

A Semantic Account of Adverbials as Discourse Connectives

Kate Forbes
Linguistics Department
University of Pennsylvania
Philadelphia, PA
forbesk@unagi.cis.upenn.edu

Bonnie Webber
Division of Informatics
University of Edinburgh
Edinburgh UK EH8 9LW
bonnie@cogsci.ed.ac.uk

Abstract

We address the question of why certain adverb and preposition phrases are only interpretable with respect to the discourse, and not just their own matrix clause. We show that, in many cases, an adverbial’s compositional semantics explains why. We close by reporting on an annotation study aimed at providing specific evidence for how adverbials are interpreted with respect to the discourse.

1 Introduction

Traditionally in linguistics, syntax and semantics provide mechanisms to build the interpretation of a sentence from its parts. Although it’s non-controversial that a *sequence* of sentences such as (1) also produces an interpretation, the mechanisms are less clear.

(1) There is a high degree of stress level from the need to compete and succeed in this ‘me generation’. As a result, people have become more self-centered over time.

Rhetorical Structure Theory (Mann and Thompson, 1988) bases sequence interpretations on “discourse” (or “rhetorical”) relations. These can be used to label interpretations of sentence sequences, and (Marcu, 2000) shows that an automated method of labeling can partially replicate human annotation. This automation relies in part on *cue phrases*, or *discourse connectives*,

whose presence is taken to signal the presence of particular discourse relations. In (1), the relevant phrase is *as a result*, and the discourse relation is not surprisingly a *result relation*.

There have also been several formal attempts to build a compositional semantics for discourse, including (van den Berg, 1996; Gardent, 1997; Schilder, 1997). However, as in RST, this work assumes all discourse connectives signal a relation between adjacent discourse units. DLTAG (Webber et al., 2001), in contrast, proposes a different role for adverbial phrases. While subordinating and coordinating conjunctions signal a relation between the interpretations of the clauses they structurally connect, adverbials relate the interpretation of their matrix clause to a clausal interpretation available anaphorically from the discourse context. These distinctions are based both on considerations of computational economy and behavioral evidence.

While DLTAG discusses adverbials that function as discourse connectives, it does not isolate this subset from the set of all adverbials. In this paper, we consider the larger class of adverb (ADVP) and preposition (PP) adverbials. We investigate the mechanisms that cause certain adverbials, which we call *discourse adverbials*, to be bound by the discourse context, while other adverbials, which we call *clausal adverbials*, seem fully interpretable with respect to just their matrix sentence (once anaphoric references to entities are resolved¹). Clausal adverbials are exemplified in (2), and discourse adverbials in (3).

¹This is (Knott, 1996)’s methodology, see Section 2.

(2) *Surprisingly/Probably/In my town*, men take care of all household finances.

(3) *As a result/Consequently/In addition*, men take care of the household finances.

Differences between these two types of adverbials cannot be attributed to syntax. Viewed in terms of a tree structure, both types of adverbials adjoin to the S node, and the resulting structure is still an S².

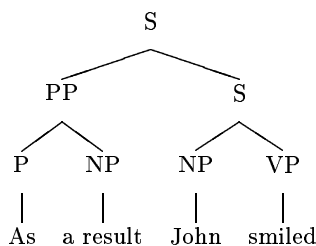


Figure 1: S-Adjoined Adverbials

It is in terms of semantics and the argument structure of adverbials that significant differences appear. We argue that both the number of semantic arguments an adverbial has and the resolution of these arguments determines whether or not it functions semantically or pragmatically as a discourse or clause adverbial. While the semantics and pragmatics discussed here will not provide a complete account of the discourse functions of all adverbials, it will show that the analysis of adverbials can be viewed modularly: certain functions can be attributed to the semantic domain, others to the pragmatic domain, and still others to larger issues of discourse structure.

There are numerous benefits of this analysis. First, it is economical, making use of pre-existing mechanisms to build the interpretation of discourse, thereby reducing the load on inference to account for discourse interpretation (c.f. (Kehler, 1995)). Secondly, it provides a theoretical grounding for (Knott, 1996)’s empirical approach to discourse connectives, in the pro-

²Most of these adverbials can adjoin to other nodes, but our semantics is applicable at any adjunction site. See (Forbes, 2002) for further discussion.

cess showing that additional adverbials should be included in the class, and that some of those included there don’t really belong. Thirdly, it expands an existing model of discourse which argues that discourse structure can be built directly on top of clause structure. It thereby bridges the gap between the clause-level structure and the shallow discourse parse produced by RST.

The remainder of this paper is organized as follows. In Section 2 we discuss prior research into the semantics of adverbials. Section 3 presents our analysis. Section 4 discusses an annotation project which uses this approach to provide specific evidence for how adverbials are interpreted with respect to discourse.

2 Prior Research

The semantics of most discourse adverbials has not been well studied. Work by (Knott, 1996) is an exception. Although the lexical semantics he proposes for adverbial cue phrases is orthogonal to the issues under discussion in this paper, he developed an experimental methodology aimed at isolating adverbials and conjunctions that appear unacceptable outside of a discourse context. Because his methodology relied completely on human judgments and did not control for factors such as intonation, his corpus contains apparent “errors of omission” as well as “errors of commission”. For example, his set of connectives includes *surprisingly* and *clearly* but not *in that event* or *as a consequence* (See (Forbes, 2002) for further discussion).

In general, investigations into PP (Verspoor, 1997) and ADVP (Ernst, 1984) treat issues of syntactic and semantic ambiguity at the clause level; when discourse adverbials are mentioned at all they are called “conjunctive adverbials” or “discourse connectives” and are specified as the domain of discourse research. We can nevertheless use this research as a guide when building the semantics of discourse adverbials, because all S-adjoined ADVP and PP can be classified along two semantic dimensions: (1) the type of modification they perform; and (2) the type of semantic object they modify.

2.1 Type of Modification

(Ernst, 1984)’s classification of modification that ADVP clausal adverbials perform is shown in Table 1. While there are no similar classifications in the literature of PP clausal adverbials, it is easy to find PP synonyms for these ADVP. Both are exemplified in the table.

Mod Type	ADVP	PP
Agent-Oriented	wisely	in a wise way
Domain	legally	in legal terms
Evaluative	surprisingly	to my surprise
Mental Attitude	willingly	in a willing way
Speaker-Oriented	frankly	to speak frankly
Epistemic	probably	in all likelihood

Table 1: Modification Types (Ernst, 1984)

While this classification may be useful for distinguishing ADVP functions, it does not help to distinguish *discourse* from *clausal* adverbials.

A broader classification of modification type can be formulated in terms of semantic features to encompass both the discourse and clausal adverbials. In such terms, adverbials supply *manner*, *spatio-temporal*, and/or *degree-of-likelihood* features to the semantic objects they modify, among other features. For example, “clearly” and “generally” may supply both degree-of-likelihood and spatio-temporal features. Such features could be helpful for annotations such as those described in Section 4, and are discussed in more detail in (Forbes, 2002).

What these features do not capture is that while adverbials like “tactfully” supply the manner of an agent performing an action, adverbials like “in addition” supply the manner in which the modified clause is related to the surrounding discourse. Accounting for this requires discussion of the semantic objects that these adverbials supply features *to*, as we now discuss.

2.2 Type of Semantic Object

Adverbials can also be classified according to the type of semantic object they modify. This classification is described only vaguely in (Ernst, 1984) with respect to ADVP, in (Verspoor, 1997) with respect to PP, and across the adverbial literature in general. While (Moore, 1993) provides a more formal analysis, we believe that

research on discourse deixis provides a better basis for classification.

2.2.1 (Ernst, 1984)

As shown in Table 2, Ernst’s semantics makes three distinctions between the modified clause: it can be *a situation*, *a state of affairs*, or *a mental state*. Specific properties of each are not discussed. Note that he makes use of the adjective (ADJ) derivative of the ADVP. α refers to the semantic object being modified by the ADVP.

Agent-Oriented	The agent is judged ADJ due to α . $\alpha = \textit{situation}$
Domain	α is relevant in ADJ domain. $\alpha = \textit{situation}$
Epistemic	α is ADJ. $\alpha = \textit{situation}$
Evaluative	α is ADJ. $\alpha = \textit{state of affairs}$
Mental Attitude	α manifests mental state. $\alpha = \textit{situation}$

Table 2: Ernst’s Semantic Interpretations

2.3 Abstract Objects

We argue here that the range of semantic objects evoked when an adverbial modifies a clause coincides with the range of semantic objects to which a discourse deictic refers.

Briefly, the term “discourse deixis” applies to uses of demonstrative pronouns such as (4)–(7), where their antecedents (derived from the bracketed material) are interpretations of entire clauses as in (4), more than one clause as in (5), discourse relations between clauses, such as the contrast relation in (6), and inferred defeasible rules, such as in (7), where “if it’s raining the sun isn’t shining” is presupposed and denied (Knott, 1996).

(4) [John took Biology 101.] *That* means he can take Biology 102.

(5) [I woke up and brushed my teeth. I went downstairs and ate breakfast, and then I went to work.] *That’s* all I did today.

(6) [If a white person drives this car it’s a “classic”. If I, a Mexican-American, drive it,

it's a “low-rider”.] *That* hurts my pride. (Dahl and Hellman, 1995)

(7) [The sun is shining although it's pouring rain.] That's a rule Bermuda always breaks.

In general, discourse deixis is used to refer to the interpretation of non-NP constituents, including (a set of) tensed or untensed clauses, verb phrases, and adverbial phrases. The nature of such interpretations goes under the rubric *Abstract Object* (AO). (Asher, 1993) organizes types of AOs as shown in Figure 2.

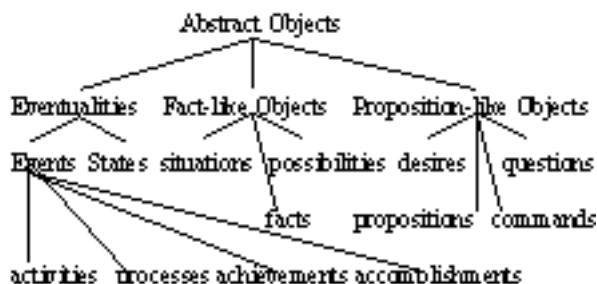


Figure 2: (Asher, 1993) Classes of AOs

Whether or not this classification of AOs is complete, we argue that AOs provide an appropriate way of understanding adverbial semantics. We end this overview of AOs with a comment on how they can be formalized in the semantics. Traditionally, only the proposition interpretation of S's is formalized. The semantic type of S is usually a truth value: true or false. Discourse deictic research and adverbial research both show that other interpretations are possible. (Dahl and Hellman, 1995), for example, use type coercion to access other AO interpretations. Type coercion is a term taken from computer science, where it defines an operation by which an expression which is normally of one logical type is re-interpreted as another (e.g., when an integer is understood as a Boolean value). Type coercion has also been used to explain linguistic phenomena, such as when an expression which is indeterminate as to logical type is 'coerced' into one particular interpretation and thus acquires a fixed type.

3 A New Approach

3.1 Data Used in this Study

Data for the following discussion was drawn from both the Brown Corpus and WSJ Corpus. Because both clausal and discourse adverbials adjoin to the S node, this syntactic property was used to extract the data investigated in this study. Using `tgrep`, all S-initial, S-adjoined ADVPs and PPs were collected from the Treebank parses of these corpora³. Counts of types and tokens for S-initial, S-adjoined ADVPs and PPs are shown in Table 3.⁴ For comparison, counts of those occurring 12 or more times are shown, labeled “>11”. This comparison shows that while high frequency S-adjoined ADVP tokens occur more frequently than high frequency S-adjoined PP tokens, overall there are many more tokens and types of S-adjoined PP in the combined corpora. This is due to the wide variability in the internal PP argument; over half of the PP tokens in these corpora occur only once.

	Tokens	Types
>11 ADVP	2559	53
>11 PP	1738	49
Total ADVP	3519	578
Total PP	8829	5982

Table 3: Counts of ADVP/PP Adverbials

3.2 Adverbials and Abstract Objects

We argue that what adverbials take as a semantic argument is the AO interpretation of the clause they adjoin to. The third column of Table 4 contains common adverbials found in the WSJ and Brown corpora. For comparison, the second column contains comparable discourse deictic reference, and the first column specifies the AO type. For illustrative purposes, consider the sentence “People made mistakes” as the one being referred to and modified. As shown, we can interpret this sentence as any number of AOs. If

³See <http://www ldc.upenn.edu/ldc/online/treebank/> for documentation of `tgrep` and Penn Treebank.

⁴Due to annotation errors, these counts are not perfectly accurate. Also topicalized modifiers of other categories are not excluded because their underlying position is not represented in Treebank.

referred to by a discourse deictic, it is the predication on the discourse deictic that determines the AO interpretation. “Happened in 1984”, for example, is a property of events, so the referent of the discourse deictic is interpreted as an event. If the clause is modified, it is the adverbial that performs the predication, thereby determining the AO interpretation.

AO	DD	Adverbial
event	that happened in 1987	in 1987,
fact	that's a fact	in fact,
proposition	that's true	in truth,
speech act	that's a frank statement	frankly,
pure textual	repeat that	again,
rule	example (7)	as a rule,

Table 4: Abstract Object Examples

More generally, notice that we can create a PP adverbial for an AO, simply by inserting it into the phrase “As a(n) ...”.

We claim that the AO interpretation of the clause they adjoin to is the *only* semantic AO argument of clausal adverbials. In contrast, discourse adverbials require more: they either take a second AO argument or implicate at least one other AO, thereby rendering them uninterpretable with respect to their matrix clause *alone*, even after resolution of any NP-referring constituents.

Discourse adverbials are thus very similar to discourse deixis in the sense that both require an AO interpretation from the prior discourse or spatio-temporal context. (Webber et al., 2001) provides strong behavioral evidence for this view. One of their most compelling examples is shown in (8). Just as in (6) the discourse deictic takes as its referent the discourse relation between clauses, so does *for example* in (8) select as its prior argument the result relation (imparted by *so*) between the two clauses.

(8) John just broke his arm. So, for example, he can't cycle to work now.

More generally, notice that a discourse deictic can always be replaced by its explicit demonstrative+AO counterpart, e.g. “that text”, “that speech act”, etc. We can usually supply the de-

ictic or the explicit demonstrative+AO counterpart as the NP argument of a PP S-modifier, thereby creating a discourse adverbial which relates two abstract objects. *For that reason, in that case, in this instance, after that* are found in the corpora and exemplify such cases.

Of course, explicit reference is not the only mechanism by which discourse adverbials are created. In the following sections we discuss these and other mechanisms in greater detail. We use a simplified version of lambda calculus such as found in (Heim and Kratzer, 1998) to represent the adverbials and their arguments.

3.2.1 ADVP adverbials

We can represent ADVP as semantically akin to ADJP and use its ADJP derivative. For example, the interpretation of *gray* (indicated as $[[\cdot]]$) is as follows, where *gray'* distinguishes the semantic property supplied by the term from the term itself, and x represents its structural argument:

$$(9) [[gray]] = \lambda x[\text{gray}'(x)]$$

The interpretation of *probably* is thus represented as follows, where x resolves to the interpretation of the modified clause⁵:

$$(10) [[probably]] = \lambda x[\text{probable}'(x)]$$

Common ADVP clausal adverbials are shown in Table 5, along with their corpus counts and their ADJP counterpart. While not all ADVP have an ADJP counterpart (i.e. *maybe*), their similarity in meaning to those that do can be used to identify their semantic arguments.

Count	Adverbial	ADJP
23	obviously	obvious
21	unfortunately	unfortunate
22	certainly	certain
19	apparently	apparent
13	probably	probable

Table 5: ADVP Clausal Adverbials

⁵In Bayesian epistemic models, the semantics of such epistemic expressions is slightly more complex, because they are interpreted with respect to possible worlds.

Note that many of these clausal adverbials have a syntactically optional indefinite PP argument, as exemplified in (11), which does not resolve to an AO interpretation.

(11) Unfortunately (for John), vandals covered John’s car in graffiti.

In general, syntactically optional arguments, also called “missing” or “hidden”, fall into two types (Mittwoch, 1982): definite and indefinite. Indefinite hidden arguments are not necessarily anaphoric with anything. Examples include the VP *eat*: “Mary ate (the food)”, and the NP *mother*: “Mary talked to the mother (of John)”. Definite hidden arguments are anaphoric to some salient entity in the discourse context. For example, the hidden argument in the *apply* VP in (12) is the earlier-mentioned grant:

(12) The due date for the grant has passed. Mary didn’t apply ϕ .

Some ADVP discourse adverbials, like some ADVP clausal adverbials, have hidden semantic arguments. The crucial difference, however, is that the hidden argument of ADVP discourse adverbials is definite, and it resolves to an AO interpretation. Examples are shown in Table 6, along with their counts.

Count	Adverbial
15	consequently
15	mainly
6	conversely
2	additionally
2	alternatively

Table 6: Hidden AO Arguments

These ADVP discourse adverbials can be represented, e.g. as follows:

(13) $[[\text{consequently}]] = \lambda y[\text{consequent}'(y, X)]$

The definite hidden argument – in this case, what y , which is bound to the interpretation of the matrix clause, is a consequence of – is represented as a free variable X , e.g. anaphoric to some AO from the previous discourse.

3.2.2 PP adverbials

Structurally, PPs are more complex than ADVP. A PP is composed of one internal (usually NP) argument, and one external argument.

Semantically, prepositions are thus binary predicates. The semantics of prepositions is represented as in (14), where P represents the property supplied by the preposition “prep”, and x and y represent its structural arguments:

(14) $[[\text{prep}]] = \lambda x \lambda y.P(x, y)$

The semantic interpretation of a PP S-modifier has the form of interpretation in (14). For example, if city' is taken to represent the interpretation of the NP “the city”, then the interpretation of the clausal adverbial “in the city” can be written:

(15) $[[\text{in}]] = \lambda y.\text{in}'(\text{city}', y)$

where y resolves to the interpretation of the modified clause.

Note that if we exchange “the city” for “consequence”, we produce the discourse adverbial “in consequence”. What distinguishes PP clausal adverbials from PP discourse adverbials, therefore, is the character of the internal NP argument.

Many of the PPs found in PP adverbials take a proper noun as internal argument (Table 7) and supply a spatio-temporal property to the event interpretation of the modified clause.

Count	Adverbial
33	in New York Stock Exchange composite trading
33	on Friday
27	in August
17	in 1987

Table 7: PP Adverbials with Proper Noun

None of these cases found in the corpora require prior context to interpret their NPs, although months and days are interpreted with respect to the spatio-temporal context.

Sometimes, the internal argument of the PP

adverbial may be anaphoric or deictic (Table 8).

Count	Adverbial
27	since then
14	after that

Table 8: PP Adverbial with Anaphor

In (16), for example, “that” refers to the event interpretation of the first clause. Thus the PP adverbial functions as a discourse adverbial, relating the AO interpretation of its matrix clause with that of a previous clause. Note the potential for ambiguity: In (17), “then” can refer to the NP “the morning” or to the event of waking early.

(16) I went to the movies. *After that*, I ran a hundred errands.

(17) In the morning, I woke up early. *Since then*, I’ve run a hundred errands.

The internal PP argument can also be an explicit demonstrative reference to an AO, as exemplified in Table 9.

Count	Adverbial
2	for that reason
10	in that case
3	in this instance

Table 9: Explicit AO Reference

Again, the referent of these NP arguments may be an AO interpretation of a prior clause, as in (18), or an NP within a prior clause, as in (19), or it may be ambiguous.

(18) John couldn’t sleep. *For that reason*, he got out of bed.

(19) Yesterday we discussed reasons for leaving. You gave me a good one. *For that reason*, I thank you.

The internal argument of a PP adverbial can also be a definite NP (Table 10).

Following (Heim and Kratzer, 1998), definite determiners are partial functions, requiring the

noun they modify to return true for only one entity in the set of individuals, but the resulting interpretation may or may not be anaphoric, in the sense that it refers back to an entity in the prior discourse. Only when they are anaphoric to an AO interpretation, will these PP adverbials function as discourse adverbials.

Count	Adverbial
71	at the same time
69	for the nine months
20	in the end
19	in the past
17	at the time
13	in the third quarter

Table 10: PP Adverbial with Definite NP

Of the entries in Table 10, the first one, “at the same time”, is always anaphoric because “same” (like “other”) requires a discourse entity of the appropriate type for it to be the same as (or other to).

Returning to the example with which we started this paper (“as a result”), it is also possible that the internal PP argument itself has a hidden semantic AO argument. Other examples are shown in Table 11⁶.

Count	Adverbial
204	in addition
167	for example
70	for instance
41	at least
28	by contrast
17	at last
16	at first
16	in particular

Table 11: Hidden AO Argument

These will receive a similar semantic treatment to corresponding ADVP adverbials such as “additionally” which contain hidden arguments.

Note that like definite, demonstrative, and pronominal arguments, the resolution of hidden

⁶In many of these examples we can supply the argument (e.g. “in addition to this”). A reviewer has noted that in German similar arguments are not optional; they are either explicit or lexically incorporated into the adverbial. For example, according to a German speaker, ‘als Folge’ or ‘als Resultat’ is infelicituous, but ‘als Folge dessen’ (‘as a result of that’) is fine, and even better is ‘demzufolge’, (‘dem’ = ‘that’; ‘zu’ = preposition meaning ‘as’; ‘folge/resultat’ = ‘consequence/result’).

arguments can be ambiguous. Consider (20):

(20) *Complications* is filled with other stories demonstrating the capriciousness of medicine. For example, Gawande once detected a case of the rare, often fatal infection necrotizing fasciitis (flesh-eating bacteria) because he happened to have seen a case a few weeks before. [The Nation, 6 May 2002, p.35]

Gawande once detecting a case of necrotizing fasciitis could be taken to be an example of any of the following: (1) the fact of *Complications* being filled with other stories demonstrating the capriciousness of medicine, (2) the set of other stories that demonstrate the capriciousness of medicine, (3) the event of other stories demonstrating the capriciousness of medicine, or (4) the capriciousness of medicine (i.e. the NP).

Moreover, like discourse deixis, adverbial AO arguments can sometimes be derived from sub-clausal constituents (e.g. “He was tired and, despite that, happy”).

The potential for such ambiguities in ADVP and PP adverbials means that unless the (hidden or explicit) anaphor can *only* be resolved to a noun phrase rather than equally well to the AO it participates in, it is likely most efficient to resolve it to the most-inclusive AO that makes sense in the context and derive more specific readings from that.

In the case of (20), if something is an example of a book being filled with stories that demonstrate *X*, then one can infer (if no evidence to the contrary) that it is an example of such stories, and from that, that it is an example of demonstrations of *X*, and from that, that it is an example of *X*.

Understanding both the mechanisms that determine which elements are accessible to function as antecedents and the elements which allow an AO interpretation to be derived, will require the production and analysis of an annotated corpus, like the one described in Section 4, and should, moreover, consider the analyses already proposed for discourse deixis. (See (Forbes, 2002) for further discussion.)

3.3 Other Factors

We end this discussion of adverbials by noting that we are not claiming that it is only due to their argument structure that adverbials appear to require discourse for their interpretation. Knott’s set of ADVP cue phrases also includes those in Table 12.

Count	Adverbial
31	actually
16	surely
6	really

Table 12: Other ADVP Adverbials

In these cases, the semantics contains only one AO argument, the interpretation of the modified clause. The truth or fact asserted by this clause is asserted to be “actual”, “sure”, or “real”.

Similarly, the PP shown in Table 13 are generally viewed as discourse connectives. Following (Heim and Kratzer, 1998), indefinite NPs and common NPs denote sets of entities, and are not in and of themselves referential. As such, their interpretation should be fully interpretable in isolation.

Count	Adverbial
24	in any case
21	for one thing
117	of course
105	in fact
49	after all

Table 13: Other PP Adverbials

We believe that the explanation for why Knott’s subjects found such adverbials to require a discourse context comes from semantic mechanisms such as (topic) focus and Contrastive Theme (Steedman, 2000) and/or pragmatic mechanisms such as Conversational Implicature (Grice, 1989). For example, adverbials in first position, as in:

(21) Nowadays we eat alot of vegetables.

may receive the same stress as other topicalised constructions, linking them to what they are being used to illustrate or contrast with in the discourse context.

Focus and implicature can also interact with discourse adverbials; in the focus literature, for example, adverbials such as “at least” are generally viewed as being sensitive to focus, though discussion is limited to non-clausal modification, such as “Most of my friends are uneducated. John at least finished high-school.” For further discussion, see (Forbes, 2002).

4 Annotation Project

Because in the clause level parse, only one of the arguments of discourse adverbials comes compositionally, the other must be retrieved from the discourse. DLTAG views this as a problem of anaphora resolution. As with other anaphora, developing algorithms capable of resolving them in way that reflects their actual distribution in discourse requires developing an annotated corpus. In this section, we discuss a project (Creswell et al.,) to develop such a corpus, and some of our initial results.

The main objective of this project is to build a corpus with discourse annotation. While not a complete representation of discourse structure, it provides a rich intermediate level between high level discourse structure and clause structure that can be reliably annotated, namely, the syntax and semantics associated with discourse connectives. We use the Penn Treebank annotated corpus, which contains naturally occurring data from a variety of sources, has been annotated for clause structure and part-of-speech, and is being annotated for predicate-argument structure. Our annotation schema is designed to build an additional layer of discourse annotation, with links to the clause level information.

Discourse annotation will occur in two stages. First, the DLTAG parser (Forbes et al., 2001) is used to parse the discourse, i.e. the structural arguments of each discourse connective. Second, human annotators correct any errors in the parse, and add annotation tags for the anaphoric arguments of the discourse adverbials. This strategy proved successful in the clause-level annotation of the Penn Treebank corpus in minimizing human effort.

Preliminary development of the project in-

volves determining an initial set of discourse connectives to be annotated, building a semantic frame for each connective, and developing annotation tags. Semantic frames, modeled after those used in predicate-argument tagging, are provided to the annotator, along with annotation tags and guidelines, to serve as a basic semantics for each connective and help the annotators determine the relevant semantic roles played by the context around them in the corpus discourses. Each frame will indicate the number of arguments each connective takes, the semantic or pragmatic relation it supplies to them, their relevant semantic or pragmatic properties, such as AO type, and semantic features of the connectives such as mentioned in Section 2 and in (Knott, 1996). The annotation tags so far developed include a *POSITION* tag, which records the position of the adverbial in the clause it modifies, a *COMBINATION* tag, which records any other connectives that co-occur with the adverbial in the clause, a *TYPE* tag, which records the syntactic type of the argument, and a *NOTE* tag, which records any unexpected behavior. So far, three annotators have each annotated 75-100 instances of three different discourse adverbial, each representing one of three types of discourse relation (concessive, resultative, additive), yielding a total of nine annotation sets.

A pilot study investigating the reliability of human annotation of the anaphoric arguments of discourse adverbials has employed a single *LOC* tag with four possible values: SS (argument is within same main clause as connective), PS (argument is within immediately prior main clause), PP (argument spans multiple main clause immediately prior), NC (argument found in non-contiguous prior clause). 25 instances of three adverbials, one from each of the above-mentioned relation types, have been inter-annotated by three different annotators. Results are shown in Table 14, where all three annotators produced the same tag.

As shown, agreement varies across adverbials, due in part to the meaning of the adverbial, and also to the size of its anaphoric argument. In general, the larger the argument, the harder it is to determine the exact unit from which it is

drawn. Resolution algorithms will have to take these issues into account.

	'as a result'	'nevertheless'	'in addition'
3-way agree	21/25	15/25	19/25

Table 14: Inter-annotator Agreement

5 Conclusion

The underlying theme of our work is that discourse is not a completely separate category from semantics and that discourse properties arise from the same lexico-syntactic substrate as clause-level semantic properties. Here we have shown that what have been called “cue phrases” are not an accidental grouping of adverbials and PPs: their discourse properties arise naturally from their semantics. Corpus annotation and analysis will then allow us to better understand the empirical realisation of this view, and to better understand the correspondences between discourse connectives and discourse relations, such as when and which discourse connectives are used, versus when and which discourse relations are inferred.

Acknowledgements

The authors thank Dr. Ellen Prince and the DLTAG and dissertation groups at UPenn for helpful criticism.

References

- Nicholas Asher. 1993. *Reference to Abstract Objects*. Kluwer, Dordrecht.
- C. Creswell, K. Forbes, E. Miltsakaki, R. Prasad, B. Webber, and A. Joshi. The discourse anaphoric properties of connectives, in progress.
- Osten Dahl and Christina Hellman. 1995. What happens when we use an anaphor. *Presentation at the XVth Scandinavian Conference of Linguistics, Oslo, Norway*.
- Thomas Boyden Ernst. 1984. *Towards an Integrated Theory of Adverb Position in English*. PhD dissertation, Indiana University.
- K. Forbes, E. Miltsakaki, R. Prasad, A. Sarkar, B. Webber, and A. Joshi. 2001. Dltag system: Discourse parsing with a lexicalized tree-adjointing grammar. *ESSLLI 2001*.
- Katherine Forbes. 2002. *Discourse Semantics of S-Modifiers*. PhD dissertation, in progress, University of Pennsylvania.
- Claire Gardent. 1997. Discourse tree adjoining grammars. Claus report nr.89, University of the Saarland, Saarbrücken.
- H. P. Grice. 1989. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA.
- Irene Heim and Angela Kratzer. 1998. *Semantics in Generative Grammar*. Blackwell.
- Andrew Kehler. 1995. *Interpreting Cohesive Forms in the Context of Discourse Inference*. PhD dissertation, Harvard University.
- Alistair Knott. 1996. *A Data-Driven Methodology For Motivating A Set Of Coherence Relations*. PhD dissertation, University of Edinburgh.
- William Mann and Sandra Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.
- Daniel Marcu. 2000. The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational Linguistics*, 25(3):395–448.
- Anita Mittwoch. 1982. On the difference between eating and eating something: Activities versus accomplishments. *Linguistic Inquiry*, 13.
- Robert C. Moore. 1993. Events, situations and adverbs. In R. Weischedel and M. Bates, editors, *Challenges in Natural Language Processing*. Cambridge University Press.
- Frank Schilder. 1997. Tree discourse grammar, or how to get attached to a discourse. In P. Dekker and M. Stokhof, editors, *Proceedings of the Second International Workshop on Computational Semantics*. Tilburg, Netherlands.
- Mark Steedman. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4):649–689.
- Martin van den Berg. 1996. Discourse grammar and dynamic logic. In P. Dekker and M. Stokhof, editors, *Proceedings of the Tenth Amsterdam Colloquium*. University of Amsterdam, ILLC/Department of Philosophy.
- Cornelia Maria Verspoor. 1997. *Contextually-Dependent Lexical Semantics*. PhD dissertation, University of Edinburgh.
- Bonnie Webber, Aravind Joshi, Matthew Stone, and Alistair Knott. 2001. Anaphora and discourse semantics. *submitted to Computational Linguistics*.