

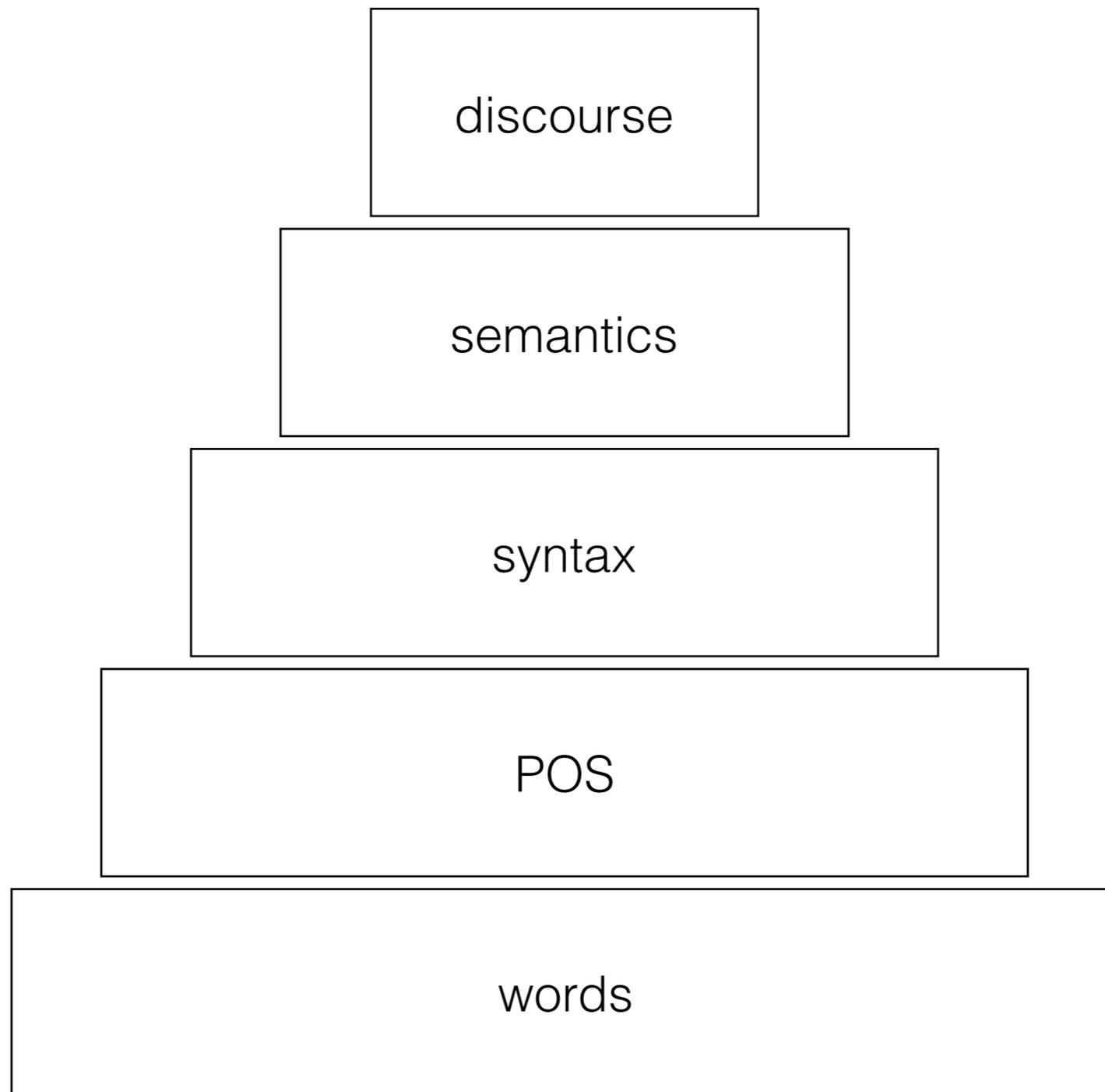


Natural Language Processing

Info 159/259

Lecture 22: Discourse (Nov. 7, 2017)

David Bamman, UC Berkeley



JJ

great

great



We are met on a great battle-field of that war.

$\lambda x.\lambda y.meet(x,y)$

$\lambda x.battlefield(x)$

$\lambda x.war(x)$

We are met on a great battle-field of that war.

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. **We are met on a great battle-field of that war.** We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we can not dedicate -- we can not consecrate -- we can not hallow -- this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us -- that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion -- that we here highly resolve that these dead shall not have died in vain -- that this nation, under God, shall have a new birth of freedom -- and that government of the people, by the people, for the people, shall not perish from the earth.

Discourse

- Discourse covers linguistic expression **beyond the boundary of the sentence.**
 - Dialogues: the structure of turns in conversation
 - Monologues: the structure of entire passages, documents



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true! That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true! That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true! That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



General Secretary
Gorbachev, if you seek
peace, if you seek prosperity
for the Soviet Union and
eastern Europe, if you seek
liberalization, come here to
this gate.

Mr. Gorbachev, open this
gate.

Mr. Gorbachev, tear down this
wall!



General Secretary
Gorbachev, if you seek
peace, if you seek prosperity
for the Soviet Union and
eastern Europe, if you seek
liberalization, come here to
this gate.

Mr. Gorbachev, open this
gate.

Mr. Gorbachev, tear down this
wall!

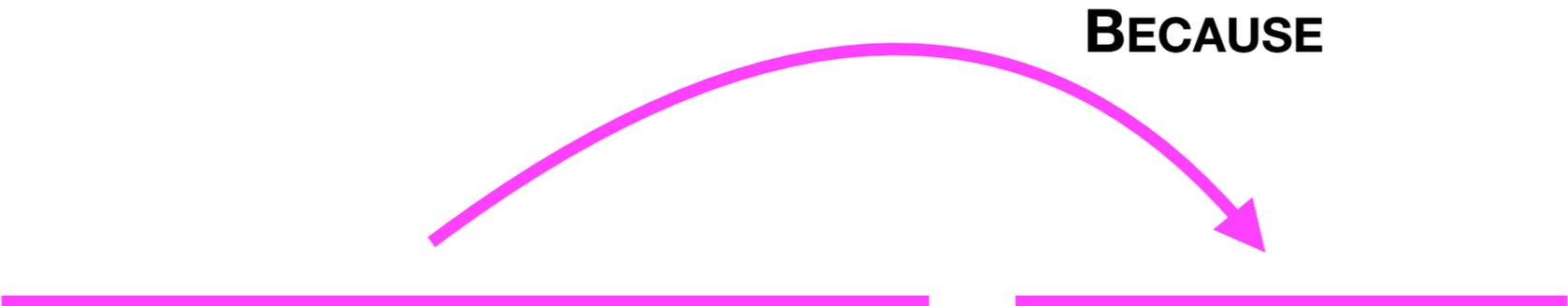


- General Secretary Gorbachev,
- if you seek peace,
- if you seek prosperity for the Soviet Union and eastern Europe,
- if you seek liberalization,
- come here to this gate.
- Mr. Gorbachev, open this gate.
- Mr. Gorbachev, tear down this wall!

What is the **structure** of discourse?

Discourse structure

BECAUSE

- 
- John hid Bill's car keys. He was drunk
 - John hid Bills car keys. He likes spinach.

Discourse structure

- Useful for:
 - Text summarization (Louis et al. 2010)
 - Evaluating writing (Burnstein et al. 2013)

Discourse structure

- Two representations of structure in discourse:
 - deep structure (RST)
 - shallow structure (PDTB)

RST

- Rhetorical structure theory (Mann and Thompson 1987)
- RST Corpus: 385 WSJ articles (from Penn Treebank) = 176K words.

Elementary discourse units

- Elementary discourse units: the “minimal building blocks of a discourse tree” (Carlson 2002)

Elementary discourse units

- Clauses
- Prosodic units
- Turns of talk
- Sentences

Elementary discourse units

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue.

This earned mixed reviews from Wall Street analysts.

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue,

which earned mixed reviews from Wall Street analysts.

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue,

earning mixed reviews from Wall Street analysts.

The 6.2% growth of Xerox Corp's third quarter net revenue on 7.3% higher revenue

earned mixed reviews from Wall Street analysts.

Elementary discourse units

- Clauses
- Prosodic units
- Turns of talk
- Sentences



RST Discourse treebank

Clauses in RST

- Subject/object clauses (gerunds, infinitives) are not EDUs:
- Making computer smaller often means sacrificing memory.
- Insurers could see claims totaling nearly \$1 billion from the San Francisco earthquake

Clauses in RST

- Some phrasal EDUs that begin with a strong discourse marker
 - because
 - in spite of
 - as a result of
 - according to

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C
- and can dip to -123 C.
- Only the midday sun at tropical latitudes is warm enough
- to thaw the ice

Relations

- **Adjacent** EDUs are linked with discourse relations.

Relations

- Relations hold between a **nucleus** and **satellite**.
 - Nucleus: more salient to discourse
 - Satellite: supporting information
- Analogous to head/dependent in dependency parsing but scoped over constituents.

Mars
experiences
frigid weather.

Surface
temperatures
average -60 C

Attribution	Reported speech (direct and indirect)	[The legendary GM chairman declared] [that his company would make “a car for every purse and purpose.”]
Background	The context or the grounds with respect to which the nucleus is to be interpreted	[Banco Exterior was created in 1929 to provide subsidized credits for Spanish exports.] [The market for export financing was liberalized in the mid-1980s, however, forcing the bank to face competition.]
Cause	Nuclear causes the satellite	[This year, a commission appointed by the mayor to revise New York's system of government completed a new charter,] [expanding the City Council to 51 from 35 members.]
Comparison	Two spans are compared along some dimension	[It was as easy] [as collecting sea shells at Malibu.]

Condition	Truth of nucleus associated with fulfillment of condition expressed in satellite	[S.A. brewing would make a takeover offer for all of Bell Resources] [if it exercises the option,] according to the commission
Contrast	Two nuclei contrast each other	[But from early on, Tiger's workers unionized,] [while Federal's never have.]
Elaboration	Further information about content of nucleus	[The projects are big.] [They can be C\$1 billion plus.]
Enablement	Satellite increases chance of nucleus being realized	["I'd like to see (Kidder) succeed.] [But they have to attract good senior bankers who can bring in the business from day one."]

Evaluation	An appraisal, estimation, rating, interpretation, or assessment of a situation.	[Employers must deposit withholding taxes exceeding \$3,000 within three days after payroll -- or pay stiff penalties --] [and that's a big problem for small businesses.]
Explanation	Satellite present factual explanations for nucleus	[But their 1987 performance indicates that they won't abandon stocks unless conditions get far worse.] ["Last time, we got rewarded for buying stocks when the panic was the worst," said John]
Joint	Elements are listed [List]	[A union, sooner or later, has to have an adversary,] [and it has to have a victory.]
Manner-Means	Satellite explains the way in which something is done.	[Magazine editors did not take the criticisms] [lying down.]

Topic-Comment	Topic is introduced and then commented upon	[As far as the pound goes,] [some traders say a slide toward support at \$1.5500 may be a favorable development for the dollar this week.]
Summary	Satellite summarizes the information presented in the nucleus.	[Stocks are bracing for a turbulent separation, following DZ's announcement that the exchanges will sever ties.] [On Friday, DZ ordered the Kuala Lumpur Stock Exchange to take appropriate action immediately to cut its links with the Stock Exchange of Singapore.]
Temporal	Nuclear occurs before satellite [Temporal-before]	[We want to make sure they know what they want] [before they come back.]
Topic-Change	sharp change in focus going from one segment to the other.	[South Africa freed the ANC's Sisulu and seven other political prisoners.... Mandela, considered the most prominent leader of the ANC, remains in prison. But his release within the next few months is widely expected.] [The Soviet Union reported that thousands of goods needed to ease widespread shortages across the nation were piled up at ports and rail depots, and food shipments were rotting because of a lack of people and equipment to move the cargo.]

Relations

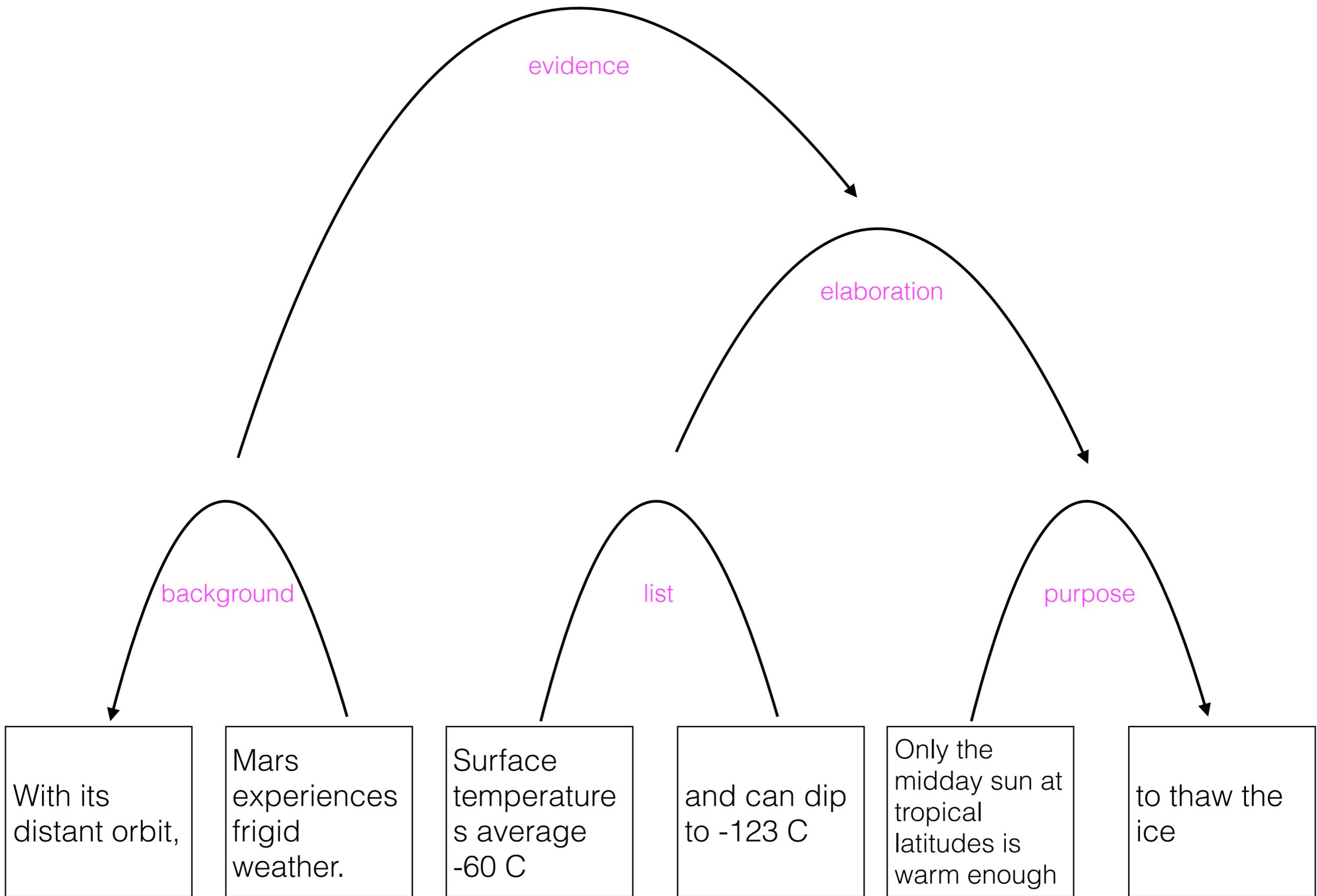
List	Billy was the mate and Long John was the quartermaster
Elaboration	The company wouldn't say , citing competitive reasons
Attribution	Analysts estimated that that sales at US stores declined
Same-Unit	[two EDUS interrupted by another EDU]
Contrast	The priest was in a bad mood but the lama was happy
Background	T is the pointer to the root of a binary tree . Initialize T

Most common relations in RST treebank

Relations

List	Billy was the mate and Long John was the quartermaster
Elaboration	The company wouldn't say , citing competitive reasons
Attribution	Analysts estimated that that sales at US stores declined
Same-Unit	[two EDUS interrupted by another EDU]
Contrast	The priest was in a bad mood but the lama was happy
Background	T is the pointer to the root of a binary tree . Initialize T

Most common relations in RST treebank



Arrows are directed from nucleus to satellite (undirected edges are nuclei)

Relations

- A given span of text will hold several (nested) relations

With its distant orbit,	Mars experiences frigid weather.	background
With its distant orbit, Mars experiences frigid weather.	Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice	evidence

RST parsing

- Shift-reduce (Marcu 1999, Sagae 2009)
- Neural-network shift reduce (Ji and Eisenstein 2014)
- Recursive neural network (Li et al. 2014)

With its distant orbit,

Mars experiences frigid weather.

Surface temperatures average -60 C

and can dip to -123 C

stack

action

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

Attribution
Background
Cause
Comparison
Condition
Contrast
Elaboration
Enablement
Evaluation
Explanation
Joint
Manner
Topic-Comment
Summary
Temporal
Topic-Change

Each relation generates three possible transition labels (for the position or the nucleus and satellite in the subtree)

- Attribution-NS
- Attribution-SN
- Attribution-NN

With its
distant orbit,

Mars
experiences
frigid
weather.

background-SN:

background relation holds between satellite on left and nucleus on right

Surface temperatures average -60 C

and can dip to -123 C

stack

action

Mars experiences frigid weather.

With its distant orbit,

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

Surface
temperatures
average -60 C

and can dip to
-123 C

stack

action

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

background-SN

With its distant orbit,

Mars experiences frigid weather.

Surface temperatures average -60 C

and can dip to -123 C

stack

action

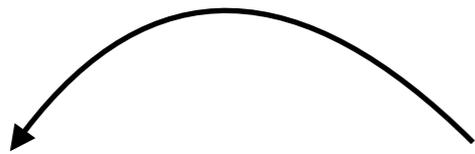
Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

background

With its distant orbit,

Mars experiences frigid weather.



Surface temperatures average -60 C

and can dip to -123 C

stack

action

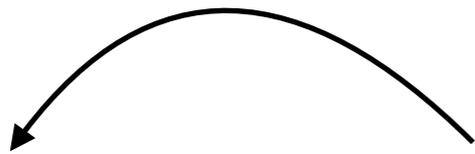
Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

background

With its distant orbit,

Mars experiences frigid weather.



and can dip to
-123 C

stack

action

Surface
temperatures
average -60 C

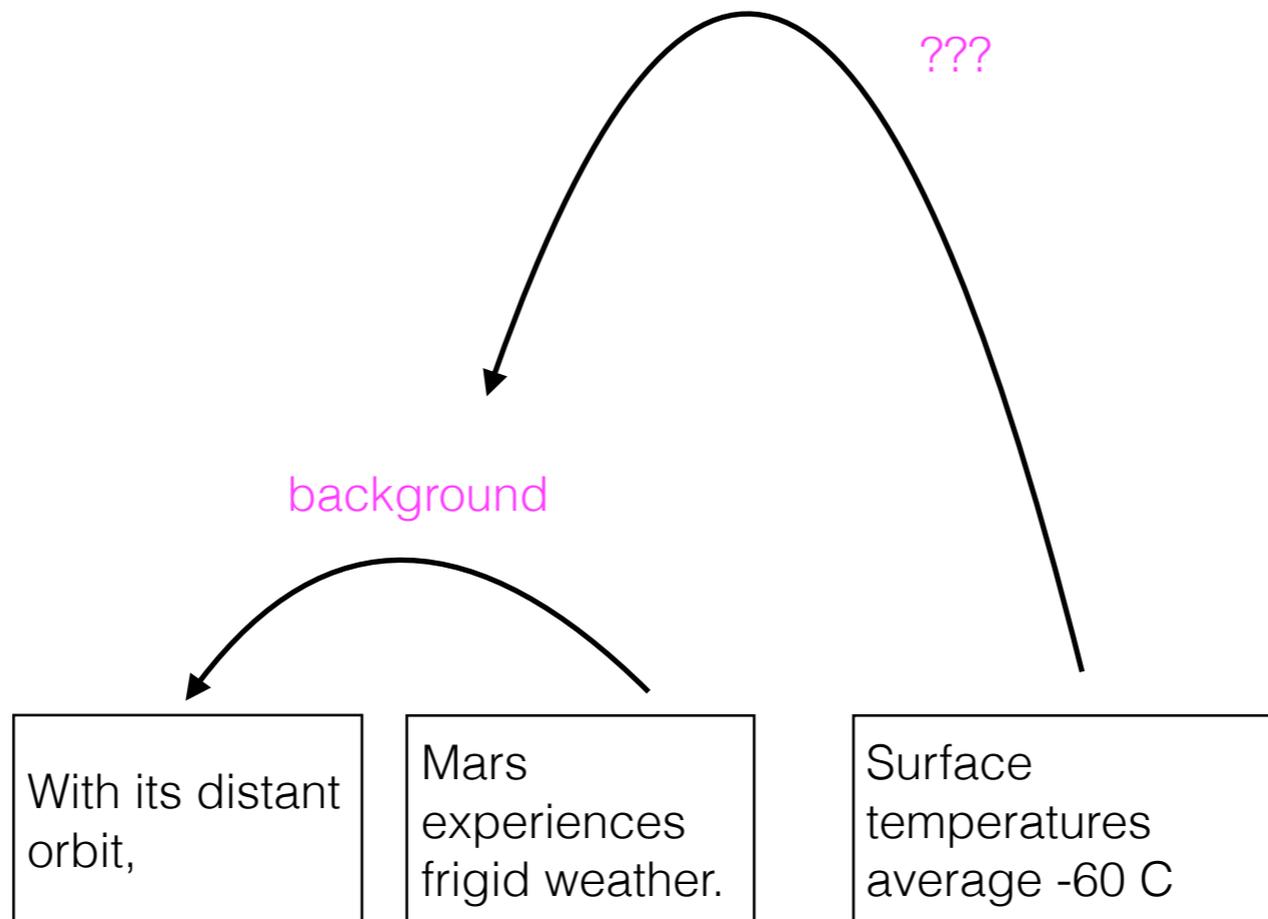
background

With its distant
orbit,

Mars
experiences
frigid weather.

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack



list-NN:

list relation holds between nucleus on left
and nucleus on right

and can dip to
-123 C

stack

action

Surface
temperatures
average -60 C

Shift: Remove EDU from
front of input buffer and
push it onto stack

Reduce(label): assert
relation between subtree at
stack₁ and subtree at stack₂:
remove stack₁ and stack₂
from stack; create a new
subtree from stack₁ and
stack₂ and place it onto the
stack

background

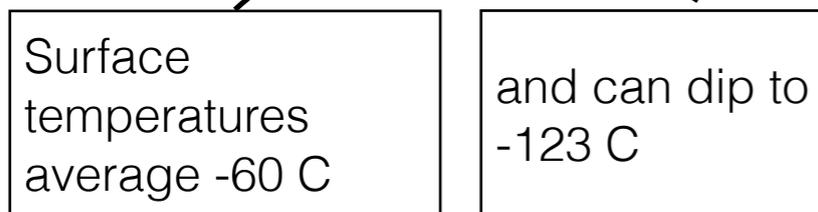
With its distant
orbit,

Mars
experiences
frigid weather.

Continue in this way until the buffer is empty and stack has a single tree

stack

list



background



action

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at $stack_1$ and subtree at $stack_2$: remove $stack_1$ and $stack_2$ from stack; create a new subtree from $stack_1$ and $stack_2$ and place it onto the stack

- RST builds a rhetorical tree over the entire span of a sentence = **deep** discourse parsing.
- SOTA for getting the labeled relations + heads right (assuming gold EDU): **61.8 F**

- The Penn Discourse Treebank encodes a different approach: **shallow** discourse parsing.
- Rather than focusing on the tree spanning the entire sentence, it focuses on recognizing the structure around **discourse connectives**.

Discourse connectives

- Discourse connectives are those words and phrases that link together the content or purpose of adjacent spans
 - John eats porridge for breakfast, **while** Mary eats muesli.
 - Eat your porridge. **Otherwise** you're not going to football practice.
 - You've eaten your porridge every day this week. **As a result** I'm going to give you the iPod I promised you.

Discourse connectives

- Discourse connectives are binary predicates that require two arguments

connective	sense	count
and	conjunction	3000
but	contrast	3308
if	hypothetical	1223
when	synchrony	989
while	contrast, synchrony	781
so	result	263
in addition	conjunction	165

Discourse connectives

- Discourse connectives are binary predicates that require two arguments
 - [John eats porridge for breakfast] **while** [Mary eats muesli]
 - [Eat your porridge.] **Otherwise** [you're not going to football practice.]
 - [You've eaten your porridge every day this week] **As a result** [I'm going to give you the iPod I promised you.]

Discourse connectives

- Those arguments can be in the same sentence as the connective, adjacent sentences, or non-adjacent sentences.
- The federal government suspended sales of U.S. savings bonds because *Congress hasn't lifted the ceiling on government debt*.
- Because *it operates on a fiscal year*, Bear Stearns's yearly filings are available much earlier than those of other firms.
- Most oil companies, when *they set exploration and production budgets for this year*, forecast revenue of \$15 for each barrel of crude produced.

Discourse connectives

- Each discourse connective also is also tagged with its **sense**. Most words almost always occur with one sense.

temporal	the situations described in the arguments are related temporally (before, after, while, since)
contingency	one argument causally influences the other (because, if, so, since)
comparison	highlights prominent differences between arguments (but, however, although)
expansion	expand discourse and move exposition forward (in particular, specifically, in other words); conjunction; list (and)

Discourse connectives

coarse sense	fine sense
temporal	Asynchronous, synchrony
contingency	Cause, pragmatic cause, condition, pragmatic condition
comparison	Contrast, pragmatic contrast, concession, pragmatic concession
expansion	Conjunction, instantiation, restatement, alternative, exception, list

Discourse connectives

- Annotation/prediction task: for all explicit discourse connectives in a text:
 - What is the discourse sense of the connective?
 - What are the spans in the text that correspond to its arguments?

Discourse connectives

I came.

and I saw.

and I conquered.

Discourse connectives

The \$6 billion that some 40 companies are looking to raise in the year ending March 31 compares with only \$2.7 billion raised on the capital market in the previous fiscal year.

In contrast, In fiscal 1984 before Mr. Gandhi came to power, only \$810 million was raised.

Discourse connectives

The gruff financier recently started socializing in upper-class circles.

For example, Last year he attended a New York gala where his daughter made her debut.

Discourse connectives

John was tired. *As a result* he left early.

John was tired. He left early.

Discourse connectives

What is the implicit relationship between sentences with no explicit discourse connector?

SOTA

- CoNLL 2015 shared task on shallow discourse parsing using PDTB
- End-to-end: input = just text; output = PDTB-style discourse annotations (implicit and explicit connectors, their arguments and senses).

Pipeline

- Connective classifier: identifying explicit discourse connectives
 - Argument extractor
 - Sense classifier
-
- For all adjacent sentences, classify its sense into its implicit relation (including NoRel).

	on PDTB Section 23					
	our parser			Lin's parser		
	P	R	F_1 (%)	P	R	F_1 (%)
Explicit connective	94.83	93.49	94.16	-	-	93.62
Explicit Arg1 extraction	51.05	50.33	50.68	-	-	47.68
Explicit Arg2 extraction	77.89	76.79	77.33	-	-	70.27
Explicit Both extraction	45.54	44.90	45.22	-	-	40.37
Explicit only sense	35.52	34.69	34.93	-	-	-
Non-Explicit Arg1 extraction	64.83	69.50	67.08	-	-	-
Non-Explicit Arg2 extraction	66.02	70.78	68.32	-	-	-
Non-Explicit Both extraction	51.20	54.89	52.98	-	-	-
Non-Explicit only sense	53.18	10.45	9.06	-	-	-
All Arg1 extraction	59.20	61.03	60.10	-	-	-
All Arg2 extraction	71.43	73.64	72.52	-	-	-
All Both extraction	48.62	50.13	49.36	-	-	-
Sense (Explicit+Non-Explicit)	31.44	30.42	29.83	-	-	-
Overall Parser	29.27	30.08	29.72	-	-	20.64

Wang and Lin (2015), "A Refined End-to-End Discourse Parser"

	WSJ Test Set				
	2015 F₁	2016 F₁	P	OPT R	F₁
Explicit Connectives	94.8	98.9	96.4	92.5	94.4
Explicit Arg1 Extraction	50.7	53.8	53.1	50.9	52.0
Explicit Arg2 Extraction	77.4	76.7	74.1	71.1	72.6
Explicit Both Extraction	45.2	45.3	44.9	43.0	43.9
Explicit Sense Micro-Average			38.6	40.2	39.4
Non-Explicit Arg1 Extraction	67.2	69.9	72.0	68.0	69.9
Non-Explicit Arg2 Extraction	68.4	71.5	73.5	69.5	71.5
Non-Explicit Both Extraction	53.1	53.5	55.0	52.0	53.5
Non-Explicit Sense Micro-Average			17.5	18.6	18.0
All Both Extraction	49.4	49.6	50.2	47.8	48.9
Overall Parser Performance	29.7	30.7	27.5	28.9	28.2

Open et al. (2016), “OPT: Oslo–Potsdam–Teesside: Pipelining Rules, Rankers, and Classifier Ensembles for Shallow Discourse Parsing”

Data

- RST Treebank [UCB Library]
<http://digitalassets.lib.berkeley.edu/UCBonly/lidc/2002T07/LDC2002T07.tgz>
- Penn Discourse Treebank 2.0 [UCB Library]
<http://digitalassets.lib.berkeley.edu/UCBonly/lidc/2008T05/>