DLTAG: Extending Lexicalized TAG to Discourse*

Bonnie Webber

Draft 1, May 27, 2003

*Material in this article derives from talks I have given and from papers co-authored by members of the DLTAG group — Cassandra Creswell, Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad and Aravind Joshi, all at the University of Pennsylvania, and myself, at the University of Edinburgh. The paper has also gained from on-going discussions among group members, in connection with the development of the Penn Discourse TreeBank, and from comments from Mark Steedman on an earlier draft of the paper. The Penn Discourse TreeBank project is partially supported by NSF Grant EIA 0224417 (Joshi).

DLTAG: Extending Lexicalized TAG to Discourse

Bonnie Webber, University of Edinburgh

1 Introduction

Lexicalized grammars such as LTAG (Joshi, 1987; Schabes, 1990) have aided the understanding of clause-level phenomena and have spurred the development of parsers that can handle large sets of lexico-syntactic categories. Lexicalized grammars are now being applied to discourse, with hopes of similar benefits.

In 1997, working with Dan Cristea (Cristea & Webber, 1997), I noticed that if one wanted to "parse" discourse in a TAG framework – a line that other researchers were also exploring (Gardent, 1997; Polanyi & van den Berg, 1996; Schilder, 1997; van den Berg, 1996), one needed to posit a *substitution* operation, such as that used in lexicalized TAG, as well as the *adjoining* operation that formed the basis of these other "discourse TAGs". Aravind Joshi and I began to explore what it would mean to have a *fully* lexicalized TAG for discourse that would allow us to examine how the insights of lexicalized grammars – that the basic elements were not simply words but *structures* that reflected a word's role(s) and syntactic/semantic scope – would carry over to discourse (Webber & Joshi, 1998).

This exploration has continued over the last five years, with several students and colleagues taking part as well – Creswell et al. (2002), Forbes et al. (2001), Forbes & Webber (2002), Forbes (2003), Miltsakaki et al. (2003), Webber et al. (November 2000, 2001, 1999a,b, to appear 2003). What we find attractive about this LTAG-based approach to low-level discourse structure and semantics is that

- it provides a uniform way for *any* lexico-syntactic elements – not just conjunctions and adverbial phrases – to contribute to both the syntax and semantics of the clause and the syntax and semantics of discourse.
- it has the potential to allow sentence processing and low-level discourse processing to be carried out together.
- it simplifies the computation of discourse semantics by recognizing a range of ways in which different discourse connectives can contribute to discourse coherence.

The paper aims to demonstrate these three attractions. Section 2 illustrates what it means to have a lexicalized TAG for discourse – a DLTAG – and how it relates to lexicalized TAG at the clause-level. It thereby shows how DLTAG provides a uniform way for lexico-syntactic elements to contribute to both the clause and the discourse. Section 3 presents a brief look at our first experiment on analysing discourse automatically with respect to DLTAG (Forbes et al., 2001). This work uses the same chart-based left-corner LTAG parser (Sarkar, 2000) for both sentence and discourse processing, taking the sequence of derivation trees produced from sentence-level analysis and outputting a derivation tree for the discourse as a whole. This is a first step towards carrying out sentence processing and low-level discourse processing in coordination with one another.

Section 4 briefly describes how looking at text from a DLTAG perspective – which requires one to associate a compositional semantic construction with each element of lexicalized syntax – has forced us to look more carefully at just how words and phrases make contributions to discourse semantics. This has revealed three different ways in which discourse connectives contribute to discourse: Some contribute a relationship between adjacent discourse elements. Some related a discourse element to something in the discourse context via anaphor resolution, while atleast one adverbial – *for example* – makes its contribution via abstracting over the nearest predication, be it clause-level or discourse-level (Webber et al., to appear 2003). Section 5 describes the *Penn Discourse TreeBank* (http://www.ircs.upenn.edu/~dltag), which aims to do for automated discourse processing what the Penn TreeBank has done for sentence-level processing. Finally, Section 6 speculates on the future of DLTAG.

2 DLTAG: Lexicalized TAG for Discourse

In any lexicalized TAG, words serve as lexical *anchors* for tree structures, one for each *minimal syntactic construction* in which the word can appear. For example, the verb *like* anchors one tree corresponding to *The boys like apples*, another corresponding to the topicalized *Apples the boys like*, a third corresponding to the passive *Apples are liked by the boys*, and others as well. All these trees share the same predicate-argument structure. Likewise, the preposition *like* anchors one tree corresponding to the NP post-modifier *apples like this one* and another tree corresponding to the VP post-modifier *Sing like a bird*, both of which again share the same predicate-argument structure. Together, all these trees comprise the *treeset* of *like*. This syntactic/semantic encapsulation is possible because of

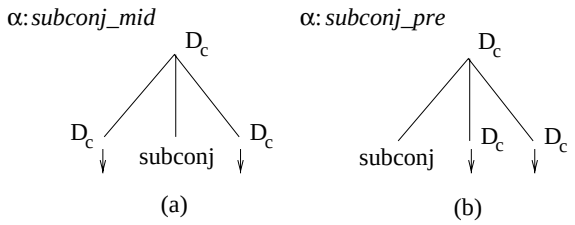


Figure 1: Initial trees (a-b) for a subordinate conjunction.

the extended domain of locality of an LTAG.

There are two kinds of *elementary* trees in an LTAG: *initial* trees that reflect basic functor-argument dependencies and *auxiliary* trees that introduce recursion and allow elementary trees to be modified and/or elaborated. Unlike the wide variety of trees needed at the clause level, we have found that an LTAG for discourse (DLTAG) only requires a few elementary tree structures, possibly because clause-level syntax exploits structural variation in ways that discourse doesn't. The root node of all elementary trees in DLTAG is a *discourse clause* (D_c) – a basic clause (either tensed or untensed) or a structure composed of discourse clauses. Leaves may be discourse clauses or lexico-syntactic elements.

Since words can contribute to the syntax and semantics of both the clause and discourse, we can treat words as having a *treeset* in the discourse-level DLTAG, as well as one in the clause-level LTAG. We first look briefly at *initial* trees in DLTAG and the interesting range of lexical items that anchor them and convey the predicate of discourse-level predicate-argument structures. We then look at *auxiliary* trees in DLTAG and the lexical items that anchor trees that elaborate the ongoing discourse.

2.1 Initial Trees in DLTAG

DLTAG associates initial trees with a variety of lexical elements that take clausal elements as arguments and convey discourse-level predication on those arguments: subordinate conjunctions and other *subordinators*; the lexical anchors of parallel constructions; some coordinate conjuctions; and even some specific verb forms.

In LTAG (XTAG-Group, 2001), subordinate conjunctions such as *if*, *although*, *since* and *so that* anchor *auxiliary* trees because they are outside the domain of locality of the verb, heading clausal or VP *adjuncts*. In DLTAG, however, it is predicates on clausal arguments that define the domain of local-

ity. Thus, at the discourse-level, subordinate conjunctions anchor *initial trees* into which clauses substitute as arguments. Figure 1 shows the initial trees for postposed subordinate clauses (a) and preposed subordinate clauses (b). In this and other figures, $\mathbf{D}_c$ stands for "discourse clause", $\downarrow$ indicates a substitution site, <subconj> stands for the particular subordinate conjunction that anchors the tree.

Similar to subordinate conjunctions are what Quirk et al. (1972) call *subordinators* – lexical items such as *in order for, in order to, to* (all possible heads for a *purpose clause*) and *by* (heading a *manner clause*). These also anchor *initial* trees in DLTAG, while anchoring *auxiliary* trees in LTAG. They differ from subordinate conjunctions in having a non-finite (untensed) clause as one argument and a finite (tensed) clause as the other one.

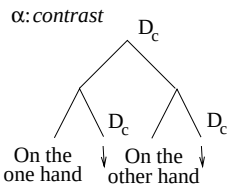


Figure 2: Initial tree for a parallel contrastive construction

DLTAG also associates initial trees with the lexical anchors of parallel constructions such as

(1) *On the one hand,* John is generous. *On the other hand,* he's hard to find.

The initial tree for this parallel construction is shown in Figure 2. It is associated with both the lexical anchors *on the one hand* and *on the other (hand)*. While in LTAG, the idiomatic prepositional phrases ("on the one hand" and "on the other") would be auxiliary trees, adjoining at the sentence-level, in DLTAG, both serve as anchors for the same initial tree. There are similar multiply-anchored initial trees for *disjunction* ("either"... "or"...), *addition* ("not only"... "but also"...), and *concession* ("admittedly"... "but"...).

There are also initial trees anchored by coordinate conjunctions that convey a particular relation between the connected units. One clear example is *so*, conveying *result*. Its initial tree is shown in Figure 3. In contrast, the coordinate conjunction *and* anchors an auxiliary tree, as discussed in the next section.

Finally, there is even one verb form – imperative *suppose* – that anchors an initial tree in DLTAG. In LTAG, imperative *suppose* anchors an *auxiliary* tree

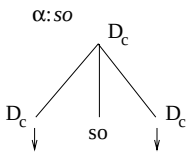


Figure 3: Initial tree for the coordinate conjunction *so*.

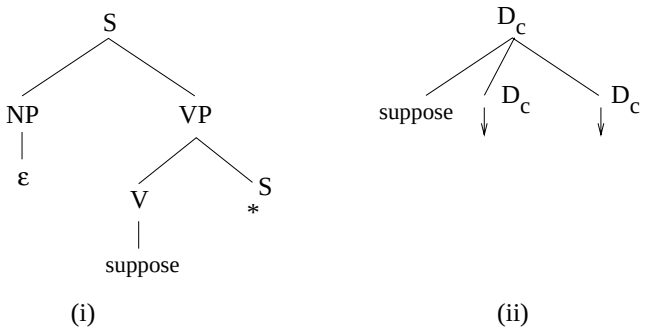


Figure 4: LTAG and DLTAG trees for imperative *suppose*.

rooted in an S-node, with an empty subject NP (ε, filled in its semantics by *you*) and an adjunction site, where it adjoins to its clausal object. The LTAG auxiliary tree for imperative *suppose* is shown in Figure 4(i), with $*$ indicating the *foot node* of the tree. In Example 2

(2) Suppose an investor wants to sell a stock, but not for less than $55. A limit order to sell could be entered at that price.

this imperative *suppose* tree would be adjoined to the clause "an investor wants to sell a stock . . . ".

In contrast, in DLTAG, imperative *suppose* anchors an initial tree with *two* substitution sites for discourse clauses – one filled by the discourse clause that is its clausal object, the other by the subsequent discourse clause (Figure 4ii) – here, "A limit order could be entered at that price". As with the subordinate conjunction *if*, the first discourse clause provides the hypothetical or counterfactual condition under which the second discourse clause follows. Of course, imperative *suppose* doesn't always play a discourse role of linking a hypothesized condition with what follows. So whether it should be interpreted as projecting an *initial* tree in the discourse, or just as a simple discourse clause, leads to ambiguity in DLTAG analyses. I will mention other sources of ambiguity in the next section.

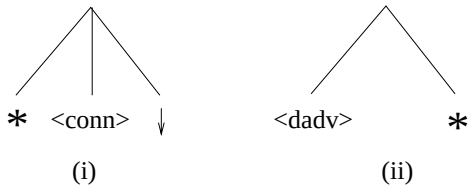


Figure 5: Auxiliary trees in DLTAG. <conn> stands for any explicit coordinating conjunction or null connective (ϕ). <dadv> stands for any discourse adverbial.

One final point here. In all our DLTAG papers to date, we have talked as if words anchor both LTAG trees and DLTAG trees. Because it is often the case (as with *suppose*) that only when a lexical item occurs in a particular structural configuration that it should be associated with a particular tree in DLTAG, it is possible that we should talk in terms of anchored LTAG trees anchoring DLTAG trees, rather than simply as lexical items doing so. This is certainly how the parser effectively operates, as I will show in Section 3.

2.2 Auxiliary Trees in DLTAG

As noted in Section 1, *auxiliary* trees in an LTAG introduce recursion and allow elementary trees to be modified and/or elaborated. DLTAG associates auxiliary trees with discourse connectives that continue a description in some way and with discourse adverbials that contribute a predicate-argument structure additional to that coming from connectives (Webber et al., 1999a,b, to appear 2003).

It is clear that descriptions of objects, events, situations, states, etc. can extend over several clauses in a discourse. This is signalled through the use of coordinate conjunctions and/or unrealized (null) connectives. Thus, in DLTAG both coordinate conjunctions and unrealized connectives anchor *auxiliary* trees – cf. Figure 5(i). When such a tree is adjoined to a discourse clause and its substitution site is filled with another discourse clause, the latter extends the description of the situation or entity conveyed by the former.[1] Such auxiliary trees are used in the derivation of simple discourses such as (3):

(3) a. John went to the zoo.

[1]The latter use of an auxiliary tree is related to *dominant topic chaining* in (Scha & Polanyi, 1988) and *entity chains* in (Knott et al., 2001).

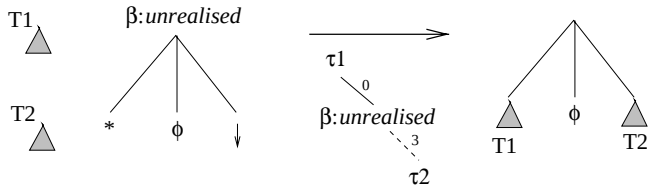


Figure 6: DLTAG derivation of Example 3

b. He took his cell phone with him.

This derivation is shown in Figure 6. To the left of the arrow ($\rightarrow$) are the elementary trees to be combined: T1 stands for the LTAG tree for clause 3a, T2 for clause 3b, and β:*unrealized*, for the auxiliary tree that connects adjacent clauses without an overt connective. In the derivation, the foot node of β:*unrealized* is adjoined to the root of T1 and its substitution site is filled by T2. The result is shown to the right of $\rightarrow$. (A standard way of indicating TAG derivations is shown under $\rightarrow$, in the form of a *derivation tree* in which solid lines indicate adjunction, and dashed lines, substitution. Each line is labelled with the address of the argument at which the operation occurs. $\tau 1$ is the derivation tree for T1, and $\tau 2$, the derivation tree for T2.)

DLTAG has a second type of *auxiliary* tree, shown in Figure 5(ii), this one anchored by a discourse adverbial such as *instead, otherwise, then, in contrast, therefore, for example,* etc. Such adverbials also anchor auxiliary trees in LTAG.[2] This is because adverbials are outside the domain of locality of the *verb*, adding information to either the S or the VP. In DLTAG, discourse adverbials are outside the domain of locality of *structural connectives* (including the null connective), contributing additional information to the discourse clause to which they are adjoined or to the complex discourse clause in which that clause is embedded. I discuss the nature of this added information in Section 4.

There are other sources of lexical ambiguity in DLTAG beyond that associated with whether a particular instance of imperative *suppose* serves to connect its embedded clause to what follows in the discourse and should thus be "supertagged" with the initial tree it anchors in DLTAG, or whether the entire clause

[2]LTAG doesn't distinguish between a discourse adverbial such as *instead* and a clausal adverbial such as *spontaneously*. Both have the same *treeset*, since they can appear at the same positions within the clause. In DLTAG however, only discourse adverbials anchor the auxiliary tree shown in Figure 5(ii). Forbes (2003) contains an extensive analysis of what causes an adverbial to be interpreted at the discourse-level, rather than at the clause-level.

headed by imperative *suppose* only functions as a simple discourse clause. Another source of ambiguity comes from the fact that adverbials can appear in one structure in which they are discourse adverbials (as in the preposed adverbials in 4a-b), and other structures in which they function simply at the clausal level (as in 4c-d).

(4) a. *Instead*, John ate an apple.

b. *Otherwise*, you can forget dessert.

c. John ate an apple *instead* of a pear.

d. Mary was *otherwise* occupied.

In these cases, the clause-level analysis serves to disambiguate whether or not the lexcial item functions at the discourse-level.

Another source of ambiguity is the fact that many of the adverbials found in second position in parallel constructions (e.g., *on the other hand*, *at the same time*, *nevertheless*, *but*) can also serve as simple discourse adverbials on their own. In the first case, they will be one of the two anchors of an initial tree, such as in Figure 2, while in the second, they will anchor the simple auxiliary tree shown in Figure 5(ii). These lexical ambiguities correlate with structural ambiguity at the discourse-level.

For example, in the following passage, *at the same time* serves at the second anchor of an initial tree expressing contrast, whose first anchor is *on the one hand*.

(5) Brooklyn College students have an ambivalent attitude toward their school. *On the one hand*, there is a sense of not having moved beyond the ambiance of their high school. This is particularly acute for those who attended Midwood High School directly across the street from Brooklyn College. . . . *At the same time*, there is a good deal of self-congratulation at attending a good college . . . (*cf25*)

However, in the following minor variation of Example 5, *at the same time* anchors an auxiliary tree that elaborates on the positive aspects of attending Brooklyn College, with *on the other hand* serving as the second anchor of the initial tree that expresses contrast.

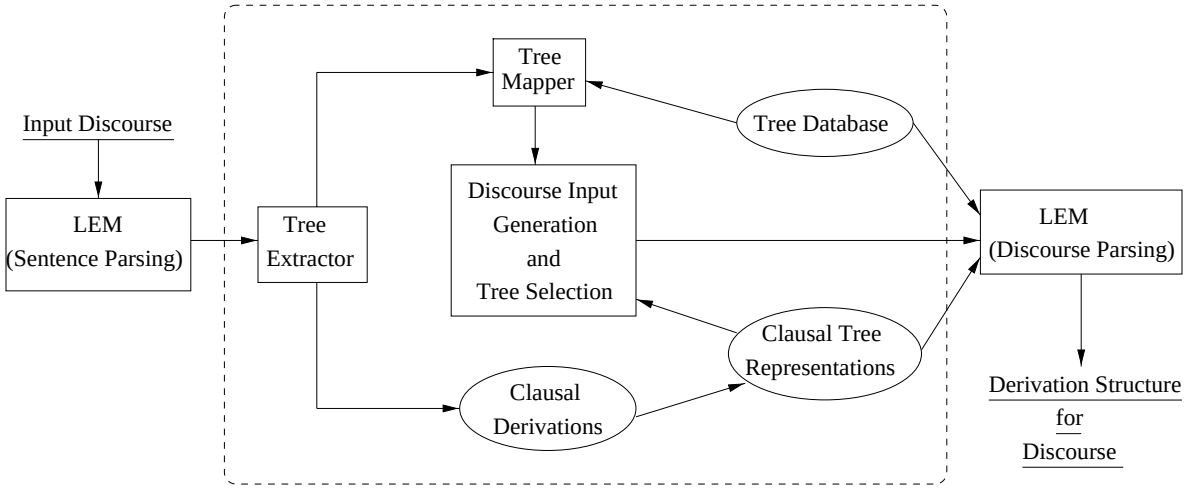


Figure 7: Two-pass sentence/discourse parsing using LEM

(6) Brooklyn College students have an ambivalent attitude toward their school. *On the one hand*, there is a good deal of self-congratulation at attending a good college. *At the same time*, they know they're saving money by living at home. *On the other hand*, there is a sense of not having moved beyond the ambiance of their high school.

DLTAG analyses do not introduce any kind of local or global discourse ambiguity that is not present in the original discourse. As with ambiguity at the clause-level, discourse ambiguity is a problem that parsers must punt on or deal with, as I will discuss briefly in the next section.

3 A Parser for DLTAG

Discourse parsing involves analyzing a discourse according to a discourse grammar – in our case, DLTAG. To date, we have carried out a single experiment with discourse parsing (Forbes et al., 2001). In this work, a single chart-based left-corner LTAG parser, LEM (Sarkar, 2000) makes two passes through the text, the first producing XTAG derivation trees for each sentence from the sequence of elementary trees associated with its words, the second producing a DLTAG derivation for the discourse as a whole from the sequence of elementary trees associated with each of its discourse connectives and with each of its clausal derivations. The flow of processing is shown in Figure 7

Given a sentence in the discourse, LEM produces a single *derivation tree* according to the XTAG English grammar (XTAG-Group, 2001), using heuristics to

- decide which elementary tree to assign to each word;
- pick the lowest attachment point between these trees.

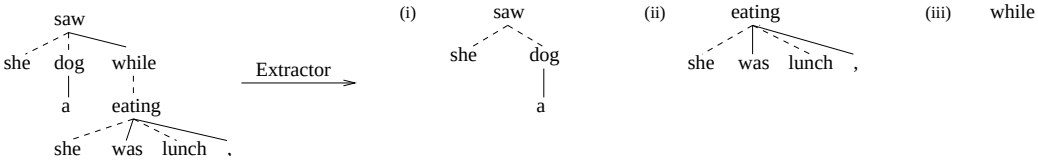
This sequence of derivation trees is input to a *Tree Extractor* (**TE**), which extracts from each one, the derivation trees for each of its clauses and any elementary trees anchored in a discourse connective. This is done in two passes – the first, to identify the discourse connectives , and the second, to detach clausal derivations from their substitution and/or adjunction nodes. The first, top-down traversal of the derivation tree considers both *lexical* and *structural* properties of each lexical item because, as noted earlier,

- lexical items that serve as discourse connectives usually serve other functions as well (e.g., *instead* as an NP post-modifier – "an apple instead of a pear"; *and* as an NP conjunction). So lexical features are alone insufficient.
- LTAG does not distinguish between clausal adverbials like *frequently* and discourse adverbials like *otherwise*. So structural features alone are also insufficient.

So from the sentence

(7) While she was eating lunch, she saw a dog.

TE extracts the two clausal derivations and one elementary tree anchored in a discourse connective shown below.



With clause-medial discourse connectives, as in

(8) Susan will *then* take dancing lessons.

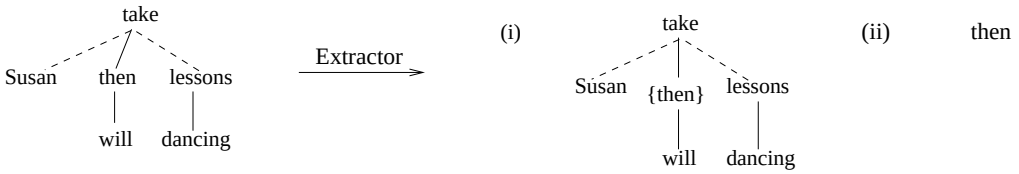


Figure 8: Application of *TE* to the derivation tree of Example 8.

the Tree Extractor makes a *copy* of the derivation and replaces the discourse connective with an *index*, to retain its clause-internal position. This is because clause-medial adverbials appear to be relevant to Information Structure (Steedman, 2000), and thus their position in the clause is important to preserve. So in Example 8, **TE** extracts a single clausal derivation and one elementary tree anchored in a discourse connective, as shown in Figure 8.

Tree Mapping applies to the output of Tree Extraction, to map *sentence-level* structural descriptors of connective elementary trees to their *discourse-level* structural descriptors. (Note that this embodies the assumption suggested at the end of Section 2.1 that it is not lexical items that anchor DLTAG trees, but rather anchored LTAG trees.)

Because LEM takes as input a sequence of lexicalised trees, the role of the next stage of the process, Discourse Input Generation (**DIG**) is, essentially, to produce such a sequence at the discourse level. To do this, **DIG** first converts clausal derivations into elementary tree representations (*discourse clauses*, D_c). These discourse clauses, along with the discourse connectives, make up the sequence of lexicalised trees. (**N.B.** Where there is no structural connective between clausal units, **DIG** inserts an auxiliary tree with an empty lexical anchor into the input sequence.)

While we have simplified the problem by assuming (incorrectly) that there is no *lexical ambiguity* associated with tree selection, the parser does acknowledge *attachment ambiguities* associated with the auxiliary trees for structural connectives, as in

(9) John is stubborn. (T1)
His sister is stubborn. (T2)
His parents are stubborn. (T3)
So they are continually arguing. (T4)

Figure 9 shows the output from **DIG** for Example 9. Its five possible derivations are shown in Figure 10, corresponding to five derived structures shown in

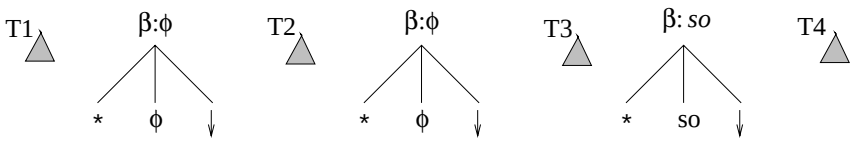


Figure 9: Trees that serve as input to LEM's discourse parsing from Example 9.

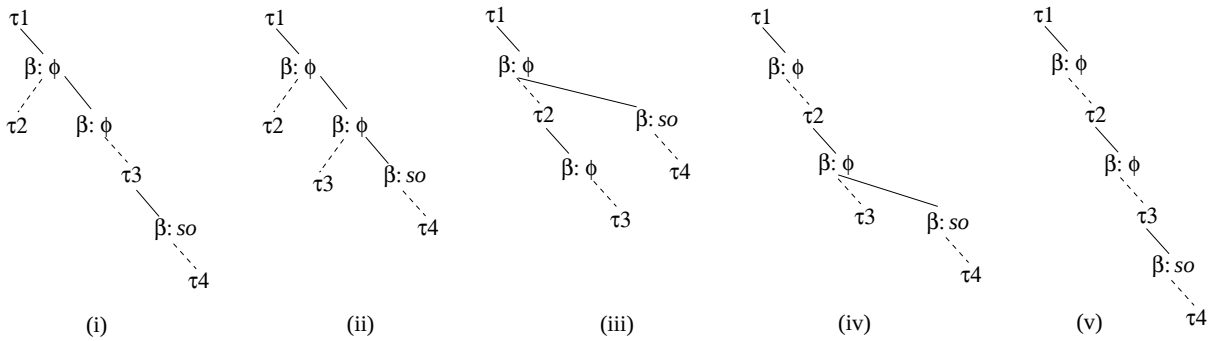


Figure 10: Potential discourse-level derivation trees for Example 9.

Figure 11. Structure (i) can be paraphrased as

> John and his sister are stubborn. His parents are stubborn. So they [his parents] are always arguing.

Structure (iv) can be paraphrased as

> John is stubborn. His sister and his parents are stubborn. So they [his sister and his parents] are always arguing.

while structures (ii), (iii) and (v) can all be paraphased as

> John and his sister and his parents are stubborn. So all of them [the whole family] are always arguing.

Currently, LEM only considers the unique derivation that satisfies the following criteria:

1. Adjunction in initial trees is only allowed at the root node.
2. For all other trees, only the lowest adjunction is allowed.

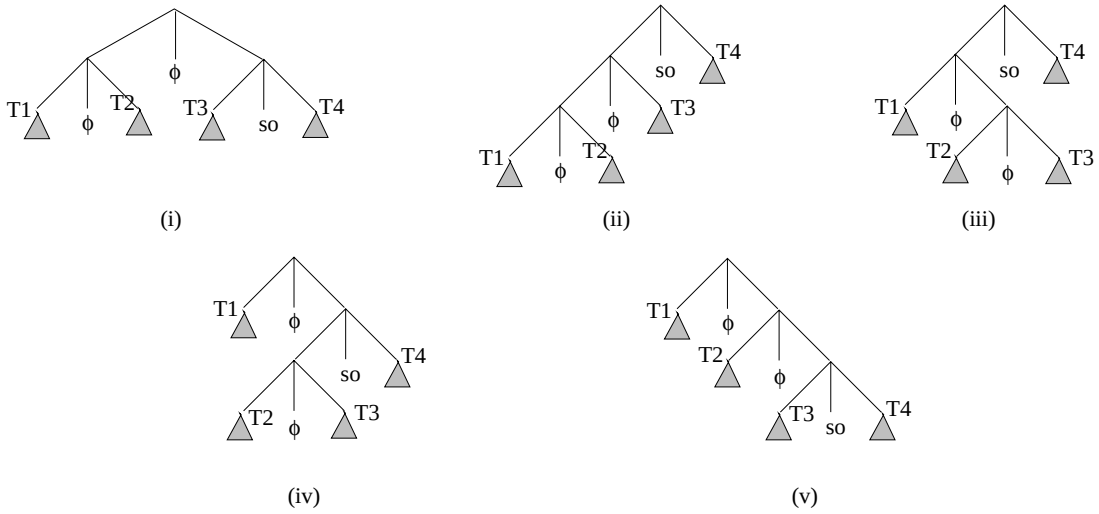


Figure 11: Derived structures for discourse parsing of Example 9.

That means LEM only associates derivation (v) and derived tree (v) with Example 9, rather than any of the others.

The next version of the parser should take on the problem of lexical ambiguity – the appropriate DLTAG tree to associate with each instance of a connective – as well as providing a more informed treatment of structural ambiguity.

There is one other problem that any parser for discourse must address – that of discourse embedded in indirect speech or a propositional attitude, as in (10) and (11).

(10) The pilots could play hardball by noting that they are crucial to any sale or restructuring because they can refuse to fly the airplanes.

(11) Epigenesists, on the other hand, believed that the organism was not yet formed in the fertilized egg. Rather, it arose as a consequence of profound changes in shape and form during the course of embryogenesis.

In both these cases, the sentential object of the verb (*note* in (10) and *believe* in (11)) must itself be analysed as a discourse, extending in the case of (11) to the next sentence as well.

Our initial solution to this problem resembles, in part, our treatment of imperative *suppose* in Example 2. I have already mentioned, in discussing imperative *suppose*, that in LTAG, verbs that take sentential objects do so in the form of

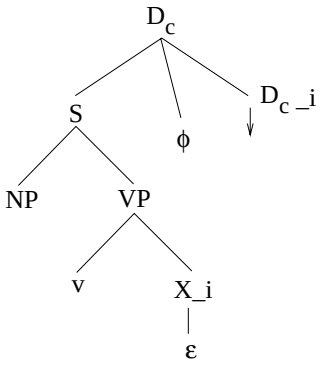


Figure 12: Proposed DLTAG initial tree for propositional attitude and indirect speech verbs

an *auxiliary* tree that adjoins to the object clause (cf. Figure 4i). In DLTAG however, we posit an *initial* tree for imperative *suppose* that takes two discourse clauses as arguments. For indirect speech and propositional attitude verbs, we follow Aravind Joshi's suggestion and posit something similar: an *initial* tree anchored by the propositional attitude or indirect speech verb that has a covert argument that is coindexed with the (overt) clausal complement introduced by the complementizer (Figure 12). So,

(12) John believes that Mary is tired.

is analyzed as

(13) John believes X_i that $[$Mary is tired$]_i$.

There is cross-linguistic evidence for such an analysis coming from Hindi[3], where the X_i may be overtly expressed, as in

(14) raam ye samajhtaa hai ki sita thakii-huii hai
Ram this believes is that Sita tired is
Ram believes this that Sita is tired

The discourse analysis of Example 10 would then involve the trees shown in Figure 13, where T1 represents the analysis of "The pilots could play hardball", T2 represents the analysis of "they are crucial to any sale or restructuring", and T2, the analysis of "they can refuse to fly the airplanes".

[3]Rashmi Prasad, personal communication

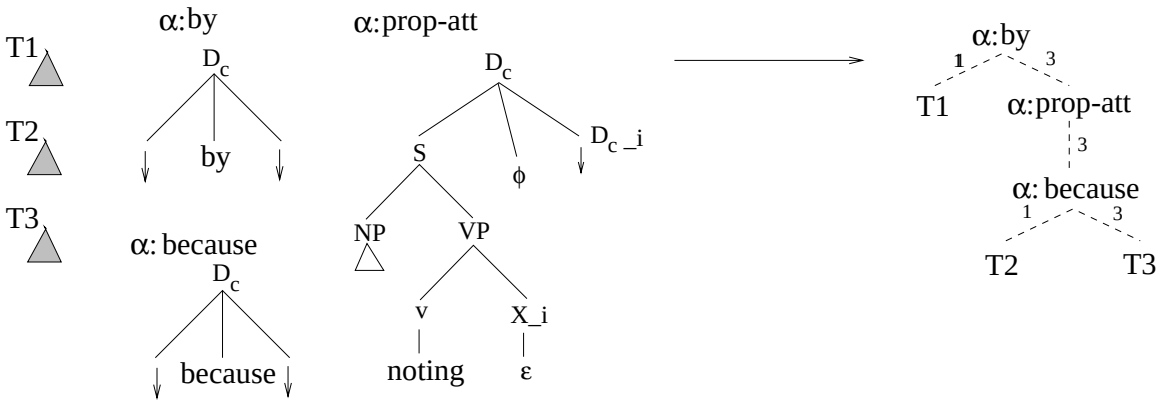


Figure 13: DLTAG derivation of Example 10

Similarly, the discourse analysis of Example 11 would involve the trees shown in Figure 14, where T1 represents the analysis of "the organism was not yet formed in the fertilized egg" and T2, the analysis of "it arose as a consequence of profound changes . . . ".

Neither this view of propositional attitude and indirect speech verbs, nor imperative *suppose*, nor the (local) ambiguity caused by discourse connectives that can appear in more than one DLTAG tree, have yet been incorporated into the parser described earlier. I expect that when they are (hopefully sooner rather than later), we will discover other aspects of low-level discourse analysis that need exploring.

4 Differences between Discourse Connectives in DLTAG

As shown in Section 2, DLTAG distinguishes between (1) structural connectives that anchor *initial* trees and convey discourse-level predicate-argument relations; (2) structural connectives (including the null connective) that anchor *auxiliary* trees and that elaborate the preceding discourse; and (3) discourse adverbials that anchor *auxiliary* trees and contribute predicate-argument relations distinct from (but that may interact with) those conveyed by structural connectives.

Webber et al. (to appear 2003) argue extensively that while structural connectives and discourse adverbials may both convey discourse-level predicate-argument relations, they get their arguments in different ways. Structural connectives get both their arguments from the discourse clauses to which they are

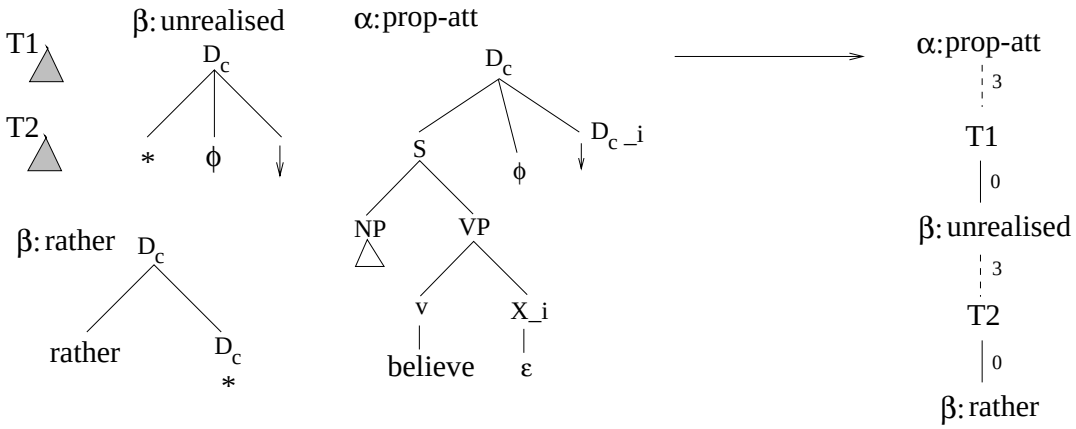


Figure 14: DLTAG derivation of Example 11

structurally connected in the discourse, as in the following[4]

(15) a. Because [**Healthcare actually owes HealthVest $4.2 million in rent and mortgage payments each month**], [**the amount due above the amount paid will be added to the three-year note.**]

b. Even though critical, [**it was just the kind of attention they were seeking.**] So [**they fired back at the Goldman Sachs objections in their own economics letter, "The BMC Report."**]

On the other hand, many discourse adverbials get only one argument from the clause or sentence to which they are adjoined and the other anaphorically from the preceding discourse as in

(16) a. While many people are happy [**to sit watching television**], [**a person who seeks adventure might** instead **try skydiving**].

b. A person who hates [**to sit watching television**] might instead [**try skydiving**].

c. [**If the light is red**], stop.
Otherwise, [**just continue down the road.**]

[4]Following the conventions used in the Penn Discourse TreeBank (Section 5), arguments are bracketted and in bold, while the connective is underlined.

d. One great difference distinguished the Soviet and German systems: [**there was no Soviet equivalent of the death camps**]. People sentenced to death in the Soviet Union were generally shot before entering the camp network. Applebaum estimates these victims at just under one million during the Stalin years. Instead, [**Soviet prisoners were expected to earn their keep by contributing to the creation of Soviet Socialism**].

While I won't repeat the arguments from (Webber et al., to appear 2003) here, it should be clear from the examples that the "lefthand" argument of a discourse adverbial can come from a non-adjacent clause, a relative clause, clauses embedded in adjuncts, etc., none of which are structurally connected to the discourse adverbial. There is also empirical evidence for this distinction between structural connectives and discourse adverbials. Creswell et al. (2002) describe an annotation experiment in which annotators were asked to identify the minimal text unit in the preceding discourse containing the source of the "left-hand" argument of the following nine connectives:

- Resultatives: *as a result, so, therefore*
- Additives: *also, in addition, moreover*
- Concessives: *nevertheless, yet, whereas*

The data came from Brown corpus, WSJ corpus, Switchboard corpus, and 58 transcribed oral histories of online Social Security Administration (SSA) Oral History Archives[5]. The results showed a variety of distribution patterns:

- *So* always took the immediately preceding sentence or sequence of sentences as its left argument.
- *Nevertheless* often took XP arguments.
- *Therefore* often took its left-hand argument from a subordinate clause.

Connectives that patterned with *so* were taken to be structural connectives, while the others were taken to get their "left-hand" argument anaphorically.

If a phenomenon is anaphoric, then one must have a procedure for resolving it. Of course, different anaphors display different patterns vis-a-vis the distribution and type of their antecedents: plural pronouns allow *split antecedents*

[5]http://www.ssa.gov/history/orallist.html

while singular pronouns do not; definite noun phrases (NPs) allow antecedents related through *bridging* while pronouns do so only rarely; the antecedents of demonstrative pronouns commonly derive from clauses, while those of personal pronouns most commonly derive from NPs; etc. In the case of a discourse adverbials, if its "left-hand" argument is anaphoric, then one needs to articulate a procedure for finding its antecedent and from that, deriving its argument.

We do not think all discourse adverbials will pattern exactly the same vis-a-vis their antecedents, so we are proceeding on a case-by-case basis. So far, we have carried out a preliminary study of the discourse adverbial *instead* (Miltsakaki et al., 2003). Here I will summarize that study and comment on how we are now proceeding.

Instead comes in two forms: (i) a bare adverbial, as in

(17) *Instead* John ate an apple.

and (ii) modified by an "of" PP, as in

(18) John ate an apple *instead of* a pear.

(19) John spent the afternoon at the zoo *instead of* at the museum.

With an "of" PP, both args of *instead* derive *structurally*: the first from the modified phrase (e.g., "an apple") and the second from the "of" PP (e.g., "a pear"). Semantically, that second argument is a salient but unchosen *alternative* to the first, with respect to the given predication. This is basic to the interpretation of *instead* in both its modified and bare forms.

As a bare adverbial, *instead* continues to get its first argument structurally, but its second argument – the salient but unchosen alternative – must be derived *anaphorically*, from the discourse context. But not every context provides alternatives:

(20) a. John ate an apple. #*Instead* he ate a pear.

b. John didn't eat an apple. *Instead* he ate a pear.

c. I told John to eat an apple. *Instead* he ate a pear.

and, as far as I am aware, there is no theory regarding what kind of phrases/clauses suggest alternatives and which don't.

To begin to discover this empirically, pairs of annotators separately examined 100 successive instances of bare *instead* in the Penn TreeBank and recorded the

minimal text span containing the antecedent of its anaphoric argument. There was agreement in 97/100 cases, and the other 3 cases were excluded from further analysis.

We then chose features to annotate that we had observed in serendipitously encountered instances of *instead*:

- clausal negation

  (21) John *couldn't* sleep. Instead, he wrote code. (**Verbal neg**)

  (22) *No one* could sleep. Instead, everyone wrote code. (**Subj neg**)

  (23) John ate *none of his spinach*. Instead, he fed it to his frog. (**Obj neg**)

- presence of a monotone-decreasing quantifier (**MDQ**)

  (24) *Few* students like to do homework. Instead, they would rather party.

  (25) Students *seldom* sleep in class. Instead, they take notes assiduously.

- presence of a modal auxiliary (**Modal**)

  (26) You *should* exercise more. Instead you sit like a couch potato.

- whether the antecedent is embedded in a higher clause (**Embed**)

  (27) *John wanted* to eat a pear. *Instead*, he ate an apple.

  (28) *Chrysler officials resisted* cutting output. *Instead*, they slapped $1000 cash rebates on vehicles.

  (29) *Paine Webber considered* recommending specific stocks. *Instead*, it just urged its clients to stay in the market.

The results are shown in Figure 15.[6]

We then investigated whether other clauses that don't serve as antecedents for *instead*, which we call "potentially competing antecedents" or "PCAs", have a similar distribution with respect to these features. As in Soon et al. (2001), we limited potentially competing antecedents to ones occurring between the anaphor and its true antecedent. Here, PCAs were finite or non-finite clauses intervening between *instead* and its true antecedent. For the 97 tokens of *instead*

Features	YES (of 97)	NO (of 97)
Verbal neg	37 (38%)	60 (62%)
Subj neg	5 (5%)	92 (95%)
Obj neg	10 (10%)	82 (85%)
MDQ	1 (1%)	96 (99%)
Modal	12 (12%)	85 (88%)
Condit	1 (1%)	96 (99%)
Embed	57 (59%)	40 (41%)

Figure 15: Distribution of Features of the Antecedent of *instead*

	Antecedents		PCAs	
Features	YES (of 97)	NO (of 97)	YES (of 169)	No (of 169)
Verbal neg	37 (38%)	60 (62%)	21 (12%)	148 (88%)
Subj neg	5 (5%)	92 (95%)	8 (5%)	161 (95%)
Obj neg	10 (10%)	82 (85%)	6 (4%)	139 (82%)
MDQ	1 (1%)	96 (99%)	0 (0%)	169 (100%)
Modal	12 (12%)	85 (88%)	17 (10%)	152 (90%)
Condit	1 (1%)	96 (99%)	0 (0%)	169 (100%)
Embed	57 (59%)	40 (41%)	14 (8%)	155 (91%)

Figure 16: Distribution of Features of the PCAs of *instead*

on which annotators agreed, this produced 169 PCAs. The distribution of the same seven features for these PCAs is shown in Figure 16.

There are some obvious differences between the antecedents and PCAs of *instead*. First, as shown in the following summary of clausal negation features

	Antecedents		PCAs	
Features	YES (of 97)	NO (of 97)	YES (of 169)	No (of 169)
Verbal neg	37 (38%)		21 (12%)	
Subj neg	5 (5%)		8 (5%)	
Obj neg	10 (10%)		6 (4%)	

clausal negation was found to be over 2.5 times more common in the antecedent of *instead* than in PCAs – 52/97 times ($\approx$ 53%) versus 35/169 times ($\approx$ 20%).

Second, focussing on the **embed** feature

	Antecedents		PCAs	
Features	YES (of 97)	NO (of 97)	YES (of 169)	No (of 169)
Embed	57 (59%)		14 (8%)	

the antecedent of the anaphoric argument of *instead* was found to be over seven times more frequently embedded in a higher verb than a PCA was – 57/97 times ($\approx$ 59%) vs 14/169 times ($\approx$ 8%).

On the other hand, for the features related to the antecedent being in a conditional (**condit**) or containing a monotonically decreasing quantifier (**MDQ**), there isn't enough data to draw any conclusions. The feature related to the antecedent containing a modal auxiliary (**Modal**) does not, as such, seem at all predictive.

Subsequent to this study, we reviewed the data and decided that this initial feature set should be refined in at least the following ways, to widen the difference between antecedents and PCAs.

1. Although the embedding feature is strongly predictive, we realised that not all embedding contexts suggest alternatives to their embedded clauses. In particular, some embedded PCAs (but no embedded antecedents of *instead*) were embedded under factive verbs like *know*. It is well-known that factive verbs presuppose the truth of their embedded clause (Kiparsky & Kiparsky, 1970), as in

[6]Antecedents could display one or more compatible features – e.g., both **Subj neg** and **Modal**.

(30) John knows that Fred eats meat.

They therefore do not provide alternatives that can serve as antecedents for *instead*

(31) John believes/*knows that Fred eats meat. Instead Fred eats tofu.

Therefore, we should annotate a feature on the embedding verb, identifying whether or not it is factive, to exclude clauses embedded under the latter as potential antecedents. Since there is only a small number of factive verbs (although they are relatively common), such a feature could be annotated automatically, with high reliability.

2. Certain verbs appear to suggest alternatives, independent of whether the clause also contains explicit negation, a monotonically-decreasing quantifier, a modal auxiliary or clausal embedding. Consider the following examples.

(32) John *doubted* Mary's resolve. *Instead*, he thought she would give up as soon as he left.

(33) NBC is contemplating *getting out of* the cartoon business. *Instead*, it may "counter-program" with shows for an audience that is virtually ignored in that time period: adults.

(34) Investors have *lost* their enthusiasm for the stock market. *Instead*, they are buying government bonds.

(35) But respectability still *eludes* Italy's politics. *Instead*, it has the phenomenon of Mr. Belusconi.

Many additional such verbs have come to our attention. They appear to fall roughly into two classes, although neither corresponds to any known thesaurus or WordNet class. The first class – including *doubt*, *refuse*, *deny*, etc. – appears to contain an element of implicit negation, and might be called *negative propositional attitude verbs*. The second class – including *stop*, *lose*, *get out of*, *change*, *drop*, *give up*, *elude*, etc. – might be called *negative state change verbs*. They indicate that the situation after the event conveyed by the clause negates some fluent of the situation before the event. As such, the negated fluent seems to be available as an alternative to the indicated change.

Given the ability of verbs in either rough class to suggest alternatives, it may be useful to annotate them. However, there does not appear to be a list of verbs belonging to either class, so we must acquire such lists at the same time we are carrying out the annotation.

3. Even more of a challenge to automatic identification, is the fact that other lexico-syntactic elements that do not fall into *a priori* classes appear able to suggest alternatives as well. In the following example from the Penn TreeBank

(36) The tension was evident on Wednesday evening during Mr. Nixon's final banquet toast, normally an opportunity for reciting platitudes about eternal friendship. *Instead*, Mr. Nixon reminded his host, Chinese President Yang Shangkun, that Americans haven't forgiven China's leaders for the military assault of June 3-4 that killed hundreds, and perhaps thousands, of demonstrators.

either the adverb "normally" or the noun "opportunity" appears to be a sufficient trigger for alternatives and hence the use of *instead*:

(37) Normally, we eat pasta on Tuesday. *Instead*, tonight we're having fish.

(38) John had the opportunity to buy a cheap used car. *Instead*, he bought a scooter.

So while it is clear that we should broaden the range of features being considered, it is not clear how to go about identifying them, except by noticing them in the context of *instead*.

Finally, I should comment on *relational features* such as whether or not the anaphoric argument of *instead* has the same surface subject as its structural argument (as in most, but not all, of the examples above), or whether the subjects of the two are related, as in example 39.

(39) In an abrupt reversal, the United States and Britain have indefinitely put off their plan to allow Iraqi opposition forces to form a national assembly and an interim government by the end of the month. *Instead*, top American and British diplomats leading reconstruction efforts here told exile leaders in a meeting tonight that allied officials would remain in charge of Iraq for an indefinite period, said Iraqis who attended the meeting.

While *relational features* appear relevant to resolving *instead*, they were not included in our original feature set. But it is clear that relational features should be included as well. The context in which all of this will happen is the Penn Discourse TreeBank, the subject of the next section.

5 Penn Discourse TreeBank

The Penn Discourse TreeBank aims to do for discourse what the Penn TreeBank has done for sentence-level processing – that is, to provide a shared resource for the development of automated techniques applicable to aspects of discourse analysis and discourse generation. The value of a TreeBank comes from the "knowledge" that has been added to it, over and beyond the sequence of sentences they start with.

Creating the Penn Discourse TreeBank (PDTB) involves manual identifying, annotating and assessing inter-annotator agreement on (a) all discourse connectives in the Penn TreeBank, and (b) the text segments from which each connective draws its arguments. While the PDTB reflects the theoretical bias of DLTAG – that certain discourse connectives draw both their arguments structurally from adjacent discourse clauses, while other ones draw only one of their arguments structurally, drawing the other anaphorically, from the previous discourse context, the instructions to annotators[7] only requires them to identify the minimal spans of text whose meaning is involved in the connective's use. These spans may cover *inter alia* an embedded clause, as in the first (anaphoric) argument to *instead* in

(40) Anne Compoccia wanted to [**be a nun**].
Instead, [**she found herself in prison for embezzling city funds**].

a previous (non-adjacent) clause, as in the first (anaphoric) argument to *otherwise* in

(41) [**If the light is red**], stop.
Otherwise, [**just continue down the road.**]

or the immediately preceding sentence or clause, as in

(42) [**There are no separate rafters in a flat roof**];
instead, [**the ceiling joists of the top story support the roofing.**]

Our discourse-level annotation is produced using WordFreak [8], a tool for discourse annotation developed by Tom Morton. To support the development of anlystic techniques, this discourse-level annotation is being linked to both the Penn Treebank syntactic annotations and the predicate-argument annotations of PropBank (Kingsbury & Palmer, 2002). Initially, we are annotating the following connectives:

[7] http://www.ircs.upenn.edu/ dltag/annotation-manual/annotation-manual.html
[8] http://www.sourceforge.net/projects/wordfreak

- discourse adverbials: *instead, otherwise, nevertheless, indeed, therefore*
- subordinate conjunctions: *because* (both alone and when preceded by *partly, in part, only, just* or *largely*); *although* (both alone and when preceded by *even*); *when* (both alone and when preceded by *just, only, even* or *largely*).
- the null connective, which does not appear lexically in the discourse. As noted earlier, in DLTAG, with a null connective, two discourse clauses are connected at the discourse level by a tree with a null anchor. This structural description will prove very useful to the annotators who will be able to annotate the arguments of the null predicate on the DLTAG parse output.

Initially we start with a set of tags corresponding to (1) the types of connectives (structural, anaphoric and null) and their positions (initial, medial, and final), and (2) the positions of the arguments of the connective. An argument of a connective can be, for example, an embedded clause, the preceding sentence, or the immediately preceding discourse. Eventually, the annotated arguments will also be linked to the appropriate segments in the PTB and PropBank within the stand-off annotation architecture. We will also annotate certain additional kinds of semantic information associated with connectives, as well as lexico-syntactic information that provides evidence for semantics

Specification of the frames associated with the argument structures of connectives, including anaphoric links, will also help the annotators during the annotation process by letting them judge quickly and accurately the relevant roles played by the surrounding context for each connective in a discourse corpus, enabling them to distinguish the arguments for a connective from the surrounding clauses. This is very similar to providing argument-adjunct frames for a verb to the annotators of PropBank.

The Penn Discourse TreeBank is not the first or only effort to annotate discourse structure. Efforts to do so started at least 10 years ago, as a way of providing empirical justification for high-level theories of discourse structure (Grosz & Sidner, 1986; Moser & Moore, 1996). Although much time and energy was devoted to the work (Di Eugenio et al., 1998), the results have not been widely used in the computational arena, certainly not like the Penn TreeBank.

The work closest to the Penn Discourse TreeBank is the resource developed by Marcu (1999, 2000) based on Rhetorical Structure Theory (Mann & Thomp-

son, 1988). RST is a theory of discourse analysis that holds that (1) adjacent units of discourse are related by a single rhetorical relation that accounts for the semantic or pragmatic (intentional) sense associated with their adjacency; (2) units so related form larger units that participate in rhetorical relations with units that they themselves are adjacent to; and (3) in many, but not all, such juxtapositions, one of the units (the satellite) provides support for the other (the nucleus), which then appears to be the basis for rhetorical relations that the larger unit participates in.

Given these principles, the two main aspects of RST annotation are (1) demarcation of the elementary discourse units that participate in relations and (2) labeling of those relations. The two are not independent. For example, a relation (attribution) postulated between the specification of a speech act (e.g., *Riordan said*) and its content specifed as direct or indirect speech (e.g., *We must expand the vision of our party*) means that a subject-verb fragment must be marked as an elementary discourse unit if the object of the verb is direct or indirect speech.

Marcu's RST-annotated corpus[9] differs from the Penn Discourse TreeBank in three main ways: First, the RST-annotated corpus does not indicate the basis for a rhetorical relation being annotated between two elementary or derived units. Even though there is a strictly ordered protocol to follow in assigning rhetorical relations (Marcu, 1999), the corpus contains no record of either the particular basis on which a rule from the protocol such as *If the relation is one of Explanation, assign relation Explanation*, is taken to hold, or why the conditions for earlier rules in the protocol are taken to have failed.

The PDTB undertakes to annotate all and only the arguments of discourse connectives. As such, the basis for each coherence relation is the higher-order predicate associated with the connective, to which the discourse units involved serve as arguments. (The precise semantic nature of that relation may be ambiguous - e.g., whether the relation conveyed by *then* is one of temporal ordering or logical consequence. But existing lexico-syntactic annotations and annotations of clausal predicate-argument relations currently in progress in the PropBank project (Kingsbury & Palmer, 2002) should provide a solid basis for disambiguation.)

Secondly, the discourse relation holding between units has to be inferred, using semantic and pragmatic information, in cases where an overt connective is missing from the discourse. While the RST-annotated corpus records inferred

[9]distributed now by the Linguistic Data Consortium, http://www.ldc.upenn.edu

relations, it omits any indication of what was used in inferring them. The PDTB annotation scheme takes two steps towards remedying this omission: (a) it is built on top of the DLTAG parse, which provides structural descriptions for the empty connectives and (b) the DLTAG parse links up to sentence-level syntactic and semantic annotation for each sentence. Identifying the empty connectives and accessing sentence-level syntactic and semantic information are crucial steps towards an automated inference of discourse relations in the absence of lexically realized connectives.

Finally, RST annotation of elementary discourse units, derived discourse units and rhetorical relations bear the entire burden of supporting language technology algorithms derived from the RST annotated corpus. The PDTB annotation effort will be an additional layer on top of text already annotated with syntactic structure (PTB) and predicate-argument relations (PropBank). These layers will be linked, and both their presence and their linkage will provide a richer substrate for the development and evaluation of practical algorithms.

We are not downplaying the importance of having an annotated corpus of coherence relations associated with adjacent discourse units. But we believe that the task of producing such a corpus can be made easier by having already identified the higher-order predicate-argument relations associated with explicit discourse connectives. They can then be factored into the calculation or removed from the calculation, as appropriate (Webber et al., to appear 2003).

6 The Future

For the next few years, the Penn Discourse TreeBank is the future of DLTAG. It will provide a Gold Standard for further parser development for DLTAG, and through its integration with the Penn TreeBank and PropBank, enable the development of data-intensive, probabilistic methods for resolving anaphoric connectives. It will undoubtedly be a source of interesting data and interesting ideas for many years to come.

References

van den Berg, M. H. (1996). Discourse grammar and dynamic logic. In P. Dekker & M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, pp. 93–111, ILLC/Department of Philosophy, University of Amsterdam.

Creswell, C., Forbes, K., Miltsakaki, E., Prasad, R., Joshi, A., & Webber, B.

(2002). The discourse anaphoric properties of connectives. In *Proceedings of the Discourse Anaphora and Anaphor Resolution Colloquium*, Lisbon, Portugal.

Cristea, D. & Webber, B. (1997). Expectations in incremental discourse processing. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL97/EACL97)*, pp. 88–95, Madrid, Spain.

Di Eugenio, B., Jordan, P. W., Moore, J. D., & Thomason, R. H. (1998). An empirical investigation of proposals in collaborative dialogues. In *Proceedings of COLING/ACL'98*, pp. 325–329, Montreal, Canada.

Forbes, K. (2003). Discourse semantics of s-modifying adverbials. Ph.D. thesis, Department of Linguistics University of Pennsylvania.

Forbes, K., Miltsakaki, E., Prasad, R., Sarkar, A., Joshi, A., & Webber, B. (2001). D-LTAG System – discourse parsing with a lexicalized tree-adjoining grammar. In *ESSLLI'2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*, Helsinki, Finland.

Forbes, K. & Webber, B. (2002). A semantic account of adverbials as discourse connectives. In *Proceedings of Third SIGDial Workshop*, pp. 27–36, Philadelphia PA.

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

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

Joshi, A. (1987). An introduction to Tree Adjoining Grammar. In A. Manaster-Ramer (ed.), *Mathematics of Language*, pp. 87–114. Amsterdam: John Benjamins.

Kingsbury, P. & Palmer, M. (2002). From treebank to propbank. In *Proceedings of the 3rd International Conference on Language Resources and Evalution (LREC)*, Las Palmas.

Kiparsky, P. & Kiparsky, C. (1970). Fact. In M. Bierwisch & K. E. Heidolph (eds.), *Progress in Linguistics*, pp. 143–173. Mouton.

Knott, A., Oberlander, J., O'Donnell, M., & Mellish, C. (2001). Beyond elaboration: The interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord, & W. Spooren (eds.), *Text Representation:Linguistic and psycholinguistic aspects*, pp. 181–196. John Benjamins Publishing.

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

Marcu, D. (1999). Instructions for manually annotating the discourse structure of texts. Available from http://www.isi.edu/~marcu.

Marcu, D. (2000). The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational Linguistics, 26(3)*, 395–448.

Miltsakaki, E., Creswell, C., Forbes, K., Prasad, R., Joshi, A., & Webber, B. (2003). Anaphoric arguments of discourse connectives: Semantic properties of antecedents versus non-antecedents. In *EACL Workshop on Computational Treatment of Anaphora*, Budapest, Hungary.

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

Polanyi, L. & van den Berg, M. H. (1996). Discourse structure and discourse interpretation. In P. Dekker & M. Stokhof (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, pp. 113–131, University of Amsterdam.

Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. (1972). *A Grammar of Contemporary English*. Harlow: Longman.

Sarkar, A. (2000). Practical experiments in parsing using tree-adjoining grammars. In *Proceedings of the 5th TAG+ Workshop*, pp. 193–198.

Scha, R. & Polanyi, L. (1988). An augmented context free grammar for discourse. In *Proceedings of the* 12th *International Conference on Computational Linguistics (COLING'88)*, pp. 573–577, Budapest, Hungary.

Schabes, Y. (1990). Mathematical and computational aspects of lexicalized grammars. Ph.D. thesis, Department of Computer and Information Science, University of Pennsylvania.

Schilder, F. (1997). Tree discourse grammar, or how to get attached to a discourse. In *Proceedings of the Second International Workshop on Computational Semantics*, Tilburg, Netherlands.

Soon, W. M., Ng, H. T., & Lim, D. C. Y. (2001). A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics, 27(4)*, 521–544.

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry, 34*, 649–689.

Webber, B. & Joshi, A. (1998). Anchoring a lexicalized tree-adjoining grammar for discourse. In *Coling/ACL Workshop on Discourse Relations and Discourse Markers*, pp. 86–92, Montreal, Canada.

Webber, B., Joshi, A., & Knott, A. (November 2000). The anaphoric nature of certain discourse connectives. In *Making Sense: from Lexeme to Discourse*, Groningen, The Netherlands.

Webber, B., Knott, A., & Joshi, A. (2001). Multiple discourse connectives in a lexicalized grammar for discourse. In H. Bunt, R. Muskens, & E. Thijsse (eds.), *Computing Meaning (Volume 2)*, pp. 229–249. Kluwer. Earlier version appears in the Proceedings of the Third International Workshop on Computational Semantics, Tilburg, 1999.

Webber, B., Knott, A., Stone, M., & Joshi, A. (1999a). Discourse relations: A structural and presuppositional account using lexicalised TAG. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics*, pp. 41–48, College Park MD.

Webber, B., Knott, A., Stone, M., & Joshi, A. (1999b). What are little trees made of: A structural and presuppositional account using lexicalised TAG. In *Proceedings of International Workshop on Levels of Representation in Discourse (LORID'99)*, pp. 151–156, Edinburgh.

Webber, B., Stone, M., Joshi, A., & Knott, A. (to appear 2003). Anaphora and discourse structure. *Computational Linguistics, 29*.

XTAG-Group, T. (2001). A Lexicalized Tree Adjoining Grammar for English. Tech. Rep. IRCS 01-03, University of Pennsylvania. See ftp://ftp.cis.upenn.edu/pub/ircs/technical-reports/01-03.