

729649 Språk och datorer (VT2025)

Semantik

Lingvistik 6

Marcel Bollmann

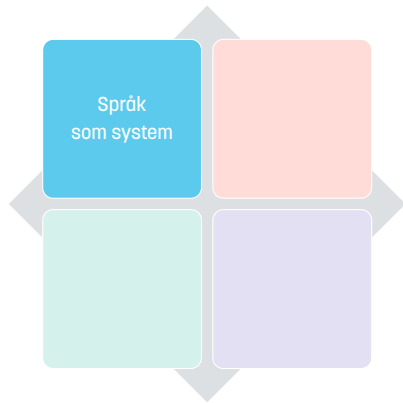
Institutionen för datavetenskap (IDA)



Baserat på material från
Mikael Svensson & Evelina Rennes.

Påminnelse: Språk som system

- **Fonologi:** läran om språkljud och hur de är organiserade
- **Morfologi:** läran om hur ord byggs upp av de "minsta betydelsebärande enheter"
- **Syntax:** läran om hur ord kombineras för att forma fraser, satser och meningar
- **Semantik:** läran om språkliga uttrycks betydelse



Denna föreläsning

■ Vad är semantik?

- Semantikens roll
- Analysnivåer

■ Lexikal semantik

- Centrala begrepp
- Semantiska relationer:
homonymi/polysemi,
synonymi/antonymi,
hyponymi/hypernyymi

■ Satssemantik

- Tematiska roller
- Formell semantik
- Sanningshalt

■ Sammanfattning

... och vad som är viktigt för tentamen

Vad är semantik?



Semantik är läran om språklig betydelse

“In semantic analysis, there is always an attempt to focus on **what the words conventionally mean**, rather than on what an individual speaker might think they mean, or want them to mean, on a particular occasion. [...] Doing semantics is attempting to spell out **what it is we all know** when we behave as if we share knowledge of the meaning of a word, a phrase or a sentence in a language.”

— Yule (2020)

Semantikens roll i lingvistik

All **språklig analys** relaterar till uttryckens betydelse.

- **Fonologi**: /bil/ och /fil/ har olika betydelser → /b/ och /f/ är fonem
- **Morfologi**: *bil* är en, *bilar* är flera → *-ar* är ett pluralsuffix
- **Morfosyntax**: betydelsekriteriet för ordklasser
- **Syntax**: *han såg [mannen med teleskopet]* och *han såg mannen [med teleskopet]* har olika betydelser → olika syntaktiska strukturer

Hur skiljer sig semantik från syntax?

- Meningar kan vara **syntaktiskt korrekta, men semantiskt udda.**

Hamburgaren åt upp pojken.

Bordet lyssnar på radio.

Hästen sitter och läser tidningen.



Kan du stänga fönstret?

- **Semantik:** satsbetydelse, icke situationsberoende, ingen koppling till deltagarna
 - *'Har du förmågan för att stänga fönstret?'*
- **Pragmatik:** avsedd betydelse, situationsberoende, deltagarnas tolkningar
 - *'Stäng fönstret!'*

Komponentanalys

	pojke	flicka	man	kvinna	häst	hingst
mänsklig	+	+	+	+	−	−
vuxen	−	−	+	+	?	+
kvinnlig	−	+	−	+	?	−

Subst [+ mänsklig]

sitter och läser tidningen.

Prototypteori

Vad är det bästa exemplet på en "fågel"?



- **Prototypteori** kategoriserar element som "centrala" eller "perifera" för en kategori.
 - t.ex. kategorin "fågel"
- Graderad modell: en *gråsparv* är en **mer prototypisk** fågel än en *pingvin*.

Semantik på olika analysnivåer

- **Lexikal semantik** avser betydelsen på **ordnivå**.
 - Betydelsen av enskilda ord
 - Semantiska relationer mellan ord
- **Satssemantik** avser betydelsen på **satsnivå**.
 - Betydelsen av hela satser
 - Semantiska relationer mellan satser eller fraser
 - Sanningshalt i uttalanden

Lexikal semantik och satssemantik

Per har precis köpt en husvagn.

- **Lexikal semantik**

- 'Per' är ett namn som typiskt avser en manlig person.
- En 'husvagn' är en typ av 'släpvagn' som har en del 'husinredning' såsom en 'säng'.

- **Satssemantik**

- 'Per' är den agerande personen som utför handlingen '*att köpa en husvagn*'.
- '*Per har precis köpt en husvagn*' medför att '*Per äger en husvagn*'.

Lexikal semantik



Målet med lexikal semantik

- Beskriva **betydelsen för varje enskild ord** i ett språk.
 - ”Vad är betydelsen av *fågel*?”
- Visa hur ordbetydelserna i ett språk är **relaterade** och **organiserade**.
 - ”Hur relaterar betydelsen av *fågel* till betydelsen av *gråsparv*?”

Begrepp och enheter inom lexikal semantik

- Ett **ortografiskt ord** är en ordform ”som man skriver den.”
 - Alla böjningsformer: *lova, lovar, lovade, lovat, lovas, lovande*
- Ett **lexem** är en abstrakt enhet av betydelse.
 - LOVA¹: styra ett fartyg upp mot vinden
 - LOVA²: ge löfte om
 - LOVA³: berömma, prisa
- Ett **lemma** är en ordform som väljs för att representera ett lexem.
 - Infinitivform *lova*
 - Den ”konventionella ordboksformen”

Se: *lova* i SAOB

Lexeme och lemma

Han lovade och har man lovat så har man.

- Samma **lexem**, samma **lemma** (lova)

Han lovade att han skulle lova upp båten.

- Olika **lexem**, båda med samma **lemma** (lova)

Kollokationer

En **kollokation** är en sekvens av ord som ofta uppträder tillsammans.

- Ibland bestämmer lexikonet vilka ord kan eller inte kan användas tillsammans:

*ruttna ägg ↔ *ruttet smör*

**härskna ägg ↔ härsket smör*

- Ofta är det bara vana eller konvention att sätta ihop vissa ord:

svag insats ↔ dålig insats

Idiom

- Ett **idiom**, eller **idiomatiskt uttryck**, är en fast fras där betydelsen inte framgår av de enskilda ordens betydelser.

Han köpte ris i säcken.

Han köpte grisen i säcken.



Tvetydighet och vaghet

- **Lexikal tvetydighet** kan uppstå om det finns flera lexem för ett ord.
 - Jag såg en torsk.
 - Jag såg en torsk i fiskdisken. (TORSK¹: fisk)
 - Jag såg en torsk på Malmskillnadsgatan. (TORSK²: sexköpare)
- **Semantisk vaghet** uppstår när ord avsiktligt är underspecificerade i någon aspekt av sin betydelse.
 - De förflyttade sig 500 meter. — underspecificerad för ”sätt”
 - De sprang/gick/åkte bil/cyklade/ålade/simmade/red 500 meter.

Semantiska relationer

- **Semantiska relationer** beskriver hur lexemens betydelser relaterar till varandra.

De vanligaste semantiska (också: *lexikala*) relationer är:

- 1 **Homonymi och polysemi**
- 2 **Synonymi och antonymi**
- 3 **Hyponymi och hypernyymi**

Homonymi och polysemi

- **Polysemer** är ord som har flera **relaterade** betydelser.
 - Vår skola instiftades på medeltiden. ← institution
 - Vår skola vann i skoltävlingarna. ← elever som bevista institutionen
 - Han bor mittemot skolan. ← byggnaden
- **Homonymer** uttalas och/eller stavas likadant men har **orelaterade** betydelser.
 - En fil är ett verktyg av härdat stål.
 - Hur skickar man en stor fil genom nätet?

Se: *skola* i SAOB

Homografer och homofoner

stavas likadant

uttalas likadant



- Vissa homonymer är både **homografer** och **homofoner**.
 - *fil* (plural *fil*) ↔ *fil* (plural *filer*)
 - *fara* (substantiv) ↔ *fara* (verb)
- Vissa homonymer är **bara homografer**.
 - *köra* (framföra fordon) ↔ *köra* (sjunga i kör)
 - *anden* (i flaskan) ↔ *anden* (i dammen)
 - *tomten* (på julen) ↔ *tomten* (vid huset)

Homonymi eller polysemi?

krona



polysem
(med varandra)



homonym
(med de andra två)

Synonymi och antonymi

- **Synonymer** är ord med **nästan samma** betydelse.
 - *kvinnadam, träffasmötas, friskraskkry*
 - Behöver inte vara synonyma i sammansättningar: *kvinnokamp* $\neq$ *damkamp*
- **Antonymer** är ord med **motsatta** betydelser.
 - Typiskt exempel: *varmkall*
 - "Motsats" kan dock tolkas på olika sätt...

Typer av antonymi

- **Komplementära/binära motsatser**

- gift–ogift (*'inte gift'*), levande–död (*'inte levande'*)

- **Gradmotsatser**

- smal–tjock, varm–kall, glädje–sorg
- lite tjock, mycket varm, men inte **lite död*

- **Reverser** är händelser som sker i olika riktningar.

- höja–sänka, klä på–klä av, hämta–lämna (*inte samma sak som 'att inte hämta'*)

- **Konverser** beskriver samma relation från olika perspektiv.

- leverera–mottaga, framför–bakom, lärare–student

Hyponymi och hypernymi

- **Hyponymer** är **under**ordnade begrepp.



- **Hyper(o)nymer** är **över**ordnade begrepp.

- Hyponymi kan bygga upp en hierarki av ord.
 - *djur* → *ryggradsdjur* → *fågel* → *tätting* → *svala* → *hussvala*
 - *fågel* är hyponym till *djur* — *fågel* är hypernym till *svala*



Satssemantik



Målet med satssemantik

- Beskriva **betydelsen för fraser och satser** i ett språk.
 - ”Vad är betydelsen av *en vacker svala* i meningen *Jag såg en vacker svala*?”
- Visa hur fraser och satser i ett språk är **relaterade**.
 - ”Hur relaterar *bollen* till *Kalle* i meningen *Kalle sparkade bollen*?”

Tematiska roller (också: semantiska roller)

- Lexikala verb delar ut **tematiska roller** till nödvändiga konstituenten i satsen.
 - *sparka* [AGENT, TEMA]
- Dessa är konstanta oavsett hur de uttrycks i satsen.

Kalle sparkade *bollen*
AGENT TEMA
subjekt direkt objekt

Bollen sparkades av *Kalle*
TEMA AGENT
subjekt rektion/
prep. objekt

Vanliga tematiska roller (I)

- **AGENT**: den som utför handlingen (med intention)
 - *Dennis skalade potatisen.*
- **PATIENT**: den/det som påverkas eller förändras av handlingen
 - *Dennis skalade potatisen.*
 - *Olle krossade stenen.*
- **TEMA**: den/det som är inblandad i, förflyttas, eller vars plats beskrivs
 - *Olle kastade stenen.*
 - *Bollen ligger i gräset.*

Obs! Vissa teoretiska ramverk gör ingen åtskillnad mellan PATIENT och TEMA, och kallar dem båda TEMA.

Vanliga tematiska roller (II)

- **UPPLEVARE:** den som har en känsla eller ett tillstånd men inte påverkar handlingen
 - Jag känner mig nervös.
 - Han hörde sirener.
- **INSTRUMENT:** den/det som åstadkommer handlingen
 - Han lämnade av henne med sin bil.
- **BELÄGENHET:** platsen där något/någon är belägen
- **KÄLLA:** platsen varifrån något/någon förflyttas
- **MÅL:** platsen som något/någon förflyttas till

Tematiska roller och satsdelar

- Tematiska roller är ofta **korrelerade med satsdelar**.
 - **AGENT** eller **UPPLEVARE** är ofta **subjekt**. (*vanligt undantag: passivsatser*)
 - **PATIENT** eller **TEMA** är ofta **objekt**.
 - **INSTRUMENT** är ofta **prepositionsobjekt**.
 - **BELÄGENHET**, **KÄLLA** och **MÅL** är ofta **rumsadverbial**.

Varför tematiska roller?

- Tematiska roller kan tydliggöra semantiska **relationen mellan verb och dess argument**.
- Vi kan **klassificera olika verbtyper** beroende på vilka roller de kan ta.
 - *sparka* kräver en AGENT och ett TEMA
- Det finns en **stor variation** i vilka olika begrepp som används för de olika rollerna.

Formell semantik

- Problem inom semantiken: **cirkuläritet** — språket används för att förklara språket.
- **Formell semantik** använder sig av matematiska och logiska verktyg för att beskriva betydelse med ett ”metaspråk.”

Alla igelkottar är vänner med en giraff.

$$\forall x [\text{IGELKOTT}(x) \Rightarrow \exists y (\text{GIRAFF}(y) \wedge \text{VÄN}(x, y))] \quad (1)$$

$$\exists y [\text{GIRAFF}(y) \wedge \forall x (\text{IGELKOTT}(x) \Rightarrow \text{VÄN}(x, y))] \quad (2)$$

- I formeln (1) kan det vara olika giraffer; i (2) är det en och samma giraff.

Scope ambiguity

Allt som glimmar är inte guld.

- Sådana meningar är exempel på **scope ambiguity** ('tvetydighet i omfattningen').
- Oftast beror det på tvetydiga referensgränser av **kvantifikatorer** (t.ex. *alla*, *varje*) eller **negationsord**.

$$\forall x[\text{GLIMMAR}(x) \Rightarrow \neg \text{GULD}(x)] \quad (3)$$

- "Om något glimmar så är det inte (= aldrig) guld." (??)

$$\neg[\forall x(\text{GLIMMAR}(x) \Rightarrow \text{GULD}(x))] \quad (4)$$

- "Det är inte så att om något glimmar så är det (alltid) guld." ← avsedd betydelse

Sanningsrelation mellan satser

- **Sanningstillstånd** kan beskriva satsens sanningshalten.
- **Entailment** (*logisk inferens*): "A är sant" medför att "B är sant"
 - *Kalle slutade röka. $\Rightarrow$ Kalle rökte.*
 - *Några killar spelar fotboll. $\Rightarrow$ Några människor spelar fotboll.*
- **Contradiction** (*motsättning*): A och B kan inte vara sanna samtidigt
 - *Mannen kontrollerar soldatens uniform. $\Rightarrow \Leftarrow$ Mannen sover.*
- I andra fall är A och B **neutrala** med avseende på sina sanningshalter.

Nästan klar!



Lingvistikdelen: Överblick och vad som är viktigt för tentamen

Morfem

Ord

Fras

Sats

Form

- morfemtyper, t.ex. fria och bundna
- rot och stam
- prefix och suffix

F3: Morfologi

Funktion (inom ord)

- avledning och böjning
- sammansättningar
- trädidiagram av morfemstruktur

F3: Morfologi

Form

- ordklass (**alla**)
- finita och infinita verb
- bestämdhet av substantiv och artikel

F2: Morfosyntax

Funktion (inom fraser)

- huvudord
- för-/efterbestämning, rektion
- typer av attribut i nominalfrasen

F4: Syntax I

Betydelse

- lexikal semantik, lexem, lemma
- semantiska relationer (**alla**) (inkl. typer av antonymi)

F6: Semantik

Form

- frastyper (**alla**)
- verbfraser i snäv och vid bemärkelse
- frasstrukturanalys

F4: Syntax I

Funktion (inom andra fraser och satser)

- satsdelar
 - subjekt och predikat
 - typer av objekt
 - subjektiv och objektiv predikativ
 - typer av adverbial

F5: Syntax II

Betydelse

- satssemantik
 - tematiska roller
 - (logisk inferens, motsättning)

F6: Semantik

Form

- huvudsats och bisats
- bisatstyper

F5: Syntax II

