

# Accounting for Discourse Relations: Constituency and Dependency



# 1

# Accounting for Discourse Relations: Constituency and Dependency

Bonnie Webber

At the start of my career, I had the good fortune of working with Ron Kaplan on Bill Woods' Lunar system (Woods et al., 1972). One day, in talking with Ron, I marvelled to him over the range of syntactic constructions I was able to implement in Lunar's ATN grammar formalism. Ron replied was that you could implement *anything* in an ATN: the point was, rather, to identify the minimal machinery required for a task. This sensible advice I subsequently sought to follow, and in this paper for Ron's festschrift, I try to apply it to understanding and comparing accounts of discourse relations.

## 1.1 Introduction

Theories of discourse that attempt to explain how the meaning of a text is more than the sum of the meaning of its component sentences have often been presented in ways that discourage easy comparison. One attempt to remedy this was made by Moore and Pollack (1992), who suggested a distinction between an *intentional* organization of discourse and an *informational* organization. Intentional organization could be described in terms of the speaker's plans with respect to his/her utterances and how the parts of the plan relate to each other (Grosz and Sidner, 1990, Lochbaum, 1998), while informational organization could be described in formal semantic and pragmatic terms, with discourse relations such as *consequence*, *cause*, *explanation*, *contrast*, *narration*,

*Festschrift for Ron Kaplan.*
Mary Dalrymple (ed.).

etc.

While this solved part of the problem, it did not contribute to understanding how theories concerned with the informational organization of discourse differ from one another. In this paper, I want to suggest that recasting them in terms of the common linguistic concepts of *constituency* and *dependency* might help us to better understand their similarities and differences.

## 1.2 Constituency and Dependency

### 1.2.1 Constituency

In syntax, *constituency* has been defined in the following terms:

1. Sentences have parts, which may themselves have parts.
2. The parts of sentences belong to a limited range of types.
3. The parts have specific roles or functions within the larger parts they belong to. (Huddleston and Pullum, 2002, 20)

Of course, different theories of syntax posit different parts, and different parts may be appropriate for different languages. Nevertheless, it is this idea of parts within parts that is basic to *constituency*, as is the idea that parts have specific roles or functions. Both ideas are integral to the compositional interpretation of constituent structure (i.e., *compositional semantics*).

Another aspect of constituency is that parts are *continuous spans*. A constituent can become *discontinuous* if (1) part of it moves somewhere else — e.g., to the front of the sentence in questions (the discontinuous PP in "Which pocket is your wallet in?"), to the rear of the sentence when the informational content of a relative clause is more than that of the main predicate (the discontinuous NP in "A man came in who was wearing a green hat"); or (2) the constituent is interrupted by a parenthetical phrase (the discontinuous VP in "I can tell you, since you ask, more about my parents."). But even in languages with free word order, the parts of one constituent will not be arbitrarily scrambled with those of others.

### 1.2.2 Dependency

There are several notions of *dependency* in linguistics. One is a *syntactic* notion related to morphology, where the morphological realization of a lexico-syntactic element depends on another element elsewhere in the sentence or clause, including realisation as an empty element or *gap*. Such a syntactic dependency may be unbounded or bounded.

In English, bounded syntactic dependencies include gender and number agreement, as in:

(1) a. These$_i$ boys$_i$ shave$_{(i)}$ themselves$_i$.

b. This$_i$ boy$_i$ shaves$_{(i)}$ himself$_i$.

Here, $i$ used to co-index the dependent element (simple subscript) and the element it depends on (parenthesized subscript). The dependency shows itself in the morphological features of the determiner, noun, verb and reflexive pronoun.

*Unbounded dependency constructions* in English (Huddleston and Pullum, 2002, 1079) contain a *gap* in a position that syntax requires to be filled, as in

(2) a. This is [the book]$_{(i)}$ which$_i$ I think Fred said he wrote $\_\_{}_i$.

b. [The other chapters]$_{(i)}$ I think Fred said he wrote $\_\_{}_i$ himself.

Here $\_\_{}_i$ indicates the gap in the object position in both (2a) and (2b), with $i$ being the index of the element it depends on. *Unbounded* refers to the arbitrary depth to which the clause containing the gap can be embedded. Syntactically, *dependency* refers to the fact that the relative pronoun in (2a) and the topicalized NP in (2b) require an associated gap. The semantic consequences are that the gap in both examples draws its interpretation from the constituent that it is co-indexed with and that the topicalised NP in (2b) would be taken as being in contrast with something else in the discourse context.

A second notion of dependency between words underlies *dependency grammar*. Hudson defines this sense of dependency in terms of *support*: A dependent word is *supported* by the word it depends on, which allows it to occur in a sentence and also constrains such features as its possible location in a sentence and its possible morphological form.[1] This generalizes the previous notion of dependency, although it is limited to word-to-word relations, rather than relations between lexico-syntactic constituents, including gaps.

While both these notions of dependency are syntactic, they can also have semantic consequences in terms of how a sentence is interpreted. On the other hand, *semantic dependency* affects only interpretation — and in particular, truth-conditional semantics. Linguistic elements that exert such influence include quantifiers (3a), negation (3b) and adjuncts (3c), and what they can influence includes both reference and truth values.

(3) a. I need *a student* to work on *every project*.

b. I do *not* love you *because you have red hair*.

[1] http://www.phon.ucl.ac.uk/home/dick/enc/syntax.htm#dependency, *The Encyclopedia of Dependency Grammar*, accessed April 2006.

c. *A woman* has been elected president *for the second time*. (Huddleston and Pullum, 2002, 719)

Their range of influence is called their *scope*, which will be either *narrow* or *wide*. For example, if *every project* is taken to have narrow scope in (3a), the term *a student* will be interpreted as referring independently of the set of projects being iterated over (meaning that one student is needed for the whole set of projects). With wide scope, *a student* will be interpreted as dependent on the set, meaning that the student that is needed depends on the project. In (3b), it is a truth value that is affected: if *not* has narrow scope, it is the proposition headed by *love* that is negated (ie, because you have red hair is the reason I *don't* love you), while if it has wide scope, it is the proposition headed by *because* that is negated (ie, because you have red hair is *not the reason* that I do love you). Finally, in (3c), if the adjunct *for the second time* is taken to have narrow scope, the referent of *a woman* is independent of the election (ie, the same woman has been elected twice), while with wide scope the referent is dependent on it (ie, a possibly different woman has been elected each time). It should be clear that scope does not correlate directly with relative linear order: within a clause, a scope-bearing element can appear to the left or right of the elements it has scope over.

The final notion of dependency is *anaphoric dependency*. Here, all or part of an element's interpretation depends on what is available in the discourse context. For example, the interpretation of *he* in (4a) depends on, and is coreferential with, the man introduced in the previous sentence.[2] In (4b), the interpretation of *another man* is dependent on the same thing, but only in part: it is a man other than that one. In (4c), the interpretation of *one* depends on set descriptions available in the context: here, *man*, while in (4d), the interpretation of the ellipsed VP depends on available predicates. In (4e), the interpretation of the demonstrative pronoun *that* depends on an *abstract object* (i.e., a fact, proposition, eventuality, claim, etc. (Asher, 1993, Webber, 1991)) available from the previous discourse — here the action of ordering a single malt. Finally, in (4f), the interpretation of the adverbial *instead* depends on available predications that admit alternatives — here, *refusing a drink*, which admits the alternative *accepting one* — i.e., instead of accepting a drink, she started talking (Webber et al., 2003).

(4) a. A man walked in. *He* sat down.

[2]This notion of *anaphoric dependency* does not cover such intra-clausal coreference as "John shaves his father" and "John shaves himself", which fit better under the notion of *(bounded) syntactic dependency*.

b. A man walked in. *Another man* called him over.

c. A tall man walked in, then a short *one*.

d. A man walked in and ordered a single malt. Then a woman *did*.

e. A man walked in and ordered a single malt. *That* showed he had good taste.

f. The woman refused a drink. *Instead* she started talking.

Only *syntactic dependency* is part of syntax. *Semantic dependency* is part of the semantic composition process that operates alongside or on the result of syntactic analysis, while *anaphoric dependency* is outwith both syntactic analysis and semantic composition (but cf. footnote 2).

Since syntactic dependencies can impact the power of a grammar, one on-going challenge in linguistics has been to understand whether a given phenomenon is a matter of syntactic, semantic or anaphoric dependency. For example, it was once thought that the relation induced by the adverbial *respectively* was a matter of syntactic dependency between elements, such as in (5a), where the two parts of the conjoined subject are paired in order with the two parts of the conjoined verb phrase (VP) — ie, John washing and Mary painting, and as in (5b), where the three parts of the conjoined subject are so paired with the three parts of the conjoined VP.

(5) a. I think that John and Mary will *respectively* wash his car and paint her boat today.

b. I think John, Mary and Kim will *respectively* wash his car, paint her boat, and clean their room today.

If these pairings come from the grammar, it would mean two crossing dependencies (one crossing point) for (5a) and three crossing dependencies (three crossing points) for (5b). More generally, a sentence with N conjoined subjects, N conjoined VPs and *respectively*, would have N crossing dependencies and N*(N-1)/2 crossing points. This would require a grammar to have more than context-free power.

Later, however, it was noted that negation removed the need for such pairing:

(6) a. I think John and Mary will wash his car and paint her boat, but probably not *respectively*.

b. I think John, Mary and Kim will wash his car, paint her boat, and clean their room, but probably not *respectively*.

This wouldn't happen if the dependency were syntactic or even semantic. Thus, the individual dependencies associated with *respectively* must

be the result of inference based on anaphoric dependency.[3]

In the following sections, I will use these notions of *constituency* and *dependency* to characterize the source of discourse relations in several theories of discourse. This will, I hope, illuminate both their similarities and differences. However, I do not have the space here for what would have to be an extended discussion of SDRT (Asher and Lascarides, 2003), given the deep involvement of *inference* in how it derives discourse relations.

### 1.3 Dependency as a Source of Discourse Relations

The theory of discourse articulated by Halliday and Hasan (1976) is one in which discourse relations can be seen to arise solely from *anaphoric dependency*. Specifically, Halliday and Hasan (1976) consider *cohesion* to be what relates parts of a discourse, where *cohesion* is defined as holding when one part cannot be effectively interpreted except by recourse to the interpretation of another part. Halliday and Hasan posit five types of cohesion, each associated with a particular set of lexical or syntactic elements:

1. elements expressing referential identity or dependence, such as pronouns and other forms of anaphora;
2. substitution, as with *one(s)*, *so* and *do so*
3. ellipsis, including nominal ellipsis (e.g., "the best hat" → "the best"), verb phrase ellipsis, and clausal ellipsis;
4. lexical cohesion, as in the reiteration of the same word, or a synonym or near synonym, or a superordinate term or generic word;
5. conjunction, as expressed through coordinating and subordinating clausal conjunctions, adverbials like *later on*, and prepositional phrases like *in that case*.

The fifth is the source of discourse relations in Halliday and Hasan (1976) and is, in fact, the sole source. That is, discourse relations arise from identifying what other part of the discourse the interpretation of a conjunctive element depends on. It should be clear that *cohesion* as a source of discourse relations is essentially *anaphoric dependency*, with three interesting features. First, there is no theoretical constraint on its locality, since any part of a text can theoretically depend on any other part. Secondly, there are no constraints on how many parts of the text a given part may depend on: in the analysis given in Halliday and Hasan (1976, Ch. 8.3), multiple cohesive links between the parts of a sentence

[3]The fact that *respectively* can be paraphrased with the demonstrative *that* — ie, "in *that* order" — provides additional evidence for this conclusion.

(including the sentence as a whole) and the previous discourse are the norm, rather than the exception. Thirdly, there are no constraints on what parts of a discourse can be linked together, so cohesive links are as likely to cross one another as to be embedded.

Note that Halliday and Hasan explicitly reject any notion of structure or *constituency* in discourse, saying for example:

> Whatever relation there is among the parts of a text — the sentences, the paragraphs, or turns in a dialogue — it is not the same as structure in the usual sense, the relation which links the parts of a sentence or a clause. (Halliday and Hasan, 1976, 6)

> We doubt whether one can demonstrate generalized structural relations into which sentences enter as the realisation of functions in some higher unit as can be done for all units below the sentence. (Halliday and Hasan, 1976, 10)

> Between sentences, there are no structural relations. (Halliday and Hasan, 1976, 27)

Thus, only the notion of *anaphoric dependency* is needed in describing the source of discourse relations in Halliday and Hasan (1976): neither *constituency* nor *scope* plays any role.

## 1.4 Constituency as a Source of Discourse Relations

In contrast with Halliday and Hasan (1976), both Mann and Thompson's *Rhetorical Structure Theory* (RST) and Polanyi's *Linguistic Discourse Model* (LDM) take *constituency* as the sole basis for discourse relations. Both provide an exhaustive top-down context-free constituency analysis of a text, associating with many (LDM) or all (RST) constituents, a formal pragmatic account in terms of discourse relations. Differences between the two lie in what they take to be a constituent and how the associate discourse relations with constituents.

### 1.4.1 Rhetorical Structure Theory

In RST (Mann and Thompson, 1988), the constituency structure of a discourse consists of instantiated *schemas* which specify discourse relations (called here *rhetorical relations*) between adjacent spans, which may be clauses or the projection of instantiated schemas. The set of *schemas* essentially defines a context-free (CF) grammar on a single non-terminal, which we can call $D$. RST has five kinds of *schemas*, which differ with respect to how they re-write $D$ in terms of (i) rhetorical relations that hold between right-hand side (RHS) sisters; (ii) whether or not the RHS has a head (called in RST, a *nucleus*); and (iii) whether there are two, three, or arbitrarily many sisters (the latter

like Kleene plus).

The first is a headed binary-branching schema in which a specific rhetorical relation such as CIRCUMSTANCE or EVIDENCE holds between the head daughter and its sister (called a *satellite*). As the order of the daughters doesn't matter, such a schema actually stands for two standard CF rules. This type of schema is illustrated in Figures 1(a) and 2(a), where the former follows the conventions used in Mann and Thompson (1988), and the latter follows more standard tree-drawing conventions. RST also allows for an N-ary version of this schema, which retains a single head daughter, but with multiple sisters, all of which bear the same rhetorical relation with the head.

The second is a headed binary-branching schema in which the rhetorical relation CONTRAST holds between two daughters of equal headedness. This schema is illustrated in Figures 1(b) and 2(b), where again, the former follows the graphic conventions used in Mann and Thompson (1988), and the latter follows more standard tree-drawing conventions.

The third is an N-ary branching schema called JOINT, in which no rhetorical relation is taken to hold between sisters that nevertheless belong together in some way, and all sisters are equal in headedness.

The fourth is a ternary-branching schema in which the head has a MOTIVATION relation to one sister and a ENABLEMENT relation to the other. As in the first type of schema, the order of the daughters doesn't matter, so that this schema actually stands for six standard CF rules, each corresponding to a different order of the three sisters. This schema is illustrated in Figures 1(d) and 2(d).

The final type of schema is a multi-headed N-ary branching rule in which each sister except the first is related to its left-adjacent sister by a SEQUENCE relation. This is illustrated in Figures 1(e) and 2(e). Since one's choice of schema commits one to a particular discourse relation (or none) holding between sisters, there is no way to say that two adjacent spans in a discourse are related, without saying what the relation is. This is one way that RST differs from the *Linguistic Data Model*, to be discussed next.

An RST analysis of a discourse is a tree whose root non-terminal covers the entire string span and where adjacent non-terminals in any cut across the tree cover adjacent string spans. A possible RST analysis of the short discourse in Example 7 is shown in Figure 3, with Figure (3a) following RST conventions and Figure (3b) following more standard tree-drawing conventions.

(7) a. You should come visit.
b. Edinburgh is lovely in early fall
c. and there are no rabbits around.

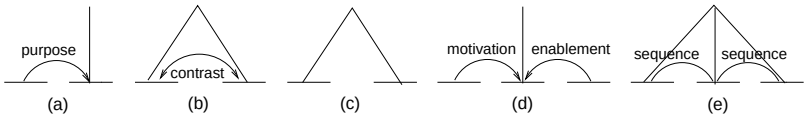


FIGURE 1 RST schema types in RST notation

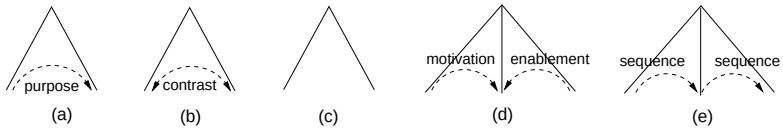


FIGURE 2 RST schema types in standard tree notation

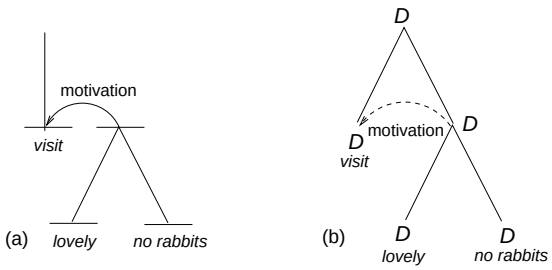


FIGURE 3 RST analysis of Example 7

Much of Mann and Thompson (1988) is concerned with delineating the conditions under which it is appropriate for someone analysing a discourse to assert that a particular relation holds between adjacent spans, which then determines what schema applies. Such conditions can be placed on (1) the informational content of the spans themselves, (2) the speaker's perceived intent with respect to each span and its role with respect to its sister(s), and (3) the intended effect of the span. RST requires an analyst to produce a single RST analysis of a discourse, so judgments must be made in cases of perceived ambiguity and in cases where more than one rhetorical relation can simultaneously be taken to hold between sisters.

What RST does not do is place conditions on anything elsewhere in the discourse. Thus it views discourse relations solely in terms of context-free *constituency*. They can hold only between sisters on the RHS of some schema, although those sisters may correspond to either terminal nodes or the non-terminal nodes of instantiated RST schema. (This latter point will become relevant in comparing RST and the *GraphBank* approach presented in Section 1.5.2.)

### 1.4.2 Linguistic Discourse Model (LDM)

The Linguistic Discourse Model (Polanyi, 1988, Polanyi and van den Berg, 1996, Polanyi et al., 2004) is a theory of discourse and discourse parsing that resembles RST in constructing an explicit tree-structured representation of discourse constituents, but differs in separating discourse structure from discourse interpretation. It does this by having three (and only three) context-free re-write rules, each associated with a rule for semantic composition:

1. an N-ary branching rule for *discourse coordination*, in which all the RHS sisters bear the same relationship to their common parent (used for elements of lists and narratives). Here there is no specific relation between sisters. The interpretation of the parent node is the information common to all its daughters.
2. a binary branching rule for *discourse subordination*, in which one sister (considered *subordinate*) elaborates an entity or situation described in the other (considered *dominant*). Here, ELABORATION is the discourse relation between the sisters, and the interpretation of the parent node is the interpretation of the dominant daughter.
3. an N-ary branching rule in which the RHS sisters are related by a logical or rhetorical relation, or by a genre-based or interactional convention. Here, the interpretation of the parent node derives

> from the interpretation of each daughter and from the relationship between them. Evidence for that relationship can come from lexical and/or syntactic information. It appears that this rule could apply without the particular relationship being fully specified, awaiting further specification later on. Such a move would not be possible in RST.

Recent work on the LDM has been concerned with the issue of automatically parsing discourse efficiently, with respect to the model. But the issue of concern here — whether the source of discourse relations in the LDM is *constituency* or *dependency* — comes down firmly for the former.

## 1.5 Mixed Approaches to Discourse Relations

In contrast with the theories presented earlier, both D-LTAG (Webber et al., 2003) and the theory underlying GraphBank (Wolf and Gibson, 2005) exploit both *constituency* and *anaphoric dependency* in their accounts of discourse relations, but in very different ways.

### 1.5.1 A Lexicalized TAG for Discourse (D-LTAG)

D-LTAG is a lexicalized approach to discourse relations (Webber et al., 2001, Forbes et al., 2003, Webber et al., 2003, Webber, 2004, Forbes-Riley et al., 2006). Lexicalization means that D-LTAG provides an account of how lexical elements (including some phrases) anchor discourse relations and how other parts of the text provide arguments for those relations.

D-LTAG arose from a belief that the mechanisms for conveying discourse relations were unlikely to be entirely different from those for conveying relations within the clause. Because the latter can be anchored on lexical items, D-LTAG was developed as a *lexicalized grammar* for discourse — in particular, a lexicalized Tree Adjoining Grammar (Schabes, 1990). A lexicalized TAG (LTAG) differs from a basic TAG in taking each lexical entry to be associated with the set of tree structures that specify its local syntactic configurations. These structures can be combined via either *substitution* or TAG's *adjoining* operation, in order to produce a complete sentential analysis. In D-LTAG, elementary trees are anchored (by and large) by discourse connectives (representing predicates), whose substitution sites (arguments) can be filled by clauses or other trees.

Elementary trees anchored by a *structural connective* (ie, a coordinating or subordinating conjunction, a subordinator such as *in order to*, *so that*, etc.), or what we call an *empty connective* are used to build

constituent structure. The compositional interpretation of this structure is in terms of discourse relations between arguments (Forbes-Riley et al., 2006). Discourse adverbials, on the other hand, exploit *anaphoric dependency* to convey a discourse relation between the *abstract object* (AO) interpretation of its matrix clause and the AO interpretation of a previous clause, sequence of clauses, or nominalization in the discourse. That discourse adverbials such as *instead*, *afterwards*, *as a result*, etc. differ from structural connectives in terms of the distribution of their arguments is demonstrated on theoretical grounds in Webber et al. (2003) and on empirical grounds in Creswell et al. (2002). An explanation for the anaphoric character of discourse adverbials is given in Forbes (2003) and Forbes-Riley et al. (2006).

Both *constituency* and *dependency* can be seen in the D-LTAG analysis of Example 8:

(8) John loves Barolo.
So he ordered three cases of the '97.
But he had to cancel the order
because he *then* discovered he was broke.

The analysis is shown in Figure 4. It involves a set of elementary trees for the connectives (*so, but, because, then*) and a set of leaves ($T1$-$T4$) corresponding to the four clauses in Example 8, minus the connectives. Through the operations of *substitution* (solid lines) and *adjoining* (dashed lines) recorded in the *derivation tree* (here shown to the right of the arrow), a *derived tree* is produced (here shown at the head of the arrow). More detail on both the representation of connectives and D-LTAG derivations is given in Webber et al. (2003). A preliminary parser for D-LTAG is described in Forbes et al. (2003).

Compositional interpretation of the derivation tree produces the discourse relations associated with *because*, *so* and *but*, while anaphor resolution produces the other argument to the discourse relation associated with *then* (ie, the ordering event), just as it would if *then* were paraphrased as *soon after that*, with the pronoun *that* resolved anaphorically. Details on D-LTAG's syntactic-semantic interface are given in Forbes-Riley et al. (2006), along with a detailed discussion of *discourse adverbials* and the various sources of their anaphoric links with the previous discourse. Empirical data on the predicate-argument structure of discourse connectives are now available in Release 1.0 of the annotated Penn Discourse TreeBank[4] (Dinesh et al., 2005, Miltsakaki et al., 2004a,b, Prasad et al., 2004, Webber, 2005). An early effort to use data in the PDTB to develop a procedure for resolving the

[4] http://www.seas.upenn.edu/~pdtb

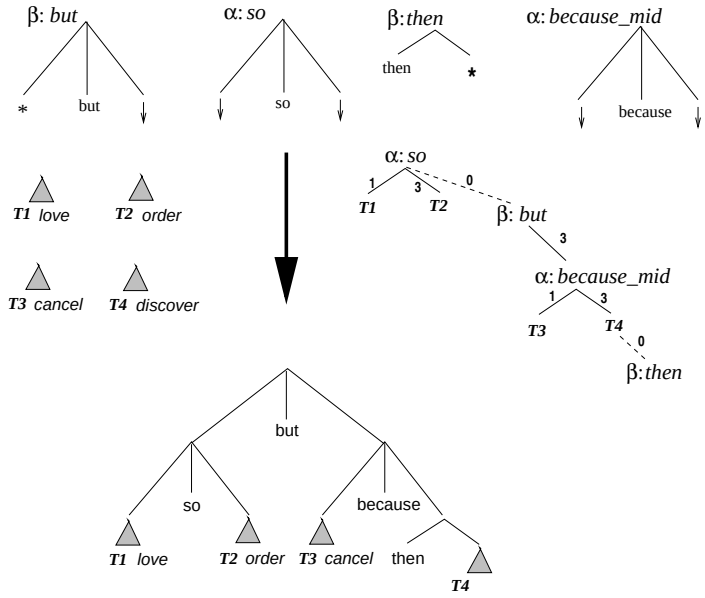


FIGURE 4 Derivation of Example 8

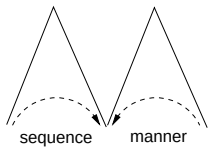


FIGURE 5 Simple multi-parent structure

anaphoric argument of the discourse adverbial *instead* is described in Miltsakaki et al. (2003), and the effect of *Information Structure* on the preferred argument of the discourse adverbial *otherwise* is described in Kruijff-Korbayová and Webber (2001).

One final note: although D-LTAG produces only trees, it is acknowledged in Webber et al. (2003) — as noted earlier by Bateman (1999) and Gardent (1997) — that a discourse unit must be allowed to participate in one constituent structure with left-adjacent material and another with right-adjacent material, as in Figure 5. While RST handles this as a special schema (cf. Figures 1(d) and 2(d)), D-LTAG would have to incorporate such structures in a more general way.

### 1.5.2 Wolf and Gibson (2005)

Wolf and Gibson (2005) present a view of discourse related to RST (Section 1.4.1), but different in one important way: Rather than analyzing a text as a recursive structure of discourse spans with discourse relations holding between sisters, Wolf and Gibson (2005) claim that discourse structure should be seen as a relatively shallow graph of discourse segments linked to one or more previous adjacent or non-adjacent segments (or segment *groupings*, see below) via discourse relations (here called *coherence relations*).

More specifically, Wolf and Gibson (2005) assume two types of basic discourse segments: clauses and attributions. The latter are related to what Huddleston and Pullum (2002) call *reporting frames* (e.g., "John asserted that ..."). Clause fragments are also treated as discourse segments if they result from the interruption of a clause by a discourse segment such as the reporting frame in (9).

(9) The economy,
according to some analysts,
is expected to improve by early next year. (Wolf and Gibson, 2005, 255, Ex. 17).

In this case, the non-adjacent clausal fragments are linked into a discourse segment through a *coherence relation* called SAME. (SAME is never used inter-sententially though to form a segment from two non-adjacent segments.)

While Wolf and Gibson make use of eleven broad classes of binary relations in their analysis, including SAME, CONDITION, ATTRIBUTION, CAUSE-EFFECT, CONTRAST, ELABORATION, and GENERALIZATION, they do not require a relation to hold between the entire content of the segments so linked together. For example, in (10)

(10) a. Difficulties have arisen in enacting the accord for the independence of Namibia
b. for which SWAPO has fought for many years. (Wolf and Gibson, 2005, Ex. 18)

an ELABORATION relation is taken to hold between the non-restrictive relative clause that constitutes (10b) and the matrix clause (10a), even though (10b) only elaborates an NP within (10a) — ie, "the independence of Namibia". This is significant because it means that the NP does not have to be analysed as a separate discourse segment, as would then the remainder of the matrix clause.

The only hierarchical structuring in Wolf and Gibson's approach comes from *grouping* a sequence of adjacent discourse segments to serve as one argument to a coherence relation whose other argument is a

previous discourse segment or *grouping*. The basis for a grouping is common attribution or common topic.[5] Within a grouping, a coherence relation can hold between segments and the same is true between a within-grouping segment and one outside the grouping. Because a grouping of segments on a common topic might itself contain groupings on common sub-topics, it appears that groupings could determine a partial hierarchical structure for parts of a text, and that grouping is a matter of *constituency*. But this is the only hierarchical structure in Wolf and Gibson's approach: unlike RST (Section 1.4.1) and the LDM (Section 1.4.2), the existence of a coherence relation between two segments does not produce a new segment that can serve as argument to another coherence relation.

Procedurally, a text is analyzed in a sequence of left-to-right passes. First, the text is segmented in a left-to-right pass, then groupings are generated, and finally, the possibility of a coherence relation is assessed between each segment or grouping and each discourse segment or grouping to its left. This produces a rather flat discourse structure, with frequent crossing arcs and nodes with multiple parents that Wolf and Gibson argue should be represented as a *chain graph* — that is, a graph with both directed and undirected edges, whose nodes can be partitioned into subsets within which all edges are undirected and between which, edges are directed but with no directed cycles.[6] This is illustrated in the discourse structure ascribed to Example 11, shown in Figure 6.

(11) 1. "The administration should now state
2. that
3. if the February election is voided by the Sandinistas
4. they should call for military aid,"
5. said former Assistant Secretary of State Elliot Abrams.
6. "In these circumstances, I think they'd win." (Wolf and Gibson, 2005, Ex. 26)

Figure 6 contains two *groupings* — one from segments 3 and 4, the other from segments 1 and 2 and the first grouping, while among the coherence relations are ones that hold within a grouping (COND) and between a segment and a grouping (ATTRIBUTION). The approach has been used to analyze a corpus of 135 news articles called the *Discourse GraphBank*, available from the LDC catalogue as LDC2005T08.

Wolf and Gibson's claim that discourse structure is best modelled as

[5] Documentation for the *Discourse Graphbank* (Wolf et al., 2003) states that a grouping should only be assumed if otherwise truth conditions are changed.

[6] A *Directed Acyclic Graph* (DAG) is a special case of a chain graph, in which each subset contains only a single node.

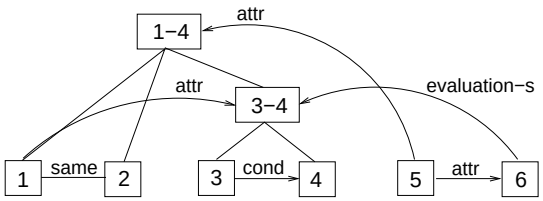


FIGURE 6 Coherence graph for Example 11

a chain graph, rather than a tree, shows that they associate discourse relations with *constituency* structure alone. However, two things suggest that *anaphoric dependency* plays a significant role. The first is the extent to which the discourse relation ELABORATION underpins their discourse structure, coupled with arguments by Knott et al. (2001) that OBJECT-ATTRIBUTE ELABORATION is actually a matter of *anaphora* — that is, *anaphoric dependency*. The second parallels the argument in Section 1.2.2 that pairings in the *respectively* construction arise from *anaphoric dependency* rather than *constituency*.

ELABORATION is the most common discourse relation in Wolf and Gibson's *Discourse GraphBank*, comprising 43.6% of the total. Many instances of ELABORATION are claimed to hold between non-adjacent discourse segments, such as between (12.4) and (12.2).

(12) 1. Susan wanted to buy some tomatoes
2. and she also tried to find some basil
3. because her recipe asked for these ingredients.
4. *The basil* would probably be quite expensive this time of year.
(Wolf and Gibson, 2005, Ex. 21)

Example (12) exemplifies OBJECT-ATTRIBUTE ELABORATION, which Mann and Thompson (1988) take as holding when one segment presents an object, and the related segment subsequently presents an attribute of that object. But as in Knott et al. (2001), the claim for that relation holding relies on the fact that *the basil* in (12.4) is anaphorically related to *some basil* in (12.2).[7]

[7]The *Discourse GraphBank* also has cases of elaboration holding between adjacent segments. But many of these involve syntactic constructions such as a relative clause (as in (10) above), an appositive (as in the relation between $1a$ and $1b$ in (13) below), or an adjunct (as in the relation between grouping (2–3) and grouping (1a–1b) in the same example).

(13) 1. $_{1a}$[Mr. Baker's assistant for inter-American affairs,] $_{1b}$[Bernard Aronson,]
2. while maintaining
3. that the Sandinistas had also broken the cease-fire,
4. acknowledged:
5. "It's never very clear who starts what." (Wolf and Gibson, 2005, Ex. 23)

Knott et al. (2001) came to their view of OBJECT-ATTRIBUTE ELABORATION as *anaphoric dependency* in trying to automatically generate passages of text similar to that found in actual museum guidebooks. They first analysed such passages in terms of RST, adhering to RST's assumption that the spans linked by a relation must either be *adjacent* or if not adjacent, that any intervening spans must also be linked to the initial span by the same relation (cf. Section 1.4.1). But they found that in passages such as (14), they had to analyse non-adjacent segments as standing in an ELABORATION relation, here (14.4) elaborating (14.2).

(14) (1) In the women's quarters the business of running the household took place. (2) Much of the furniture was made up of chests arranged vertically in matching pairs (...). (3) Female guests were entertained in these rooms, which often had beautifully crafted wooden toilet boxes with fold-away mirrors and sewing boxes, and folding screens, painted with birds and flowers.
(4) Chests were used for the storage of clothes ....

This, however, violated RST's adjacency assumption. In analysing all the cases where they saw ELABORATION relations holding between non-adjacent segments, they noticed that in each case, the elaborating segment re-introduced and described an entity that had been mentioned earlier in the discourse.[8] Subsequent segments then continued that description until another previously mentioned entity was re-introduced and elaborated.

To explain this, Knott et al. (2001) assumed that discourse was locally structured as an *entity chain* — ie,

> a sequence of Rhetorical Structure (RS) trees, each constructed just as in RST, but minus the ELABORATION relation. These trees can either be simple trees consisting of just one text span, or more complex trees with several layers of hierarchy. In each case, we can define the *top nucleus* of the tree to be the leaf-level text span which is reached by following the chain of nuclei from its root. A legal *entity chain* whose *focus* is entity **E** is one where the top nucleus of each tree is a fact about **E**.

and globally structured as a sequence of entity chains, as in Figure 7, where the focussed entity in each chain is mentioned in a proposition somewhere within the previous $N$ chains. For the current discussion, the most important thing to notice is that the links between entity chains are based on *anaphoric relations*, and that such links can and

[8] The entity could have been introduced by one or more noun phrases, or by one or more clauses. In the latter case, reference was via a demonstrative pronoun (*this* or *that*) or definite or demonstrative NP.

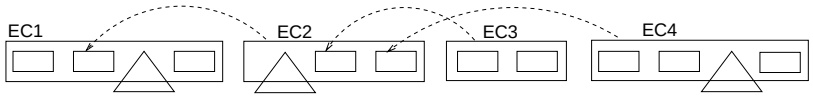


FIGURE 7 A legal sequence of entity-chains (Knott et al., 2001, Fig. 4)

will cross (Figure 7).

As noted, the *Discourse GraphBank* contains many instances of ELABORATION, many holding between non-adjacent segments (including ones at a considerable distance apart), as in the following example from file 62 (segment numbering as in the source file):

(15) 0 First lady Nancy Reagan was saluted Monday night for her sense of style and her contribution to the American fashion industry.
1 Mrs. Reagan was presented with a Lifetime Achievement Award by the Council of Fashion Designers of America at its eighth annual awards ceremony,
. . .
46 Also among those recognized at the ceremony was designer Geoffrey Beene,
47 who was saluted for making "fashion as art."
. . .
48 Performer Liza Minelli,
49 one of many women attending the ceremony who dressed in Mrs. Reagan's favorite color,
. . .
54 Actress Audrey Hepburn,
55 wearing a red gown designed by Givenchy,
56 presented a Lifetime Achievement Award to photographer Richard Avedon,
. . .

Here annotators have recorded ELABORATION relations between segment 46 and segment 1, between segment 48 and segment 1, between segment 54 and segment 1, and between several later segments and segment 1. But each of these appears to be based on an anaphoric bridging relation between the people mentioned in segments 46, 48, 54 and elsewhere, and the attendees at the awards ceremony mentioned in segment 1. (Recall that Wolf and Gibson do not require a relation to hold between the entire contents of linked segments, as in Example 10.) Other anaphoric relations are discernable in other instances of ELABORATION. Thus what is being captured here by Wolf and Gibson is a discourse relation based on *anaphoric dependency* rather than *constituency*.

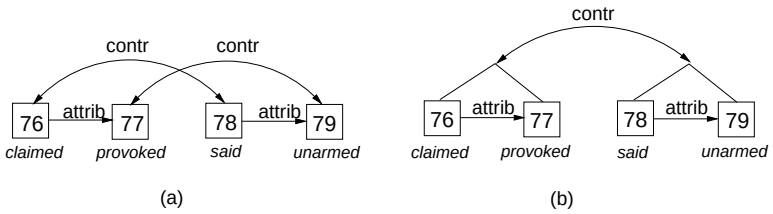


FIGURE 8 Coherence graph for Example 16

A second such case can be found in constructions that resemble *respectively* at the discourse level, as in Example 16 taken from file 105 of the *Discourse GraphBank* (segment numbering as in the source file), which has been annotated with the discourse relations shown in Figure 8(a).

(16) 76 Washington claimed
77 the downings were provoked by aggressive and threatening actions by armed Libyan aircraft,
78 while Libya said
79 its jets were unarmed and on a reconaissance flight over international waters.

If from the perspective of constituency, this were analysed as the simpler structure in Figure 8(b), then as with the *respectively* construction, the individual pairings would be a matter of anaphoric dependency, derived only when appropriate. While such examples might not occur frequently in discourse, the *respectively* construction shows there are alternative explanations for crossing dependencies that don't involve arbitrary inter-leaving of *constituency structure*. And once instances of discourse relations are understood to arise from *anaphoric dependency*, it calls into question Wolf and Gibson's leap upward in complexity from trees to *chain graphs* as a model for discourse structure.

## 1.6 Conclusion

I started this paper by mentioning Ron's advice that one should seek to identify the minimal machinery required for a task. Here, the task was to understand and compare the source of discourse relations in five different theories: Halliday and Hasan's theory of discourse cohesion, Mann and Thompson's Rhetorical Structure Theory, Polanyi's Linguistic Data Model, Wolf and Gibson's theory of discourse graphs, and the lexically-based theory that I've been involved in, D-LTAG. What I have tried to show is that the paired notions of *constituency* and *dependency* provide a useful way of understanding some of the significant similari-

ties and differences between these theories. If I have succeeded, I may also have convinced the reader that they are equally important in understanding discourse relations.

## Acknowledgments

I would like to thank all those who took the time to read and comment on earlier versions of this manuscript — Mark Steedman, Rashmi Prasad, Nikhil Dinesh, Eleni Miltsakaki, Annie Zaenen and two anonymous reviewers. The paper has benefitted from their comments, and any remaining errors are my own.

## References

Asher, Nicholas. 1993. *Reference to Abstract Objects in Discourse*. Boston MA: Kluwer.

Asher, Nicholas and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge UK: Cambridge University Press.

Bateman, John. 1999. The dynamics of 'surfacing': An initial exploration. In *Proceedings of International Workshop on Levels of Representation in Discourse (LORID'99)*, pages 127–133. Edinburgh.

Creswell, Cassandre, Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2002. The discourse anaphoric properties of connectives. In *4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC)*. Lisbon, Portugal.

Dinesh, Nikhil, Alan Lee, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2005. Attribution and the non-alignment of syntactic and discourse arguments of connectives. In *ACL Workshop on Frontiers in Corpus Annotation*. Ann Arbor MI.

Forbes, Katherine. 2003. *Discourse Semantics of S-Modifying Adverbials*. Ph.D. thesis, Department of Linguistics, University of Pennsylvania.

Forbes, Katherine, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber. 2003. D-LTAG System: Discourse parsing with a lexicalized Tree-Adjoining Grammar. *Journal of Logic, Language and Information* 12.

Forbes-Riley, Katherine, Bonnie Webber, and Aravind Joshi. 2006. Computing discourse semantics: The predicate-argument semantics of discourse connectives in D-LTAG. *Journal of Semantics* 23:55–106.

Gardent, Claire. 1997. Discourse tree adjoining grammars. Claus report nr.89, University of the Saarland, Saarbrücken.

Grosz, Barbara and Candace Sidner. 1990. Plans for discourse. In P. Cohen, J. Morgan, and M. Pollack, eds., *Intentions in Communication*, pages 417–444. MIT Press, Cambridge MA.

Halliday, Michael and Ruqaiya Hasan. 1976. *Cohesion in English*. Longman.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press.

Knott, Alistair, Jon Oberlander, Mick O'Donnell, and Chris Mellish. 2001. Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord, and W. Spooren, eds., *Text Representation:Linguistic and psycholinguistic aspects*, pages 181–196. John Benjamins Publishing.

Kruijff-Korbayová, Ivana and Bonnie Webber. 2001. Information structure and the semantics of "otherwise". In *ESSLLI'2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*, pages 61–78. Helsinki, Finland.

Lochbaum, Karen. 1998. A collaborative planning model of intentional structure. *Computational Linguistics* 24(4):525–572.

Mann, William and Sandra Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3):243–281.

Miltsakaki, Eleni, Cassandre Creswell, Kate Forbes, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2003. Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents. In *EACL Workshop on Computational Treatment of Anaphora*. Budapest, Hungary.

Miltsakaki, Eleni, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2004a. Annotating discourse connectives and their arguments. In *NAACL/HLT Workshop on Frontiers in Corpus Annotation*. Boston MA.

Miltsakaki, Eleni, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2004b. The Penn Discourse TreeBank. In *LREC*. Lisbon, Portugal.

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

Polanyi, Livia. 1988. A formal model of the structure of discourse. *Journal of Pragmatics* 12:601–638.

Polanyi, Livia, Chris Culy, Martin H. van den Berg, Gian Lorenzo Thione, and David Ahn. 2004. A rule based approach to discourse parsing. In *Proceedings of 5th SIGdial Workshop on Discourse and Dialogue*. Cambridge MA.

Polanyi, Livia and Martin H. van den Berg. 1996. Discourse structure and discourse interpretation. In P. Dekker and M. Stokhof, eds., *Proceedings of the Tenth Amsterdam Colloquium*, pages 113–131. University of Amsterdam.

Prasad, Rashmi, Eleni Miltsakaki, Aravind Joshi, and Bonnie Webber. 2004. Annotation and data mining of the Penn Discourse TreeBank. In *ACL Workshop on Discourse Annotation*. Barcelona, Spain.

Schabes, Yves. 1990. *Mathematical and Computational Aspects of Lexicalized Grammars*. Ph.D. thesis, Department of Computer and Information Science, University of Pennsylvania.

Webber, Bonnie. 1991. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes* 6(2):107–135.

Webber, Bonnie. 2004. D-LTAG: Extending Lexicalized TAG to discourse. *Cognitive Science* 28:751–779.

Webber, Bonnie. 2005. A short introduction to the Penn Discourse TreeBank. In *Copenhagen Working Papers in Language and Speech Processing*.

Webber, Bonnie, Alistair Knott, and Aravind Joshi. 2001. Multiple discourse connectives in a lexicalized grammar for discourse. In H. Bunt, R. Muskens, and E. Thijsse, eds., *Computing Meaning (Volume 2)*, pages 229–249. Kluwer.

Webber, Bonnie, Matthew Stone, Aravind Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. *Computational Linguistics* 29:545–587.

Wolf, Florian and Edward Gibson. 2005. Representing discourse coherence: A corpus-based study. *Computational Linguistics* 31:249–287.

Wolf, Florian, Edward Gibson, Amy Fisher, and Meredith Knight. 2003. A procedure for collecting a database of texts annotated with coherence relations. Documentation accompanying the *Discourse GraphBank*, LDC2005T08.

Woods, William, Ron Kaplan, and Bonnie Nash-Webber. 1972. The lunar sciences natural language information system: Final report. Tech. Rep. 2378, Bolt Beranek and Newman, Cambridge MA.