# Preparatory course for beginning M.Sc. students: Pragmatics 1: Discourse and Reference

Caroline Sporleder

Universität des Saarlandes

Wintersemester 2009/10

06.10.2009

# Overview of Pragmatics Lectures

### Today's Lecture (Caroline Sporleder):

- What is Pragmatics?
- Dimensions of Discourse Structure (linguistic, intentional, informational, focus)
- Models of Discourse
- Referring Expressions

#### Thursday's Lecture (Magdalena Wolska):

- Grice's Maxims of Conversation
- Speech Acts
- Presuppositions
- Dialogue

#### **Background Reading**

Daniel Jurafsky & James H. Martin: Speech and Language Processing, Chapters 18 & 19

What is Pragmatics?

### What is Pragmatics?

#### **Semantics**

context-independent meaning of utterances

- lexical semantics (meaning of words): hypernymy-hyponymy (dog vs. animal), homonymy/word-senses (bank vs. bank) . . .
- meaning of sentences, propositions, truth values . . . (e.g. The dog barks  $\Rightarrow \exists x (dog(x) \land bark(x))$ )

#### **Pragmatics**

context-dependent meaning of utterances

- linguistic context: discourse, dialogue
- situational context: discourse participants, time, location etc.



#### **Deixis**

- He has a new job.
- Stop doing that!
- The weather was nice yesterday.

#### **Deixis**

- He has a new job. (Who has a new job?)
- Stop doing that!
- The weather was nice yesterday.

#### Deixis

- He has a new job. (Who has a new job?)
- Stop doing that! (Doing what?)
- The weather was nice yesterday.

#### **Deixis**

- He has a new job. (Who has a new job?)
- Stop doing that! (Doing what?)
- The weather was nice yesterday. (When was the weather nice?)

#### **Implicatures**

meaning implied in discourse context

A: Do you have the time?

#### **Implicatures**

meaning implied in discourse context

A: Do you have the time?

B: \*Yes.

#### **Implicatures**

meaning implied in discourse context

A: Do you have the time?

B: \* Yes.

#### Intended Meaning:

Do you happen to know the time and if so could you please tell me?

#### **Implicatures**

meaning implied in discourse context

A: Do you know what time it is?

B: I think I just heard the milkman.

#### **Implicatures**

meaning implied in discourse context

A: Do you know what time it is?

B: I think I just heard the milkman.

#### Intended Meaning:

I think I just heard the milkman. We both know that the milkman usually comes at a quarter to eight, so it must be a quarter to eight.

#### **Implicatures**

meaning implied in discourse context

A: Where is Bill?

B: There's a blue Ford outside Sue's house.

#### **Implicatures**

meaning implied in discourse context

A: Where is Bill?

B: There's a blue Ford outside Sue's house.

#### Intended Meaning:

There's a blue Ford outside Sue's house. We both know that Bill drives such a car, so I assume he's at Sue's.

Discourse and Discourse Structure

### What is a Discourse?

#### Discourse:

a coherent sequence of utterances.

### What is a Discourse?

#### Discourse:

a coherent sequence of utterances.

How is "coherence" defined?

### Example: Coherence

Reconstruction work will begin next month on a shrine in the Iraqi city of Samarra. There is so much anger in Burma right now, particularly about the brutal treatment of the monks. I visited the worst-hit areas in the north of Ghana and neighbouring Togo, which are more used to battling drought than floods.

#### Coherent?

### Example: Coherence

Greek officials hope the new site will boost the country's long campaign for the return of the Elgin Marbles. Crowds of bystanders watched the first of the monuments lifted by cranes at the 2,500-year-old Parthenon. Greece has begun moving the ancient sculptures from the Acropolis in Athens to a new home - a museum at the foot of the hilltop citadel.

#### Coherent?

#### What is a Discourse?

- a sequence of utterances
- but: an arbitraty collection of well-formed utterances is not necessarily a "discourse"
- ⇒ utterance have to cohere ("hang together")
  - topics which are related
  - events which are connected (e.g. cause-result, temporal succession)
  - utterances have to fulfil a purpose in discourse

#### Dimensions of Discourse Structure

Four interdependent aspects of discourse structure:

- Linguistic Structure: linguistic manifestation of discourse structure, e.g., cue words, intonation, gesture, referring expressions etc.
- Informational Structure: how do the different segments of a discourse relate to each other?
- Intentional Structure: each discourse segment fulfils a purpose (why does a speaker/writer make a given utterance in a given form?)
- Focus Structure: which entities are salient at a given point in discourse?

John hid Peter's car keys. He was drunk.

John hid Peter's car keys. He was drunk.

Mary likes vegetables but she hates tomatoes.

John hid Peter's car keys. He was drunk.

Mary likes vegetables but she hates tomatoes.

Dan insulted Tom and then HE hit him.

John hid Peter's car keys. He was drunk.

Mary likes vegetables but she hates tomatoes.

Dan insulted Tom and then HE hit him.

It was John who hid Peter's car keys.

#### Linguistic form

often an indicator of discourse structure:

- discourse connectives (but, because):
  - $\Rightarrow$  reflect how sentences are related to each other (contrast, explanation etc.)
- referring expressions (she, Mary, a girl, the girl who likes ice-cream . . . )
  - $\Rightarrow$  reflect the status of an entity in the discourse (salient, not-salient, new, old, inferred etc.)

John hid Peter's car keys. He was drunk.

John hid Peter's car keys. He was drunk.

 $\Rightarrow$  The fact that John was drunk explains why he hid Peter's car keys.

John hid Peter's car keys. He was drunk.

 $\Rightarrow$  The fact that John was drunk explains why he hid Peter's car keys.

Mary likes chocolate, Maggie likes crisps

John hid Peter's car keys. He was drunk.

 $\Rightarrow$  The fact that John was drunk explains why he hid Peter's car keys.

Mary likes chocolate, Maggie likes crisps

 $\Rightarrow$  The fact that Maggie likes crisps contrasts with Mary's liking of chocolate.

### Intentional Structure

John hid Peter's car keys. He was drunk.

#### Intentional Structure

John hid Peter's car keys. He was drunk.

Possible intention: explain to listener why John hid Peter's keys (and why Peter was consequently late for work)

#### Intentional Structure

John hid Peter's car keys. He was drunk.

Possible intention: explain to listener why John hid Peter's keys (and why Peter was consequently late for work)

Another Possible intention: outline to listener what consequences John's drunkenness has (and why something must be done about his binge drinking)

#### Focus Structure

Susan would like to go on a holiday. But she needs to find somebody to do her work while she's away. She can't think of anybody to do that. She considered Mike but he's a bit unreliable. Yesterday he forgot to turn up for an important meeting with a client. The client was very annoyed and said she would never do business with Susan's company again.

#### Focus Structure

Susan would like to go on a holiday. But she needs to find somebody to do her work while she's away. She can't think of anybody to do that. She considered Mike but he's a bit unreliable. Yesterday he forgot to turn up for an important meeting with a client. The client was very annoyed and said she would never do business with Susan's company again.

#### Focus Structure

Susan would like to go on a holiday. But she needs to find somebody to do her work while she's away. She can't think of anybody to do that. She considered Mike but he's a bit unreliable. Yesterday he forgot to turn up for an important meeting with a client. The client was very annoyed and said she would never do business with Susan's company again.

#### Focus Structure

Susan would like to go on a holiday. But she needs to find somebody to do her work while she's away. She can't think of anybody to do that. She considered Mike but he's a bit unreliable. Yesterday he forgot to turn up for an important meeting with a client. The client was very annoyed and said she would never do business with Susan's company again.

#### Discourse Coherence

A discourse is perceived as coherent if

- hearer can determine speakers intentions
- hearer can work out informational structure of utterances (based on linguistic clues or inferred by knowledge of speakers intentions)
- focus and linguistic structure fit with intentional and informational structure

John hid Peter's car keys. He likes spinach.

• no linguistic cues for informational structure

- no linguistic cues for informational structure
- possible intentional structure: speaker wants to convey why
  John hid the keys (i.e., because somebody promised to give
  him spinach for it and John would do everything for spinach)

- no linguistic cues for informational structure
- possible intentional structure: speaker wants to convey why
  John hid the keys (i.e., because somebody promised to give
  him spinach for it and John would do everything for spinach)
  - ⇒ informational structure: explanation

- no linguistic cues for informational structure
- possible intentional structure: speaker wants to convey why
  John hid the keys (i.e., because somebody promised to give
  him spinach for it and John would do everything for spinach)
   informational structure: explanation
- another possible intentional structure: speaker wants to convey what an idiot John is (i.e., not only did he hide Peter's keys but he also likes spinach which no normal person does)

- no linguistic cues for informational structure
- possible intentional structure: speaker wants to convey why
  John hid the keys (i.e., because somebody promised to give
  him spinach for it and John would do everything for spinach)
   informational structure: explanation
- another possible intentional structure: speaker wants to convey what an idiot John is (i.e., not only did he hide Peter's keys but he also likes spinach which no normal person does)
  - ⇒ informational structure: continuation

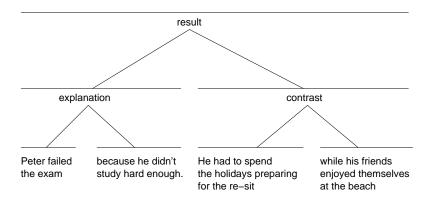
# Modelling Discourse Structure: Rhetorical Structure Theory

### Modelling Discourse Structure

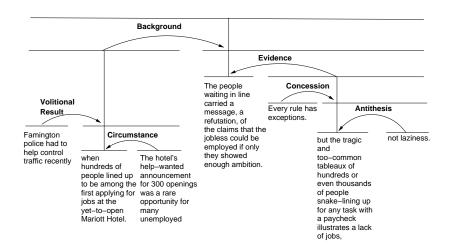
#### Rhetorical Structure Theory (Mann & Thompson, 1987)

- theoretical framework for describing discourse structure (informational structure)
- elementary discourse units (usually clauses) are linked by pre-defined set of 24-30 rhetorical relations
  - ⇒ hierarchical discourse structure (cf. syntax trees)
- RST website: http://www.sfu.ca/rst/

### Example: Simplified RST



### Example: Proper RST



### So, what is it useful for?

- text generation
- text understanding
- text summarisation
- question answering

Referring Expressions

# Referring Expressions vs. Real World Entities



The Treachery of Images, René Magritte, 1928-29

### Referring Expressions vs. Real World Entities

Referent: real world entity to which a linguistic expression refers.



Referring Expression: linguistic expression (usually a noun phrase) used to refer to a referent

George W. Bush, George Bush jnr., the former President of the United States, he, that man, Dubya, . . .

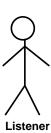
Reference: the process of referring to a referent with a referring expression

#### Discourse Model

**Real World** 

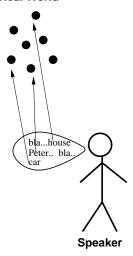


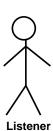




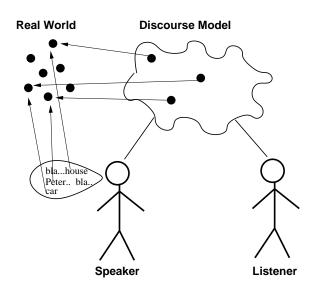
#### Discourse Model

#### **Real World**





#### Discourse Model



#### He claims record

The 22-year-old computer science undergraduate from Bath is claiming a world record for the longest distance ridden on a unicycle in 24 hours.

A unicycling student covered exactly 282 miles at Aberystwyth University's athletics track.

Sam Wakeling was aiming to beat the existing record of 235.3 miles.

#### He claims record

The 22-year-old computer science undergraduate from Bath is claiming a world record for the longest distance ridden on a unicycle in 24 hours.

A unicycling student covered exactly 282 miles at Aberystwyth University's athletics track.

Sam Wakeling was aiming to beat the existing record of 235.3 miles.

#### Unicycling student claims record

A student is claiming a world record for the longest distance ridden on a unicycle in 24 hours.

Sam Wakeling covered exactly 282 miles at Aberystwyth University's athletics track.

The 22-year-old computer science undergraduate from Bath was aiming to beat the existing record of 235.3 miles.

#### Unicycling student claims record

A student is claiming a world record for the longest distance ridden on a unicycle in 24 hours.

Sam Wakeling covered exactly 282 miles at Aberystwyth University's athletics track.

The 22-year-old computer science undergraduate from Bath was aiming to beat the existing record of 235.3 miles.

#### Reference and linguistic form

the linguistic form reflects the saliency of the referent.

#### Reference and linguistic form

the linguistic form reflects the saliency of the referent.

#### Typically:

- new discourse referents are introduced by indefinite NPs
- known/old discourse referents are referred to by definite NPs or pronouns

#### Reference and linguistic form

the linguistic form reflects the saliency of the referent.

#### Typically:

- new discourse referents are introduced by indefinite NPs
- known/old discourse referents are referred to by definite NPs or pronouns
- $\Rightarrow$  I saw a cat. The cat/It was black.

#### Reference and linguistic form

the linguistic form reflects the saliency of the referent.

#### Typically:

- new discourse referents are introduced by indefinite NPs
- known/old discourse referents are referred to by definite NPs or pronouns
- $\Rightarrow$  I saw a cat. The cat/It was black.

#### But:

- Peter walked towards the house. The door was open.
- He is going to the US for a year. (A to B when C walks by)

referent	discourse-new	discourse-old
hearer-new		
hearer-old		

referent	discourse-new	discourse-old
hearer-new	brand-new	
hearer-old		

**brand-new**: new discourse referent, representing an unknown entity (a man)

referent	discourse-new	discourse-old
hearer-new	brand-new	
hearer-old		

**brand-new**: new discourse referent, representing an unknown entity (a man)

referent	discourse-new	discourse-old
hearer-new	brand-new	_
hearer-old	unused	

**brand-new**: new discourse referent, representing an unknown entity (a man)

unused: new discourse referent, representing a known entity
(Queen Elisabeth)

referent	discourse-new	discourse-old
hearer-new	brand-new	<del>-</del>
hearer-old	unused	evoked

**brand-new**: new discourse referent, representing an unknown entity (a man)

unused: new discourse referent, representing a known entity
(Queen Elisabeth)

**evoked**: referring to an entity which was mentioned before in the discourse (*the 22-year old*) or is present in the situational context (*you*)

referent	discourse-new	discourse-old
hearer-new	brand-new	<del>-</del>
hearer-old	unused	evoked

**brand-new**: new discourse referent, representing an unknown entity (a man)

unused: new discourse referent, representing a known entity
(Queen Elisabeth)

**evoked**: referring to an entity which was mentioned before in the discourse (the 22-year old) or is present in the situational context (you)

**inferrable**: new discourse referent which is related to a known entity. (*Peter walked towards the house.* The door was open.)

#### Example

Chris spent yesterday afternoon in a café. The waitress told him to try the hot chocolate but he ordered a coffee instead. Later he watched a movie with Tom Cruise.

### Reference resolution

The postman stroked the dog. Suddenly he bit him. Who bites whom?

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.

Who had a heard attack?

### Reference resolution

The postman stroked the dog. Suddenly he bit him. Who bites whom?

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.

Who had a heard attack?

### Applications:

- Information extraction
- Question-Answering
- Summarisation
- Machine Translation
- . .

### Co-Reference

- referring expressions (the Queen, the bus, a cat, he . . . ) refer to real world entities
- referring expressions, which refer to the same entity are co-referent

### Co-Reference

#### coreference chain:

a set of referring expressions in a text/discourse which are co-referent

### Anaphor:

an expression referring to a preceding expression (antecedent)

Muriel saw a cat. It was black.

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

#### Coreference Chains:

• {Sophia Loren, she, the actress, her, she}

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }
- {a thunderstorm}

Sophia Loren says she will always be grateful to Bono. The actress revealed that the U2 singer helped her calm down when she became scared by a thunderstorm while travelling on a plane.

- {Sophia Loren, she, the actress, her, she}
- {Bono, the U2 singer }
- {a thunderstorm}
- {a plane}

## Coreference Resolution

Difficulties:

### Coreference Resolution

#### Difficulties:

### Coreference Resolution

#### Difficulties:

- same form 

  ⇒ same referents
   (the cat, Michael Jackson the singer vs. Michael Jackson the British general)

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

 $\Rightarrow$  prosody: she = Mary

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

 $\Rightarrow$  prosody: she = Mary

Jane warned Mary she was in danger.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

 $\Rightarrow$  prosody: she = Mary

Jane warned Mary she was in danger.

 $\Rightarrow$  lexical semantics (warned): she = Mary

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

 $\Rightarrow$  prosody: she = Mary

Jane warned Mary she was in danger.

 $\Rightarrow$  lexical semantics (warned): she = Mary

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.

Jane told Peter he was in danger.

 $\Rightarrow$  Agreement (gender, number etc.): he = Peter

Peter said that John is running the business for himself.

 $\Rightarrow$  syntactic constraints: himself = John

The cat did not come down from the tree. It was scared.

 $\Rightarrow$  selectional preferences:  $it = the \ cat$ 

Jane told Mary she was in danger.

 $\Rightarrow$  salience (subject position): she = Jane

Jane told Mary SHE was in danger.

 $\Rightarrow$  prosody: she = Mary

Jane warned Mary she was in danger.

 $\Rightarrow$  lexical semantics (warned): she = Mary

Tony Blair met President Yeltsin. The old man had just recovered from a heart attack.

 $\Rightarrow$  world knowledge: the old man = Yeltsin

## Co-reference Resolution Techniques

#### Before 1990 ...

- co-reference resolution = pronoun resolution
- rule-based (hand-crafted rules)

### After 1990 ...

- corpus-based (co-occurrence statistics, machine learning)
- co-reference resolution for non-pronominal expressions (definite NPs, bridging)

### **Demos**

```
Lappin & Leass (1994):
http://www-appn.comp.nus.edu.sg/%7Erpnlpir/cgi-bin/
JavaRAP/JavaRAPdemo.html
Mitkov (2002)
http://clg.wlv.ac.uk/demos/MARS/
```



### What to take home from this lecture ...

• meaning is often context-dependent

### What to take home from this lecture . . .

- meaning is often context-dependent
- ② linguistic form is to some extent influenced by context

### What to take home from this lecture . . .

- meaning is often context-dependent
- Iinguistic form is to some extent influenced by context
- **3** discourse is a coherent sequence of utterances

### What to take home from this lecture . . .

- meaning is often context-dependent
- Iinguistic form is to some extent influenced by context
- discourse is a coherent sequence of utterances
- discourse has linguistic structure, intentional structure, focus structure, and informational structure