

# Einführung in die Computerlinguistik: Semantik

Manfred Pinkal / Sebastian Pado

# Semantik ... für Dialogverarbeitung [1]

*Wann fängt heute die CL-Vorlesung an?*

*Sie beginnt um 11.*

## Semantik ... für Dialogverarbeitung [2]

- *Ist der Flug um 2 nach London pünktlich?*
- *Ihr Flug nach London Heathrow geht pünktlich um 14:05 ab und steht zum Einsteigen bereit.*
- *Und der nächste?*
- *Der nächste Flug nach London Gatwick geht um 15.20.*
- *Gibt es auch einen nach Heathrow?*

# Semantik ... für Maschinelle Übersetzung

- *Geht es um 2 bei Ihnen?*
- *Is 2 p.m. ok for you?*
- *Can we meet at 2 in your office?*

## Semantik ... für Textverstehen und Fragebeantwortung

*Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 Transportflugzeuge des Typs Hercules C130J erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.*

- *Wieviel kosten die Flugzeuge, die Lockheed an Großbritannien verkauft hat?*

# Semantik: Grundlegende Fragestellungen

- Wie kann Bedeutungsinformation grundsätzlich dargestellt werden? (Semantische Repräsentation)
- Insbesondere: Wie werden Wortbedeutungen erfasst? (lexikalische Semantik)
- Wie wird die Bedeutung von komplexen Ausdrücken (Sätzen) auf der Grundlage von Wortbedeutungen und der syntaktischen Struktur berechnet? (Semantikkonstruktion)
- Wie ergibt sich die spezifische Bedeutung einer Äußerung im Kontext? Wie verändert eine Äußerung den Kontext? (Diskurssemantik)
- Insbesondere: Wie wird Mehrdeutigkeit aufgelöst? (Ambiguitätsresolution)
- Wie kommt man von der spezifischen Äußerungsbedeutung zur relevanten Information? (Inferenz, semantische Auswertung)

# Semantische Phänomene [1]

- Quantoren und Skopus
  - (1) *Zwei Fremdsprachen spricht jeder Linguist.*
  - (2) *Alle sitzen in einem Boot*
  - (3) *Jeder sitzt in einem Boot*
  - (4) *Ein Teilnehmer von jedem Seminar hält zwei Referate.*

## Semantische Phänomene [2]

- Anaphern
  - (5) *Hans hat Peter nicht begrüßt. Er ist beleidigt.*
  - (6) *Hans mag seinen Hund, obwohl er ihn manchmal beißt.*



## Semantische Phänomene [3]

- Ellipsen

(9) *Hans verwöhnt seinen Hund, und Peter seine Katze.*

(10) *Hans verwöhnt seinen Hund, und Peter auch.*

(11) *Und der nächste?*

# Semantische Phänomene [4]

- Metapher und Metonymie
  - (12) Börsenkurse auf Talfahrt
  - (13) Der Betzenberg tobt
  - (14) Ich stehe in der Einfahrt
  
  - (15) Fahren wir nach Stuttgart!
  - (16) Treffen wir uns nach Ostern!
  - (17) Treffen wir uns nach Stuttgart!

# Lexikalische Semantik

*Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 **Transportflugzeuge** des Typs Hercules C130J erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.*

- *Wieviel kosten die **Maschinen**, die Lockheed an Großbritannien verkauft hat?*

# Lexikalische Semantik

*Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 **Abfangjäger** erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.*

- *Wieviel kosten die **Flugzeuge**, die Lockheed an Großbritannien verkauft hat?*

# WordNet

- eine große lexikalische Datenbank für das Englische (z.Zt. >150.000 Wörter; dt. Version: GermaNet, >100.000 Wörter)
- Hierarchische Struktur des Lexikons, Ober- und Unterbegriffe (Hyponyme/Hyperonyme)
- Einheiten von WordNet sind „synsets“: Mengen von synonymen Ausdrücken, die Lesarten von Wörtern charakterisieren...

## Beispiel: engl. *case*

- {*case*, *carton*}
- {*case*, *bag*, *suitcase*}
- {*case*, *pillowcase*, *slip*}
- {*case*, *cabinet*, *console*}
- {*case*, *casing* (the enclosing frame around a door or window opening)}
- {*case* (a small portable metal container)}

# WordNet: Oberbegriffe im Nominalbereich

Table 1  
List of 25 unique beginners for WordNet nouns

<i>{act, action, activity}</i>	<i>{natural object}</i>
<i>{animal, fauna}</i>	<i>{natural phenomenon}</i>
<i>{artifact}</i>	<i>{person, human being}</i>
<i>{attribute, property}</i>	<i>{plant, flora}</i>
<i>{body, corpus}</i>	<i>{possession}</i>
<i>{cognition, knowledge}</i>	<i>{process}</i>
<i>{communication}</i>	<i>{quantity, amount}</i>
<i>{event, happening}</i>	<i>{relation}</i>
<i>{feeling, emotion}</i>	<i>{shape}</i>
<i>{food}</i>	<i>{state, condition}</i>
<i>{group, collection}</i>	<i>{substance}</i>
<i>{location, place}</i>	<i>{time}</i>
<i>{motive}</i>	

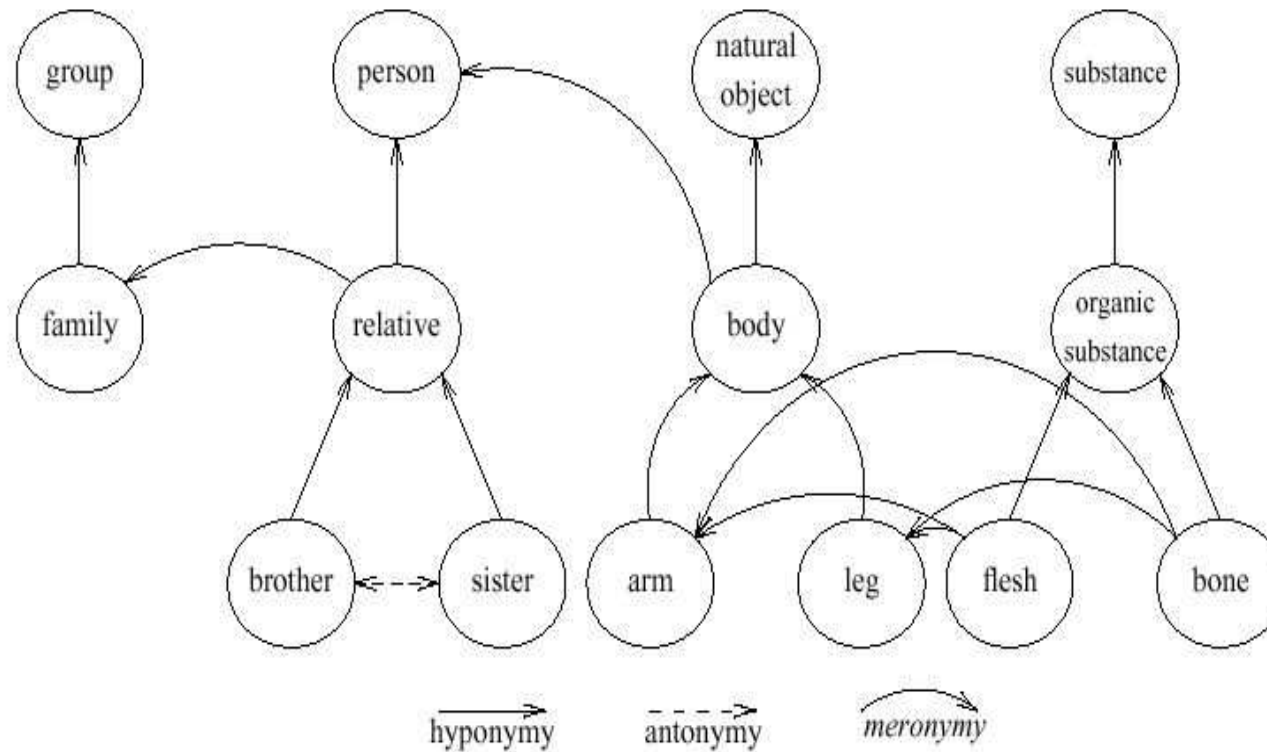
## Weitere Relationen

- Meronymie
  - Teil-Ganzes (*Ast - Baum*)
  - Gruppe - Element (*Baum - Wald*)
  - Stoff - Objekt (*Holz - Baum*)
- Antonymie
  - *Mann - Frau*
  - *Vater - Mutter*
- Funktion
  - Messer - schneiden
  - Wolle - stricken



# WordNet-Teilhierarchie, Beispiel

Figure 2. Network representation of three semantic relations among an illustrative variety of lexical concepts



# Lexikalische Semantik: Prädikat-Argument-Relationen

Der US-Flugzeughersteller Lockheed hat von Grossbritannien **den Auftrag** fuer 25 Transportflugzeuge des Typs Hercules C130J erhalten. Vertreter des Verteidigungsministeriums bezifferten **den Wert** des Auftrags mit umgerechnet 2,5 Mrd. DM.

- Wieviel **kosten** die Flugzeuge, die Großbritannien bei Lockheed **bestellt** hat?

## Thematische Rollen

- *Maria mag den Film*
- *Der Film gefällt Maria*
  
- *Maria hat Peter das Buch gegeben*
- *Peter hat das Buch von Maria erhalten*
  
- *Peter hat Maria Blumen geschenkt*
- *Maria hat Peter mit Blumen beschenkt*

# Thematische Rollen

- Thematische Rollen beschreiben die Mitspieler in einem Ereignis- oder Handlungszusammenhang, unabhängig von deren grammatischer Oberflächenrealisierung.

# Thematische Rollen, Beispiele

- Agens
- Thema/ Patient/ Object
- Rezipient
- Instrument
- Ausgangspunkt (source)
- Ziel (goal)
- Benefizient
- Experienter

## Thematische Rollen: Satzanalysen

- *[Peter]<sub>ag</sub> hat [Maria]<sub>rec</sub> [das Buch]<sub>pat</sub> gegeben*
- *[Maria]<sub>rec</sub> hat [das Buch]<sub>pat</sub> [von Peter]<sub>ag</sub> erhalten*
- *[Maria]<sub>rec, go</sub> hat [das Buch]<sub>pat</sub> [von Peter]<sub>ag, so</sub> erhalten*
- *[Maria]<sub>ag, go</sub> hat [Peter]<sub>ag, so</sub> [das Buch]<sub>pat</sub> genommen*

# Syntax-Semantik-Schnittstelle

- Option 2: Explizite Zuordnung von Argumentpositionen und syntaktischen Komplementen im Lexikon:

<i>geben:</i>	Subj:	Agens
	Obj (akk):	Thema
	Obj (dat):	Rezipient
<i>erhalten:</i>	Subj:	Rezipient
	Obj (akk):	Thema
	PObj (von):	Agens

# Thematische Rollen

- Ermöglichen eine systematische Beschreibung des Komplement-Argument-Bezuges
- Erlauben eine abstraktere, bedeutungsnähere Darstellung von Äußerungs- und Dokumentinformation, die die gezielte Suche erleichtert und Inferenzen unterstützt.



# Problem

- Es gibt kein kleines, abgeschlossenes, universell anwendbares Inventar von Rollen, mit denen sich der gesamte Wortschatz beschreiben lässt.

# FrameNet

- Einheiten für die systematische Beschreibung der Wortsemantik sind Frames: Situationstypen mit frame-spezifischen Mitspielern
- Rollen („Frame-Elemente“) sind Frame-spezifisch definiert. Es gibt Frame-übergreifende Generalisierungen von Rollen-Relationen. Diese sind aber nicht obligatorisch.

## Beispiel:

### Commercial Transaction Frame:

- Target-Wörter:
  - buy, sell, pay, spend, cost, charge,
  - price, change, debt, credit, merchant, broker, shop
  - tip, fee, honorarium, tuition
- Frame-spezifische Rollen:
  - Money, Goods, Buyer, Seller
  - Price, „temporal mode of exchange“, Change

# Frame Elements: Commercial Transaction Frame

<u>FE</u>	<u>Example (in italics)</u>
Buyer	<i>Pat</i> <b>bought</b> a new guitar.
Seller	Pat <b>bought</b> a guitar <i>from Kim</i> .
Payment	Kim <b>sold</b> the guitar <i>for \$250</i> .
Goods	Kim <b>sold</b> <i>the guitar</i> for \$250.
Rate	The plumber <b>charges</b> <i>\$20 an hour</i> .
Unit	The plumber <b>charges</b> <i>by the hour</i>

# Frame Elements: Commitment

## FE

Speaker

Addressee

Message

Topic

Medium

## Example (in italics)

*Kim* **promised** to be on time.

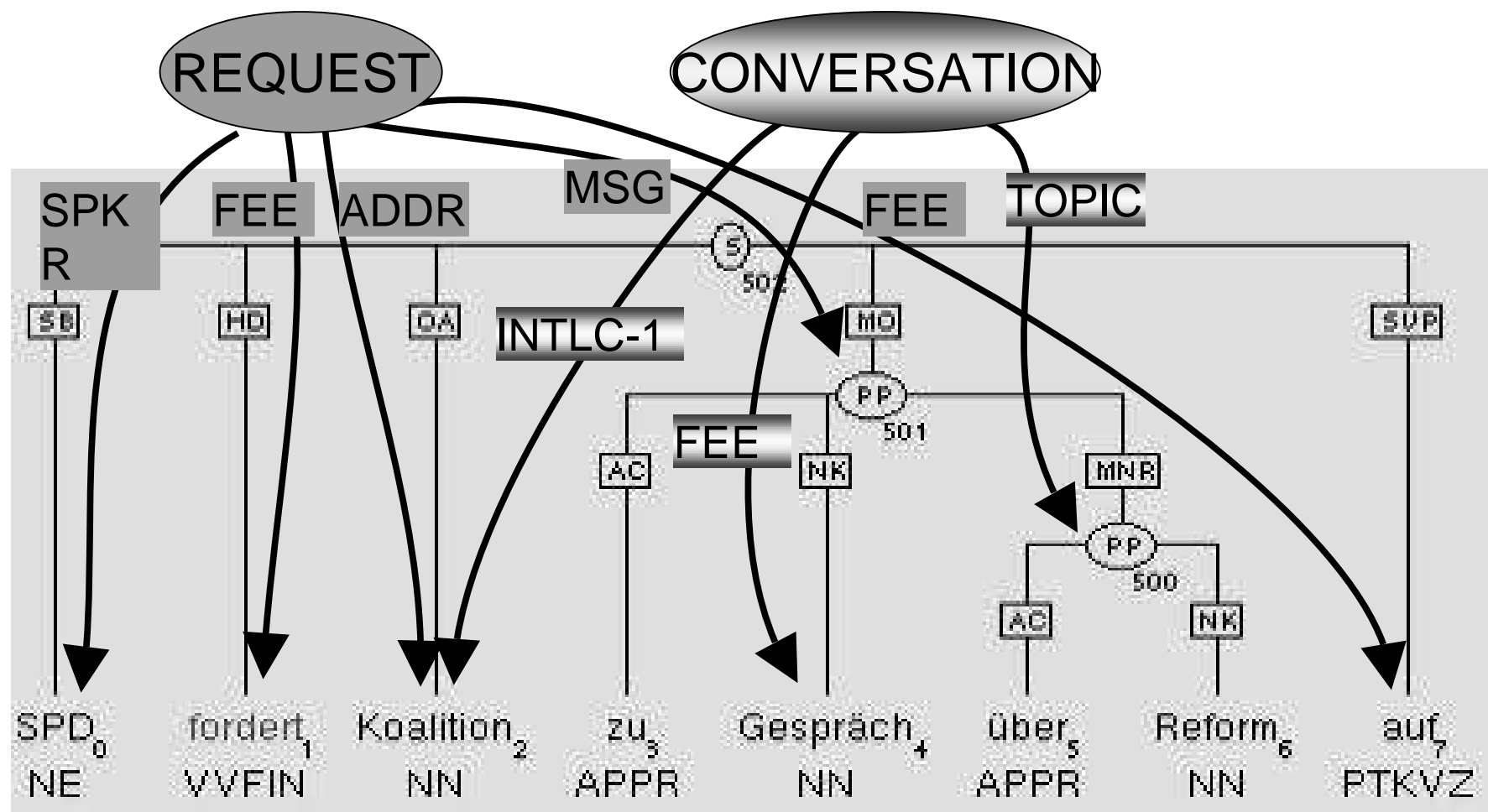
Kim **promised** *Pat* to be on time.

Kim **promised** *to be on time*.

The government broke its  
promise *about taxes*.

Kim **promised** *in writing* to  
sell Pat the house.

# Annotation scheme: an example



# Frame-übergreifende Strukturen

- Oft sind alle oder einige Frame-Elemente in verwandten Frames identisch.
  - Beispiel: Alle Kommunikationsframes, z.B. COMMITMENT, CONVERSATION; REQUEST; STATEMENT teilen sich die Frame-Elemente:  
Speaker, Addressee, Message, Topic, Medium
- Frame-Hierarchie: Abstrakte „Meta-Frames“ vererben Eigenschaften

# Das FrameNet-Projekt, Berkeley

- FrameNet-Datenbank mit ca.500 frames
- Lexikon mit zur Zeit 6000 Einträgen
- Angestrebt: 10-15000 Einträge



# SALSA

- The Saarbrücken Lexical Semantics Annotation and Analysis Project
  - Manuelle Annotation eines deutschsprachigen Korpus: FR/TIGER Korpus mit 1.5 mill. Textwörtern
  - Automatische Annotation eines sehr großen Korpus: Wahrig Korpus, ca. 1 Mrd. Textwörter
  - Entwicklung statistische und Lernverfahren zur automatischen Annotation und zur Extraktion lexikalisch-semantischer Information