

Discourse Connectives: Structural, Anaphoric or What?

Bonnie Webber and Rashmi Prasad

University of Edinburgh

University of Pennsylvania

September 17, 2009

1 Introduction

- Motivation
- Background

2 Relevant Evidence

- Intra-sentential connectives
- Inter-sentential discourse connectives
- Paired Connectives

3 Conclusion

What's the problem?

The arguments to discourse connectives are not always adjacent.

On a level site you can provide a cross pitch to the entire slab by **raising one side of the form (step 5, p. 153)**, but for a 20-foot-wide drive this results in an awkward 5-inch ($20 \times 1/4$ inch) slant across the drive's width. Instead, **make the drive higher at the center**.

⇒ Instead of raising one side of the form, make the drive higher at the center

What's the problem?

The arguments to discourse connectives appear to establish non-standard relations within a sentence – eg, between a relative clause and the main clause

wsj_1171

UAL, which closed on the Big Board Monday at \$178.375 a share, traded in the third market afterward as low as \$158 a share.

or a headless relative to the main clause

wsj_0961

whatever losses are incurred in the pursuing of price stability would surely be more than made up in increased output thereafter.

What's the problem?

Discourse connectives with different senses can have essentially the same arguments, thereby establishing extra relations between those arguments.

wsj_1453

But while the merits of diversification shine through when times are tough, there's also a price to pay: ...

But while the merits of diversification shine through when times are tough, there's also a price to pay: ...

while ⇒ Comparison.Concession

also ⇒ Expansion.Conjunction

What's the problem?

Multiple discourse connectives can occur in a single clause, that establish distinct relations with different parts of the discourse.

On a level site you can provide a cross pitch to the entire slab by **raising one side of the form (step 5, p. 153)**, but for a 20-foot-wide drive this results in an awkward 5-inch ($20 \times 1/4$ inch) slant across the drive's width. So instead, **make the drive higher at the center**.

... , **but for a 20-foot-wide drive this results in an awkward 5-inch ($20 \times 1/4$ inch) slant across the drive's width. So instead, **make the drive higher at the center**.**

Why is this a problem?

If all discourse connectives contribute to discourse structure, this would make discourse structure very complex.

With only one argument to many discourse connectives well-defined syntactically (AKA, the *bound argument*), there is significant variability in the location and character of the other argument.

Perhaps this is just a matter of salience and probability: Anything is possible as an argument, but some things are more salient and hence more likely than others.

Alternatively, there may be theoretical (ie, **categorical**) differences between connectives that

- tell upon the location and character of the other argument;
- mean a less complex discourse structure.

What evidence can be brought to bear on this? What does such evidence imply? What follows from it?

Assumptions about discourse structure

Discourse connectives convey relations between elements of text interpretable as **abstract objects** (eg, events, states, facts, beliefs, propositions, etc.) relations.

Adjacency between sentences or clauses can also imply relations between them:

wsj_0640

The numbers show that “we don’t have a monolithic economy,” said Isaac Lagnado, council research director. “There are a lot of have and have-not markets.”

just as in N-N compounds such as *aluminum soup pot cover*.

If such relations define a level of discourse structure, this still doesn’t require all **discourse connectives** to contribute to discourse structure: Only some of them might do so.

Assumptions about discourse structure

This is the case in syntax, where not all relations between elements in a sentence are taken to contribute to syntactic structure – verbs and their arguments, yes; binding and anaphora, no –

Any woman who_i John $_j$ asks 0_i to come to his $_j$ party with [her $_i$ husband] $_k$ and [another man/woman] $_{k,i}$ is likely to accept.

These often complex relations don't count towards the complexity of syntactic structure.

In characterizing the formal complexity of discourse structure, it is possible that not all discourse relations should be counted either.

Background on discourse structure

Early work on discourse **assumed** a particular structure for discourse, and instead focussed on the source of that structure or against alternative structural assumptions.

- Some approaches assumed discourse formed a **tree** structure and argued whether this arose from task decomposition [Dale, 1992], rhetorical relations [Mann & Thompson, 1988], genre-specific schemas [McKeown, 1985], discourse intentions [Grosz & Sidner, 1986].
- Some approaches assumed that certain types of discourse have a simple **linear** topic structure – e.g. [Sibun, 1992], [Hearst, 1997], [Barzilay & Lee, 2004].

Background on discourse structure

More recently, Wolf & Gibson (2005) have claimed that if one considers all possible discourse relations, including relations between entities and *abstract objects* (e.g. `ATtribution` and `ENTITY ELABORATION`), discourse can be seen to have the structure of a **chain graph** of directed and non-directed edges, many of which will cross.

Background on discourse structure

We started our **lexicalized** approach to discourse relations (DRels) >10 years ago [Webber & Joshi, 1998] with these assumptions:

- DRels were associated with discourse connectives and with adjacency.
- DRels could only hold between text spans interpretable as **abstract objects** (AOs).
- Discourse connectives were either **structural** or **anaphoric**.
- **Structural connectives** had a structural connection to both their arguments, while **anaphoric connectives** were connected structurally to one argument and anaphorically to the other.
- Only DRels from structural connectives or adjacency contributed to the complexity of discourse structure.
- That structure was probably a tree.

Background on discourse structure

But if discourse connectives were either structural or anaphoric, one still needs to specify which were which.

We assumed that structural connectives were ones that linked clauses within a sentence:

- coordinating conjunctions
- subordinating conjunctions
- subordinators (eg, *in order to*, *so that*)

or formed paired adverbials within or across sentences (eg, *On the one hand ... On the other (hand) ...*; *Not only ... But also ...*).

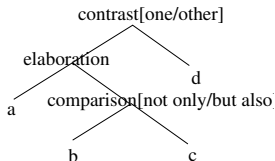
Single discourse adverbials (eg, *instead*, *for example*, *however*, *in that case*), we assumed to be anaphoric.

Evidence from Crossing

Evidence for this categorical distinction was that, for connectives taken to be **structural**, their args could not **cross** without changing the interpretation.

Embedded [Webber et al, 2003, ex. 8]

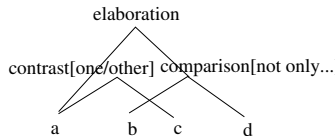
- On the one hand, Fred likes beans.
- Not only does he eat them for dinner.
- But he also eats them for breakfast and snacks.
- On the other hand, hes allergic to them.



Evidence from Crossing

Crossed [Webber et al, 2003, ex. 10]

- On the one hand, Fred likes beans.
- Not only does he eat them for dinner.
- On the other hand, hes allergic to them.
- But he also eats them for breakfast and snacks.

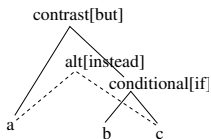


Evidence from Crossing

As with other anaphora, there was no such constraint on the arguments to connectives such as *instead* taken to be **anaphoric**.

Crossing [Webber et al, 2003, ex. 12]

- a. High heels are fine for going to the theater.
 - b. But **wear comfortable shoes**
 - c. **if instead you plan to go to the zoo.**
-
- a. High heels are fine for **going to the theater.**
 - b. But wear comfortable shoes
 - c. if instead **you plan to go to the zoo.**



Evidence from Explicit Deictic/Anaphoric Adverbials

Additional evidence for a categorical distinction was taken to come from discourse adverbials with explicit deixis or anaphora:

in that case, at this point, then, thus, now

Other discourse adverbials could be argued to have an implicit deictic or anaphor [Forbes et al, 2005]:

- in addition \Rightarrow in addition to that
- furthermore \Rightarrow further on to that
- as a result \Rightarrow as a result of that
- at the same time \Rightarrow at the same time as that
- by/in contrast \Rightarrow by/in contrast with that
- likewise \Rightarrow in like ways as that
- later on \Rightarrow later on from that

But neither crossing nor the existence of implicit deictic or anaphoric component yields a procedure for deciding whether a given connective is **structural** or **anaphoric** (or whether it belongs to some other class of connective).

- Evidence is incomplete: Not all connectives seem to have a deictic or anaphoric component.
- Evidence is inconclusive: Even with an implicit deictic or anaphoric component doesn't guarantee it's resolved in the same way.
- Evidence is sparse: Even in a large corpus, there may be little or no evidence of how combinations of connectives behave vis-a-vis crossing.

Nevertheless, we want to characterize connectives theoretically and empirically, in terms of both

- **where** they can and do get their arguments (both with respect to sentential structure and discourse structure);
- **what**, if anything, constrains their arguments (other than being spans interpretable as AOs).

For this, we need ways of finding key examples – possibly ones suggested by data available in annotated corpora. But those corpora aren't big enough.

The **Linguist's Search Engine** [Resnik & Elkiss, 2005] could help find such examples in unannotated corpora by their lexico-syntactic patterns, but the LSE is limited to searching for individual sentences.

In the next section of the talk, I'll point to kinds of evidence that would be relevant.

Relevant Evidence

Relevant Evidence could come from:

- Discourse connectives, both of whose arguments are in the same sentence;
- Sentence-initial coordinating conjunctions
- Paired connectives with the same arguments

Assumption: If there are different types of discourse connectives (eg, **structural** and **anaphoric**), a particular connective will belong to only a single type, although it may have non-connective uses as well (eg, *so*, *when*) or multiple senses as a connective (eg, *still*).

Evidence from connectives with the same args

Multiple connectives within the same sentence, like multiple verbs within the same sentence, can have the same arguments or different ones.

If multiple verbs within the same sentence have the same arguments, the verbs must be in construction together (eg, conjoined or embedded).

John prepared and ate his dinner.

John prepared to eat his dinner.

By analogy, one could say that if multiple discourse connectives in the same sentence have the same arguments, either the connectives must be in construction together or at most one of them can get both its arguments through structure.

wsj_1453

But while the merits of diversification shine through when times are tough, there's also a price to pay: ...

But while the merits of diversification shine through when times are tough, there's also a price to pay: ...

wsj_0239

... even if you hate heights, you can still balloon.

... even if you hate heights, you can still balloon.

Evidence from connectives with the same arguments

If we take subordinating conjunctions as **structural** since their other argument is predicted by syntactic structure (modulo the effects of attribution and conjunction), then there is evidence in the WSJ corpus for

- *also* (wsj_1453)
- *later on* (wsj_0640)
- *nevertheless* (wsj_2070)
- *still* (wsj_0186, wsj_0239, wsj_1416, wsj_1866, wsj_2048, wsj_2161)

being anaphoric.

Evidence from connectives with different arguments

Now consider multiple connectives within the same sentence with different arguments within the sentence, as in:

wsj_1171

UAL, which closed on the Big Board Monday at \$178.375 a share, traded in the third market afterward as low as \$158 a share.

wsj_0961

whatever losses are incurred in the pursuing of price stability would surely be more than made up in increased output thereafter.

If, as with binding, we disallow discourse adverbials from creating additional structure within the sentence, then they too must be anaphoric. The WSJ corpus provides evidence for the following as anaphoric:

- *afterward* (wsj_1171)
- *also* (wsj_0725, wsj_0820, wsj_1622, wsj_1659, wsj_2048, wsj_2104, wsj_2346)
- *earlier* (wsj_1026, wsj_2381)
- *instead* (wsj_0155)
- *later* (wsj_0609, wsj_1261)
- *nevertheless* (wsj_0444)
- *nonetheless* (wsj_0296, wsj_1480)
- *previously* (wsj_0464, wsj_1802, wsj_1837, wsj_1844, wsj_1986)
- *still* (wsj_1938)
- *then* (wsj_0909)

Evidence from S-initial coordinating conjunctions

Within a sentence, when they connecting clauses or VPs, we take the coordinating conjunctions (*and*, *but*, *nor*, *or* and *so*) to be **structural** discourse connectives.

What about when they occur sentence-initially?

CC	Total	S-initial
And	3000	648
But	3308	2119
Nor	31	18
Or	98	28
So	263	111

Evidence from S-initial coordinating conjunctions

If S-initial coordinating conjunctions are taken to be **structural** connectives, do they display any behavior that would distinguish them from **anaphoric** connectives?

The only possible differences could be in

- where within discourse structure they get their non-bound argument
- where within sentential structure they get their non-bound argument.

Evidence from S-initial coordinating conjunctions

Even discarding *attribution*

wsj_0003

“There’s no question that some of those workers and managers contracted asbestos-related diseases,” said Darrell Phillips, vice president of human resources for Hollingsworth & Vose. “But you have to recognize that these events took place 35 years ago.”

the non-bound argument to a S-initial coordinating conjunction isn’t necessarily left-adjacent.

Evidence from S-initial coordinating conjunctions

Conj	Total	Adjacent	1-2	3-4	5-6	>6 sentences
And	648	506	117	19	3	2
But	2119	1666	343	68	15	27
Nor	17	11	5	1	0	0
Or	28	25	3	0	0	0
So	111	75	24	8	1	3
Also	118	75	29	7	4	2
However	176	132	37	4	0	2
In addition	164	82	64	13	2	3
Meanwhile	140	38	48	19	11	24

S-initial discourse adverbials display similar patterns. So **distance** can't be what distinguishes **structural** from **anaphoric** discourse connectives.

Evidence from S-initial coordinating conjunctions

Perhaps it could be the **nature** of any intervening material?

In syntax, coordinating conjunctions express “a relation between two or more elements of syntactically equal status [Huddleson & Pullum, 2002].

Perhaps something similar holds of inter-sentential coordinating conjunctions.

Can we show that intervening material does not have *equal status* to the arguments, while the same constraint doesn't hold for some/all discourse adverbials?

To be determined!

Evidence from S-initial coordinating conjunctions

Structural and anaphoric connectives might also be distinguishable in terms of the syntactic nature of what could serve as their non-bound argument.

In the WSJ corpus, everything that can serve as the argument to a S-initial coordination conjunction can also serve as the argument to a discourse adverbial.

The reverse is not true: Several things that commonly serve (ie, >4 times in the WSJ Corpus) as the argument to a discourse adverbial are rarely, if ever, found with s-initial coordinating conjunctions – eg, indefinite relative clauses and appositives.

Non-bound args in indefinite non-restrictive relative clauses

wsj_0209

The researchers said they have isolated a plant gene **that prevents the production of pollen**. **The gene thus can prevent a plant from fertilizing itself**.

wsj_0795

Much of the moderation came from declining energy prices, **which have since turned up a bit**, analysts said. Consequently, Michael Darby, undersecretary for . . . , said **inflation probably will edge up from the third-quarter rate in the final three months of 1989**.

wsj_1829

These units were handling calls both from people in the San Francisco area and from computers themselves, which are set to dial Digital automatically when trouble arises. They then run remotely controlled self-diagnostic programs.

Non-bound args in indefinite restrictive relative clauses

wsj_0400

However, analysts expect the Fed to buy Treasury bills **that were auctioned yesterday in the secondary market**.
The Treasury **also** held a hastily scheduled \$2 billion sale of 51-cash management bills yesterday.

wsj_0450

Two days earlier, his attorney met in a Park Avenue law office with a cartoon dealer who expected **to sell 44 of the most important stolen strips to Mr. Russell for \$62,800**. **Instead**, New York City police seized the stolen goods, and Mr. Krisher avoided jail.

Non-bound args in (NP) appositive clauses

wsj_0093

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

Non-bound args in (NP) appositive clauses

wsj_0518

The proposal is just part of legislation that would require businesses to provide health benefits, an idea **that is strongly opposed by small business who say it would just compound the insurance-cost problems.**

But small-business lobbyists say **they support the idea, included in the Kennedy-Waxman bill, of new laws or regulations requiring greater use of community rating, . . .**

N.B. These do occur with S-initial coordinating conjunctions, but they are rare (1 instance) and probably better analyzed as event relatives.

Evidence from paired connectives

Some discourse adverbials can be found paired with both coordinating and subordinating conjunctions (ie, ones considered structural).

but otherwise

For some reason I have no sound for you tube videos but otherwise sounds work as normal. [forum.egypt.com/...]

because otherwise

Obama: Pass This Giant Spending Bill, Because Otherwise We're Just Continuing the Old Ways.
[campaignspot.nationalreview.com/...]

Evidence from paired connectives

so finally

Our nest is empty so finally there is time to blog.
[jedpatemptynest.blogspot.com/...]

but finally

Its been a long time coming — but finally we're out of beta.
[derivadow.com/...]

Suggest that even without an explicit conjunction, the same sense is conveyed.

Evidence from paired connectives

otherwise

For some reason I have no sound for you tube videos. Otherwise sounds work as normal.

Obama: Pass This Giant Spending Bill. Otherwise We're Just Continuing the Old Ways.

finally

Our nest is empty. Finally there is time to blog.

Its been a long time coming — Finally we're out of beta.

Perhaps it is an implicit connective providing structure, with the discourse adverbial providing an additional relation anaphorically.

Conclusion

It is useful to understand, both empirically and theoretically, where and how discourse connectives get their unbound argument.

There appears to be evidence, both intra-sententially and inter-sententially, for making a distinction between anaphoric and structural connectives (ie, it's not just a continuum of salience).

However, all the possible evidence isn't in.

But there is still evidence be gathered from annotated corpora.

There is a great need to be able to find targetted examples in as yet unannotated corpora.

References

- Regina Barzilay and Lillian Lee (2004). Catching the Drift: Probabilistic Content Models, with Applications to Generation and Summarization. *Proc. 2nd Human Language Technology Conference and Annual Meeting of the North American Chapter, Association for Computational Linguistics*, pp. 113-120.
- Robert Dale (1992). *Generating Referring Expressions*. Cambridge: MIT Press.
- Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi and Bonnie Webber (2005). Attribution and the (Non-)Alignment of Syntactic and Discourse Arguments of Connectives. *Proceedings of 2nd Workshop on Frontiers in Corpus Annotation*, Ann Arbor MI.
- Katherine Forbes-Riley, Bonnie Webber and Aravind Joshi (2005). Computing Discourse Semantics: The Predicate-Argument Semantics of Discourse Connectives in D-LTAG. *Journal of Semantics* 23(1), pp. 55–106.

- Barbara Grosz and Candace Sidner (1986). Attention, Intention and the Structure of Discourse. *Computational Linguistics*, 12(3):175–204.
- Marti Hearst (1997). TextTiling: Segmenting Text into Multi-paragraph Subtopic Passages. *Computational Linguistics*, 23(1):33–64.
- Alan Lee, Rashmi Prasad, Aravind Joshi, Nikhil Dinesh, and Bonnie Webber (2006). Complexity of dependencies in discourse. *Proc. 5th International Workshop on Treebanks and Linguistic Theories (TLT)*, Prague.
- Mann, W. and Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.
- Kathleen McKeown (1985). *Text Generation*. Cambridge: Cambridge University Press.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber (2008). The Penn Discourse TreeBank 2.0. *Proc. 6th Int'l Conference on Language Resources and Evaluation*.

- Philip Resnik and Aaron Elaissari (2005). The linguist's search engine: an overview. *Proc. ACL 2005 Interactive poster and demonstration sessions*, pp. 33–36.
- Penni Sibun (1992). Generating Text without Trees. *Computational Intelligence* 8(1):102–122.
- Bonnie Webber and Aravind Joshi (1998). Anchoring a Lexicalized Tree-Adjoining Grammar for Discourse. *Coling/ACL Workshop on Discourse Relations and Discourse Markers*, Montreal Canada, pp. 86-92.
- Bonnie Webber, Matthew Stone, Aravind Joshi and Alistair Knott (2003). Anaphora and Discourse Structure. *Computational Linguistics* 29(4):545–587.
- Florian Wolf and Edward Gibson (2005). Representing Discourse Coherence: A Corpus-based Study. *Computational Linguistics* 31:249–287.