

# Computational Discourse

## Elements and Approaches

Taal- en spraaktechnologie

Fall 2009

Jennifer Spenader

# Structure of lecture

1. What makes a text or dialogue coherent?
  - What makes it a *discourse* rather than just a collection of sentences?
2. What types of relationships exist between segments in a discourse?
  - What theories of discourse structure exist?
    - Grosz & Sidner (1986)
    - RST (Mann & Thompson 1988)
3. Resources for studying discourse structure
  - RST Corpus
  - PDTB 2.0
4. Why is this useful?
  - What NLP applications can benefit from discourse structure?

# Characteristics of Discourse

What makes a group of sentences coherent?

Buren Obama zetten huis te koop (NOS.nl)

**14-09-09**

Buren van president Obama in Chicago hebben hun huis te koop gezet. Het is volgens Amerikaanse media voor het eerst dat een huis naast de woning van een zittend president te koop staat.

De Obama's wonen nu in Washington, maar hebben hun huis in Chicago aangehouden. Hun buren daar zeggen dat ze kleiner willen wonen nu de kinderen de deur uit zijn.

# Linguistic theories of coherence

- Halliday and Hasan (1976)
  - A text is coherent because words and expressions in one sentence can only be interpreted by understanding the words and expressions in the surrounding context
  - Texts are coherent because they `talk' about the same things

**Buren Obama zetten huis te koop (NOS.nl)**

**14-09-09**

**Buren van president Obama** in Chicago hebben **hun huis** te koop gezet. Het is volgens Amerikaanse media voor het eerst dat **een huis naast de woning van een zittend president** te koop staat.

**De Obama's** wonen nu in Washington, maar hebben **hun huis in Chicago** aangehouden. **Hun burens** daar zeggen dat ze kleiner willen wonen nu de kinderen de deur uit zijn.

# AI-theorists

- Hobbs (1979)
  - Coherence is more than just being “about” the same entities, e.g.
    - **John took a train from Paris to Istanbul. He likes spinach.**
  - **Coherence is created by much deeper processes**
  - **Certain coherence relations hold between text segments**
  - **There are only a limited number of ways to continue a discourse relevantly**

Buren Obama zetten huis te koop (NOS.nl)

**14-09-09**

Buren van president Obama in Chicago hebben hun huis te koop gezet. → **Interesting because** Het is (volgens Amerikaanse media) voor het eerst dat een huis naast de woning van een zittend president te koop staat.

→ **Actually** De Obama's wonen nu in Washington, maar hebben hun huis in Chicago aangehouden. → **The reason why the neighbors are selling is because** Hun burens daar zeggen dat ze kleiner willen wonen nu de kinderen de deur uit zijn.



Buren Obama zetten huis te koop (NOS.nl)

**14-09-09**

Buren van president Obama in Chicago hebben hun huis te koop gezet. → **Interesting because** Het is (volgens Amerikaanse media) voor het eerst dat een huis naast de woning van een zittend president te koop staat.

→ **Actually** De Obama's wonen nu in Washington, **maar** hebben hun huis in Chicago aangehouden. → **The reason why the neighbors are selling is because** Hun burens daar zeggen dat ze kleiner willen wonen nu **because** de kinderen de deur uit zijn.

# Elementary discourse relations

- De Obama's wonen nu in Washington, **maar** hebben hun huis in Chicago aangehouden.
  - “Contrast” relation
    - The first sentence suggests something that the second sentence then denies
    - **explicitly** marked with “maar”
- Hun burens daar zeggen dat ze kleiner willen wonen nu **because** de kinderen de deur uit zijn.
  - “Effect-cause” relation
    - **implicitly marked** (no marking)

# Abstract discourse relations

- Most theories attempt to give an analysis to an entire text in terms of explicit and implicit discourse relations
- These theories all assume:
  - there are some elementary or basic discourse units
    - these are the terminal elements
  - relations hold between adjacent elements
  - these relations can be joined into larger relations
  - you can make a kind of tree structure for an entire text

## Theories differ:

- How are the elementary discourse units delimited?
- What types of relations are assumed between elementary discourse units?
- Are there both **informational** and **intentional** discourse structures? Or is there really no distinction?
- How do discourse elements relate to syntactic and lexical structure

# Major theories of global discourse coherence

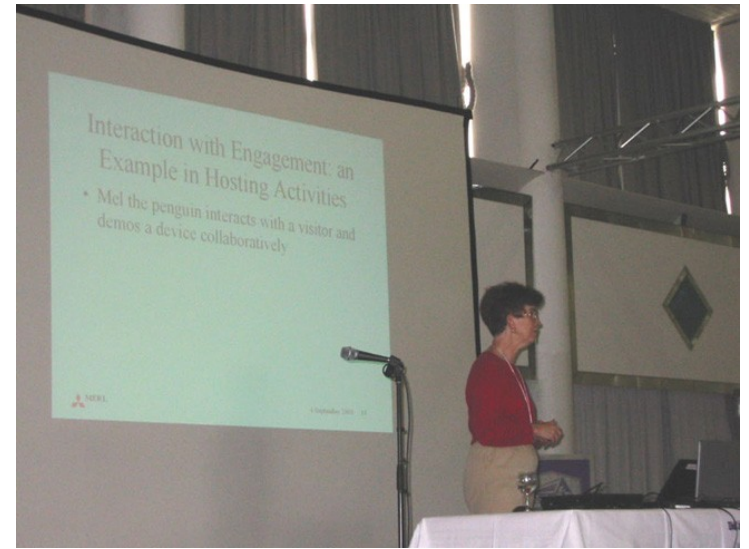
- Global Coherence
  - Grosz & Sidner (1986)
  - RST Rhetorical Structure Theory (Mann & Thompson 1988)
    - (SDRT (Asher 1993) Segmented Discourse Representation Theory)
  - L-DTAG (PDTB 2.0)

# Intentions

- The intention of a message is what gives us a clue to the structure of a discourse or dialogue
  - Intentions don't correlate exactly with a particular linguistic signal, i.e.
  - Could you open a window?
  - It's very stuffy in here.
    - Same intention, very different "signal"



## Grosz & Sidner (1986)



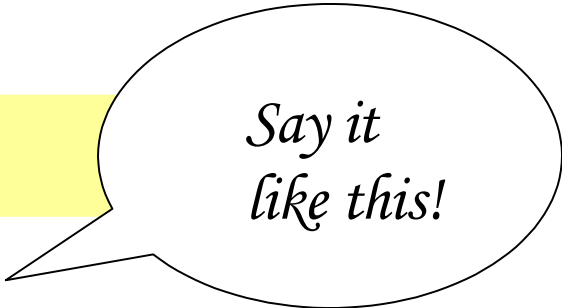
- Grosz & Sidner's discourse theory describes the processing of the entire discourse
- First “computational” theory of discourse structure
- Grosz & Sidner's theory can be considered **intention based**
  - The concept of “intention” plays a role throughout the theory
    - a determination to act in a certain way (Mirriam-Webster)
    - means the structure of the discourse is seen as choices made by the speaker for how they should present information

# Three components

- G & S identify three different types of information
  1. Linguistic structure
  2. Intentional structure
  3. Attentional structure



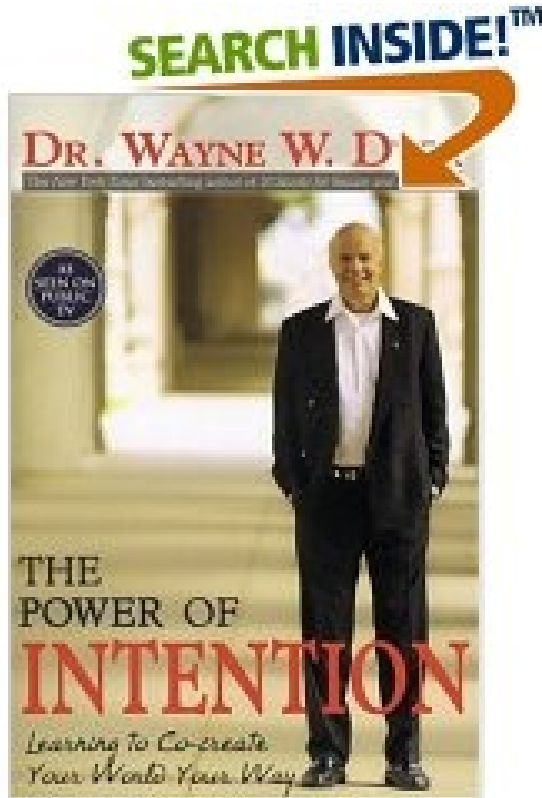
# Linguistic structure



*Say it  
like this!*

- Linguistic structure = the actual text
  - What exact words are used, the exact syntax, etc.
- The text is divided into **discourse segments (DS)**
  - Often: sentences or clauses, but not necessarily so
- The linguistic structure doesn't contain elements like concepts, inferences, etc.
- How we decide what counts as a discourse segment is a complex problem

# Intentional structure



What each segment is trying to do:

- inform us? convince us of something? give an explanation?

A discourse always has one main intention:

- Discourse Purpose or **DP**

Every discourse segment has an intention as well:

- **Discourse Segment Purpose**

# Attentional State



- Attentional state includes what the **focused items** in the discourse are at a given moment,
  - the Obama's, the neighbors, the neighbor's house, etc.
- These items are stored or organized in Focus Spaces
- Focus Spaces represent the Attentional Structure (**FS**)
- Includes all salient concepts
- NOTE: Discourses have attentional states, not discourse participants

# At what level do you find each structure?

Linguistic structure →

DS

Discourse  
segments

Intentional structure →

DSP

Discourse  
segment  
purposes

Attentional structure →

FS

Focus  
structures

# Three separate but isomorphic structures

- Each discourse segment (DS) is ruled by one DSP
- Each focus space is tied to a discourse segment with its associated DSP
- Focus spaces are collected in a stack
- The state shows dynamic relationships
- Manipulations with the state are governed by dominance relationships.
- Relationships between DSP are given by a dominance hierarchy which is static.





## Example from Lascarides (1999)

I would like for you to assemble the compressor.

- It seems natural to treat each sentence as a discourse segment.

I suggest you begin by attaching the pump to the platform

Fine. Now let's see if it works.



## After segmenting

**DS0** I would like for you to assemble the compressor.

**DS1** I suggest you begin by attaching the pump to the platform

**DS2** Fine. Now let's see if it works.

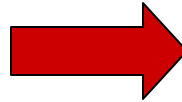
- What is the intention of DS0?
- DSP0 Agents wants the user to put together a compressor
- We can also guess that this is the intention of the entire discourse
- Now we need focus spaces...





## Keep track of attentional state with stack

**DS0** I would like for you to assemble the compressor.



**DSP0** Agent want to get user to correctly set up compressor

**DS1** I suggest you begin by attaching the pump to the platform

**DS2** Fine. Now let's see if it works.

Focus stack

Intentional Structure

Compressor,  
DSPO FSO

DSPO

FSO gets "pushed" on the stack

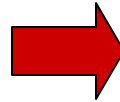


## Purpose of the DS1

**DS0** I would like for you to assemble the compressor.

- What is the purpose of the DS1?

**DS1** I suggest you begin by attaching the pump to the platform



DSP1: Agent wants to get the user to do a part of the assembly

**DS2** Fine. Now let's see if it works.



## Dominance relationship

**DS0** I would like for you to assemble the compressor.

**DS1** I suggest you begin by attaching the pump to the platform

**DS2** Fine. Now let's see if it works.

DSP1: Agent wants to get the user to do a part of the assembly

DSP 1 is dominated by DSP0  
DS1 is "embedded" in DS0

The embedding of DS's is decided by the dominance relationship between the segments on an intentional level



# Dominance relationship

**DS0** I would like for you to assemble the compressor.

We know this because we understand the task-structure, and can therefore identify the purpose of each segment!

**DS1** I suggest you begin by attaching the pump to the platform

DSP1: Agent wants to get the user to do a part of the assembly

**DS2** Fine. Now let's see if it works.

DSP 1 is dominated by DSP0  
DS1 is "embedded" in DS0

The embedding of DS's is decided by the dominance relationship between the segments on an intentional level



# Dominance relationship

**DS0** I would like for you to assemble the compressor.

**DS1** I suggest you begin by attaching the pump to the platform

**DS2** Fine. Now let's see if it works.

DSP1: Agent wants to get the user to do a part of the assembly

Focus stack

Intentional Structure

Pump,  
Platform  
DSP1 FS1

Compressor,  
DSPO FSO

DSPO

DSP1

FS1 gets pushed on the stack on top of FSO



## Purpose of the DS2

**DS0** I would like for you to assemble the compressor.

- What is the purpose of the DS2?

**DS1** I suggest you begin by attaching the pump to the platform

**DS2** Fine. Now let's see if it works. 

DSP2: Agent wants user to check if the compressor has been put together correctly



## Cue phrase signal

**DS0** I would like for you to assemble the compressor.

The purpose of DS2 is dominated by DS0 rather than DS1

“Fine” is a **linguistic signal**, a **cue phrase**

**DS1** I suggest you begin by attaching the pump to the platform

The cue phrase signals that we should “**pop**” the focus space for DS1 (FS1) because we have satisfied its intention.

**DS2** Fine. Now let's see if it works.

DS1 is now closed off here



# Popping the stack

**DS0** I would like for you to assemble the compressor.

DSP0 dominates DSP2

Focus stack

Intentional Structure

**DS1** I suggest you begin by attaching the pump to the platform

DSP0

DSP2

**DS2** Fine. Now let's see if it works.

Compressor,  
DSPO FSO

DSP0

DSP1

The pump can't be an antecedent for "it"

FS1 is popped from the stack



## Notice:

- **Discourse segmenting** is hard to do without an idea about what the intentional structure is
- The **embedding** of a segment is determined from the intentional structure
- **All three structures are partially isomorphic with each other**
  - If you know something about one, you know something about the others
- **Intentional structure** seems to be primary

## Notice also:

- **Attentional structure**
  - This limits possible dominance relationships. New discourse segments can only have a relationship with something on top of the stack
    - Notice: this is then a theoretical claim that can be empirically tested
- **Linguistic signals**
  - steer push or pop operations on the stack

# Intention-based

- The stack manipulation is controlled by the dominance hierarchy. This means the entire discourse structure that is built up is grounded in how intentions are fulfilled
- Grosz & Sidner have worked with “task-oriented dialogue”
  - Their theory may work best with this type of dialogue
  - Question: How well does this model describe e.g. small talk?

# Two types of relationships between intentions

- 1. Dominance
  - DSP1 dominates DSP2
  - DSP1 dominates DSP2 if it is necessary to satisfy DSP2 in order to be able to satisfy DSP1
- 2. Satisfaction precedes
  - DSP1 satisfaction precedes DSP2 if DSP1 has to be satisfied before DSP2

## Example of “satisfaction precedes”

- How do you register for a course at SU? **DSO**
  - 2. Fill in the registration form. **DS1**
  - 3. Send the form to “Admissions”. **DS2**
- 
- Filling in the form necessarily precedes sending in the form
  - $DSP0 \text{ dom } DSP1$ ,  $DSP1$  satisfaction precedes  $DSP2$ ,  $DSP0 \text{ dom, } DSP2$

# How to do a G&S analysis?

Partially adapted from Lascarides (1999)

1. Decide what you will consider a discourse segment (DS)
2. Decide what the underlying purpose of the discourse segments are (DSP)
3. What relationship holds between each DSP and the other DSPs in the discourse
  - Domain information, plans etc can be used here as well
4. Divide the discourse into discourse segments that reflects this
5. Manipulate the stack if necessary, I.e. pop? If not, push the focus space for the DS on the stack.

# How has G & S's theory been used?

- Barbara Grosz: studied the relationship between discourse structure and prosody, goal: improve speech synthesis
- Candy Sidner: works with developing better systems for “collaborative dialogue systems” between humans and machines, using a modified version of G&S (1986)
- Text-generation: especially work with cue-phrases

## Shortcomings of G&S (1986)

- G&S recognize only two types of relationships between segments, **dominance** and **satisfaction precedes**
  - not clear how this relates to syntactic structure or lexical cues
  - These two categories seem to be too general, I.e. it is not a natural or intuitive task to categorize relationships between segments at this high level
  - It is not clear if we need to distinguish an intentional and an information level



# Rhetorical Structure Theory



- RST (Mann & Thompson 1988) was developed to be used in text analysis
- RST tries to explain the coherence of a text, and describes the text itself, rather than the processing of the text
- Has been used in NLG (natural language generation) and automatic summarization applications

# Swedish cuisine

(1) I love to cook Swedish cuisine.

My recipe for Swedish meatballs is legendary.

(2) I love to cook Swedish cuisine.

My recipe for strawberry cheese cake is legendary.



← More coherent...  
why?

← Less coherent...  
Why?

# Elaboration

- Swedish meatballs are a kind of Swedish cuisine, but strawberry cheesecake is not.

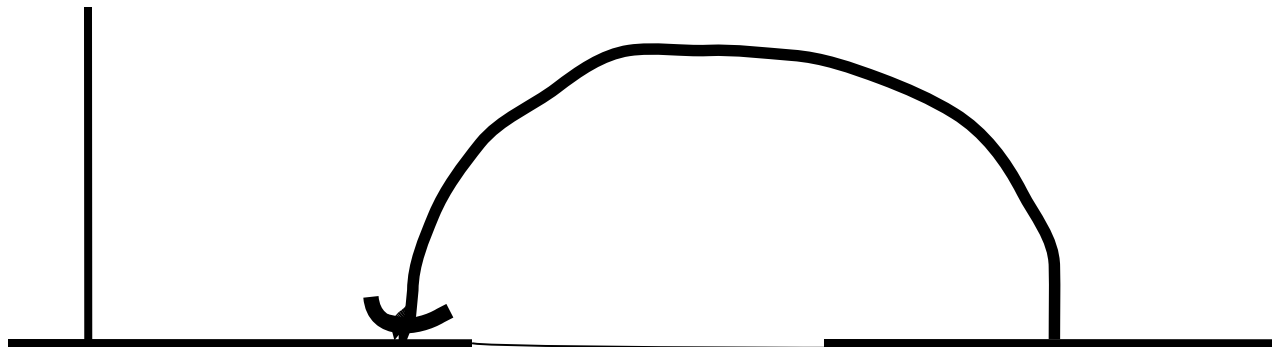
1. I love to cook Swedish cuisine.

4. My recipe for Swedish meatballs is legendary

Sentence 2 gives more information about the same theme introduced in the first sentence

(in G&S's theory we'd say that 1. Dominates 2.)

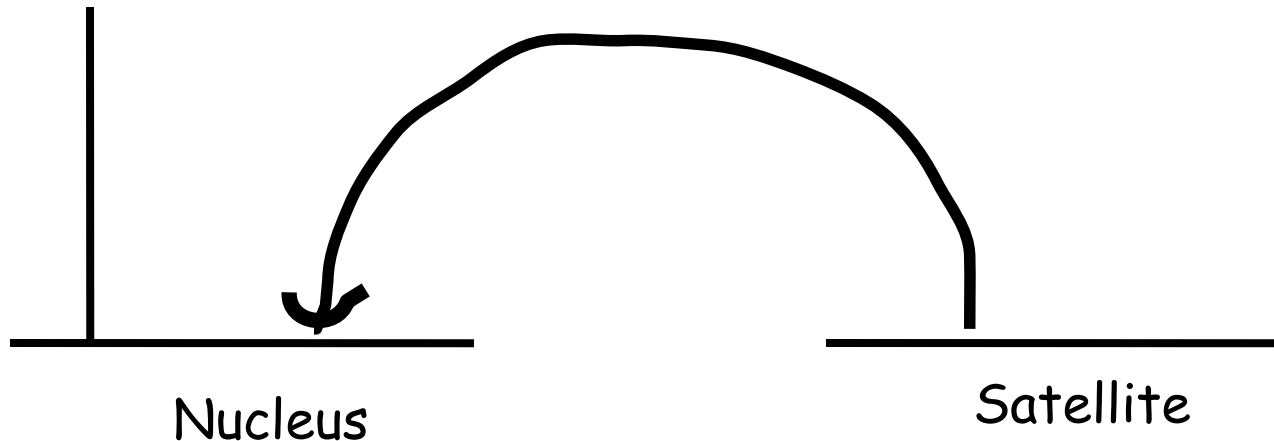
# Elaboration schema



1 I love to cook  
Swedish  
cuisine.

2 My recipe for  
Swedish  
meatballs is  
legendary.

# General schema



If you take remove the satellite, the discourse should still be coherent.

If you remove the nucleus, the discourse should lose coherency

(! A prediction we can test!)

# RST – important concepts

4 important objects

2. Nucleus

- Nucleus (N) “More central to the authors purposes”

3. Satellite (S)

4. Relations

5. Schemes

# Relations

- Relations hold between segments (text spans)
- What relations exist is an open question
- Relations are explicitly defined according to the requirements for the
  - Nucleus (N)
  - Satellite (S)
  - Combination of nucleus and satellite
  - And according to their **effect**

# Detailed relation

(from Mann & Thompson, 1988, p. 251)

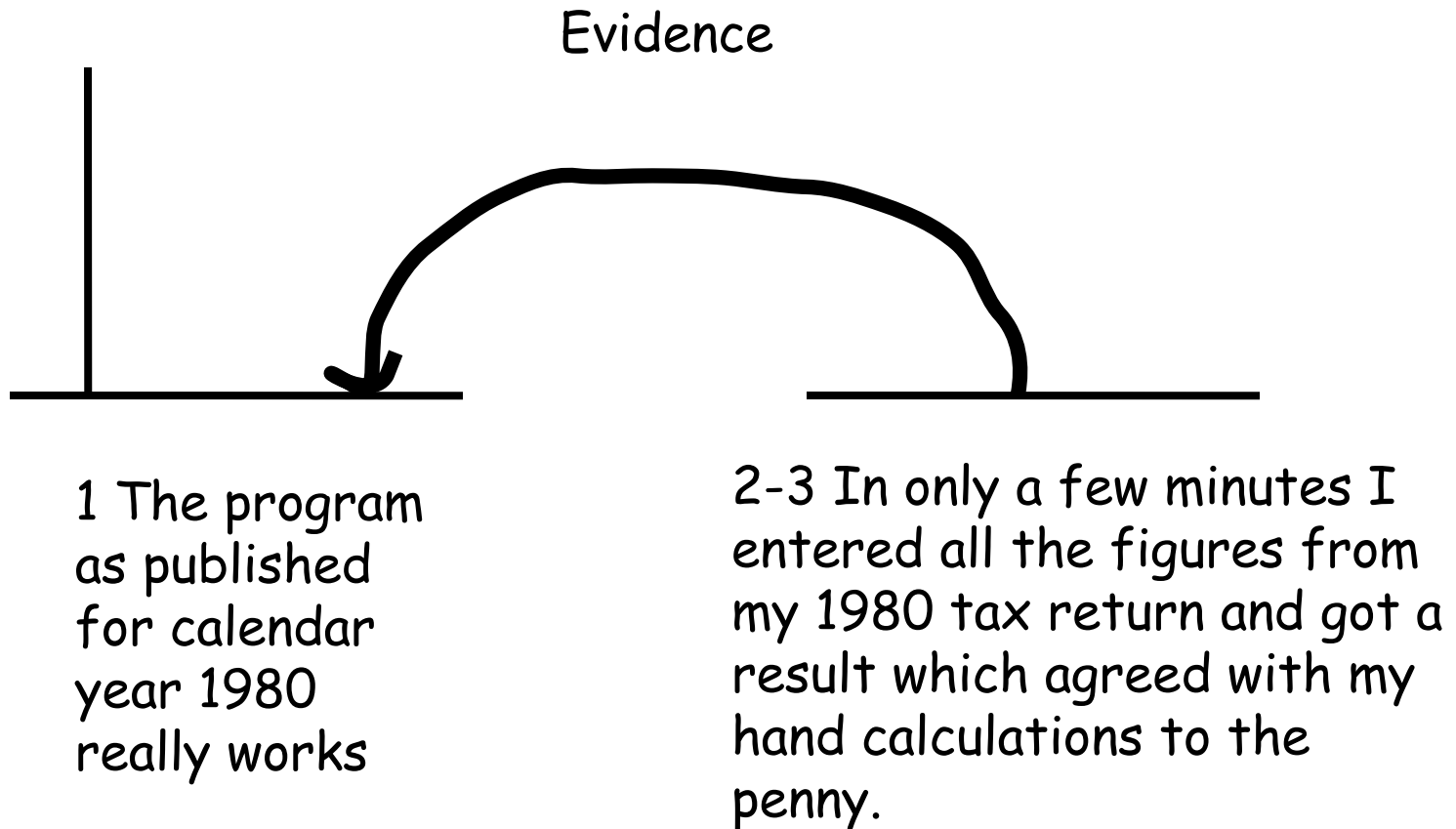
<b>Relation name:</b>	Evidence
<b>Constraints on N:</b>	R (Reader) might not believe N to a degree satisfactory to W (Writer)
<b>Constraints on S:</b>	R believes S, or will find it credible
<b>Constraints on the N + S combination:</b>	R's comprehending S increases R's belief in S
<b>The effect</b>	R's belief of N is increased
<b>Locus of effect:</b>	N



## Example of *Evidence*

1. The program as published for calendar year 1980 really works.
2. In only a few minutes, I entered all the figures from my 1980 tax return and got a result which agreed with my hand calculations to the penny.

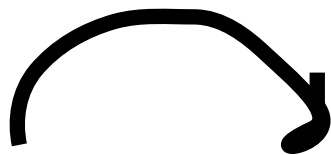
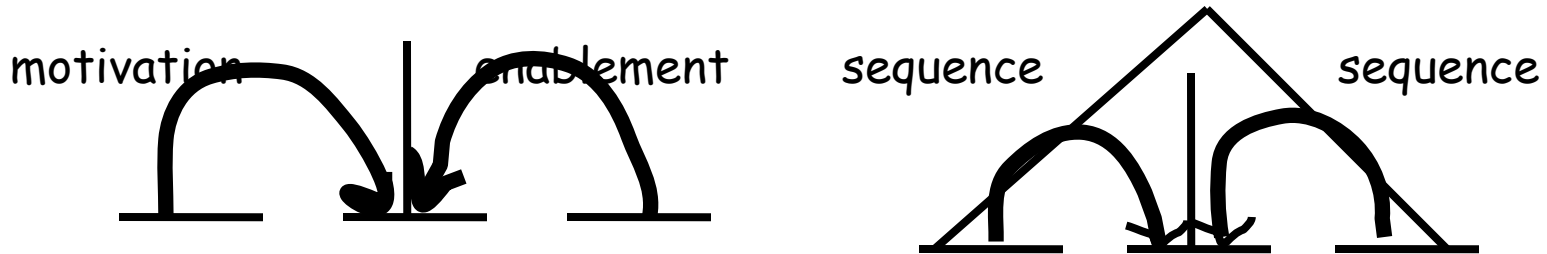
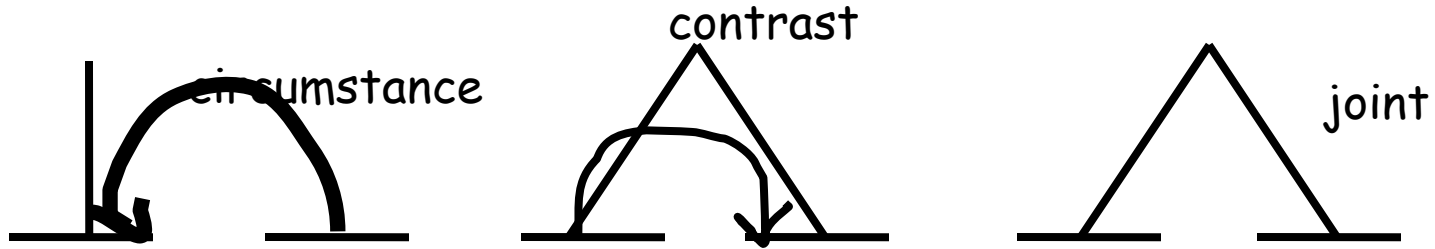
# Evidence schema



# Other relations

Relation name	Nucleus	Satellite
<b>Background</b>	Text whose understanding is being facilitated	Text for facilitating understanding
<b>Elaboration</b>	Basic information	Additional information
<b>Preparation</b>	Text to be presented	Text which prepares the reader to expect and interpret the text to be presented

# 5 schemas



Arrows:  
relations



"nuclear  
span(s)"



## Lactose & Lactase

### Lactose and Lactase

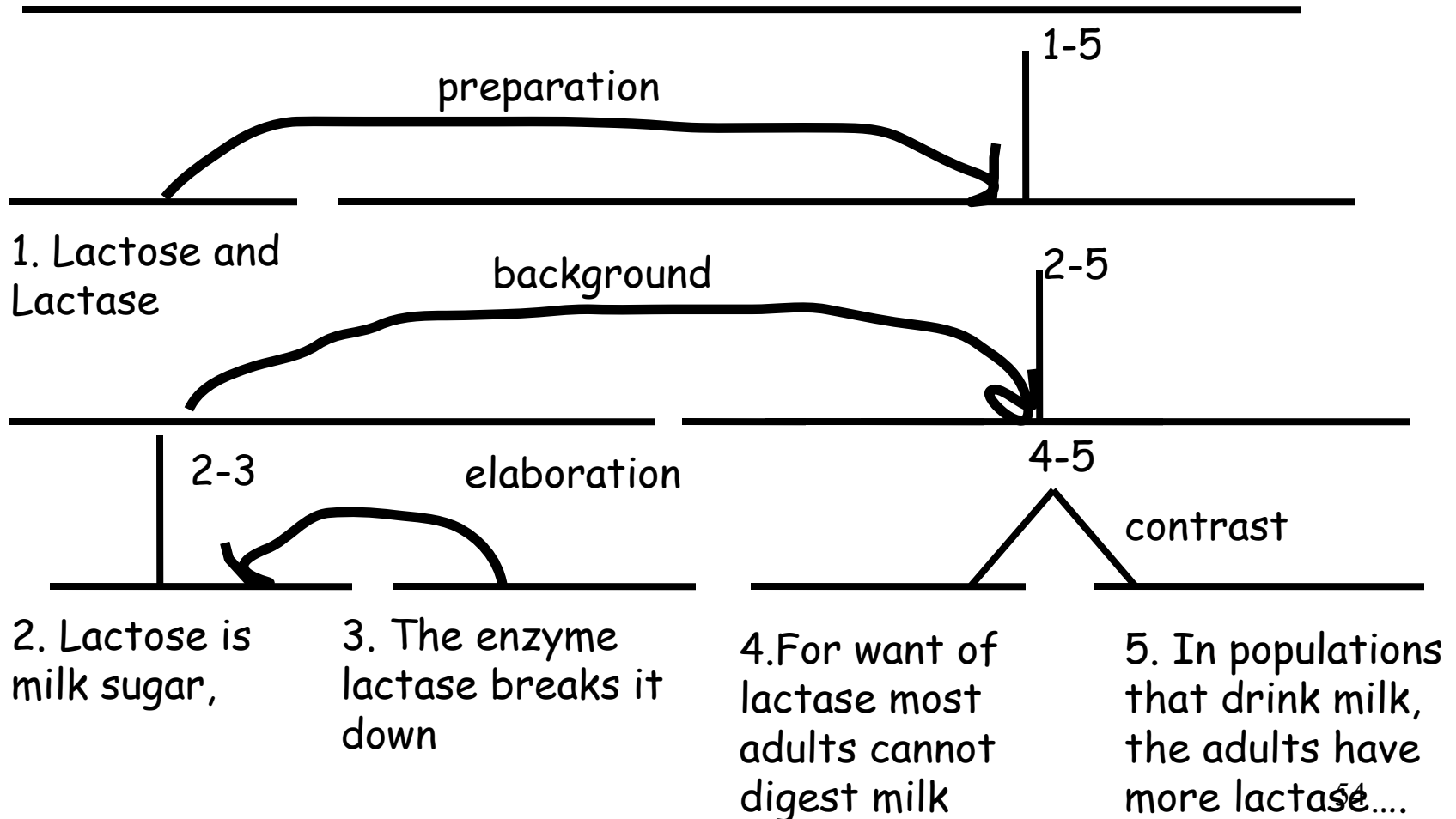
Lactose is milk sugar; the enzyme lactase breaks it down.

For want of lactase most adults cannot digest milk.

In populations that drink milk, the adults have more lactase, perhaps through natural selection.



# Lactose and lactase



# How do you do an RST analysis?

- Segment it
- Apply schemas
  - Divide into more general categories
- Determine relations
  
- Mann & Thompson emphasize that it is possible to have multiple interpretations of the same text

# Comparing RST to G&S

- G&S is very “AI”-like: reasoning about the state of “agents” in the dialogue is core to the theory
  - Understanding a discourse is done by understanding the discourse purpose
  - The why and how of understanding the DP is not part of the theory
- RST is more descriptive, an analysis tool
  - Yet definitions of relations still refer to speaker intentions



# Problems with RST

- RST doesn't claim that one text has one definitive analysis
  - this basic assumption has to be ignored to make it computational useful, but what if the RST people are right?
- The RST notion of nuclearity doesn't exist in most other theories of discourse structure
  - It doesn't relate directly to e.g. subordination and coordination so it is very abstract
  - Recently challenged in Stede (2008)
  - Analysis beyond local relationships is very, very difficult (true of all theories though!)

# Corpora with rhetorical structure annotation

- **RST Discourse Treebank**
  - (Carlson et al.)
  - 385 articles from the Wall Street Journal portion of the Penn Treebank
- **PTDB 2.0, Penn Discourse Treebank**
  - (Marcus et al. 1993)
  - entire Wall Street Journal Treebank annotated with discourse information (2304 articles)
  - Based on D-LTAG, a lexically grounded theory of discourse structure

```
( Nucleus (leaf 53) (rel2par span) (prom 53) (text As a result, industry
operated out of small, expensive, highly inefficient industrial units.) )
    ( Satellite (span 49 52) (rel2par result) (prom 49)
        ( Nucleus (leaf 49) (rel2par span) (prom 49) (text In the
past, the socialist policies of the government strictly limited the size of
new steel mills, petrochemical plants, car factories and other industrial
concerns) )
            ( Satellite (span 50 52) (rel2par purpose) (prom 50 51)
                ( Nucleus (leaf 50) (rel2par List) (prom 50) (text to
conserve resources)
                    ( Nucleus (span 51 52) (rel2par List) (prom 51)
                        ( Nucleus (leaf 51) (rel2par span) (prom 51) (text
and restrict the profits ) )
                            ( Satellite (leaf 52) (rel2par elaboration-object-attribute-
e) (prom 52) (text businessmen could make.) )
                                )
                            )
                        )
                    )
                )
            )
        )
    )
```

# RST analyses span entire texts

```
( Root (span 1 19) (prom 2)
  ( Nucleus (span 1 15) (rel2par span) (prom 2)
    ( Nucleus (span 1 5) (rel2par span) (prom 2)
      ( Nucleus (span 1 4) (rel2par span) (prom 2)
        ( Nucleus (span 1 3) (rel2par span) (prom 2)
          ( Nucleus (span 2 3) (rel2par span) (prom 2)
```

# PDTB 2.0

- Identifies which explicit connectives are signaling rhetorical relations between adjacent clauses
  - not trivial: most connectives are ambiguous between rhetorical and non-rhetorical meaning
- Lexically based,
  - fairly theory neutral
- Appropriate discourse connectives are *added* to related clauses that do not have an explicit connective
- No embedded relations, all local discourse relations
- Allows comparisons between discourse structure and syntactic structure because the Penn Treebank is fully syntactically analyzed

## Explicit connectives and their arguments (PDTB 2.0)

*In the past, the socialist policies of government strictly limited the size of new steel mills, petrochemical plants, car factories and other industrial concerns to conserve resources and restrict the profits businessmen could make. As a result, **industry operated out of small, expensive, highly inefficient industrial units.** (0629)*

## Implicit connectives and their arguments (PDTB 2.0)

Motorola is fighting back against junk mail. *So much of the stuff poured into its Austin, Texas, offices that its mail rooms there simply stopped delivering it.*

Implicit = so Now, thousands of mailers, catalogs and sales pitches go straight into the trash. (0989)

**Implicit RESULT** marked by adding “so”

# Non-insertability of implicit connectives

- Three reasons why adjacent clauses can't have an implicit connective inserted
  - Case 1: **AltLex** A discourse relation is inferred but inserting an implicit connective is redundant because other lexical information (in the form of a non-connective expression) is signaling the same relation

*New rules force thrifts to write down their junk to market value, then sell the bonds over five years. **AltLex = (result) That's why Columbia just wrote off \$130 million of its junk and reserved \$227 million for future junk losses.***



## Non-insertion (2)

- Case 2: **EntRel**: Coherence is created by an entity-based relation
- *Hale Milgrim, 41 years old, senior vice president, marketing at Elecktra Entertainment Inc., was named president of Capitol Records Inc., a unit of this entertainment concern. EntRel Mr. Milgrim succeeds David Berman, who resigned last month.*

## Non-insertion (3)

**NoRel:** Neither discourse nor entity-based relation is inferred.

*Jacobs is an international engineering and construction concern.* **NoRel** Total capital investment at the site could be as much as \$400 million, according to Intel.

<b>PDTB Relations</b>	<b>No. of tokens</b>
Explicit	18459
Implicit	16224
AltLex	624
EntRel	5210
NoRel	254
<b>Total</b>	<b>40600</b>

# Have the theories/resources been applied to any practical tasks?

- Text analysis based applications
  - Text simplification
  - Automatic summarization
    - Nuclear are more important than satellites, more prominent relations more important than embedded relations, etc.
  - Natural language generation
- Anaphor resolution
  - Discourse structure constrains possible antecedents
  - VP-ellipsis antecedents may also be constrained by type of coherence relation
- Developing dialogue systems
- Useful for testing theories of dialogue

# Summary

- Many different elements help make a text or dialogue “coherent”
- Should be very useful for many NLP applications, but very difficult to work with/identify/study
  - it looks like syntactic structures are not isomorphic for different coherence relations
  - it looks like discourse structures might not be limited by tree structures
  - actually automatically recognizing how a discourse fits together requires programs with all the world knowledge of a person...
    - BUT: we can get part of the way with simpler heuristics

# Two readings

- Grosz & Sidner (1986)
  - classic paper. You should be able to do a G&S style analysis for the exam
  - only need to read sections 1-3
- Sporleder & Lascarides (2005)
  - practical attempt to automatically identify coherence relations with “cue phrases” and machine learning
  - (not so easy because most cue phrases are ambiguous)
  - you should know what their method was, and how good their results were, and what they see as the key advantages of their method