The Discourse Anaphoric Properties of Connectives

Cassandre Creswell, Kate Forbes, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi[*], Bonnie Webber[†]

[*]University of Pennsylvania
3401 Walnut Street
Philadelphia, PA 19104-6228, U.S.A.
creswell@babel.ling.upenn.edu, {forbesk, elenimi, rjprasad, joshi}@linc.cis.upenn.edu

[†]Division of Informatics University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, Scotland
bonnie@cogsci.ed.ac.uk

Abstract

Discourse connectives can be analyzed as encoding predicate-argument relations whose arguments derive from the interpretation of discourse units. These arguments can be anaphoric or structural. Although structural arguments can be encoded in a parse tree, anaphoric arguments must be resolved by other means. A study of nine connectives, annotating the location, size, and syntactic type of their arguments, shows connective-specific patterns for each of these features. A preliminary study of inter-annotator consistency shows that it too varies by connective. Results of the corpus study will be used in the development of resolution algorithms for anaphoric connectives.

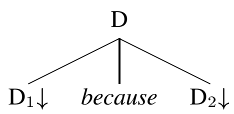
1. Introduction

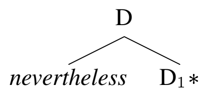
The theoretical background of our study of discourse connectives is Discourse Lexicalized Tree Adjoining Grammar (DLTAG) (Webber et al., 2001; Webber et al., 1999). DLTAG is an extension of LTAG in which *elementary trees*, anchored by discourse connectives, combine to create a discourse structure. That is, DLTAG is a grammar for discourse rather than for sentences.

As in other TAGs, there are two types of elementary trees, initial and auxiliary. Initial trees encode basic predicate-argument relations; auxiliary trees encode recursion. Discourse connectives can be analyzed as encoding predicate-argument relations whose arguments are the interpretations of discourse segments. A lexicalised grammar at the discourse level can capture these relations.

As at the sentential level, arguments to these discourse relations can be structural or anaphoric. The difference can be most easily seen in the case of multiple connectives (Webber et al., 2000). In (1), *because*, a structural connective at the discourse level, is the predicate expressing the causal relation between two eventualities, P = RAISE IRE (SALLY, FRIENDS) and Q = ENJOYS (SALLY, CHEESEBURGER). The two arguments must appear in the same elementary tree, shown in (2). In contrast, the connective *nevertheless* in S_3 finds only a single argument structurally Q = ENJOYS (SALLY, CHEESEBURGER). Its left-hand argument is derived anaphorically from S_1. The tree for *nevertheless* is found in (3); here, the discourse clause to which the *nevertheless* tree adjoins, D_1, is the sole structural argument.

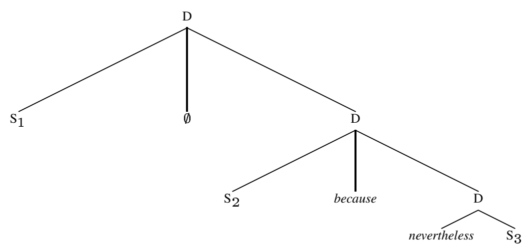
(1) a. S_1: Sally rarely eats meat and subscribes to *Vegetarian Times*.
S_2: Lately, she's raised the ire of her vegan friends
S_3: because she nevertheless enjoys the occasional bacon cheeseburger.

(2)


(3)


The full derived tree for this discourse is shown in (4).

In DLTAG, a sentence without a structural discourse connective is attached to the discourse structure via an *auxiliary tree* anchored in a lexically-empty discourse connective that conveys *continuation* of the description of the larger tree to which it is attached. Although a more specific relation may be inferred, the relation provided by the syntax alone is semantically underspecified, analogous to the semantics of noun-noun compounds.

(4)


Although the arguments to structural connectives are encoded directly in a parse tree, the non-structural argument to an anaphoric connective must be resolved by other means. This is similar to the case of bound versus free pronouns.

Discourse connectives share many properties with other types of discourse anaphora. For example, their anaphoric arguments may be found intra- or inter-sententially, as in (5) and (6), respectively.

(5) A person who seeks adventure might, for example, try skydiving. [(Webber et al., 2000)'s ft.8 (i)]

(6) Some people seek adventure. For example, they might try skydiving.

Because discourse connectives are some of the clearest indicators of discourse structure, annotating the arguments of the relations they convey provides information both about those arguments and about the range of possible discourse structures. Such an annotation study is described in the next section.

2. Corpus Study

This work is part of a larger discourse annotation project whose main goal is to provide a large, reliably annotated corpus for further scientific research and development of NLP applications. For each overt or null discourse connective, our goal is to identify and mark the minimal textual unit in the preceding discourse which contains the source of its left-hand argument. The current work focuses primarily on the arguments of anaphoric discourse connectives, as discussed in Section 1. The success of the overall project will contribute to our ability to understand and deal with an important aspect of discourse meaning, i.e., discourse relations.

2.1. Corpus Annotation

The work we report here is a first attempt to better articulate this research problem. As such, annotation is necessarily experimental and explorative. We start with a set of nine connectives picked from three semantic classes: resultatives (*as a result of, so, therefore*), concessives (*nevertheless, yet, whereas*) and additives (*also, in addition, moreover*). They are all adverbials that may modify phrasal constituents or the entire clause. *Whereas* also serves as a subordinate conjunction. We can verify its status as a subordinate conjunction if we apply the *reversibility test* (Quirk et al., 1972). Subordinate conjunctions introduce clauses that can be preposed with respect to the matrix clause, as in (7). Example (8) from the Brown corpus attests to the fact that *whereas* clauses can be preposed.

(7) a. Mary went to the party although she was tired.
b. Although she was tired, Mary went to the party.

(8) Whereas persons of eighth grade education or less were more apt to avoid or be shocked by nudity , those educated beyond the eighth grade increasingly welcomed and approved nudity in sexual relations . (Brown)

Whereas can also be found as an adverbial conjunction (Knott, 1996), both historically, as in Example (9), and in contemporary usage.

(9) The first bridge known to have been covered wholly or in part, – and perhaps the most interesting one, connected Newbury (now Newburyport) with Salisbury Point. Its building was first proposed in 1791, when a group of citizens, mostly Newburyport men, petitioned the General Court for an act of incorporation. This document began: "No. 1 Newbury Port, May 30th, 1791.

"Whereas, a Bridge over Merrimack River, from the Land of Hon'ble Jonathan Greenleaf, Esquire, in Newbery, to Deer Island, and from said Island to Salisbury, would be of very extensive utility, by affording a safe Conveyance to Carriages, Teams and Travellers at all seasons of the year, and at all Times of Tide .

In our corpus, *whereas* is mainly used as a subordinate conjunction, with the exception of the historical (now, legal) use of *whereas* which appeared in our corpus twice, as in Example (9).

For each of the nine connectives, seventy-five tokens (a total of 675 tokens) were extracted from a variety of corpora: Brown, Wall Street Journal, Switchboard and 58 transcribed oral histories from the online Social Security Administration Oral History Archives (SSA).[1] The 675 tokens were split in three groups (each group containing a connective from each semantic class) and annotated by three annotators (225 tokens per annotator).

Each token was annotated with tags that encoded information about (a) the connective's left argument (ARG), and (b) the clause containing the connective (CONN). Table 1 shows the ARG and CONN tag(sets) in the top and bottom box respectively. Both ARG and CONN were annotated with a REF tag that encoded an ID number which was the same for both in a single token. ARG was further tagged with a TYPE tagset that identified the size of the argument. The tags under TYPE were as follows: MAIN if the argument was contained in a full sentence (including subordinate clauses); MAIN-MULT if the argument was contained in a sequence of sentences; SUB if the argument was contained in a subordinate clause; and XP if the argument was contained in a phrasal constituent. The variation in the size of the argument was thus specified as a structural description.

This set of tags should enable us to identify statistically useful information about the type of the antecedent of anaphoric connectives, which will help us formulate constraints for anaphora resolution. In particular, the distinction between MAIN/MAIN-MULT and SUB/XP combined with the LOC tag (discussed in Section 2.3) will help us determine optimal structural descriptions for the connectives that will be useful for systems such as the DLTAG parser (Forbes et al., 2001). For example, connectives found to take only contiguous MAIN/MAIN-MULT arguments can be associated with a tree taking two structural arguments, thus maximizing compositional semantic representations derived directly from the syntax of discourse.

The clause containing the connective, CONN, was annotated with two tagsets: COMB and POSITION. COMB was used to identify punctuation marks (PERIOD, COMMA, etc.), coordinating conjunctions ('AND' and 'BUT'), and adverbial connectives ('YET', 'SO', etc.) that can co-occur with the connective. Information about co-occurrence with punctuation and other (mainly structural) connectives will also be useful for determining structural descriptions of connectives. In DLTAG, *and* and *but* are structural connectives anchoring elementary trees. That is, both their arguments must be realized structurally. Co-occurrence with *and* and *but* may be an indication that a connective cannot take both its arguments structurally without crashing the derivation or being assigned computationally complex structural descriptions. For the purposes of anaphora resolution, co-occurrence with punctuation com-

[1] The Brown, Wall Street Journal and Switchboard corpora are available from LDC, http://www.ldc.upenn.edu. The SSA corpus is available at http://www.ssa.gov/history/orallist.html

ARG		
	REF	ID #
	TYPE	MAIN= sentence MAIN-MULT= multiple sentences SUB = subordinate clause XP= phrasal constituent (NONE)= no left argument
CONN		
	REF	ID #
	COMB	PERIOD COMMA COLON SEMI-COLON DASH 'AND' 'BUT' CONN
	POSITION	INITIAL MEDIAL FINAL

Table 1: Annotation tagsets

bined with the results of the argument-size (TYPE) annotation will guide automated search for anaphoric arguments. Also, certain types of punctuation, e.g., dashes and parentheses, may indicate that the text containing the argument of the connective is not adjacent to the clause containing the connective. Co-occurrence with other connectives also raises the question of the semantics of the combined connective and its relationship to the semantics of the individual contributors, as for example, in the combination *and in addition* or *yet nevertheless*.

For CONN, we also defined a POSITION tagset which identified the position of the connective in its clause (INITIAL, MEDIAL, FINAL). As we have suggested in prior work (Forbes et al., 2001), the position of the connective in the clause will help us formulate constraints relevant to the information structure of the clause. Information structure is also relevant to anaphora resolution (Kruijff-Korbayová and Webber, 2001).

The complete set of tags we initially defined is given in Table (1). During the annotation, five more tags were added, which are not shown in this table but are discussed in the next section and appear in Table 2.

2.2. Annotation Results

Table (2) shows the results of the preliminary annotation for the nine connectives. The table contains percentages of the tags TYPE, COMB, and POSITION along with the actual number of occurrences of the tags in brackets. In the COMB tagset, a connective could combine with more that one of the categories of the group, so no percentages are given as the numbers do not add up to 75 for each category.

For most connectives there is a strong tendency for the left argument to be identified locally (in the structural sense) – either in the immediately preceding sentence or in immediately preceding sequence of sentences, in most cases the preceding paragraph. Most notably, *so* always takes a sentence or a sequence of sentences as its left argument, indicating that it may tentatively be treated as a structural connective. *In addition*, *yet*, *moreover*, *as a result* and *also*, tend to take their left argument locally but they demonstrate a larger syntactic variety of potential arguments such as subordinate clauses or phrasal constituents. *So*, *nevertheless* and *moreover* are likely to take larger discourse segments as arguments. Larger discourse segments appear to lead to vagueness in resolving anaphora – cf. Section 3. For example, it was often difficult to determine the extent of the left-hand argument of *nevertheless*, which could also be a phrasal intra-sentential constituent (XP). The connective *therefore* often takes its left-hand argument from a subordinate clause. In the ARG tagset, two additional tags were added during the annotation. The category OTHER was added by one annotator to describe cases where the left argument of the connective could not be identified. The category NONE was added only for *whereas*. It signifies that both arguments are to the right of the connective and therefore there is no left-hand argument. In fact, the presence of this tag indicates that *whereas* is a subordinate conjunction: only subordinate clauses can be fronted with respect to the main clause. This category is not relevant for the annotation of anaphoric connectives. The tag ZERO from the COMB tagset is, also, mainly relevant to *whereas*. It describes cases where the conjunction combines with no punctuation marks or other connectives. Rarely, the ZERO tag applies to adverbial connectives as in the case of *also*, shown in (10). However, in most cases, the presence of this tag indicates that the connective is a subordinate conjunction. Subordinate conjunctions do not combine with a punctuation mark or other connectives when the subordinate clause appears after the main clause. Finally, we found it useful to make special tags for combinations with a complementiser (COMP) and a subordinate conjunction (SUB). *As a result*, for example, quite often appears in complement clauses. This creates ambiguity in the interpretation, discussed in Section 4.1.

(10) The Controller's charge of rigging was the latest development in an investigation which *also* brought these disclosures Tuesday : ...

Regarding the position of connectives, *so* appears only in initial position. This supports the claim that *so* is a structural connective because structural connectives (subordinate and coordinate conjunctions) are restricted to the initial position. *Also*, on the other hand, frequently appears in medial positions, while the semantically similar *in addition* prefers the initial position.

The results of this preliminary annotation are promising and already reveal interesting distribution patterns. To further revise the annotation tags and guidelines and, crucially, test inter-annotator reliability, we focused our attention on three connectives *as a result*, *in addition* and *nevertheless*. Another twenty five tokens of each of the three connectives were extracted to add up to a total of hundred per connective and give an indication of intra-annotator consistency. The annotation of the complete set of three hundred tokens for the three connectives appears in Table (3). Comparison of Tables 2 and 3 shows that the relative percentages of each tag remained stable, indicating that the anaphoric arguments of each of these connectives display patterns that can be recognized via a large-scale annotation project, and be used to lead to reliable annotation algorithms. What remains to be shown is that this annotation is reliable, such that the same patterns are perceived across annotators.

CONNECTIVE	IN ADDITION	SO	YET	NEVERTHELESS	MOREOVER	THEREFORE	AS A RESULT	WHEREAS	ALSO
TYPE									
MAIN	65.3% (49)	45% (34)	53.3% (40)	37.3% (28)	42.7% (32)	25.3% (19)	78.6% (59)	46.7% (35)	69.3% (52)
MAIN-MULT	18.7% (14)	55% (41)	33.3% (25)	36% (27)	45.3% (34)	21.3% (16)	18.7% (14)	4% (3)	9.3% (7)
SUB	5.3% (4)	0	2.7% (2)	9.3% (7)	8% (6)	31%(24)	2.7% (2)	16% (12)	12%(9)
XP	10.7% (8)	0	10.7% (8)	17.3% (13)	4%(3)	21.3% (16)	0% (0)	1.3% (1)	4% (3)
(NONE)	-	-	-	-	-	-	-	32% (24)	-
(OTHER)	-	-	-	-	-	-	-	-	5.3% (4)
COMB									
PERIOD	65	33	33	47	68	28	55	26	49
COMMA	9	22	14	5	2	1	0	36	7
SEMICOLON	1	2	8	0	0	0	3	5	0
DASH	1	0	4	0	0	0	0	0	0
AND	12	2	8	1	4	41	14	0	7
BUT	0	0	0	17	1	0	1	0	4
YET	0	0	0	2	0	0	0	0	0
SO	0	0	0	0	0	4	0	0	0
ZERO	0	0	0	3	2	0	0	8	1
COMP	0	0	0	0	0	0	1	0	8
SUB	0	0	0	0	0	1	1	0	0
POS									
INITIAL	92% (69)	100% (75)	98.7% (74)	78.6% (59)	82.7% (62)	88% (66)	90.7% (68)	100% (75)	17.3% (13)
MEDIAL	8% (6)	0% (0)	1.3% (1)	18.7% (14)	17.3% (13)	12% (9)	2.7% (2)	0% (0)	80% (60)
FINAL	0% (0)	0% (0)	0% (0)	2.7% (2)	0% (0)	0% (0)	6.6% (5)	0% (0)	2.7% (2)

Table 2: Annotation results for 9 connectives

CONNECTIVE	IN ADDITION	NEVERTHELESS	AS A RESULT
TYPE			
MAIN	63% (63)	36% (36)	68% (68)
MAIN-MULT	19% (19)	35% (35)	26% (26)
SUB/COMP	10% (10)	10% (10)	5% (5)
XP	8% (8)	18% (18)	0% (0)
OTHER	0% (0)	0% (0)	1% (1)
COMB			
PUNCTUATION	101	78	80
DASH	1	0	0
AND	12	1	17
BUT	0	2	1
CONN	0	2	0
COMP	0	0	10
SUB	0	0	1
POS			
INITIAL	94% (94)	82% (82)	91% (91)
MEDIAL	6% (6)	16% (16)	3% (3)
FINAL	0% (0)	2% (2)	6% (6)

Table 3: Annotation results for 3 connectives

LOC	SS=same sentence PS=previous sentence PP=previous paragraph NC=non-contiguous

Table 4: Values for ARG-TAG LOC

2.3. Inter-Annotator Agreement

Our studies in the prior section suggest that a human can identify and find patterns in the arguments of the connectives studied. The study presented in this section suggests that this identification and the patterns found are reliable. To test the reliability of our annotation, three additional annotators annotated twenty-five of the original hundred tokens of each of the three connectives (*in addition*, *as a result*, *nevertheless*), yielding a total of four annotations of twenty-five tokens of each of these connectives. Each connective and its anaphoric argument were, as in the prior study, assigned an ID. However, in order to focus on the ability of multiple annotators to agree on the unit from which the anaphoric argument is derived, we employed only one tag, LOC. Each annotator labeled the anaphoric argument with one of the four possible values of this tag shown in Table 4.

The LOC tag defines the sentence as the relevant atomic unit from which the anaphoric argument is derived. A *sentence* is minimally a main clause and all (if any) of its attached subordinate clauses. The semantic argument of the connective could thus be derived from the single sentence containing the connective (SS), the single prior sentence (PS), a sequence of adjacent sentences (PP), or a sequence of sentences not contiguous to the clause containing the connective (NC). In other words, we did not ask the annotators to distinguish sub-clausal constituents or subordinate clauses, we did not distinguish the exact boundaries of sequences of sentences when we marked more than one sentence as the argument, and we did not distinguish whether a non-adjacent argument comprised one clause or a sequence of them. In this sense, the LOC tag can be viewed as an abstraction of the TYPE tag; however, it adds the additional information of whether the anaphoric argument is contiguous to the clause containing the connective. Reasons for employing the LOC tag will be discussed in Section 3.

The inter-annotation results produced using the LOC tag are shown in Table 5. The first column indicates the connective, and the remaining columns contain the percentage of tokens in which a particular pattern of agreement was found for each connective. The first column represents the case in which all four annotators produced the same tag, the second column represents the case in which three out of four annotators produced the same tag, the third column represents the case where two out of four annotators produced the same tag but the remaining two annotators had different tags, and the fourth column represents the case where two annotators produced one tag, and the other two annotators produced another tag. That there is no "0" column reflects the fact that in every case, there was some agreement among annotators, e.g. there was no case in which each annotator selected a different tag.

Connective	**4/4**	**3/4**	**2/4**	**<2, 2>/4**
in addition	76% (19)	16% (4)	4% (1)	4% (1)
as a result	84% (21)	12% (3)	0	4% (1)
nevertheless	52% (13)	36% (9)	0	12% (3)

Table 5: Inter-Annotator Agreement

Nevertheless was more difficult to annotate than *in addition* or *as a result*. As the project expands, we will probably continue to find more and less difficult annotation cases. However, four-way inter-annotator agreement is greater than 50% in every case, and majority agreement (three-way or better) is 88% for *nevertheless*, 92% for *in addition*, and 96% for *as a result*. We conclude that the anaphoric arguments of discourse connectives can be reliably annotated.

In the next section, we discuss how a detailed investigation of annotator disagreements can be used to develop a resolution algorithm for the anaphoric arguments of discourse connectives.

3. Towards a Resolution Algorithm

A closer look at 1) how the annotations vary in the inter-annotation study and 2) the results of the more complex annotations in the individual annotation studies, reveals certain issues relevant to developing a resolution algorithm, as discussed below.

As mentioned above, we employed the LOC tag instead of the TYPE tag in the study of inter-annotator agreement. By additionally asking each annotator to record the boundaries of the units she identified as the "exact" unit from which the anaphoric argument was derived, we were able to derive the values for the TYPE tags from each of the four annotations. For the purposes of inter-annotator agreement we found that "exact match" was not a useful comparison, due to differences in the implicit guidelines each annotator was individually following. However, the "exact match" comparison, combined with the data from the first study, is useful for elucidating these differences and understanding why they arise. The implicit differences between the annotations fall into two main categories, the size of the argument and the syntactic form of the argument. Both concern the annotator's understanding of the properties of the unit that are necessary to derive the semantic argument of the connective. Consider the discourse in (11).

(11) John was happy. As a consequence, he smiled. As a result, Mary smiled.

(12) John is a male American. He is six feet tall. He has brown hair. As a result, he was drafted.

(13) John is overworked, and as a result, tired.

When deciding on the anaphoric argument of *as a result*, one annotator might decide that Mary's smiling is the result of John's smiling, and so tag the argument as PS. Because John's smiling is a consequence of his being happy, however, another annotator might tag the argument as PP, e.g. as including both the first and second clause. Similarly, consider the discourse in (12). When deciding on the anaphoric argument of *as a result*, one annotator might decide that John's being a man is the cause of his being drafted (females not being drafted in America historically), and thereby tag the argument as NC because John's being six feet tall and having brown hair is an elaboration (or a parenthetical aside) of his being a male American. However, another annotator might tag the argument as PP, e.g. as including the first three clauses. Finally, consider the sentence in (13). When deciding on the anaphoric argument of *as a result*, one annotator might decide that because *as a result* modifies an adjective on the right, its left argument should be (using the TYPE tag) an XP, e.g. "overworked". Another annotator might interpret "tired" as a small clause, or a clause with a deleted subject and verb, and so he might tag the entire clause "John is overworked" as the anaphoric argument of *as a result* using the MAIN tag. (Note that this issue is avoided when the LOC tag SS is employed.)

What all of these cases have in common is the question of how large to make the argument. What they also have in common, however, is that in each case it is possible to select a *minimal unit* as the argument, and allow the relations between that unit and the surrounding context to complete the interpretation. In (11), if the annotator selects "As a consequence, he smiled" as the argument of *as a result*, the relation between John being happy and smiling will not be lost, for *as a consequence* will take as its anaphoric argument the semantic interpretation of "John was happy". Similarly, in (12) if the annotator selects "John is a male American" as the argument of *as a result*, the relation between John being a male American, being six feet tall, and having brown hair will not be lost, for the empty connective signalling basic elaboration will link these arguments to the first clause structurally[2].

An additional complication that arises in the annotation of examples like (13) is the role of the lower-level syntactic annotation. In the Penn Treebank, from which the majority of our data is drawn, there is no principled parsing of such cases, in that it is left to the annotator to decide whether a particular use of a gerund, adjective, etc. should be parsed as a clause with missing elements when it is modified by an adverbial discourse connective. Therefore, we cannot reliably invoke the syntactic parse to decide when to label the left argument as a clause or an XP. We could, however, draw an analogy with coordinating conjunctions, which are commonly parsed with two XP arguments (e.g. John is happy and tired), although at the semantic level, two propositions are arguably being conjoined. If we allow the syntactic XP unit to represent a semantic proposition, then we can invoke the *minimal unit* heuristic here too. This would have the additional benefit of retaining parallelism in the syntactic form of the arguments of the connectives in such constructions.

Another potential heuristic in resolving the arguments of anaphoric connectives is their ability to combine with particular structural connectives, such as *but* and *and*. An auxiliary tree anchored with one of these connectives must be adjoined to its left-hand argument. Another connective, like *nevertheless*, *therefore*, and *moreover*, adjoined into this structure at the same point will frequently take as its own anaphoric 'lefthand' argument the structural connective's lefthand argument (e.g., (14). [3]

(14) He believed that <ARG> the Federal Security administrator had the authority and the responsibility for actions taken throughout the agency,</ARG> <CONN>and therefore he should be apprised of them

[2]Note that these same issues arise for a series of elaborations followed by *in addition*, and in the same way a *minimal unit* can be selected.

[3]But not always, as the examples that motivate the distinction between anaphoric and structural connectives demonstrate.

and should play a part in the decisions.</CONN> (www.ssa.gov/history/ajaoral3.html)

A similar heuristic could be used for determining the size of the lefthand argument. In particular, when the right argument is a constituent smaller than a full clause (e.g. the second of two conjoined VPs), the left argument appears to consistently be the same size (e.g. the first of two conjoined VPs), as in (15).

(15) Jasper arrived late and therefore got no dinner.

An investigation of the variations in "exact match labeling" using the LOC tag and the individual labeling using the TYPE and COMB tags shows that if these heuristics had been employed, many of the 22/75 cases of less than four-way agreement would have become four-way agreement. These minimal unit and connective combination cases, however, are distinguished from other issues that arise during the annotation of anaphoric arguments of discourse connectives, in that they are not cases of "true ambiguity" because principled heuristics can be introduced to resolve them. There are true cases of ambiguity, where such heuristics are not possible. One such case is discussed in the following section.

4. Remaining issues

4.1. Ambiguity in Complement Clauses

Cases of 'true ambiguity' in identifying the left argument of a connective were found in connectives contained in complement clauses, mostly complements of verbs of *saying*. A connective in a complement clause may connect either the complement clause with the preceding sentence or with the main clause containing the verb of saying. To illustrate the point, consider example (16). This example is ambiguous between analyses (17) and (18).

(16) Moritz said Monday his leg feels fine and , as a result , he hopes to start practicing field goals this week .

(17) Moritz said Monday [that his leg feels fine and, as a result, he hopes to start practicing field goals next week.]

(18) [Moritz said Monday his leg feels fine] [and, as a result, he hopes to start practicing field goals this week.]

In (17), the left argument of *as a result* is the first complement claues and is annotated as SS (same sentence) as both the argument clause and the connective clause are the conjoined object of the matrix clause verb. In (18), the connective clause forms a full sentence by itself. On this interpretation *as a result* was not part of what Moritz said but was added by the writer. More generally, connectives appearing after a complement clause can take as their left argument either the complement clause itself on the interpretation that both the left argument and the connective clause are part of the complement, or the matrix clause and the complement clause combined.

4.2. Low Attachment

As stated above, a reason we used the LOC tag in inter-annotator agreement was because the TYPE tag did not distinguish contiguous from non-contiguous arguments. This is an important distinction to make, because such arguments cannot be modeled structurally, thus indicating that they must be resolved anaphorically.

Because anaphoric connectives do not retrieve their left argument structurally, the clause containing them must attach to the prior discourse via an empty structural connective. In the DLTAG parser (Forbes et al., 2001), we currently employ the procedure of always attaching this empty connective to the leaf of the right frontier of the growing tree. If we could identify the anaphoric argument through a resolution mechanism, we could attach this empty connective to the clause containing the argument (at the top level), thus building the resolution into the tree.

5. Conclusions and Future Work

We have reported the results of a preliminary corpus analysis of (primarily) anaphoric discourse conenctives and the location and type of their left-hand arguments. We will use this study and the annotation manual we have been developing, as the starting point for a more extensive study that will create a layer of annotations on top of both the Penn Tree Bank (syntactic) annotations and PropBank (semantic) annotations (Kingsbury and Palmer, 2002), in order to begin to capture more semantic properties of the sources of anaphoric arguments. This should increase the possibility of developing a resolution algorithm for anaphoric discourse connectives that is both highly sensitive and highly specific to the phenomena at hand.

6. References

Kate Forbes, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber. 2001. D-LTAG system: Discourse parsing with a lexicalized tree adjoining grammar. In *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*.

Paul Kingsbury and Martha Palmer. 2002. From treebank to propbank. In *Third International Conference on Language Resources and Evaluation, LREC-02, Las Palmas, Canary Islands, Spain*.

Alistair Knott. 1996. *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Ph.D. thesis, University of Edinburgh.

Ivana Kruijff-Korbayová 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.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London:Longman.

Bonnie Webber, Alistair Knott, Matthew Stone, and Aravind Joshi. 1999. Discourse relations: A structural and presuppositional account using lexicalized TAG. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, Maryland*, pages 41–48. College Park MD.

Bonnie Webber, Alistair Knott, and Aravind Joshi. 2000. Multiple discourse connectives in a lexicalized grammar for discourse. In *Proceedings of the Third International Workshop on Computational Semantics*, Tilburg, The Netherlands.

Bonnie Webber, Aravind Joshi, Matthew Stone, and Alistair Knott. 2001. Anaphora and discourse semantics. *(submitted to) Computational Linguistics*.