaum.

Mann, W. C. and S. A. Thompson (1988). Rhetorical structure theory: Towards a functional theory of text organization. *Text 8*(3), 243–281.

Marcu, D. (1999, May). Instructions for manually annotating the discourse structures of texts. Unpublished manuscript, USC/ISI.

Moore, J. and M. Pollack (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics 18*(4), 537–544.

Moser, M. and J. D. Moore (1996a). On the correlation of cues with discourse structure. Unpublished manuscript.

Moser, M. and J. D. Moore (1996b). Toward a synthesis of two accounts of discourse structure. *Computational Linguistics 22*(3), 409–419.

Moser, M., J. D. Moore, and E. Glendening (1996). Instructions for Coding Explanations: Identifying Segments, Relations and Minimal Units. Technical Report 96-17, University of Pittsburgh, Computer Science.

Nakatani, C. H. (1996). Discourse structural constraints on accent in narrative. In J. P. H. van Santen, R. W. Sproat, J. P. Olive, and J. Hirschberg (Eds.), *Progress in Speech Synthesis.* Springer Verlag.

Nakatani, C. H., B. J. Grosz, D. D. Ahn, and J. Hirschberg (1995). Instructions for annotating discourses. Technical Report TR-25-95, Harvard University Center for Research in Computing Technology.

Poesio, M. (2000a, May). Annotating a corpus to develop and evaluate discourse entity realization algorithms: issues and preliminary results. In *Proc. of the 2nd LREC*, Athens, pp. 211–218.

Poesio, M. (2000b, July). *The GNOME Annotation Scheme Manual.* University of Edinburgh, HCRC and Informatics. Available from `http://www.hcrc.ed.ac.uk/ ~ gnome`.

Poesio, M., H. Cheng, R. Henschel, J. M. Hitzeman, R. Kibble, and R. Stevenson (2000, October). Specifying the parameters of Centering Theory. In *Proc. of the 38th ACL*, Hong Kong.

Reichman, R. (1985). *Getting Computers to Talk Like You and Me*. Cambridge, MA: The MIT Press.