
Discourse coherence

Ted Gibson, courtesy of Florian Wolf
March 15, 2005

1. Bill hid John's car keys. He was drunk. It was not the first time that John got that drunk.

2. ?Bill hid John's car keys. He likes spinach. She said it works well.

What makes (1) coherent and (2) incoherent?

Overview

1. Coherence relations – basics / working vocabulary
 - What is the informational structure of a discourse? (discourse = collection of sentences that are in some relation to each other)
2. Implications of coherence – pronoun resolution
 - How does discourse structure influence other aspects of language processing? Example: pronoun resolution
3. Current theories of discourse coherence
 - How can we represent discourse structure? (cf. lecture on sentence structure)
4. Constraints on building discourse structures
 - How can we determine discourse structure? (cf. lecture on parsing)

1.) Coherence relations – basics / working vocabulary

Working vocabulary

Coherence relations in their basic form have a long history:

Aristotle
(4th cent. BC)

Hume
(18th cent.)

Boccaccio
(14th cent.)

Working vocabulary

1. Causal relations:

Cause-Effect:

Working vocabulary

1. Causal relations:

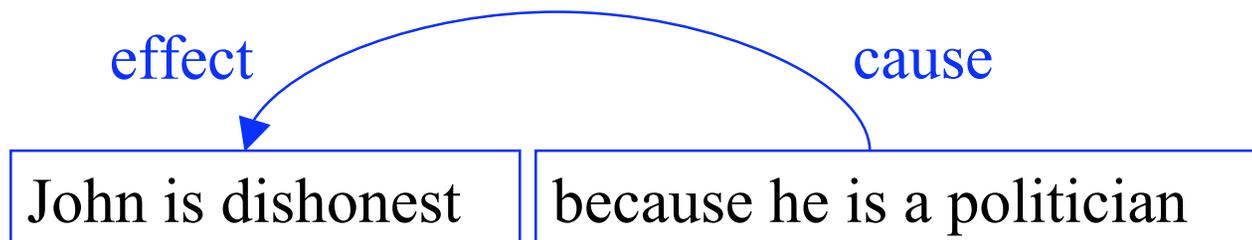
Cause-Effect:

John is dishonest because he is a politician

Working vocabulary

1. Causal relations:

Cause-Effect:



Working vocabulary

1. Causal relations:

Violated Expectation:

John is honest although he is a politician

Working vocabulary

1. Causal relations:

Violated Expectation:

John is honest although he is a politician

Working vocabulary

1. Causal relations:

Violated Expectation:

John is honest although he is a politician

Working vocabulary

1. Causal relations:

Violated Expectation:

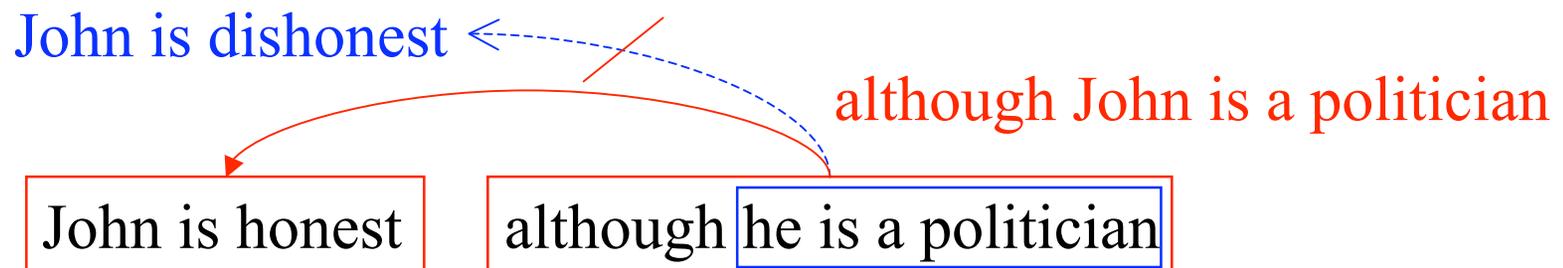
John is dishonest ← John is a politician

John is honest although he is a politician

Working vocabulary

1. Causal relations:

Violated Expectation:



Working vocabulary

1. Causal relations:

Condition:

Working vocabulary

1. Causal relations:

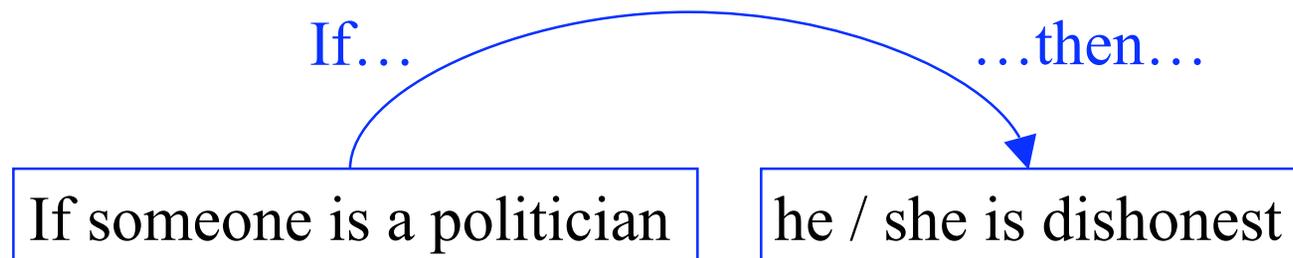
Condition:

If someone is a politician he / she is dishonest

Working vocabulary

1. Causal relations:

Condition:



Working vocabulary

2. Resemblance relations:

Parallel:

Working vocabulary

2. Resemblance relations:

Parallel:

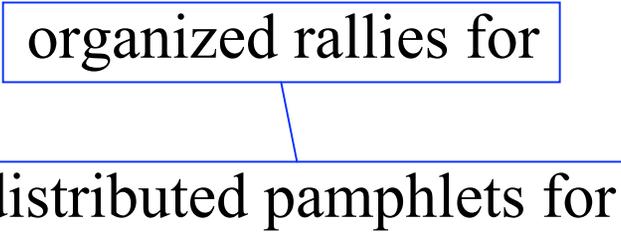
John organized rallies for Clinton,
and Fred distributed pamphlets for him

Working vocabulary

2. Resemblance relations:

Parallel:

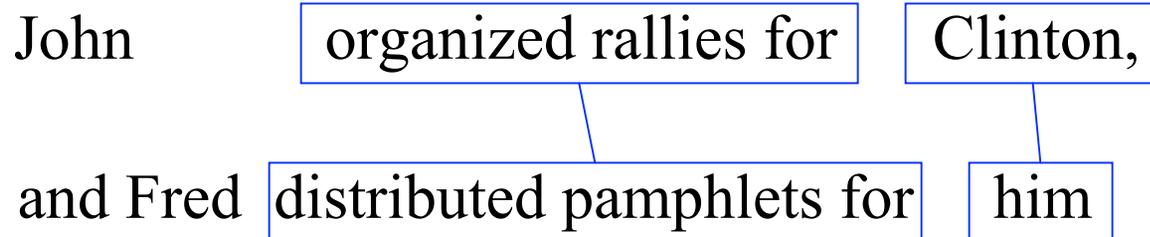
John organized rallies for Clinton,
and Fred distributed pamphlets for him



Working vocabulary

2. Resemblance relations:

Parallel:



Working vocabulary

2. Resemblance relations:

Contrast:

Working vocabulary

2. Resemblance relations:

Contrast:

John supported Clinton,

and Fred cheered for Bush

Working vocabulary

2. Resemblance relations:

Contrast:

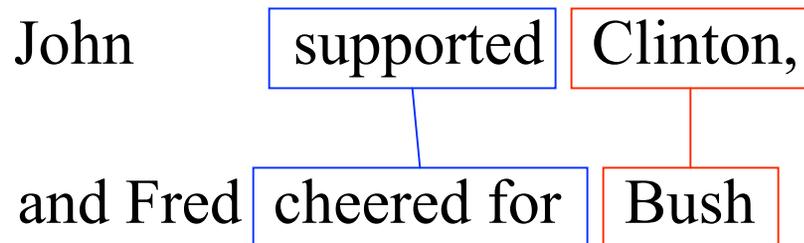
John supported Clinton,
and Fred cheered for Bush

The diagram illustrates a contrast between two verbs. The word 'supported' is enclosed in a blue rectangular box. A vertical line descends from the bottom center of this box to the top center of another blue rectangular box that encloses the phrase 'cheered for'. This visual connection highlights the relationship between the two verbs in the context of the sentences provided.

Working vocabulary

2. Resemblance relations:

Contrast:



Working vocabulary

3. Elaboration relations:

Working vocabulary

3. Elaboration relations:

A political heavyweight resigned in Washington today.

D. Rumpsteak's resignation came as no surprise.

Working vocabulary

3. Elaboration relations:

A political heavyweight resigned in Washington today.

D. Rumpsteak's resignation came as no surprise.

details



2.) Implications of coherence – pronoun resolution

Pronoun resolution

- Pronouns = expressions that depend on something that is mentioned elsewhere
- Cannot be looked up in a dictionary
- Pronouns get their meaning from context - pronouns and coherence?
- Three accounts
 - Centering Theory
 - Parallel Preference
 - Causality-based
- Unifying account and experiment
 - **How we interpret pronouns falls out our attempt to make a discourse coherent**

Centering (Grosz et al 1995)

First, **John** met Susan. Then **he** met Fred.

#First, John met **Susan**. Then **she** met Fred.

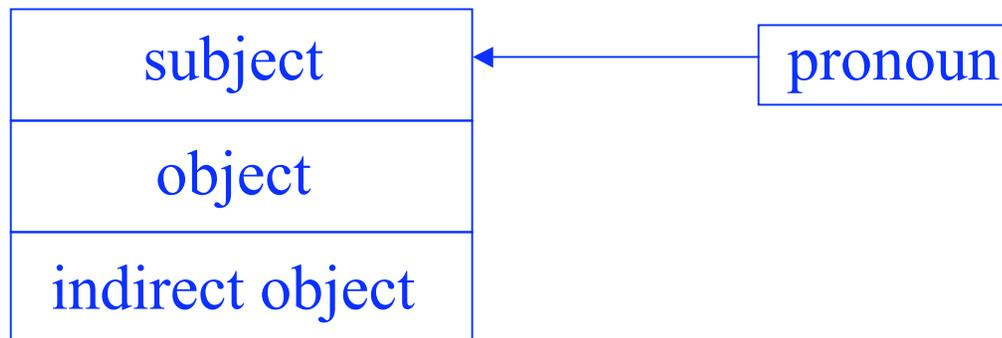
(Under normal declarative intonation, no stress on the pronouns.)

Centering

How Centering explains the contrast:

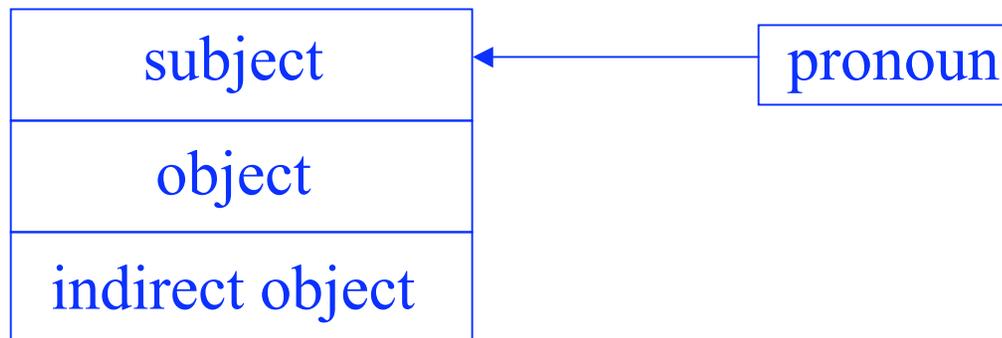
Centering

How Centering explains the contrast:



Centering

How Centering explains the contrast:



- implemented: Lappin & Leass 1994
- (somewhat) tested behaviorally: Gordon et al 1993
- Also: Wundt 1911

Centering

But:

John complimented **Susan**, and Fred congratulated **her**.

#**John** complimented Susan, and Fred congratulated **him**.

Parallel Preference (Smyth 1994)

- Subject pronoun → subject antecedent
- Object pronoun → object antecedent
- Implemented: Kameyama 1986
- Tested behaviorally: Chambers & Smyth 1998

Parallel Preference

John complimented **Susan**, and

Fred congratulated **her**.

Parallel Preference

John complimented **Susan**, and

Fred congratulated **her**.

John complimented Susan, and Fred

congratulated **him**.

Parallel Preference

John complimented **Susan**, and

Fred congratulated **her**.

John complimented Susan, and Fred

congratulated **him**.

First, **John** met Susan.

Then **he** met Fred.

Parallel Preference

John complimented **Susan**, and

Fred congratulated **her**.

John complimented Susan, and Fred

congratulated **him**.

First, **John** met Susan.

Then **he** met Fred.

First, John met **Susan**.

Then **she** met Fred.

Parallel Preference

But:

The **city council** denied the demonstrators the permit because **they** feared violence.

The city council denied the **demonstrators** the permit because **they** advocated violence.

(Winograd 1972)

Causality-based approach

- establish causal inference path to resolve the pronoun
- implemented: Hobbs et al 1990; Appelt et al 1993
- tested behaviorally: Wolf, Gibson & Desmet (2004) – details later

Causality-based approach

- Limitation of causality-based approach: does not apply to parallelism contrast:

Susan complimented **John** and similarly Fred congratulated **him**.

Susan complimented John and similarly Fred congratulated **her**.

Existing accounts - summary

sentence / judgment	Cent. Theory	Par. Pref.	Causality- based
John complimented Susan and Fred congratulated her .	no	yes	no
# John complimented Susan and Fred congratulated him .	no	yes	no
First, John met Susan. Then he met Fred.	yes	yes	no
#First, John met Susan . Then she met Fred.	yes	yes	no
The city council denied the demonstrators the permit because they advocated violence.	no	yes	yes
The city council denied the demonstrators the permit because they feared violence.	yes	no	yes

Existing accounts - summary

No earlier account can explain full range of data

→ how can we bring the accounts together?

Kehler 2002

Use different approaches to pronoun processing
under different coherence relations

Kehler 2002

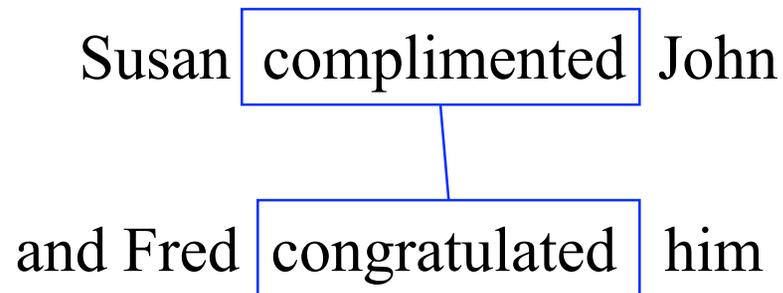
- **Resemblance** - recognize similarities and contrasts between corresponding sets of parallel entities:

Susan complimented John

and Fred congratulated him

Kehler 2002

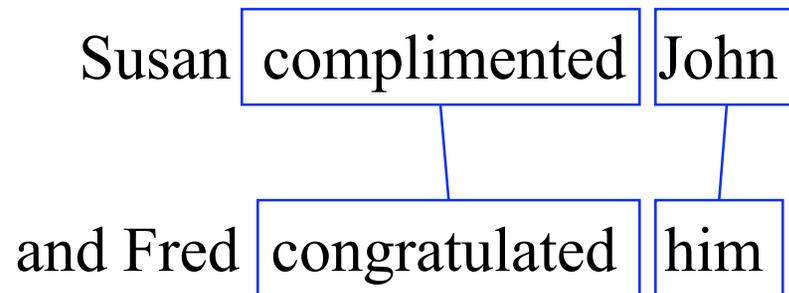
- **Resemblance** - recognize similarities and contrasts between corresponding sets of parallel entities:



- Supported by semantic and structural similarities

Kehler 2002

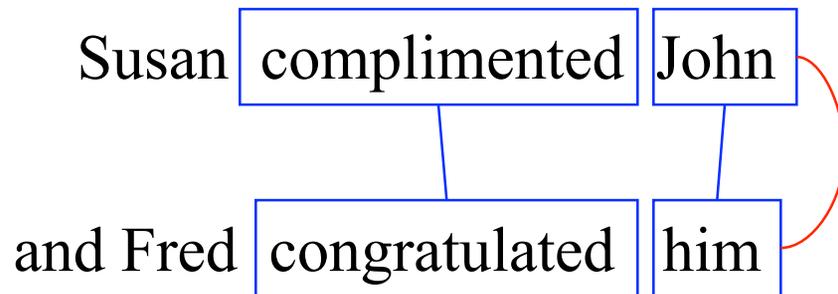
- **Resemblance** - recognize similarities and contrasts between corresponding sets of parallel entities:



- Supported by semantic and structural similarities

Kehler 2002

- **Resemblance** - recognize similarities and contrasts between corresponding sets of parallel entities:



- Supported by semantic and structural similarities

Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here

Susan defeated John and so Fred congratulated her.

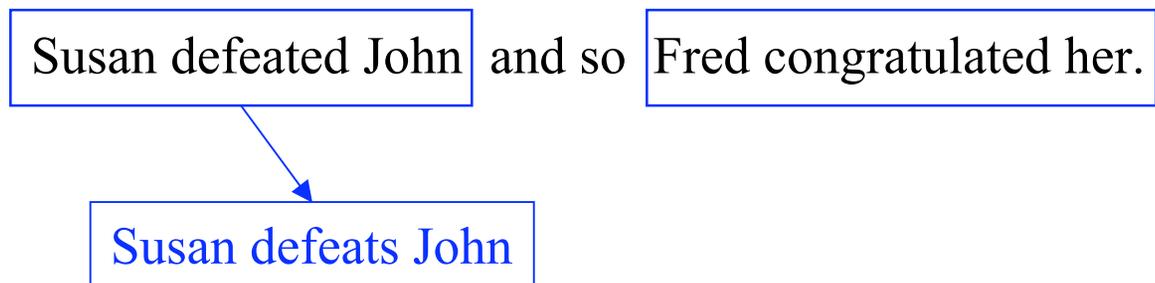
Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here

Susan defeated John and so Fred congratulated her.

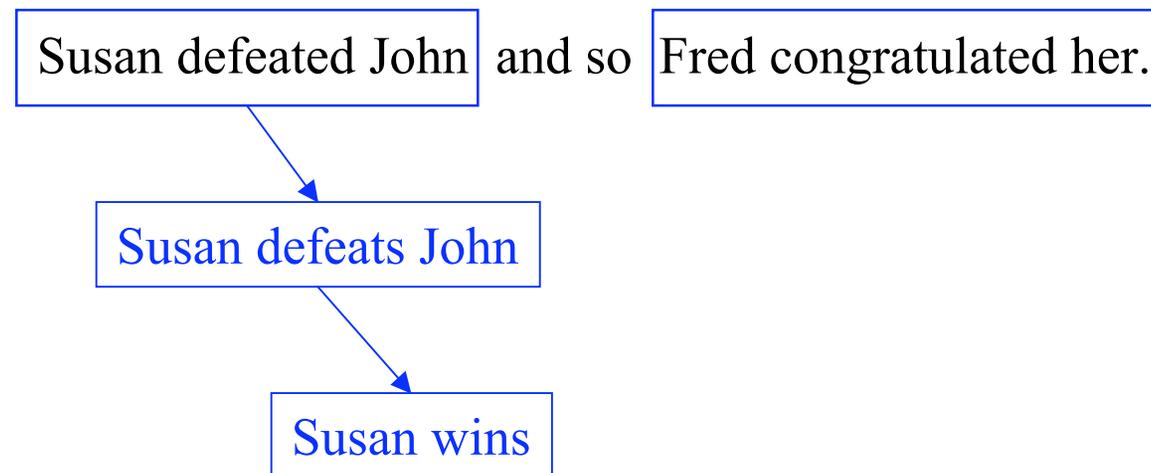
Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here



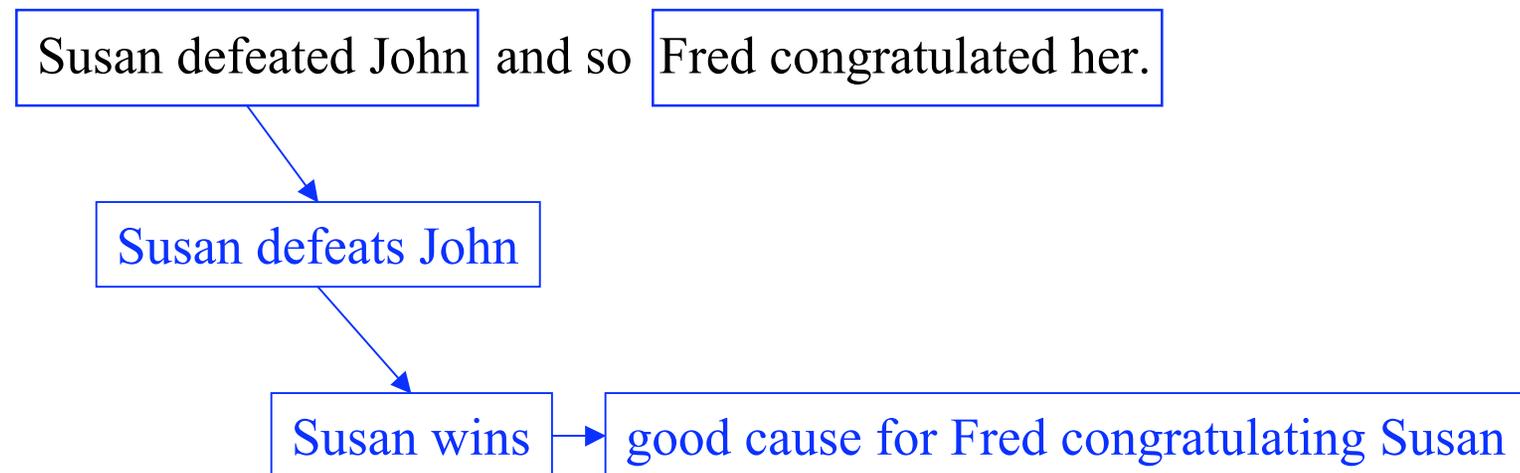
Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here



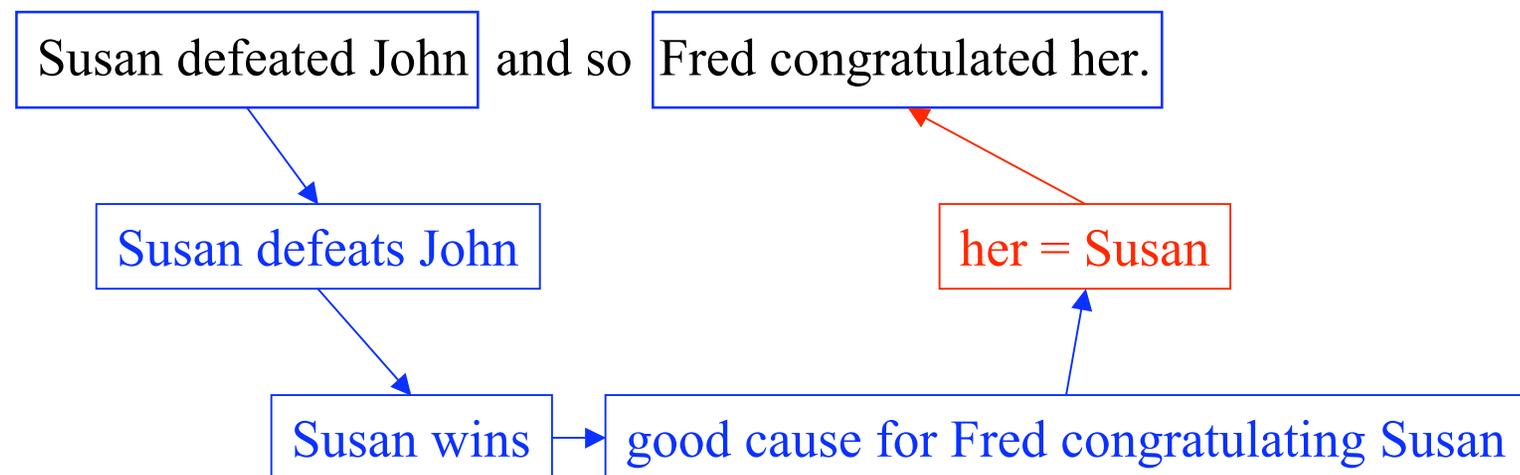
Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here



Kehler 2002

- **Cause-Effect** – establish causal inference path between utterances
- Sentence structure does not help here



Kehler 2002

Resemblance

- look at sentence structure
- parallel preference approach for pronouns

Kehler 2002

Resemblance

- look at sentence structure
- parallel preference approach for pronouns

Cause-Effect

- do not look at sentence structure
- causality-based approach for pronouns

Reading time experiment

Resemblance, parallel:

Susan complimented **John** and similarly Fred congratulated **him**.

Resemblance, non-parallel:

Susan complimented John and similarly Fred congratulated **her**.

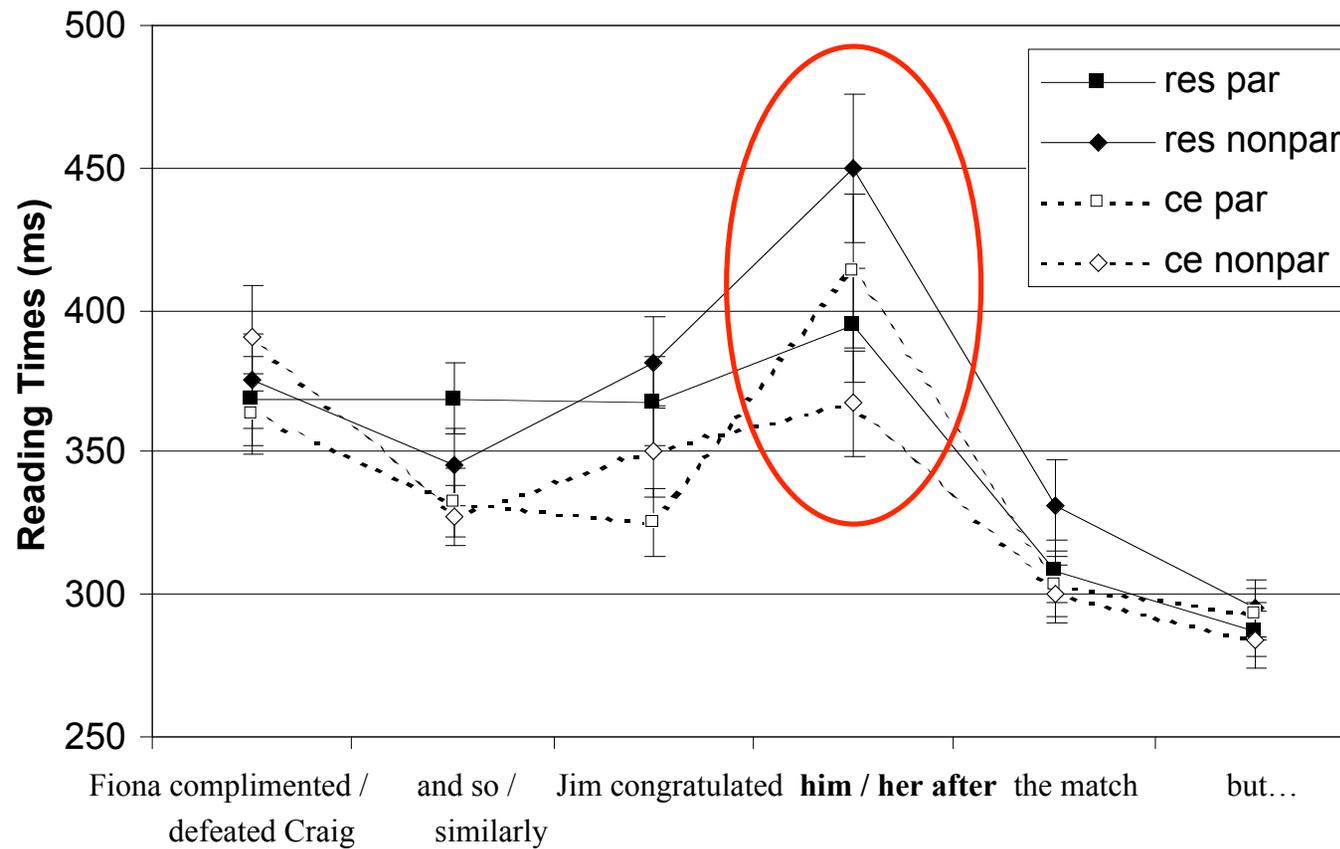
Cause-Effect, parallel (hard):

Susan defeated **John** and so Fred congratulated **him**.

Cause-Effect, non-parallel (easy):

Susan defeated John and so Fred congratulated **her**.

Reading time data



Pronouns - summary

- Purely structural approaches to pronoun resolution are not enough
- Coherence is an important factor in pronoun resolution:
 - Resemblance → parallel preference
 - Cause-Effect → causal inference

3.) Current theories of discourse coherence

Basic considerations

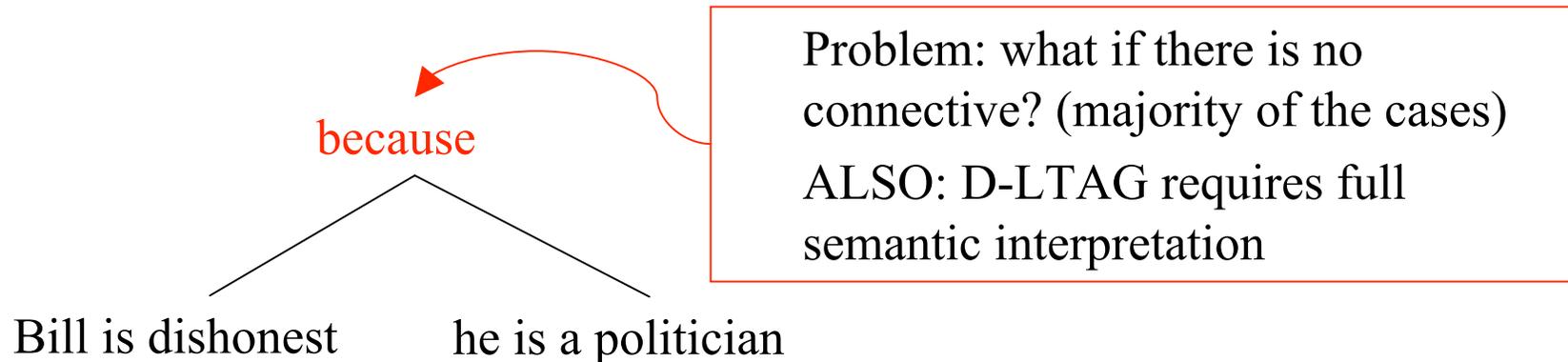
- What kind of discourse do we want to account for?
 - Conversations
 - Task-oriented dialogs
 - **Monologues**
- Basic considerations for an account of discourse coherence
 - Intention of the discourse or informational structure?
 - How many different coherence relations does one need? (cf. parts of speech)
 - Appropriate data structure for representing discourse coherence?
 - Is there something like a “discourse grammar”?

Intentional or informational structure

- Intentional structure
 - Grosz & Sidner 1986; cf. Centering Theory
 - speaker intentions determine discourse structure
 - speaker intentions = key to discourse understanding
 - Problem: determining intention: **Too hard**
- Informational structure
 - cf. slides in the beginning: *cause-effect, resemblance, elaboration, etc*
 - Some real-time evidence for validity: pronoun experiments (difference *cause-effect / resemblance*)

Informational structure accounts

- **D-LTAG (Discourse Lexicalized Tree Adjoining Grammar) – Forbes et al 2001**
- Coherence structure tree
- Uses connectives (*because, although, while, etc*)



Informational structure accounts

- **RST (Rhetorical Structure Theory) – Mann & Thompson 1988; Marcu 2000**
 - Tree structure
 - Large number of coherence relations (some versions of RST accounts have more than 400):
 - Cause-Effect:
 - volitional
 - non-volitional
 - Elaboration:
 - object-attribute
 - whole-part
 - process-step

Informational structure accounts

- **Hobbs 1985**
 - No tree structure
 - Small number of coherence relations – cf. slides in the beginning

Still open questions

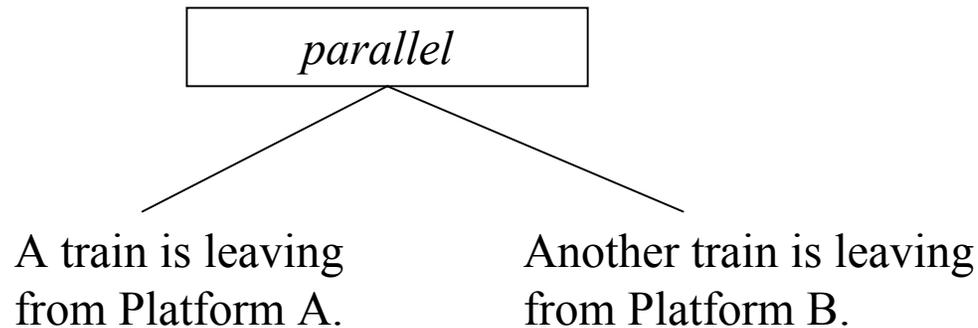
- **Large or small number of coherence relations?**
 - Hovy & Maier 1998: small sets include large sets
 - Use small set
 - more abstract and generalizable
 - easier to code
 - Hobbs, not RST
- **Trees or no trees?**

Trees or no trees?

1. A train is leaving from Platform A.
2. Its destination is Milan.
3. Another train is leaving from Platform B.
4. Its destination is Zürich.

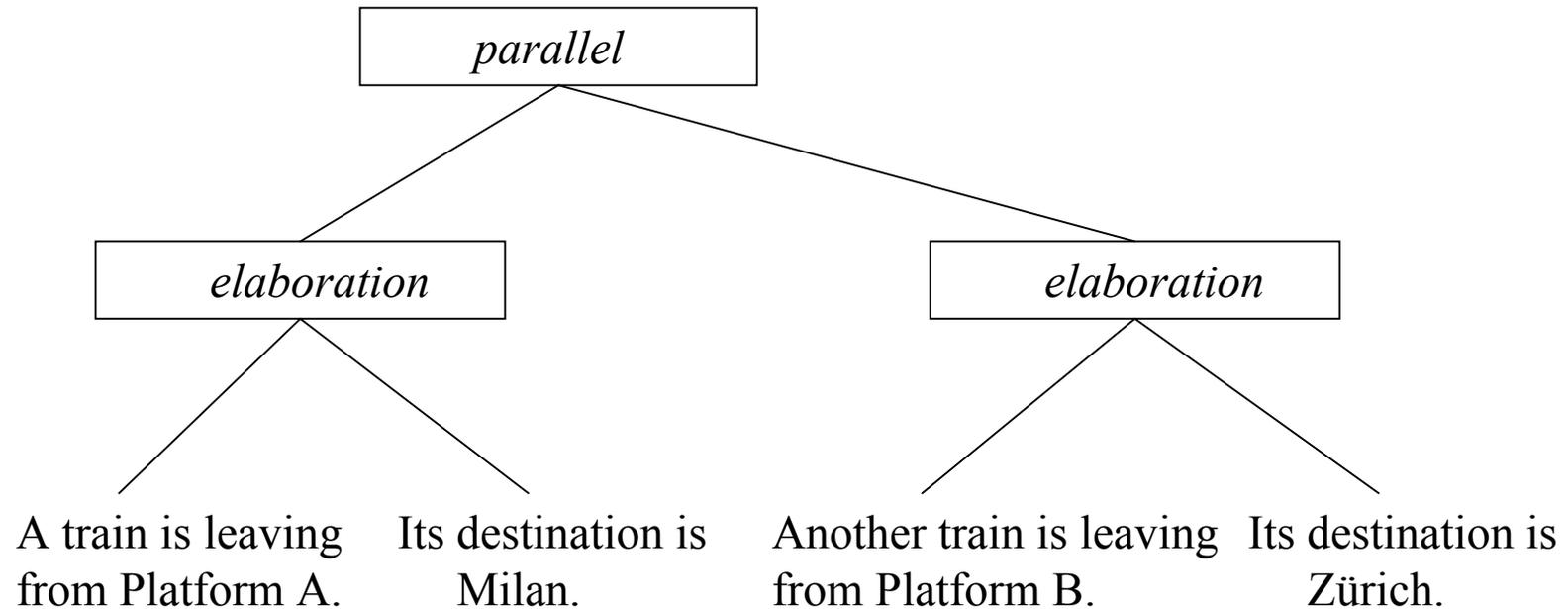
Trees or no trees?

1. A train is leaving from Platform A.
2. Its destination is Milan.
3. Another train is leaving from Platform B.
4. Its destination is Zürich.



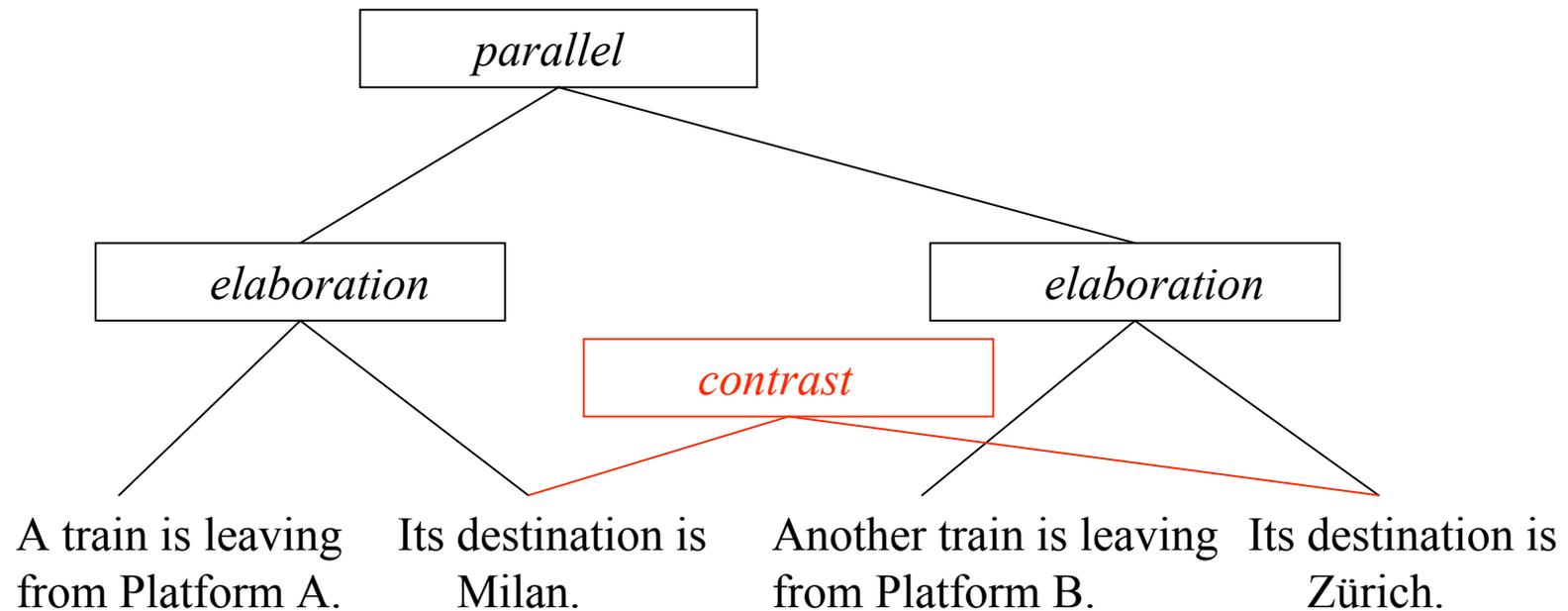
Trees or no trees?

1. A train is leaving from Platform A.
2. Its destination is Milan.
3. Another train is leaving from Platform B.
4. Its destination is Zürich.



Trees or no trees?

1. A train is leaving from Platform A.
2. Its destination is Milan.
3. Another train is leaving from Platform B.
4. Its destination is Zürich.



(20) Example text (modified from SAT practicing materials)

0. Schools tried to teach students history of science.
1. At the same time they tried to teach them hwo to think logically and inductively.
2. Some success has been reached in the first of these aims.
3. However, none at all has been reached in the second.

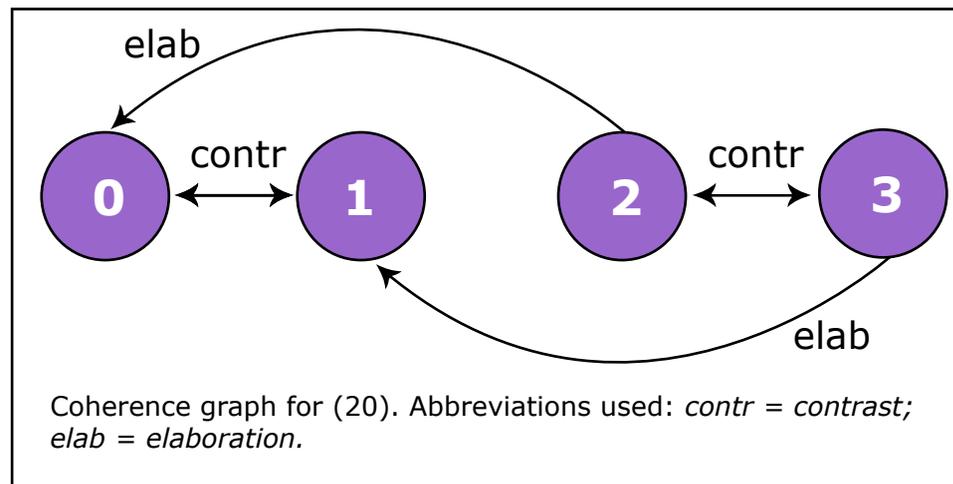


Figure by MIT OCW.

(21) Example test (constructed)

0. Susan wanted to buy some tomatoes
1. and she also tried to find some basil
2. because her recipe asked for these ingredients.
3. The basil would probably be quite expensive at this time of the year.

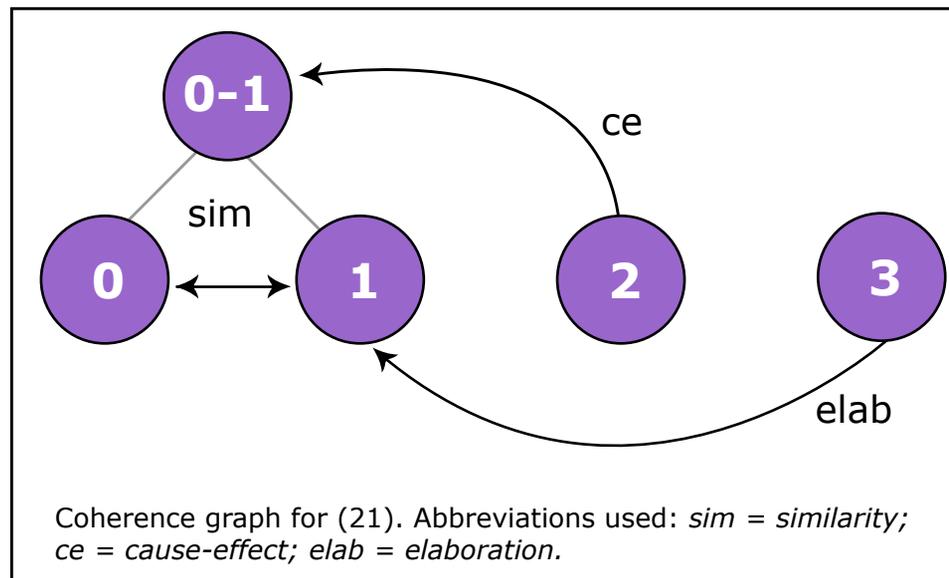


Figure by MIT OCW.

(25) Example test (from ap890103-0014; AP Newswire 1989 corpus; Harman & Liberman (1993))

0. "Sure I'll be polite,"
1. promised one BMW driver
2. who gave his name only as Rudolf.
3. "As long as the trucks and the timid stay out of the left lane."

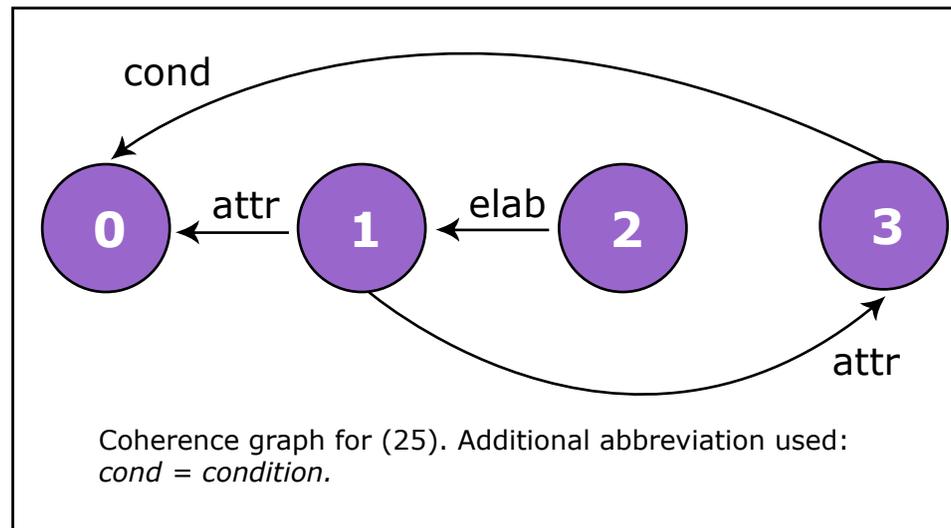


Figure by MIT OCW.

No trees: Lots of crossed dependencies

Correlation between text length, and number of crossed dependencies, from 135 texts (Wolf & Gibson, in press)

Graph removed for copyright reasons.

No trees: Lots of nodes with multiple parents

Correlation between text length, and number of nodes with multiple parents, from 135 texts (Wolf & Gibson, in press)

Graph removed for copyright reasons.

No trees...

Images of tree and chainsaw removed for copyright reasons.

Discourse grammar?

- No trees (Wolf & Gibson, in press)
 - Lots of crossed dependencies: On average, 12.5% of the arcs need to be deleted in order to form graphs without crossed dependencies
 - Lots of nodes have multiple parents: 41.2% on average.
 - No context-free phrase structure grammar ($VP \rightarrow V NP$, $NP \rightarrow Det N$, etc)
- Are there other constraints?
 - Linguistic constraints
 - Signal phrases / connectives
 - Anaphoric relations
 - Lexical chains
 - Discourse graph patterns
 - Connect any node with any other node?
 - Certain sub-patterns more likely than others?

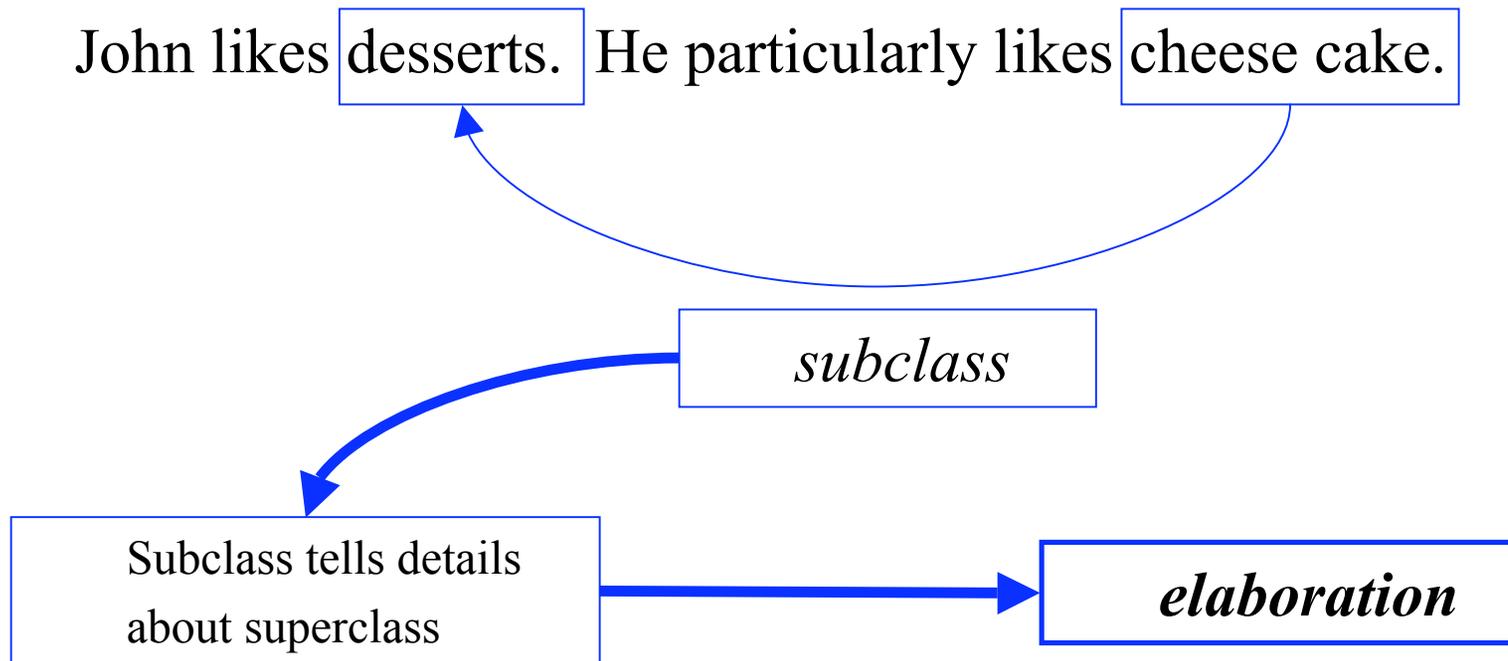
4.) Constraints on building discourse structures

Linguistic constraints

- Signal words / connectives
 - *Because, although, while, since, etc*
 - BUT: only about 15-20% of coherence relations are signaled by connectives (Schauer 2000)
- Anaphoric relations
 - IF anaphoric relation THEN coherence relation, and vice versa (Cristea et al 1999)

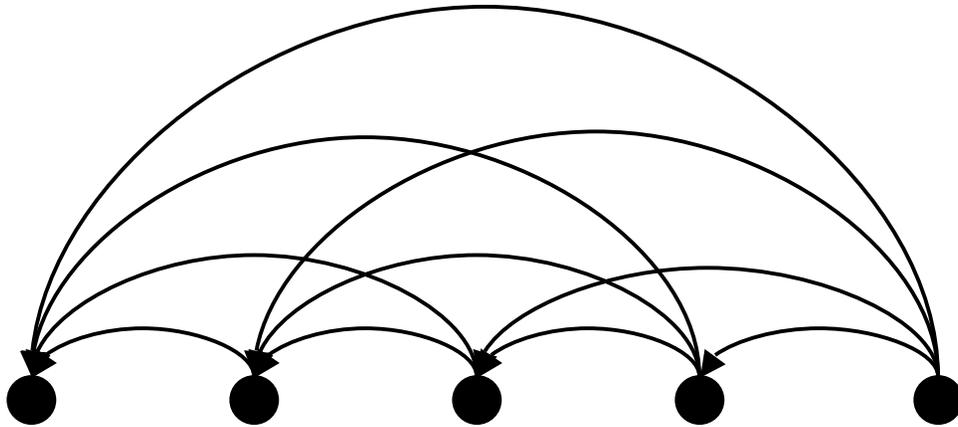
Linguistic constraints

- Lexical chains (Barzilay 1997)



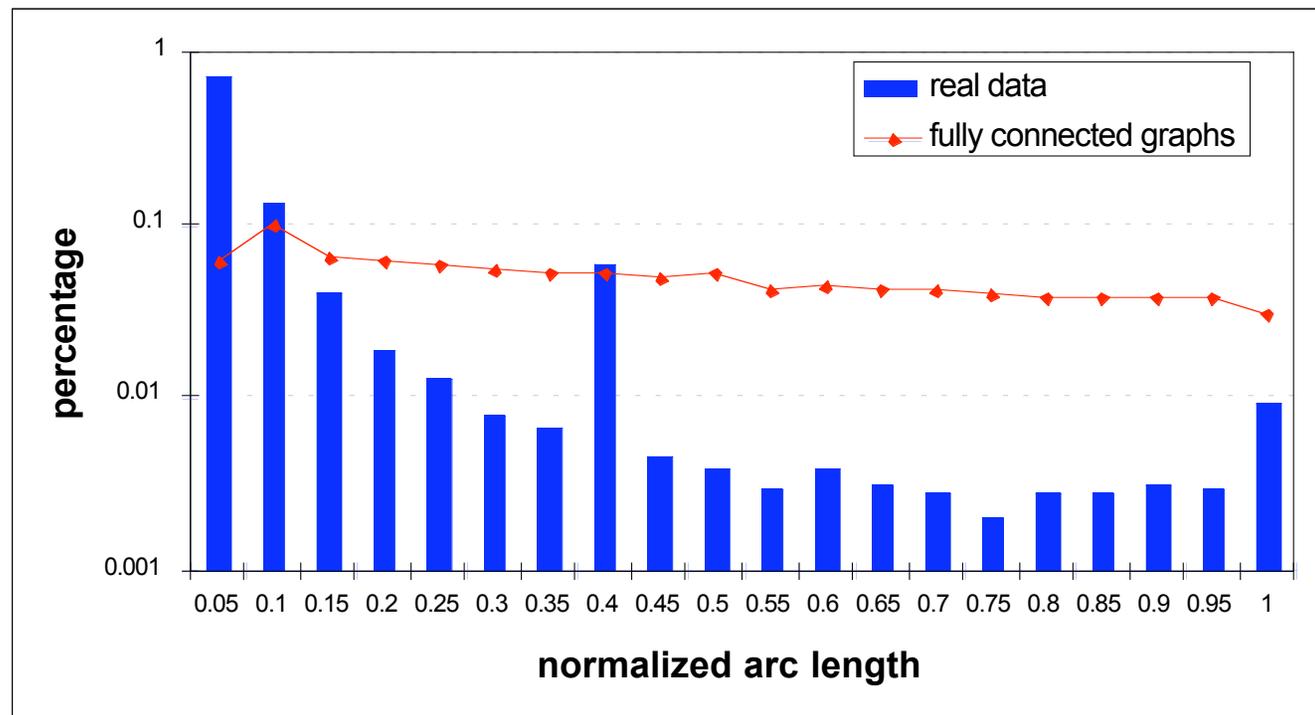
Graph structure patterns

- Connect any node with any other node?



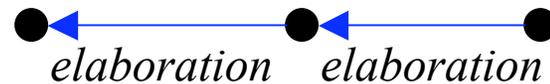
Graph structure patterns – arc length

- Short-distance relations are preferred
- Statistics from coherence structures of 135 texts
- Normalized arc length: divide absolute arc length by maximum possible arc length, given position in text



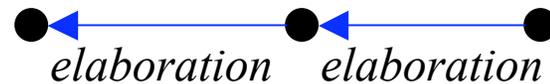
Graph structure sub-patterns

- Are there certain patterns that are more frequent than others?

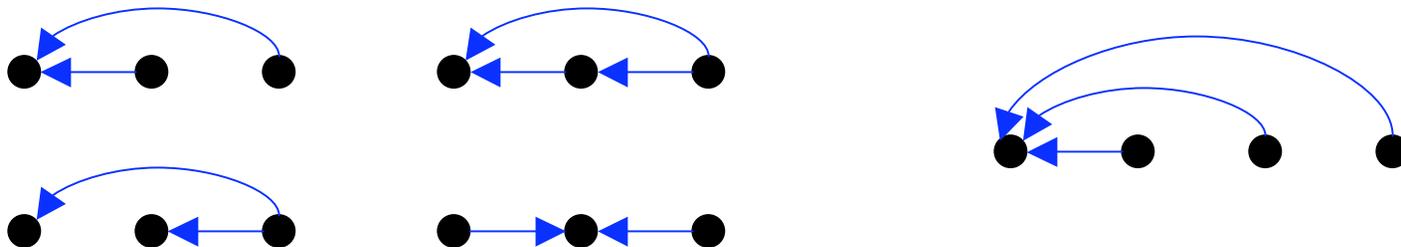


Graph structure sub-patterns

- Are there certain patterns that are more frequent than others?



- Not:



Summary

Summary

- Coherence influences other linguistic processes (e.g. pronoun processing)
- Accounts of discourse structure
 - Informational structure of monologues
 - Small set of coherence relations
 - Data structures: directed graphs instead of trees
- Determining discourse structure
 - Linguistic cues
 - Connectives
 - Anaphoric links
 - Lexical chains
 - Structural patterns
 - Arc lengths
 - Frequent sub-patterns