Penn Discourse Treebank:
Annotation Tutorial

Institute for Research in Cognitive Science
University of Pennsylvania

Project Supervisors: Aravind Joshi and Bonnie Webber
Annotation Supervisors: Eleni Miltsakaki and Rashmi Prasad

September 27, 2004

Contents

1 Introduction ........................................... 4
   1.1 Discourse Relations and Discourse Connectives .................. 4
   1.2 Corpus ........................................... 6
   1.3 Goals of PDTB ..................................... 6
   1.4 What are Explicit Connectives .................................. 7
   1.5 What are Implicit Connectives ................................ 7
   1.6 TASK A: Annotation of Explicit Connectives .................... 8
       1.6.1 Set 1 connectives ................................ 8
       1.6.2 Set 2 connectives ................................ 8
       1.6.3 Set 3 connectives ................................ 8
       1.6.4 Set 4 connectives ................................ 8
       1.6.5 Set 5 connectives ................................ 9
       1.6.6 Set 6 connectives ................................ 9
       1.6.7 Set 7 connectives ................................ 9
       1.6.8 Set 8 connectives ................................ 9
       1.6.9 Set 9 connectives ................................ 9
       1.6.10 Set 10 connectives ................................ 9

1
1.7 TASK B: Annotation of Implicit Connectives ........................................ 9

2 Guidelines for Explicit Connectives ..................................................... 10
  2.1 Words/phrases that look like connectives ........................................ 10
    2.1.1 Word/phrase has a single non-propositional argument ............... 10
    2.1.2 Word/phrase has only a single propositional argument ............ 11
    2.1.3 Word/phrase has two arguments but one or both are not situations . 12
    2.1.4 Relative Pronouns that look like Connectives .......................... 12
  2.2 Words/phrases as preposed predicate complements ............................ 13
  2.3 Connectives post-modified with “of” ............................................. 14
  2.4 Pre-modified connectives ............................................................ 14
  2.5 Co-occurring connectives ........................................................... 15
  2.6 Non-sentence/clause initial connectives ......................................... 15

3 Guidelines for Arguments of Explicit Connectives ................................. 15
  3.1 Types and Sizes of Arguments ..................................................... 16
    3.1.1 Single Sentences ............................................................. 16
    3.1.2 Multiple sentences ......................................................... 16
    3.1.3 Clauses ........................................................................... 17
    3.1.4 Nominalizations .............................................................. 19
  3.2 Inclusion of all predicate arguments and modifiers within a connective’s argument ................. 20
  3.3 Arguments and supplements: the minimality principle .......................... 21
  3.4 What kinds of things can be Supplements ....................................... 23
    3.4.1 Clauses can be supplements .............................................. 23
    3.4.2 Clause-internal arguments and modifiers cannot be supplements ..... 23
    3.4.3 Parentheticals can be supplements ......................................... 24
  3.5 Numbering of arguments for subordinate conjunctions .......................... 24
  3.6 Head Nouns in Relative Clause arguments ....................................... 25
  3.7 Exclusion of connectives from their own argument annotations .......... 25
  3.8 Discontinuous annotation .......................................................... 26
  3.9 Discourse Deixis ........................................................................ 27
  3.10 Ambiguities .............................................................................. 27
  3.11 Conventions for relative pronouns .................................................. 28
3.12 Conventions for punctuation .............................................. 28
3.13 Conventions for complementizers ........................................ 28

4 Guidelines for Particular Connectives .................................... 29
  4.1 Coordinating conjunctions in VP coordinations .................... 29
  4.2 As ................................................................. 29
  4.3 So ................................................................. 30
  4.4 Nor ................................................................. 30

5 Guidelines for Implicit Connectives ...................................... 30
  5.1 Naming implicit connectives ........................................... 30

6 Instructions for WordFreak .................................................. 30
  6.1 Tool Interface ...................................................... 31
  6.2 Launching the Tool .................................................. 32
  6.3 Opening and Loading Files for Annotation .......................... 33
  6.4 Text View .......................................................... 33
  6.5 Tree View .......................................................... 35
  6.6 Annotations in the Chooser .......................................... 35
  6.7 Viewing and selecting Annotations ................................... 36
  6.8 Adding Annotation nodes for connectives by searching with the tool .................................................. 37
  6.9 Adding annotation nodes for connectives manually ................ 38
    6.9.1 Using Tree View for text selection ............................ 38
    6.9.2 Extending annotations ....................................... 38
  6.10 Marking the Arguments and Supplements for a connective ........ 38
  6.11 Discontinuous text selection ...................................... 39
  6.12 How to correct mistakes or remove annotations ................ 39
  6.13 Saving your annotations ......................................... 40
  6.14 Closing files ...................................................... 40
  6.15 Annotation procedure for Task A ................................... 40
  6.16 Annotation Procedure for Task B ................................. 41
  6.17 Other remarks and notes ......................................... 41
  6.18 Troubleshooting .................................................. 42
6.18.1 I am having launching WordFreak away from IRCS .......................... 42
6.18.2 I removed an annotation from the Chooser but the text continues to be highlighted in the WordFreak window ................................................. 42
6.18.3 The annotation nodes in the Chooser are jumbled up or doubled up ........ 42
6.18.4 I am having trouble saving files ................................................... 42
6.18.5 I get an error message when I try to add an annotation ...................... 43

7 Practice Session ................................................. 43
7.1 Text 1 .......................................................... 43
7.2 Text 2 .......................................................... 44

8 Appendix ......................................................... 45
8.1 Propositional Situations as Arguments of Connectives ............................. 45
8.2 Elements of a Clause ............................................. 46
8.2.1 Predicates .................................................. 47
8.2.2 Predicate Arguments ......................................... 47
8.2.3 Modifiers ..................................................... 47

1 Introduction

1.1 Discourse Relations and Discourse Connectives

An important task for several natural language applications is to identify the discourse relations that hold between parts of a text as well as the particular textual spans that are related via the discourse relation. Our goal in the Penn Discourse TreeBank Annotation Project (PDTB) is to mark such discourse relations and the text spans that they relate.

One of the ways that discourse relations are marked in a text is by the use of explicit markers, the discourse connectives, and the textual spans that they relate, their arguments. Discourse connectives are thus linguistic devices used by speakers to indicate how material in one clause, sentence or segment relates to that in another. Grammatically, they can be coordinators (and, but, or etc.), subordinators (because, although, etc.), or adverbials (consequently, however, etc.). In the discourse in (1), there are five discourse connectives that express relations between the contents of the four given sentences. On the one hand and on the other hand are discourse connectives that are used in parallel to express a relation of contrast between the interpretation of the first two sentences, (1a) and (1b), and the interpretation of the second two sentences, (1c) and (1d). The discourse connective So in (1b) establishes a relation of consequence between the interpretation of the first and second sentence. The first and the second sentence are therefore
the two arguments of the connective So in (1b). Similarly, So in (1d) expresses a consequence relation between the interpretation of (1c) and (1d), its two arguments. Finally, the connective then establishes a temporal sequence relation between the events described in (1b) and (1d). As this final case shows, the arguments of a discourse connective need not always be adjacent.

(1)  
   a. **On the one hand**, [John loves Barolo].
   b. So [he went and ordered three cases].
   c. **On the other hand**, [he didn’t have much money].
   d. So then [he had to cancel the order].

In some cases, discourse relations that hold between adjacent text spans must be inferred. For example, in the discourse in (2), a causal relation is inferred between the interpretation of the first and the second sentence although no explicit connective is used to express it. In such cases, we will identify the presence of an implicit connective between the first and the second sentence. The first and the second sentence are identified as the arguments of the implicit connective.

(2)  
   a. You should never lend any books to John.
   b. He never returns them.

The arguments of discourse connectives (explicit connectives as well as implicit connectives), are propositional situations. i.e., things such as events (3), states (4), and properties (5).

(3)  Pierre won the card game.
(4)  Blake has a broken leg.
(5)  Blake is more than 7 feet tall.

Propositional situations can also be embedded within other situations, as shown in the following examples, where the italics represent the embedded situation:

(6)  Blake knows that **Pierre won the card game**.
(7)  Pierre believes that **Blake is more than 7 feet tall**.

This brings us to the next point, namely, that arguments of discourse connectives or propositional situations can be realized in texts in a number of different ways, i.e., in different syntactic realizations, and an important part of the annotation task in PDTB is to locate connective arguments in their different realizations. In particular, arguments of discourse connectives can be single sentences, multiple sentences, or clauses and phrases that are less than a sentence. So, for example, while in (9) and (12), the connectives **as a result** and **instead** relate full sentences, this is not the case for the other examples. In (8), **although** relates two clauses within the same sentence; in (10), one argument of **then** is a relative clause that modifies
a noun, while the other argument is the sentence that syntactically contains both the connective as well as the relative clause argument (but which will be excluded from the annotation of the sentence-delimited argument);\(^1\) in (11), one (the left) argument of otherwise is only the first clause of the first sentence (the second clause being [stop], containing an implicit “you” as the subject); in (13), the left argument of So contains the first two sentences; and finally, (14) contains a noun phrase (or “nominalization”) as one of the arguments of After.\(^2\)

(8) [Mary went to the party] although [she was tired].
(9) [John was very happy]. As a result, [he smiled].
(10) [A person [who seeks adventure] might, then, try skydiving].
(11) [If the light is red], stop. Otherwise, [you can go on].
(12) [John won’t eat fruit]. Instead, [he eats candy bars and potato chips].
(13) [Bill smokes. He also drinks a lot]. So, [he died].
(14) After [the sale of the company], [all the board of directors retired to Hawaii].

The annotation is divided into two separate tasks, TASK A and TASK B, based on the above distinction between explicit and implicit connectives.

### 1.2 Corpus

The corpus consists of online texts from the Wall Street Journal. There are 2318 texts, each text corresponding to a single file. The 2500 texts are organized into 25 sections, numbered from 00 through 25. There are approximately 100 texts/files in each section. The total number of words in the entire corpus is 1011787.

### 1.3 Goals of PDTB

The goal of the Penn Discourse Treebank Project is to produce a large scale and reliably annotated corpus which will encode discourse relations associated with discourse connectives, including implicit connectives. The annotations of the Penn Discourse Treebank will be linked to the existing syntactic annotations of the Penn TreeBank and the soon to be available annotations of the PropBank which encode semantic information associated with the verbs of all the sentences included in the Penn Treebank corpus.

\(^1\) In other words, the two arguments of then are identified as (a) [A person might try skydiving] and (b) [who seeks adventure]. Note that the discourse argument in (a) includes neither the connective nor the other discourse argument in (b) even though both are contained in (a) syntactically.

\(^2\) We will discuss “nominalizations” in more detail later, in Section ???. But for now, note that Noun Phrases are arguments of connectives ONLY WHEN they denote propositional situations.
1.4 What are Explicit Connectives

Explicit connectives fall under three main categories:

- **Subordinate conjunctions**: Subordinate conjunctions include connectives like *because, so that, although, when, while*, etc., and introduce subordinate clauses which can be tensed or non-tensed. The arguments of such connectives tend to be found locally, either adjacent on either side of the connective or both following the connective. An example of this category, with the two possible positions of the arguments is shown in (15).

  (15) a. [John failed the exam] **because** [he was lazy].
  b. **Because** [he was lazy], [John failed the exam].

- **Coordinate conjunctions**: Coordinate conjunctions include connectives *and, but, and or* (for our purposes we will also include *so* in this category since certain uses of this connective are for expressing coordination). Like adverbial subordinate conjunctions, the arguments of coordinate conjunctions tend to be found locally and adjacent to the connective. Unlike subordinate conjunctions, however, it is not possible for both their arguments to follow the connective. An example is given in (16).

  (16) a. [John is very smart] **but** [he failed his exam].
  b. # **But** [he failed his exam], [John is very smart].

- **Adverbial connectives**: Some examples of adverbial connectives are *therefore, as a result, however, nevertheless, in addition, moreover, instead, then*, etc., and they may appear in initial, medial, or final position of a sentence. One of the two arguments of such connectives is identified in the clause that contains them. The other argument, however, may or may not be found in an adjacent textual span. For example, the arguments of *then* in (17) are both adjacent to the connective but, this may not always be the case, as we saw earlier in (1).

  (17) Mary finished the report. Then, she went home.

Note that the characteristics of the explicit connectives described above are given as descriptive tendencies and not as annotation guidelines. One important benefit of building an annotated corpus of discourse relations is precisely to further investigate the properties of discourse connectives and their arguments in naturally occurring text.

1.5 What are Implicit Connectives

We use the term *implicit connective* for convenience to describe relations which are *inferred* between two textual segments but which are not expressed with an explicit connective. In the annotation of the Penn
Discourse Treebank, implicit connectives and their arguments will be identified within the boundaries of a paragraph. Also, we will not identify implicit connectives within sentences. Discourse relations not explicitly marked may hold across paragraph boundaries and within the boundaries of a sentence but these will not be annotated in the initial phase of the project.

1.6 TASK A: Annotation of Explicit Connectives

Annotation Task A will proceed in several phases. In each phase, a “specific set of connectives” will be provided for annotation. Your task is to “search for and identify all instances of these connectives” in the corpus and “identify and mark their arguments”. All aspects of the annotation are handled with the discourse annotation tool **WordFreak**. Detailed instructions of how to use the tool and execute this annotation task are given in Section 6 of this tutorial.

The sets of connectives that have been isolated for annotation until the current phase of the project are as follows. Your supervisors will tell you which of these sets you should annotate.

1.6.1 Set 1 connectives

- **Adverbials**: Instead, Otherwise, Therefore, As a result, Nevertheless
- **Subordinate conjunctions**: Because, Although, Even though, When, So that

1.6.2 Set 2 connectives

- **Adverbials**: then, on the other hand, however, in fact, further/furthermore
- **Subordinate conjunctions**: while, if/only if/even if, since/ever since

1.6.3 Set 3 connectives

- **Adverbials**: indeed, for example
- **Subordinate conjunctions**: as soon as, as long as, unless, after, until
- **Coordinating conjunctions**: and, or, but

1.6.4 Set 4 connectives

- **Adverbials**: though, yet, so, on the contrary, conversely
- **Subordinate conjunctions**: whereas, as, insofar as, till
- **Coordinating conjunctions**: nor
1.6.5 Set 5 connectives

- **Adverbials**: consequently, besides, thus, nonetheless, afterwards, finally, by contrast, in sum
- **Subordinate conjunctions**: once, for

1.6.6 Set 6 connectives

- **Subordinate conjunctions**: before, lest, much as
- **Adverbials**: simultaneously, in addition, thereby, accordingly, thus, overall, in the meantime, meanwhile

1.6.7 Set 7 connectives

- **Adverbials**: thus, in other words, still, previously, as an alternative, specifically, in particular, hence
- **Subordinate conjunctions**: except, else

1.6.8 Set 8 connectives

- **Adverbials**: earlier, later, regardless, for instance, in the end, on the other side, by comparison, alternatively, in short
- **Subordinate conjunctions**: now that

1.6.9 Set 9 connectives

- **Adverbials**: rather, ultimately, moreover, likewise, next, similarly, in contrast, thereafter, by then, additionally

1.6.10 Set 10 connectives

- **Subordinating conjunctions**: as though
- **Adverbials**: also, on the whole, plus, as well, separately, in turn

1.7 TASK B: Annotation of Implicit Connectives

In this task, you are required to annotate only implicit connectives. You will be assigned a set of texts which you must annotate sequentially. Each text needs to read from beginning to end, and all all implicit connectives must be annotated on a sentence-by-sentence basis.
Implicit connectives, as we explained in earlier sections, are identified between adjacent sentences that are not related by an explicit connective. Your task, as you’re reading the text, is thus, to determine for every adjacent pair of sentences as you’re reading the text, whether or not an explicit connective relates them. If not, you are required to establish an implicit connective between those sentences and then suggest an explicit connective that could have been used to express the relation between the two sentences. (A special window in the annotation tool allows you to type in the name of an explicit connective that best expresses the inferred relation.) In (18), for example, an explicit connective that reflects the inferred interpretation of the relation between the two sentences would be because. You may encounter cases in which none of the expressions that we have identified as discourse connectives will successfully express the inferred relation. In such cases, you may type in any adverbial or other phrase that naturally links the two sentences. You may also be provided with a (non-exhaustive) set of connectives that you can use as a guide to think of appropriate connectives.

(18) [You should never lend John any books]. [He never returns them].

At least for the first phase of the PDTB annotation we will annotate only the relations between two sentences within a single paragraph (when the second sentence does not contain an explicit discourse connective). For matters of technical convenience, you will select the period at the end of the first sentence and mark that as the connective. For example, in (18) above, we select the period at the end of the first sentence and use it as a placeholder for an implicit connective.

2 Guidelines for Explicit Connectives

This section contains further guidelines and instructions for identifying and annotating explicit connectives. Guidelines are provided both for cases where you should not identify something as a connective as well as for cases where you should identify something as a connective.

Identifying a word or phrase as a connective is closely related to the identification of the type of each of the arguments that the word or phrase relates. The test is simple: If the word/phrase does not relate propositional situations, then it is NOT a discourse connective. In other words, BOTH the arguments of the word/phrase should project a propositional situation.

2.1 Words/phrases that look like connectives

2.1.1 Word/phrase has a single non-propositional argument

These should be ignored and not annotated as discourse connectives.

If the tool has highlighted these word/phrases, remove their annotation nodes (see Section 6.12).
• Example of **however** when it is not a discourse connective. It takes a single argument which is an adjective.

(19) But economic arguments, **however** [solid], won’t necessarily im press Hong Kong’s 5.7 million people.

• Example of **however** as a discourse connective relating two things that are both propositional situations:

(20) [Such problems will require considerable skill to resolve]. **However**. [neither Mr. Baum nor Mr. Harper has much international experience].

### 2.1.2 **Word/phrase has only a single propositional argument**

♠ These should be ignored and not annotated as discourse connectives.

These cases will be typically encountered in Task B, where you will not have a pre-determined set of connectives and where you will be required to identify connectives yourself as you’re reading the text.

• Such cases are often found with adverbials. Not all adverbials found within sentences function as discourse connectives. When deciding whether an adverbial is a connective or not ask yourself if they express a relation between two propositional situations. The adverbial **therefore** is a discourse connective because it expresses a relation of consequence between two situations that are both available in the text. For example, in (21) below, **therefore** expresses that the event situation in the interpretation of the second clause is a consequence of the event situation in the interpretation of the first clause.

(21) [John did not finish the report]. **Therefore**. [we will postpone today’s meeting].

In contrast, the adverbial **strangely** in (22) should not be identified as a discourse connective. What this adverbial does is to assert that the situation expressed in the clause it modifies belongs to the set of **strange** situations. The predicate **being strange** does not express a relation between two propositions. Any relation that you may infer between the two given sentences should be associated with an implicit connective.

(22) [John was very hungry]. **Strangely**. [he only ordered a fruit salad].

**TEST**: can you say the sentence in isolation and understand what it means (ignoring the absent meaning of the pronoun)? or do you reject it? For example, contrast the following: you come out of the building with a friend and remark, “Strangely, it’s sunny today” vs. “However, it’s sunny today”.

11
The use of “however” is quite obviously more irksome than “strangely”. With “strangely” you are of course contrasting ‘it being sunny’ with something else, but this something is not dependent on something that should have been uttered in the prior discourse! With “however”, on the other hand, your friend will probably look at you questioningly and then you’ll say: “I mean, the forecast last night said rain. However, its sunny today”.

2.1.3 Word/phrase has two arguments but one or both are not situations

♠ These should be ignored and not identified as discourse connectives

- Examples of **and** and **for example** relating two noun phrases.

(23) Dr. Talcott led a team of researchers from [the National Cancer Institute] **and** [the medical schools of Harvard University and Boston University].

(24) Dr. Talcott led a team of researchers from the National Cancer Institute and the medical schools of [Harvard University] **and** [Boston University].

(25) These mainly involved such areas as [materials] – [advanced soldering machines], **for example** – and medical developments derived from experimentation in space, such as artificial blood vessels.

2.1.4 Relative Pronouns that look like Connectives

Some word/phrases that look like connectives, such as **when**, are in fact Relative Pronouns and should not be identified as discourse connectives. Relative pronouns introduce relative clauses that modify noun phrases. As such, they are similar to the cases seen in Section 2.1 since one of their arguments, the modified noun phrase is not a situation.

♠ Relative pronouns should not be identified as discourse connectives.

If the tool has highlighted these relative pronouns, remove their annotation nodes (see Section 6.12). It is not always trivial to distinguish relative pronouns from discourse connectives. Ask for help when in doubt.

- Examples of **when** as a relative pronoun and not a discourse connective

(26) Atsushi Muramatsu, executive vice president and chief financial officer of Nissan, said, "The company has experienced a remarkable turnaround in terms of profitability since the fiscal year ending [March 1987], **when** [the sharp and rapid appreciation of the yen caused many difficulties].
(27) Equitable of Iowa Cos., Des Moines, had been seeking a buyer for the 36-store Younkers chain since [June], when [it announced its intention to free up capital to expand its insurance business].

(28) Georgia-Pacific’s sales climbed to $9.5 billion last year, compared with $6 billion in [1983], when [Mr. Hahn took the reins].

- Examples of when as a discourse connective

(29) When [it’s time for their biannual powwow], [the nation’s manufacturing titans typically jet off to the sunny confines of resort towns like Boca Raton and Hot Springs].

(30) While [many of the risks were anticipated] when [Minneapolis-based Cray Research first announced the spinoff in May], the strings it attached to the financing hadn’t been made public until yesterday.

2.2 Words/phrases as preposed predicate complements

Some connective-like words and phrases are actually predicate complements in direct or indirect questions but appear dislocated from the position where they would normally be in the declarative counterparts.

♣ These should not be annotated as connectives.

- For example, when is not a connective when it is a wh-phrase within indirect question complement of “say”:

(31) The maker of chemical and industrial materials didn’t say how much it would pay or when [it would make the transactions].

- Other examples of when as a wh-phrase:

(32) Following a problem-solving formula used by teams throughout Federal Express, members of the Natick team monitored their morning routine, carefully noting where and when [the work group’s resources were used effectively] and where they were idle, waiting for others upstream in the process to send packages their way.

(33) Mr. Driscoll didn’t elaborate about who the potential partners were or when [the talks were held].
2.3 Connectives post-modified with “of”

Some of the connectives may be followed by a preposition like of (e.g., instead of, as a result of, because of). The post-modifying preposition may be followed either by a noun phrase or by a non-tensed verb.

♠ These post-modified connectives should not be annotated.

If the tool has highlighted these connectives, remove their annotation nodes (see Section 6.12).

- Examples of connectives post-modified by the preposition of

(34) yields "may blip up again before they blip down" because of recent rises in short-term interest rates.

(35) In 1988, Kidder eked out a $46 million profit, mainly because of severe cost cutting.

(36) Instead of [teaming up], [GE Capital staffs and Kidder investment bankers have bickered].

2.4 Pre-modified connectives

Sometimes connectives can be pre-modified by adverbs like only, even etc...

♠ These connectives, unlike the post-modified connectives in Section ??, should be identified as discourse connectives, but in addition, the highlighted connective should be extended to include its modifiers.

Extension can be done in two ways: (a) delete the annotation node for the connective (Section 6.12) and make a new selection and annotation node including the modifiers (Section 6.9), or (b) extend the existing annotation to include the modifiers (Section 6.9.2).

- Examples of modified connectives:

(37) a. original connective: Workers described "clouds of blue dust" that hung over parts of the factory, even though exhaust fans ventilated the area.

b. extended connective: Workers described "clouds of blue dust" that hung over parts of the factory, even though [exhaust fans ventilated the area].

(38) a. original connective: On Wall Street men and women walk with great purpose, noticing one another only when they jostle for cabs.

b. extended connective: On Wall Street men and women walk with great purpose, noticing one another only when [they jostle for cabs].

(39) a. original connective: The 48-year-old Mr. Corr was hired largely because he was credited with returning Trans World Airlines Inc. to profitability while he was its president from 1986 to 1988.
b. **extended connective**: [The 48-year-old Mr. Corr was hired] *largely because* [he was credited with returning Trans World Airlines Inc. to profitability while he was its president from 1986 to 1988].

### 2.5 Co-occurring connectives

Sometimes, distinct discourse connectives can occur adjacent to each other in the same sentence/clause. ♠ These should be annotated as distinct connectives.

- The following example shows *and* and the modified connective *even when* occurring next to each other. There should thus be two distinct annotations for each of the two connectives:

  (40) *And even when* specific projects are described in prospectuses, the money often is used elsewhere, according to analysts.

### 2.6 Non-sentence/clause initial connectives

As noted before in Section 1, one of the arguments of a connective is the sentence or clause that contains it. Typically, connectives appear at the beginning of the clause or sentence, but many times, they may appear medially or finally.

♠ Connectives in all positions should be identified and annotated.

- Examples of connectives in clause-medial position:

  (41) In the last year or so, *however*, this exclusive club has taken in a host of flashy new members.

  (42) Its plan, *instead*, is to spin off the remainder of its real estate unit and to possibly do the same with its mining and energy assets.

- Examples of connectives in clause-final position:

  (43) But they will have disappointing returns or even losses if interest rates rise *instead*.

  (44) International competition for the few Soviet goods that can be sold on world markets is heating up, *however*.

### 3 Guidelines for Arguments of Explicit Connectives

Arguments of discourse connectives are identified when they denote propositional situations. As discussed in Section 1.1, situations are things like events, states and properties.
This section contains further guidelines and instructions for the annotation of discourse connective arguments, starting with a specification of the different syntactic realizations of propositional situations.

3.1 Types and Sizes of Arguments

Syntactically, propositional situations as arguments of connectives can be realized in different ways. Each type is discussed below.

3.1.1 Single Sentences

♠ The most easily identifiable case is when the argument of a connective is a single sentence.

We define sentence as the textual span that contains a main clause and all its associated subordinate (tensed and non-tensed) and coordinated clauses. Orthographically, sentence boundaries correspond with the period appearing at the end.

The examples below give examples of single sentences as arguments of some discourse connectives. All the arguments are single sentences. However, note that the sentences are simple, i.e., containing only a single clause (such as the left argument of in addition) as well as complex, i.e., containing a main clause together with complement clauses (such as the tensed complement clause of “said” in the left argument of but), or subordinate clauses (such as the tensed subordinate clauses introduced by the subordinator once in the right argument of in addition and the non-tensed subordinate clause introduced by the subordinator for in the left argument of however), or coordinated clauses (such as in the right argument of but).

(45) ['The company said it was delaying construction because of current market conditions]. But [the Memphis, Tenn., facility wasn't to begin turning out product until 1993, so the decision may reveal a more pessimistic long-term outlook as well].

(46) ['Some of the large amounts are being raised by small firms]. In addition, [once money is raised, investors usually have no way of knowing how it is spent].

(47) ['The U.S., claiming some success in its trade diplomacy, removed South Korea, Taiwan and Saudi Arabia from a list of countries it is closely watching for allegedly failing to honor U.S. patents, copyrights and other intellectual-property rights]. However, [five other countries – China, Thailand, India, Brazil and Mexico – will remain on that so-called priority watch list as a result of an interim review, U.S. Trade Representative Carla Hills announced].

3.1.2 Multiple Sentences

♠ An argument of a connective can be a sequence of sentences as defined above.

- For example, in (48) below, the first argument of the connective also is (48a), which which spans two sentences.

16
(48) a. [One diplomat described the tenor of Secretary of State Dean Rusk’s speeches as “inconclusive”. But he hastened to add that, if United States policies were not always clear, despite Mr. Rusk’s analysis of the various global danger points and setbacks for the West, this may merely mean the new administration has not yet firmly fixed its policy].

b. [A certain vagueness may also be caused by tactical appreciation of the fact that the present council meeting is a semipublic affair, with no fewer than six Soviet correspondents accredited].

Note that the connective itself occurs sentence-medially in the second argument but is excluded from the annotation so that the argument annotation is discontinuous. This is the convention followed for sentence-medial connectives - see Section 2.6 and Section ??.

3.1.3 Clauses

- A clause is a grammatical unit that contains at least a predicate (see below), typically a verb, and an explicit or implicit subject. See the Appendix (8.2) for more discussion of clauses and predicates. Most importantly, a clause expresses a propositional situation.

♠ An argument of a connective can be a single clause that may constitute a whole sentence (both arguments of 49) or part of a sentence (italicized arguments in 50-56). is less than a sentence.

(49) [Some current rates exceed those on CDs]. For instance, [a CD-type annuity from North American Co. for Life & Health Insurance, Chicago, offers 8.8% interest for one year or a 9% rate for two years].

(50) [Annuities are rarely a good idea at age 35 because of the withdrawal restrictions. But [at age 55, "they may be a great deal,"] says Mr. Wilson, the Columbia, S.C., planner.

(51) If [rates unexpectedly rise], [the increasing return on the money fund will partly offset the lower-than-expected return from the bond fund].

(52) Workers described "clouds of blue dust" [that hung over parts of the factory], even though [exhaust fans ventilated the area].

(53) Program traders are fond of predicting that if [they are blocked in the U.S.], [they will simply emigrate to foreign stock markets].

(54) Since [the preferred stock is cumulative], Amdura said [it will pay all omitted dividends], which range from 1.19 to 4.88 a share, when debt-reduction requirements have been met.

(55) Since the preferred stock is cumulative, Amdura said [it will pay all omitted dividends, which range from 1.19 to 4.88 a share, when [debt-reduction requirements have been met]].
Kevin Logan, chief economist with the Swiss Bank Corp., said [that both PPI and CPI climbed around 4 1/2% year-to-year in September]. He argued [that both CPI and PPI have] in fact [decelerated since spring].

- **Clauses can be non-tensed.** Non-tensed clauses also express propositions and they may contain an implicit subject. For example, in (57) the non-tensed clause *waving his hand* contains the predicate ‘wave’. One of its arguments is the direct object ‘his hand’ which is expressed explicitly. The other argument, the subject ‘John’ is expressed implicitly via a link to the subject of the main clause.

  (57)  John walked across the hall, (while) waving his hand cheerfully.

- Non-tensed clauses can also be arguments of connectives.
  - Examples of a non-tensed “subordinate” clause with implicit subject as connective argument (first argument in (58) and second argument in (59)):

    (58)  [Knowing a tasty – and free – meal] when [they eat one], the executives gave the chefs a standing ovation.

    (59)  [Margin rules determine the minimum amount of cash an investor must put up] when [buying a security].

  - Example of non-tensed “complement” clause with implicit subject as connective argument (first argument of instead):

    (60)

  (61)  Even when it became clear this spring that demand wasn’t coming up to expectations, Chrysler officials ”resisted” [cutting output] because Cherokee and Wagoneer are ”very profitable vehicles,” the spokesman said. Instead, [Chrysler officials in late May slapped $1,000 cash rebates on the vehicles, including the first such incentives on the popular four-door Cherokee since Chrysler bought Jeep in 1987].

  - Example of non-tensed “complement” clause with explicit subject as connective argument (first argument of because):

    (62)  Everybody considered [Einstein’s contribution to be a breakthrough] because [he discovered the theory of relativity].

- **Small Clauses.** In some cases, the verb of a clause can also remain unexpressed, and even such clauses can be arguments of connectives.

In the following example, the interpretation of the second clause argument of and is actually only the manager (can talk) to the general manager, with the verbal part in parentheses is the unexpressed portion of the clause.
(63) It’s probably true that many salaried put in unproductive overtime just for the sake of solidarity, that the system is so hierarchical that [only the assistant manager can talk to the manager] and [only the manager to the general manager], and that Sony was chary of letting a young, short-term American employee take on any responsibility.

- Multiple clauses.
  - Multiple clauses in sentences can also be arguments of connectives. In the following example, the first argument of nevertheless contains 5 clauses and the second argument contains 2 clauses. The verbs projecting the clauses are underlined in the example:

(61) [The average maturity for funds open only to institutions, considered by some to be a stronger indicator because those managers watch the market closely, reached a high point for the year – 33 days]. Nevertheless, said Brenda Malizia Negus, editor of Money Fund Report, [yields "may blip up again before they blip down" because of recent rises in short-term interest rates].

3.1.4 Nominalizations

A nominalization is a noun phrase that has a systematic syntactic correspondence with a clausal predication which includes a head noun morphologically related to a corresponding verb, and that projects a situation semantically. For example, refusal to obey is a nominalization since it corresponds to he refuses to obey. and since the head noun refusal is morphologically related to the verb refuse.

- Nominalizations can also be arguments of connectives.

In the example in (65), the sale of the company can be a connective’s argument because we can paraphrase the noun phrase as a clausal predication projecting a situation, such as the company was sold.

(65) After [the sale of the company], [all the board of directors retired to Hawaii].

When the situational interpretation and the syntactic paraphrase is not available as a morphological transformation of the head noun in the noun phrase to a verb in the paraphrased clause, a noun phrase CANNOT be a connective’s argument.

For example, in [John left] after [Mary], “Mary” cannot be identified as a discourse argument of a connective even though the sentence means John left after Mary (left). This is because the clausal paraphrase to Mary left is not available from the noun phrase head. Neither can “lunch” be a discourse connective’s argument in [Some people always sleep] after [lunch], since a possible paraphrase to they eat lunch cannot be made from the head noun.
3.2 Inclusion of all predicate arguments and modifiers within a connective’s argument

Discourse relations hold between propositions, the heart of which is the predicate and its arguments. For this reason,

♠ when annotating arguments of connectives you should include a predicate and make sure that you do not exclude its arguments and adverbial modifiers.

For example, the left argument of the connective because in (66) could be either Everybody considered Einstein’s contribution to be a breakthrough or just Einstein’s contribution to be a breakthrough, depending on your interpretation. In the first case, the verb projecting the situation is “considered” and its arguments are the subject of the sentence “Everybody” and the entire non-tensed clausal complement of the verb. Both of these arguments must be included in your argument annotation if your situational interpretation for the because leads you to pick the “consider” verb. In the second case, the verb projecting the situation is “to be” and its arguments are the subject of the complement clause “Einstein’s contribution” and the object “a breakthrough”. Again, both these arguments must be included if “being a breakthrough” is the predicate projecting your situational interpretation for the because relation.

(66) [Everybody considered Einstein’s contribution to be a breakthrough] because [he discovered the theory of relativity].

Similarly, in the following example, the prepositional modifier “In Geneva” must be included in the selection of ARG2 for however.

(67) wrong: However, in Geneva, [they supported Iran’s proposal]

(68) correct: However, [in Geneva, they supported Iran’s proposal]

All connectives that are associated with the clause selected as argument (except for the connective that is itself being annotated) must also be included in the selection of the argument, as shown in the following examples (connective to be included in argument in italics):

(69) For subordinating conjunctions

a. wrong: The group says standardized achievement test scores are greatly inflated because [teachers often “teach the test” as Mrs. Yeargin did], although [most are never caught].

b. correct: The group says standardized achievement test scores are greatly inflated [because teachers often “teach the test” as Mrs. Yeargin did], although [most are never caught].

(70) For adverbials:

a. wrong: [The offer is in limbo], however, because [LIN has agreed to merge its cellular-phone businesses with BellSouth Corp].

b. correct: [The offer is in limbo, however], because [LIN has agreed to merge its cellular-phone businesses with BellSouth Corp].
(71) For coordinating conjunctions:
   a. **wrong**: *But* [industry and OPEC officials agree that a handful of members still have enough unused capacity to glut the market and cause an oil-price collapse a few months from now if OPEC doesn’t soon adopt a new quota system to corral its chronic cheaters]. **As a result**, [the effort by some oil ministers to get OPEC to approve a new permanent production-sharing agreement next month is taking on increasing urgency].
   b. **correct**: [*But* industry and OPEC officials agree that a handful of members still have enough unused capacity to glut the market and cause an oil-price collapse a few months from now if OPEC doesn’t soon adopt a new quota system to corral its chronic cheaters]. **As a result**, [the effort by some oil ministers to get OPEC to approve a new permanent production-sharing agreement next month is taking on increasing urgency].

3.3 Arguments and supplements: the minimality principle

In identifying arguments of discourse connectives, you need to select a situation or situations (or rather the linguistic expressions of the situations) that is *minimally necessary* for the interpretation of the discourse relation established by the discourse connective.

In many cases, you might be undecided about how much to include in your argument. For example, it may be the case that you are able to identify what is minimally necessary for the relation but you may feel that there is some associated text that “though not crucial for the relation, is nevertheless quite closely related to the main argument and contributes some relevant, additional information that may support your selection of the argument in some way”.

---

♠ If you want to select something that is more than minimally necessary, mark it as a supplement. The tool provides annotation holders for both arguments (ARG) and supplements (SUP).

♠ You don’t need to necessarily mark everything that is a supplement. It depends very much on whether you perceive the related part as contributing to the discourse relation. Relatedly, you don’t have to fill the SUP slot for a connective. It is fine to just mark the arguments of a connective and leave the SUP slot empty.

The examples below illustrate some reasoning behind some types of supplementary annotations that will give you an idea of how to think of the indirect, supportive relations between arguments and supplements.

- Chain reasoning:

  (72) *Foreigners complain* [that they have limited access to government procurement in Japan], **in part because** [Japanese companies unfairly undercut them].

For the above, you may want to include the embedding verb in ARG1 because there is a causal chain that includes the embedding verb, the lower verb and the predicate in ARG2. The interpretation is
“the reason why foreigners have limited access... is because the Japanese companies undercut them and that the whole thing is the reason why foreigners complain.” In other words, it seems that an inferred discourse relation holds between the embedding verb and its complement.

For these cases, where you may feel that there is a reasoning chain involving the embedding verb at one end of it, you should annotate as arguments only the clauses local to the connective (i.e., the lower end of the reasoning chain). For example, you would annotate the complement clause of “complain” as ARG1 and the clause introduced by in part because as ARG2.

The idea that you must keep in mind for this is there is an independent discourse relation that can be established between the embedding verb and the subsequent discourse, including the complement clause.

If you still feel that you would like to annotate the embedding relation, mark the main clause part as the SUP.

- Subjective reasoning:

Similar to the above chain reasoning example, you may want to include the embedding clause for the verb “said” in ARG1 below because its exclusion would make a statement of causality too strong between the marked arguments, appearing factual when in fact it was (possibly intended to be presented as) subjective.

(73) Some dealers said [the dollar was pressured slightly] because [a number of market participants had boosted their expectations in the past day and were looking for an index above 50, which indicates an expanding manufacturing economy].

In particular, you might disagree with showing the causal relationship as holding between “the dollar was pressured slightly” and “a number of market participants ... economy” because this might lead to the relationship being interpreted as a fact when that is not necessarily the case.

Again, for such cases, annotate only the arguments of the causal relationship, i.e., excluding the main verb clause. The subjective or objective nature of the expressed causal relationship will be derived by some other interpretation component.

Subjectivity here is not associated with the semantics of the connectives and therefore it should not affect the selection of its arguments. In other words, some other interpretation component should take care of the fact that the causal relationship is presented subjectively.

- Other examples of argument-supplement annotation (supplement in italics):

(74) [Since the reforms went in place, for example], [no state has posted a higher rate of improvement on the Scholastic Aptitude Test than South Carolina], although [the state still posts the lowest average score of the about 21 states who use the SAT as the primary college entrance examination].
(75) [Lawyers and their clients who frequently bring business to a country courthouse can expect to appear before the same judge year after year]. [Fear of alienating that judge is pervasive], says Maurice Geiger, founder and director of the Rural Justice Center in Montpelier, Vt., a public interest group that researches rural justice issues. As a result, says Mr. Geiger, [lawyers think twice before appealing a judge's ruling, are reluctant to mount, or even support, challenges against him for re-election and are usually loath to file complaints that might impugn a judge's integrity].

3.4 What kinds of things can be Supplements

Even though the argument-supplement distinction is available to you, it is important to note that not anything can be a supplement. Below, we describe and illustrate the types of text elements that are allowed to be marked as supplements.

3.4.1 Clauses can be supplements

♦ You can mark all clausal segments (and therefore, also, single or multiple sentences) as supplements.

- Examples of some clausal segments as supplements (supplements in italics):
  
  - participial clauses:

    (76) [Ringers often skip off to the local pub], [leaving worship for others below] when [their changes are completed]

  - clauses as quoted speech:

    (77) [stations are fuming] because [the show's distributor, Viacom Inc., is giving an ultimatum]: [Either sign new long-term commitments to buy future episodes or risk losing "Cosby" to a competitor].

3.4.2 Clause-internal arguments and modifiers cannot be supplements

♦ You cannot mark anything within the "minimal" clause as a supplement. Therefore, if you select a predicate of a clause for projecting your situational interpretation for the discourse relation, then modifiers of that predicate or modifiers of the predicate's arguments, as well as arguments of the predicate cannot be marked as supplements.

- Examples of text segments that cannot be supplements (incorrectly marked supplements in italics):

  - Prepositional phrase modifiers in sentences
(78) wrong: [For 10 yrs], [Genie Driskill went to her neighborhood bank] because [it was convenient].

(79) correct: [For 10 yrs, Genie Driskill went to her neighborhood bank] because [it was convenient],

- Noun phrase supplements in sentences
- Adverbial modifiers of sentences

3.4.3 Parentheticals can be supplements

parentheticals are elements that are neither a complement nor a modifier of any other element within a sentence.

♠ Parentheticals can be marked as supplements.

- Example of parenthetical as supplement (supplement in italics):

(80) [The machine-gun-toting guards were gone] – [for the first time in five months] – when [he arrived at the embassy two days later].

3.5 Numbering of arguments for subordinate conjunctions

As noted in Section 1.4, the arguments of subordinate conjunctions, i.e., the main clause and the subordinate clause, can be either on either side of the connective or both following the connective.

♠ Always mark as ARG2 the subordinate clause (and anything else associated with it) introduced by the subordinate conjunction and as ARG1 the main clause (and anything else associated with it) independent of the order of the clauses.

- Example of Argument marking for because when the clauses are on either side of the connective:

(81) [The federal government suspended sales of U.S. savings bonds] because [Congress hasn’t lifted the ceiling on government debt].

   ARG1 = main clause = The federal government suspended sales of U.S. savings bonds
   ARG2 = subordinate clause = Congress hasn’t lifted the ceiling on government debt

- Example of Argument marking for because when the clauses are both on the same side of the connective:

(82) Because [the drought reduced U.S. stockpiles], [they have more than enough storage space for their new crop].

   ARG1 = main clause = they have more than enough storage space for their new crop
   ARG2 = subordinate clause = the drought reduced U.S. stockpiles
3.6 Head Nouns in Relative Clause arguments

♦ Exclude the head noun of the relative clause if the relative clause is the argument of some connective

The apparently missing subject in the currently annotated relative clause below is in fact represented as an empty subject in the syntactic annotation and is also coindexed with the head noun. As a result, we will be able to retrieve the missing subject in the relative clause.

- Examples showing wrong and correct annotation for relative clause head nouns (head noun in italics):

  (83) a. **wrong**: Workers described [*clouds of blue dust*] that hung over parts of the factory, **even though** [exhaust fans ventilated the area].

  b. **correct**: Workers described [*clouds of blue dust*] [that hung over parts of the factory], **even though** [exhaust fans ventilated the area].

3.7 Exclusion of connectives from their own argument annotations

♦ When annotating the arguments of some connective, do not include the connective itself in any of its own argument annotations.

If the connective is clause-medial, do discontinuous annotation by selecting the separated parts of the text with the tool. WordFreak allows you to select distinct non-adjacent pieces of text and then glue them together. See Sections 3.8 and 6.11 for more guidelines, examples, and tool instructions on such discontinuous annotation.

- The following examples show the wrong and correct annotations with respect to the inclusion of connectives in their arguments:

  (84) a. **wrong**: [*Although* the purchasing managers’ index continues to indicate a slowing economy], [*it isn’t signaling an imminent recession*], said Robert Bretz, chairman of the association’s survey committee and director of materials management at Pitney Bowes Inc., Stamford, Conn.

     b. **correct**: **Although** the purchasing managers’ index continues to indicate a slowing economy], [*it isn’t signaling an imminent recession*], said Robert Bretz, chairman of the association’s survey committee and director of materials management at Pitney Bowes Inc., Stamford, Conn.

  (85) a. **wrong**: The average maturity for funds open only to institutions, considered by some [to be a stronger indicator] **because** those managers watch the market closely], reached a high point for the year – 33 days.
b. **correct**: The average maturity for funds open only to institutions, considered by some to be a stronger indicator because those managers watch the market closely, reached a high point for the year – 33 days.

(86) a. **wrong**: But [Santa Fe, currently trading at 18 7/8, isn’t likely to realize private market values by selling assets], because the tax against it would be onerous. [Its plan, instead, is to spin off the remainder of its real estate unit and to possibly do the same with its mining and energy assets].

b. **correct**: But [Santa Fe, currently trading at 18 7/8, isn’t likely to realize private market values by selling assets], because the tax against it would be onerous. [Its plan, instead, is to spin off the remainder of its real estate unit and to possibly do the same with its mining and energy assets].

(87) a. **wrong**: That makes them a reasonable option for investors who will accept some risk of price fluctuation in order to make a bet [that interest rates will decline over the next year or so]. Buyers can look forward to double-digit annual returns if they are right. But they will have disappointing returns or even losses if [interest rates rise instead].

b. **wrong**: That makes them a reasonable option for investors who will accept some risk of price fluctuation in order to make a bet [that interest rates will decline over the next year or so]. Buyers can look forward to double-digit annual returns if they are right. But they will have disappointing returns or even losses if [interest rates rise] instead.

3.8 Discontinuous annotation

In many cases, you might find that the text you want to select for annotating an argument is discontinuous. Wordfreak allows you to mark discontinuous (or non-adjacent) parts of the text as belonging to a single argument. See Section 6.11 for instructions for doing discontinuous annotation.

- Examples showing discontinuous annotation:

(88) [Even so], according to Mr. Salmore, [the ad was ”devastating”] because [it raised questions about Mr. Courtier’s credibility].

(89) But [Santa Fe, currently trading at 18 7/8, isn’t likely to realize private market values by selling assets], because the tax against it would be onerous. [Its plan, instead, is to spin off the remainder of its real estate unit and to possibly do the same with its mining and energy assets].

(90)
(91) [The funds’ share prices tend to swing more than the broader market]. [When the stock market dropped nearly 7% Oct. 13], for instance, [the Mexico Fund plunged about 18% and the Spain Fund fell 16%].

3.9 Discourse Deixis

Certain expressions in language, called discourse deictics can be used to refer to events. For example, in John arrived early today. THAT surprised us, “THAT” is a discourse deictic that refers to the event of John arriving early. As the following example shows, a discourse connective can relate to the event referred to by discourse deictics.

For discourse deictic referring events as arguments of connectives, select the discourse dectic expression (together with the verbs that occur with it) as the argument instead of selecting the expression denoting the event to which the discourse deictic refers.

- Therefore, selecting that’s is sufficient for the first argument of because in the following example, even though the event to which that refers is in the previous sentence:

(92) Airline stocks typically sell at a discount of about one-third to the stock market’s price-earnings ratio – which is currently about 13 times earnings. [That’s] because [airline earnings, like those of auto makers, have been subject to the cyclical ups-and-downs of the economy].

3.10 Ambiguities

In some cases, arguments are truly ambiguous. You must use your best judgement for these. There are no guidelines for such cases.

- Examples of ambiguities:

(93) a. [The group says standardized achievement test scores are greatly inflated] because [teachers often ”teach the test” as Mrs. Yeargin did].

b. [The group says standardized achievement test scores are greatly inflated] because [teachers often ”teach the test”] as Mrs. Yeargin did.

(94) a. [Workers described ”clouds of blue dust” that hung over parts of the factory], even though [exhaust fans ventilated the area].

b. Workers described ”clouds of blue dust” [that hung over parts of the factory], even though [exhaust fans ventilated the area].
3.11 Conventions for relative pronouns

♠ When you select the content of a relative clause as an argument of some connective, include the relative pronoun if the relative clause contains a relative pronoun, such as “who”, “which”, “where”, “that”:

- Example showing wrong and correct annotation w.r.t. relative pronoun inclusion (relative pronouns in italics):

  (95)  a. wrong: John talked to Mary, who [was unhappy] because [she had to leave so soon].
  b. correct: John talked to Mary, [who was unhappy] because [she had to leave so soon].

3.12 Conventions for punctuation

♠ Exclude all end-punctuations from annotation of arguments as well as connectives.

- Examples:

  (96) Her recent report classifies the stock as a “hold”. But it appears to be the sort of hold [one makes] while [heading for the door].

  (97) Meanwhile, [stations are fuming] because, many of them say, [the show’s distributor, Viacom Inc., is giving an ultimatum: Either sign new long-term commitments to buy future episodes or risk losing ”Cosby” to a competitor].

- Since the punctuation marks are often attached to words, you may sometimes have a problem leaving them out of your annotation selection in the tool. See the tool instructions in Sections 6.9.1 and 6.9.2 for handling this problem.

3.13 Conventions for complementizers

♠ When you select the content of a complement clause as an argument of some connective, include the complementizer, i.e., “that” in the argument annotation

- Example showing wrong and correct annotation w.r.t. complementizer inclusion (complementizer in italics):

  (98)  a. wrong: John said that [Jane didn’t finish the project] because [she was sick all week].
  b. correct: John said [that Jane didn’t finish the project] because [she was sick all week].
4 Guidelines for Particular Connectives

4.1 Coordinating conjunctions in VP coordinations

♠ If you are annotating coordinating conjunctions such as and, but, or etc., do not annotate the connectives if there arguments are VP coordinates.

Even though the arguments of such coordinating conjunctions are propositional situations as we discussed above, we will not annotate these cases since the complete annotation can be automatically retrieved from the syntactic annotations.

- The example below shows and in a VP coordination which SHOULD NOT be annotated. The bracketing shows that the arguments of and are the two VP coordinates connected by and and which share the same subject, Mary.

(99) **wrong:** Mary [finished her food] **and** [left].

- Note that you are still supposed to annotate coordinating conjunctions when the subjects of both coordinated clauses are explicitly expressed, as below:

(100) **correct:** [Mary finished her food] **and** [she left].

4.2 As

*As* has multiple functions.

♠ Annotate *as* as a discourse connective only if it expresses a temporal or causal interpretation.

- Examples of *as* as a discourse connective to be annotated since it expresses a temporal (101) and causal (102) relation:

(101) [He tripped over a bunch of plugs] **as** [he was leaving the room].
(102) [W.U. had major losses] **as** [its telex business faltered in the face of competition from facsimile machines].

- Examples of *as* not to be annotated since it does not express either a temporal or causal relation:

(103) **As** she puts it, there's no hope.
(104) We do as we are told, **as** is the rule.
4.3 So

So expresses a consequence relation. However, it’s not always easy to identify the consequence, i.e., ARG2. ♠ If you find it hard to identify the ARG2 consequence argument of so, talk to your supervisors and make a comment in the comment box.

- Examples of so where ARG2 consequence argument is not easy to identify:

  (105) She flunked the exam.
  So, what is next?

  (106) You said she didn’t work hard.
  So, if you believe this, you must be right.

4.4 Nor

Nor can be found by itself:

(107) [This has nothing to do with you]. Nor [will it ever].

or as a multi-word connective, as neither...nor:

♠ Sometimes, it may be hard to identify the ARG1 of nor. In this case, leave the ARG1 slot empty and make a comment in the comment box: “ARG1 hard and left out”, or something similar.

♠ Do not annotate neither...nor in a VP coordination:

(108) In doing so, he neither rejected a socialist planned economy nor embraced the free market.

5 Guidelines for Implicit Connectives

5.1 Naming implicit connectives

- Naming connectives should be as ;;CONN=some_connective. for anything else that needs to be added end named connective with ;; and then continue. SO ;; marks a field. first field can ONLY contain named connective and nothing else.

6 Instructions for WordFreak

The annotation tool used in this project is WordFreak, developed by Tom Morton and Jeremy Lacivita at the University of Pennsylvania.
6.1 Tool Interface

Figure 1 presents a snapshot of the tool interface.

![WordFreak Window](image1)
![Chooser Window](image2)

Figure 1: Snapshot of WordFreak Interface

The WordFreak tool has two distinct windows:

- The **WordFreak** window (Figure 1(a)) allows you to open, load, view, select and remove texts for annotation. There are three views of the files that you can use in this window: **Untitled Project**, **Text**, and **Tree**. Figure 1(a) shows the view in **Untitled Project**: here, you can open, load and remove files from the tool’s memory. In the figure, one file has been opened into the tool (but not yet loaded - see Section 6.3). With the **Text** view, you can see the text in the file after it has been loaded into memory. And with the **Tree** view, you can see a different representation of the text as a tree structure (see Section 6.5 for more on this).

When the tool launches, the **Untitled Project** view comes up as the default view. If the **Text** and **Tree** view buttons are not visible, you can set these views by checking them from the **Viewer** pull-down menu, as shown in Figure 2:
The **Chooser** window (Figure 1(b)) allows you to add connectives and arguments in the annotation files. You can view each individual annotation in this window and make changes to existing annotations or delete annotations. Figure 1(b) shows the chooser window without any annotations. As you start annotating a file, your annotations will start to appear in this window. More on this in Section 6.6.

The remaining sections discuss each feature and task in detail. If you have any further questions, please don’t hesitate to ask your supervisors.

### 6.2 Launching the Tool

- To launch the tool, type in the following URL in your internet browser (you can also make an internet shortcut to this on your desktop to avoid typing in the url each time you use the tool):
  

- This will bring up a webpage that will ask you for a username. Select a username and click on the **Launch** button. USE THE SAME USERNAME EVERYTIME you use the tool.

- If you are using the tool for the first time on your PC, this will start the tool, but it will first give security warnings and ask if you still want to start the tool. Ignore these warnings and click **Start**.

- As you progress in your annotation, you might be allowed to work away from IRCS, say, at home or in your home department. For this, you will need the following to use the tool.

32
– The machine must have JAVA 1.4.0 or any later version. Check to see which version you have on the machine you’re going to use. You will need to install it if you do not have the correct version of JAVA. You can install it yourself from http://java.sun.com or ask your system administrator if you do not have permissions.
(If the machine has some old version, then it may not come with the WebStart application which is what you need to launch the tool)
The above is for Windows OS. if you’re on a MAC, then you need Mac OS X with a 10+ version.
As noted above, you need JAVA Webstart to launch the tool, and this comes only with the Mac OS X 10+ versions.

6.3 Opening and Loading Files for Annotation

The default view of the tool when you launch it is the Untitled Project view. This is where you open, load, save and close files.

• Open the file you will annotate by clicking the Add button at the top of the vertically arranged buttons on the right of the window.
This will bring up another window for browsing the machine to find the file to be opened. Go to the folder where the file is and click Open.

– If you’re opening the file into the tool for the first time, it will ask whether you want to create an annotation file - click yes. If you open a file for which you previously created an annotation file, it will not ask this again since there is already an annotation file for it.
– If you want to undo all your annotations for a particular file and start all over again, the best way to do this is to go to the folder where the files are (but not from within the tool - do this independently from your desktop) and delete the annotation file. Annotation files have a .ann extension. DO NOT delete the text file. (Text files have no extensions.)

• The selected filename (as well as the filename for the annotation file) will appear in the main white window at the top of the Wordfreak window. The result of opening a file this way is shown in Figure 1(a).

• Click the Load button on the right. Doing this loads the text into the tool memory (this takes a few seconds so wait for this. Watch the clock on the file - it needs to turn to a green check mark (see Figure 3).

6.4 Text View

• Loading a file puts the text into the tool’s memory and you can now view the text and start your annotations. Click on the Text button in the Wordfreak window which will show you the text you’re
going to annotate. If there is no Text button in your window, select Text from the Viewer pull-down menu (Figure 2). When the button shows up, click on it to view the text. Figure 4 shows the text view for a loaded file. Note that this text contains no annotations (the Chooser window is empty).

Figure 4: Text View of Loaded File
6.5 Tree View

In the Tree view (if you don’t see this button, select it by checking Tree from the Viewer pull-down menu - Figure 2), each token in the text is displayed on a single line. A token is a *single space*, a *single word*, or a *single punctuation mark*.

The Tree view can be used if you have trouble selecting some part of the text in the Text view. More on this in Section 6.7.

6.6 Annotations in the Chooser

- Loading a file also opens up the Chooser window. This is the window in which the annotation is done. If you’re loading a file for the first time, the chooser window will show up with no annotations, as in Figure 4.

- As you start doing your annotations for connectives and their arguments, annotation nodes will appear in the Chooser window for each connective.

  The anchor of each node is the connective called *CONN* and has the following four sub-nodes associated with it: *ARG1* for ARGUMENT 1, *ARG2* for ARGUMENT 2, *SUP1* for SUPPLEMENT 1, and *SUP2* for SUPPLEMENT 2.

- If you’re opening a file on which you did some annotations previously, the chooser window will show up with the earlier annotations. Figure 5 shows existing annotations in a file.

- The buttons at the top of the Chooser window perform the following actions:

  - < : move to previous annotation.
  - > : move to next annotation.
  - + : Add annotation (THIS BUTTON DOES NOT WORK).
  - - : remove annotation.
  - < | : Grow the annotation on the left by one token (a token is a space, word, or a single punctuation mark).
  - > | : Shrink the annotation on the left by one token.
  - | < : Shrink the annotation on the right by one token.
  - | > : Grow the annotation on the right by one token.

  **Connective** : Create an annotation node for a connective

  **Argument** : Create an annotation node for an argument (DOES NOT NEED TO BE USED)

  **Supplement** : Create an annotation node for a supplement (DOES NOT NEED TO BE USED)
The Comment slot at the bottom of the Chooser is where you can put in (a) your comments and (b) add names for implicit connectives for Task B.

- Each distinct entry in this slot must be enclosed around three semicolons (“;;;”). The following is an example of an entry in the comment slot:
  
  comment: ||;;CONN=because;;;not sure about this;;; |

- In the comments, DO NOT use double quotes!! This is very important.

6.7 Viewing and selecting Annotations

- Annotations in the Wordfreak window: are seen as colored text. The following color scheme is used for the different annotation types:
  - Discourse Connectives: RED
  - Arguments: BLUE
  - Supplements: GREEN
• **Annotations in the Chooser window:** are seen as nodes, which as described in the previous section are anchored on the **CONN** button that is the connective marker.

• **Scrolling through or Selecting Annotations:** can be done in both windows in any of the following ways:
  
  – Click at the top of any window and use the **UP** and **DOWN** arrow keys on your keyboard.
  
  – Click at the top of the Chooser window and use the left (“<”) and right (“>”) arrow keys at the top of the window.
  
  – Click directly on the annotated text in either window.

• **Highlighting of selected annotations:** When any annotation is selected, it gets highlighted in both windows, but in different ways:
  
  – In the Chooser window, only the selected annotation gets highlighted in **BLUE**. For example, if you select the **ARG1** annotation, only the **ARG1** text will be highlighted in the Chooser. The text for the annotation as well as the corresponding button at the top of the window is highlighted. This can be seen in Figure 5 with the **ARG1** text and the **Argument** button both highlighted.

  – In the Wordfreak window, the entire annotation node gets highlighted (i.e., the connective anchor as well as all its associated nodes) with the selected annotation in **BLUE** and the rest of the annotations in the node in **YELLOW**.

• **WHAT YOU MUST NOT DO!!!!** : When clicking on any annotation in the **Chooser** window, **do not** click on the annotation buttons, only on the text! If you click on the button, it may rewrite or delete the existing annotations without any warnings.

6.8 **Adding Annotation nodes for connectives by searching with the tool**

• Search for a connective in the text in the following way:
  
  – Pull down the **Annotation** menu at the top of the WordFreak window and select **RegEx Annotation**.... This will open a small window with two slots:

    * In the **RegEx** slot, type the name of the connective. You can **search** for multiple connectives at the same time by using “|” symbol to separate the different forms. So, for example, if you’re searching for **but** and **and** as the connectives, type “**but|and**”.

    * In the **Type** slot, type **"CONN"**.

    * If the **Match case** option is checked, **uncheck** it.

  – Click **ok**.
– After you do the above, all instances of the connective(s) that were found in the text will appear in the **Chooser** window as annotation nodes, in the order that they appear in the text. In the **WordFreak** window, all the instances of the connective(s) will be colored red in the text.

6.9 Adding annotation nodes for connectives manually

- In some cases, as for the explicit and implicit connective annotation in TASK B, you may need to annotate a connective by highlighting it in the text yourself. After you’ve highlighted the word or phrase corresponding to the connective, click on the **Connective** button in the Chooser. This will add an annotation node for that connective.

6.9.1 Using Tree View for text selection

- Sometimes the tool may give you problems when you’re trying to select the word or phrase you want. It may either forcibly extend to something more than what you want or less. In such cases, you can go to the **Tree** view in the Wordfreak window and select the text from there. In the Tree view, you can select a single token for annotation by clicking on it or multiple tokens by first clicking on the leftmost (uppermost in the Tree view) token and then clicking on the other tokens one by one while holding down the **CONTROL** key. After you have completed highlighting the phrase you want, go back to the Text view, verify that the phrase has been correctly selected, and then add the annotation in the Chooser.

The problem with selecting connectives manually occurs especially with selection of the **PERIOD** for implicit connectives, since end-periods are often attached to the last word of a sentence and the tool may not work properly in separating the word and punctuation tokens.

6.9.2 Extending annotations

- If some part of the text is highlighted and if you want to extend to reduce the selected span, you can also use the **Grow right**, **Grow left**, **Shrink right** and **Shrink left** buttons in the Chooser (see Section 6.6 for a description of these features). You can do this by selecting as much as you can with the text you want, and then use the buttons to shrink the span if the tool selected more than you wanted, or to extend the span if the tool selected less than you wanted.

6.10 Marking the Arguments and Supplements for a connective

- After you have selected some instance of a connective to annotate and want to mark its arguments and/or supplements, select each argument/supplement by dragging the mouse across the text you want. Always do this selection left-to-right on the text span. Then click the **ARG1**, **ARG2**, **SUP1**, or the **SUP2** button in the **Chooser** window depending on what it is you’re marking. Be careful
to add the ARG and SUP annotations to the correct connective. The selected connective can be seen highlighted in blue in the Chooser.

6.11 Discontinuous text selection

In WordFreak, you can do annotation of discontinuous text for connectives (CONN), arguments (ARG) as well as supplements (SUP).

- To do this, hold down the CONTROL key, and make as many selections as you like in a left-to-right manner while holding down the CONTROL key. After you are done selecting all the discontinuous parts of the annotation, release the control key and click on the appropriate button in the Chooser window (CONN, ARG, or SUP) to record your annotation.

- Check your annotation in the Chooser window to make sure that all the parts of your annotation are appended together and appear in the correct order.

For example, if you wanted to select John and is very tired below as the discontinuous parts of an ARG or SUP annotation, you would hold down the CONTROL key, select John, then select is very tired, release the CONTROL key, and click on ARG/SUP in the Chooser. The ARG annotation in the Chooser should show as John is very tired.

(109)  John, however, is very tired. (ARG or SUP annotation)

Similarly, if you wanted to select On the one hand and On the other hand below as the discontinuous parts of a CONN annotation, you would hold down the CONTROL key, select On the one hand, then select On the other hand, release the CONTROL key, and click on CONN in the Chooser. The CONN annotation in the Chooser should show as On the one hand On the other hand.

(110)  On the one hand, John wants to go sailing. On the other hand, he cannot swim. (CONN annotation)

6.12 How to correct mistakes or remove annotations

- If you want to redo the annotation for some connective, select the argument that you want to redo in the Chooser and click on the “−” button at the top. This will remove the argument from the Chooser as well as from the text (highlighting will disappear). Redo the annotation as you did before.

- if you want to remove an entire annotation node, you can do the same as above, but FIRST remove the ARG and SUP annotations and ONLY THEN remove the CONN annotation.
6.13 Saving your annotations

- Pull down the File menu at the top of the WordFreak window and select **Save -> Save all.**
  The **Save-all** option is the ONLY OPTION YOU HAVE. YOU CANNOT USE ANY OTHER WAY OF SAVING. For example, DO not click on **Save Project**, DO NOT save files by clicking on them individually, and DO NOT save files by clicking on REMOVE and waiting for the SAVE prompt to come up. All these ways will lead to a complete loss of your annotations.

- **Save frequently!**

- **Save all** the files that you open for annotation. Even those that did not contain any connectives.

6.14 Closing files

- Have you saved your file? Continue only after saving (Section ??).

- Go to the **Untitled Project** view in the WordFreak window.

- Click **Remove** from the vertically arranged buttons on the right.

6.15 Annotation procedure for Task A

Wordfreak allows you to search for specific connectives in the text that is loaded into the tool. For each instance of the connective in the text, it creates an annotation node for it in the Chooser (see Section 6.8). In this task, you will have been assigned a set of connectives (the sets of connectives are given in Section 1.6) and a set of text files. In this task, you are required to annotate all the connectives in your set in all the text files. Be sure to do one connective at a time. That is, pick a connective, identify and annotate it for all files, then move on to the next connective and annotate it for all files again, and so on. Thus, for the connectives in your set, you will be reopening each file as many times as the number of connectives in your set.

The procedure to follow for Task A annotation using the tool is as follows. See the references for further details on each action.

1. **Launch the tool** - Section 6.2

2. **For each connective from the set given to you:**

   (a) **For each file from the set given to you** (proceed through the files in sequential order):
      i. **Open and Load the file** - Section 6.3
      ii. **Search for instances of connective in text** - Section 6.8
iii. For each connective instance,
   A. Identify the connective. That is, determine whether the highlighted word/phrase is a discourse connective or not. If it is NOT a discourse connective, remove its annotation from the Chooser (Section 6.12). If it IS a discourse connective, then,
   B. Identify and Mark the arguments of the connective and the supplements of the arguments - (Section 6.10).

6.16 Annotation Procedure for Task B

WordFreak allows you to select connectives manually in the text and create annotation nodes for it in the Chooser (see Section 6.9.

In this task, you will have been assigned a set of text files. Your task is to look for, identify, select and annotate all implicit connectives in the text. Work through the files sequentially.

The procedure to follow for Task B annotation is as follows:

1. Launch the tool - Section 6.2

2. For each text file given to you:
   (a) Open and Load the file - Section 6.3.
   (b) While Reading the text from beginning to end, annotation should progress sentence by sentence as well as sentence-pair by sentence-pair. For each new sentence that you read,
      i. If there is no explicit connective in the sentence for connecting it to the previous sentence (in the same para), Highlight and Select the end period of the previous sentence as the implicit connective. See Section 6.9 for how to select the implicit connective.
         A. Mark the arguments and supplements for the implicit connective (Section 6.10).

6.17 Other remarks and notes

- For every annotation in the text, you can see some information associated with it at the bottom of the WordFreak window. Depending on whether the annotation is for a connective, or an argument, or supplement, this is what you will see:
  - name of file
  - name of tagger (annotator)
  - the ID number for the connective/argument/supplement
  - the offsets for the string (offset counted from the beginning of the file)
6.18 Troubleshooting

6.18.1 I am having launching WordFreak away from IRCS

- if you’re using Windows, you need to have Java-1.4.0 or some later version. If you’re on Mac, you need Mac OS X version 10.0 or some later version.

- If your operating system specs and JAVA specs are fine, then the following may be the problem:

  1. When you launch the tool, the tool program needs to be opened with some application on the machine. This application MUST be Java Webstart (called “jazaws” on Windows). If the tool is being opened with some other application, then the tool won’t launch. So check to make sure that the tool is being opened with Webstart.

  you can check this in the following way: when you launch the tool program gets saved as a local copy (called “wordfreak.jnlp”) on your machine somewhere. find out where it’s getting saved (possible location is C: temp in Windows OR search for ‘wordfreak’ or some similar reduced expression) and look at its properties (right-click on file) to see which application is being used to launch it. if it’s not Java webstart, then set it to start with that. REBOOT your machine.

6.18.2 I removed an annotation from the Chooser but the text continues to be highlighted in the WordFreak window

1. Refresh the Text: pull down the Viewer menu in the WordFreak window, uncheck Text (upon which the window will collapse a bit), and then check it back again.

2. If the above don’t work, Save and Close the file and Relaunch the tool.

6.18.3 The annotation nodes in the Chooser are jumbled up or doubled up

1. Refresh the Chooser: pull down the Annotation menu in the WordFreak window and expand the Set Annotation option. You will see that the Discourse option is checked. Check Token (upon which the Chooser window will collapse a bit), and then redo the above, but this time, check back the Discourse option. Your annotation nodes should have been cleaned up.

2. If the above don’t work, Save and Close the file and Relaunch the tool.

6.18.4 I am having trouble saving files

1. Did you remember to follow the instructions in Section 6.13?

2. If the above was not the problem, talk to your supervisors
6.18.5 I get an error message when I try to add an annotation

where the error message is: **cannot combine annotations with the same parent**

- this is related to the problem in Section 6.18.2.

- Check to see whether there is any highlighted (but unannotated) text that includes the text you’re actually trying to annotate:

- When you click on this highlighted text, either the “Argument” or the “Connective” button in the Chooser window will get highlighted, and for which there will be no corresponding annotation in any of the annotation nodes in the Chooser.

- If this is the case, then it means that you had made the selection for the highlighted text at some point and then removed it, but the tool somehow didn’t record your removal.

- Click on the – button in the chooser to tell the tool to completely remove the selection from memory. After you click the – button, the highlighting will disappear. If not, try text refresh (Section 6.18.2).

7 Practice Session

This section contains two texts containing several explicit connectives. Identify all the explicit connectives and implicit connectives and mark their two arguments (just enclose them in brackets). We will discuss your choices and any questions you may have during the tutorial.

7.1 Text 1

<s> Austin, Texas <s> – A Texas halfback who doesn’t even know the team’s plays, Eldon Moritz, ranks fourth in Southwest Conference scoring after three games.

<s> Time stands still every time Moritz, a 26-year-old Army Signal Corps veteran, goes into the field. <s> Although he never gets to play while the clock is running, he gets a big kick – several every Saturday, in fact – out of football.

<s> Moritz doesn’t even have a nose guard or hip pads but he’s one of the most valuable members of the Longhorn team that will be heavily favored Saturday over Oklahoma in the Cotton Bowl.

<s> That’s because he already has kicked 14 extra points in 15 tries. <s> He ran his string of successful conversions this season to 13 straight before one went astray last Saturday night in the 41-8 slaughter of
Washington State.

<s> Moritz is listed on the Longhorn roster as a right halfback, the position at which he lettered on the 1956 team. <s> But ask coach Darrell Royal what position he plays and you'll get the quick response, "place-kicker".

<s> A 208-pound, 6-foot 1-inch senior from Stamford, Moritz practices nothing but place-kicking. <s> Last year, when he worked out at halfback all season, he didn't get into a single game.

<s> "This year, coach Royal told me if I'd work on my place-kicking he thought he could use me," said Moritz. <s> "So I started practicing on it in spring training.

<s> Moritz was bothered during the first two games this year by a pulled muscle in the thigh of his right (kicking) leg and, as a result, several of his successful conversions have gone barely far enough.

<s> Moritz said Monday his leg feels fine and, as a result, he hopes to start practicing field goals this week. <s> He kicked several while playing at Stamford High School, including one that beat Anson, 3-0, in a 1953 district game.

7.2 Text 2

<s> A measure of how hot the stock was, can be found in what happened to it on the market as soon as trading began.

<s> The stock was sold in the underwriting at a price of $12.50 a share. <s> The first over-the-counter trade Wednesday afternoon at Eppler, Guerin & Turner, the managing underwriter, was at $17 a share.

<s> And from that the stock moved right on up until it was trading Thursday morning at around $22 a share.

<s> But the Morton Foods issue was hot long before it was on the market. <s> Indeed, from the moment the reports of the coming issue first started circulating in Dallas last January, the inquiries and demand for the stock started building up.

<s> Letters by the reams came in from investment firms all over the nation, all of them wanting to get a part of the shares that would be sold (185,000 to the public at $12.50, with another 5,000 reserved for Morton Foods employees at $11.50 a share).
There was even a cable in French from a bank in Switzerland that had somehow learned about the Dallas stock offering. "We subscribe 500 shares of Morton Foods of Texas," it said translated. But E.G.T. could not let the Swiss bank have even 10 shares.

After it allotted shares to 41 underwriters and 52 selling group members from coast to coast there were not many shares for anyone.

But the result of it all was, E.G.T. partner Dean Guerin believes, an effective distribution of the stock to owners all over the nation.

"I feel confident the stock will qualify for the 'national list'," he said, meaning its market price would be quoted regularly in newspapers all over the country.

He was also pleased with the wide distribution because he thought it proved again his argument that Dallas investment men can do just as good a job as the big New York investment bankers claim only they can do.

But what made the Morton Foods stock issue such a hot one?

The answer is that it was a combination of circumstances.

First, the general stock market has been boiling upward for the last few months, driving stocks of all kinds up. As a result, it is not easy to find a stock priced as the Morton issue was priced (at roughly 10 times 1960 earnings, to yield a little over 5 per cent on the 64-cent anticipated dividend).

8 Appendix

8.1 Propositional Situations as Arguments of Connectives

In this section, we give some examples of propositional situations as arguments of connectives. Propositional situations can be typically characterized as events, states, belief propositions, facts, or properties (or generalizations about individual entities). Each type of situation is exemplified below. These examples are meant as a rough helpful guide for you to look for and identify the arguments of connectives.

1. **Events** as Situations

   (111) Pierre won the card game.
(112) Faniry threw herself at Pierre with new accusations.
(113) Blake ran down the stairs.
(114) Blake fell through the wooden stairs.

2. **States** as Situations

   (115) Blake is in the hospital.
   (116) Blake has a broken leg.
   (117) Faniry was extremely jealous of Pierre
   (118) The cat is on the mat.
   (119) The Colonel owns the farm.

3. **Generalizations** as Situations

   (120) Blake is more than 7 feet tall.
   (121) Pierre speaks French.
   (122) Faniry often fights with Pierre.
   (123) Blake is often in love. (stative generalization)
   (124) Pierre often knows the answer. (stative generalization)

4. **Embedded Situations**

   (125) Pierre knows that *Faniry is jealous of him.*
   (126) Blake regrets that *he ran down the stairs.*
   (127) Pierre believes that *Faniry has changed.*
   (128) Blake thought that *the stairs were strong.*

8.2 **Elements of a Clause**

A Clause as the expression of a situation is a grammatical unit that contains a **predicate**, **arguments of the predicate**, and **modifiers**.
8.2.1 Predicates

Predicates are typically verbs (129), but can also be realized as other grammatical categories, especially when the only verb in the clause is the auxiliary verb be (or one of its forms, is, are), such as adjectives (130), prepositional phrases (131), and even noun phrases (132).

(129) Blake has bought the book.
(130) Blake is tall.
(131) Blake is in the hospital.
(132) Blake is the president of our institute.

Note that the predicates as marked above all include the auxiliary verbs. This is the convention that we will follow. For example, in the sentence, John could not have been drinking in the morning, the entire verbal complex could not have been drinking is the predicate. Note also that the negation has been included. Other elements that appear within the verbal complex, however, such as adverbs (e.g., “then” in could then have not been drinking), should not be regarded as part of the predicate.

8.2.2 Predicate Arguments

The arguments of a predicate are elements of the clause that are necessary elements of the situation expressed by the predicate. Arguments are distinguished into Subject Arguments and Object Arguments. A predicate has a single subject argument, but the number of object arguments in any given clause depends on the particular situation, and thus, on the predicate itself.

For example, in the simple situation clause (133), the verb ‘is sleeping’ is the predicate expressing a ‘sleeping’ event. The subject of the clause, ‘John’, is the only argument of the predicate, i.e., it is the individual entity of which the sleeping event is predicated.

(133) John is sleeping.

Some predicates, such as ‘sleep’ above, require only one argument for their interpretation. Others, such as ‘eat’, require two arguments, the ‘eater’ often expressed as the subject, and the item that gets eaten, often expressed as the object. Yet other predicates require three arguments, such as the verb ‘give’, for example, which requires an argument expressing the giver, often the subject, an argument expressing the thing given, often the direct object, and an argument expressing the recipient, often the indirect object (indirect objects in English have the appearance of a prepositional phrase).

8.2.3 Modifiers

Both the predicate as well as the arguments of the predicate can be modified, as shown in the following examples. Note that connectives can also be modified, as shown in the last example.
• Verb modification:
  (134) John often eats cakes.

• Argument modification:
  (135) John gave Mary a history book.
  (136) John was surprisingly nice.

• Modifier modification
  (137) John very often eats cakes.

• Connective modification
  (138) John will give Mary the book only if she promises to return it.