Introduction

Francis **Bond**

Department of General Linguistics,
Palacký University, Olomouc, Czechia

<bond@ieee.org>

KOL LEXI 2026
Slides are open source (CC BY 4.0)

Roadmap

Faculty
of Arts

Welcome!

- In this course we will introduce you to the study of meaning
  - ▶ How meaning is built up from words and phrases
  - ▶ How meaning depends on context
- We will ground the analysis with real examples from the Sherlock Holmes stories and the War with the Newts
  - ▶ I try to make this as enjoyable as possible

Overview

- The course is co-taught by Markéta Dančová
  (marketa.dancova01@upol.cz)
- The timetable is available in Moodle.
- I will talk for about an hour, then we will work together
- Assessment
  - ▶ Regular attendance (80 %)
  - ▶ Active participation in class and completing of assignments
    (required reading)
  - ▶ Assignment (Pass/Fail)
  - ▶ Oral examination (you must pass the assignment to take the exam)

Textbook and Readings

The text books are

- Saeed, John (2015) *Semantics*. 4rd Edition. Wiley-Blackwell.
- Čermák, František (2010) *Lexikon a sémantika* Praha : Nakladatelství Lidové noviny, ISBN: 978-80-7422-020-3
- If you want to know more about semantics I recommend
  - ▶ Paul Kroeger (2022) *Analyzing meaning: An introduction to semantics and pragmatics*. 3rd edition. Language Science Press. DOI: 10.5281 (Open Source)
  - ▶ Lyons, John (1977) *Semantics*. Cambridge University Press

Self Introduction I

- BA in Japanese and Mathematics
- BEng in Power and Control
- PhD in English on *Determiners and Number in English contrasted with Japanese, as exemplified in Machine Translation*
- 1991-2006 NTT (Nippon Telegraph and Telephone)
  - ▶ Japanese - English/Malay Machine Translation
  - ▶ Japanese corpus, HPSG grammar and ontology (Hinoki)
- 2006-2009 NICT (National Inst. for Info. and Comm. Technology)
  - ▶ Japanese - English/Chinese Machine Translation
  - ▶ Japanese WordNet
- 2009-2022 NTU (Nanyang Technological University)
  - ▶ Abui, Chinese, Malay, Multilingual Wordnets (OMW)
  - ▶ HPSGs for Chinese, Indonesian, …
  - ▶ Multilingual Meaning Banks (Treebank + Sensebank)

Self Introduction II
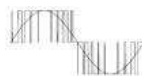
- 2022- UPOL (University Palacký Olomouc)
  - ▶ Czech, Cantonese and Multilingual Wordnets
  - ▶ ChainNet for metaphor and metonymy
  - ▶ WSD using LLMs
  - ▶ Codex of HPSG implemented grammars
  - ▶ Using HPSG to measure LLM syntactic diversity

Please tell me why you want to study lexical semantics, what languages you know, how much semantics you know already, …

Roadmap

Faculty
of Arts

What is Semantics

- Very broadly, semantics is the study of meaning
    - ▶ Word meaning
    - ▶ Sentence meaning
    - ▶ Contextual meaning (pragmatics)
- Why do we want to study meaning?
    - ▶ It underlies our understanding of the world
    - ▶ It is fundamental to our thinking, but we don't consciously know what we are doing
- What kind of knowledge does it take for a speaker to produce language and for a hearer to comprehend language?

Layers of Linguistic Analysis

1. Phonetics & Phonology
2. Morphology
3. Syntax
4. **Semantics**
5. Pragmatics
6. Stylistics

Do people share a common conceptual system? I

- What is a *high school*?
- What color is *blue*?
- What does *verb* mean?
- What is *carrot cake*?
- What color are *traffic lights*?

Do people share a common conceptual system? II

Japanese traffic lights are green (as required by international agreements). However they are typically called 青い *aoi* "blue", the same word as the color of the sky. Historically this color historically covered both green and blue "grue", with 緑 *midori* "green" being a later addition. For this reason, the Japanese government decided in 1973 to change the color of the go light to the bluest possible hue of green!

The Japanese traffic light blues: Stop on red, go on what?

Verbing Wierds Language

From *Calvin and Hobbes* by Bill Watterson (*Homicidal Psycho Jungle Cat* (1994) ISBN 0836217691)

- We can change words from one part of speech to another, with a slight change in meaning, …

Word Meaning and Sentence Meaning

- We store information about words in our **mental lexicon**
  - ▶ It is still unclear what exactly a word is!
- Words can be combined to form an infinite number of expressions
  - ▶ This building up of meaning is referred to as **composition**
  - ▶ If the meaning of the whole can be deduced from the parts then it is **compositional**

Reference and Sense

- Words **refer** to things in the world (like `unicorn`s)
- The meaning of a word across different contexts is often referred to as its **sense**
    - ▶ Same word can refer to different things
        - English: *I put my money in the bank*
        - English: *I fell asleep at the river bank*
    - ▶ Same basic concept can have different boundaries
        - French: *mouton* "sheep/mutton"
        - English: *sheep* vs *mutton*
        - Japanese: *hato* "dove/pigeon"
        - English: *dove* vs *pigeon*

Representing meaning

- One of our goals will be to represent meaning
- There are various ways to do this
  - ▶ Syntactic trees
  - ▶ Logical forms
  - ▶ Thesauri and Ontologies
  - ▶ Translation
  - ▶ Paraphrasing

  Can you think of others?
- At the end of this course you should be able to use these to describe many aspects of word meaning

Further Reading

- Introduction What does it mean to mean?
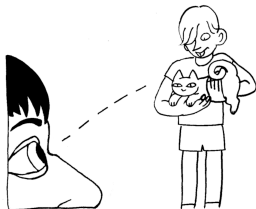   - Saeed: § 1

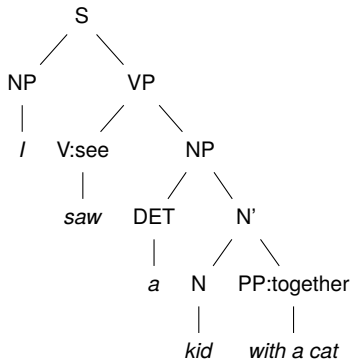# Language is normally under-specified

We get **words**:

I saw a kid with a cat.

We want **meaning**:

I saw a kid with a cat$_1$


```
see(I, kid: past); with(kid,
cat)

see ⊂ perceive
kid ~ child
with ⊂ together
```
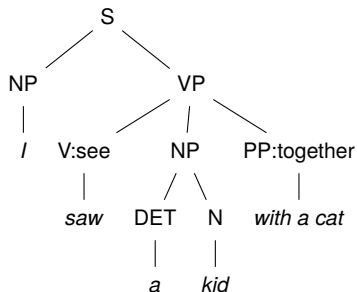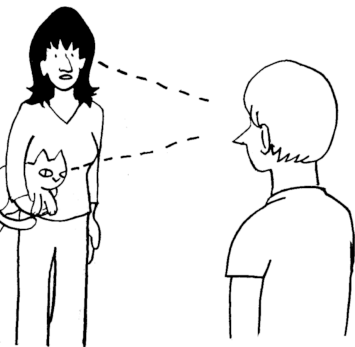
I saw a kid with a cat$_2$

S
NP
I
VP
V:see
saw
NP
DET
a
N
kid
PP:together
with a cat

see(I, kid: past) with(I, cat)

see ⊂ perceive
kid ∼ child
with ⊂ together

I saw a kid with a cat₃


```
              S
            /   \
         NP       VP
          |      /  \
          I  V:saw   NP
              |      /  \
            saw   DET    N'
                   |    /   \
                   a   N    PP:together
                       |       |
                      kid   with a cat
```
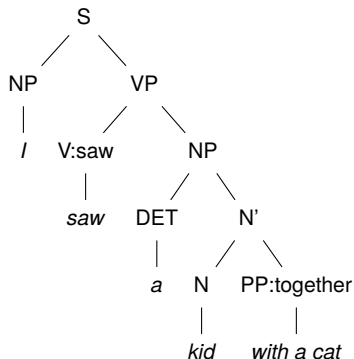
saw(I, kid: pres); with(kid, cat)

saw ⊂ cut
kid ∼ child
with ⊂ together

I saw a kid with a cat$_4$


```
              S
        /          \
      NP            VP
      |         /    |      \
      I    V:saw    NP    PP:instrument
           |      /   \         |
          saw  DET  N:goat  with a cat
                |      |
                a     kid
```
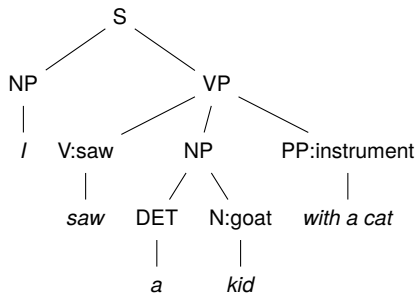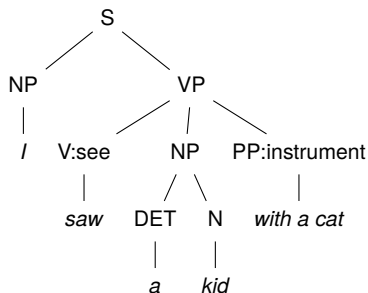
saw(I, kid: present) with(I, cat)

saw ⊂ cut
kid ∼ young goat
with ⊂ together

I saw a kid with a cat$_5$

see(I, kid: past) with(I, cat)

see $\subset$ perceive
kid $\sim$ child
with $\subset$ instrumental

We can also use translations I

(1) 我　看到了　　一个 抱着　　猫　 的 孩子
wǒ kàndàole yīgè bàozhe māo de háizi.
I　saw　　　one holding cat 's child
I did see a child holding a cat

(2) 我 抱着　　猫　 看到了　　一个 孩子
wǒ bàozhe māo kàndàole yīgè háizi
I　holding cat saw　　　one child
I holding a cat did see a child

(3) 我 锯　一个 孩子 和　他/她　 的 猫
wǒ jù　yīgè háizi hé　tā/tā　 de māo
I　saw one　child and he/she 's　cat
I saw (cut with a saw) a child and their cat

We can also use translations II

(4) 我　和　一只　猫　鋸　一只　小　山羊
　　　 wǒ hē　yīzhǐ māo jù　yīzhǐ xiǎo shānyáng
　　　 I　and one　cat　saw (cut　with a　　　　saw)
　　　 I and a cat saw a young goat

(5) 我　用　一只　猫　看到了　一个　孩子
　　　 wǒ yòng yīzhǐ māo kàndàole yīgè háizi
　　　 I　use　one　cat　saw　　　 one　child
　　　 Using a cat, I did see a child

Paraphrase Try to paraphrase — reword in English or translate to another language
aim to be unambiguous, even if slightly disfluent

Roadmap

Faculty
of Arts

Referential or Representational?

One view of meaning is to define it in terms of how it constrains reality.

- Picture the worlds in which these sentences are true:

  (6) *I patted the dog.*

  (7) *I did not pat the dog.*

Assuming that they were uttered at the same time, they are incompatible because they cannot refer to the same situation: the **referential** view.

But we can represent the same reality in different ways:

  (8) *Ich habe Hunger* "I have hunger"

  (9) *I am hungry*

Representational theories are interested in how we represent reality, and how our representations are influenced by conceptual structures conventionalized in language.

Referential View

Refers to

Speaker → Says → Expression → Denotes → Referent
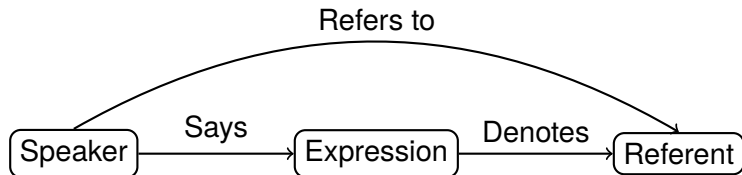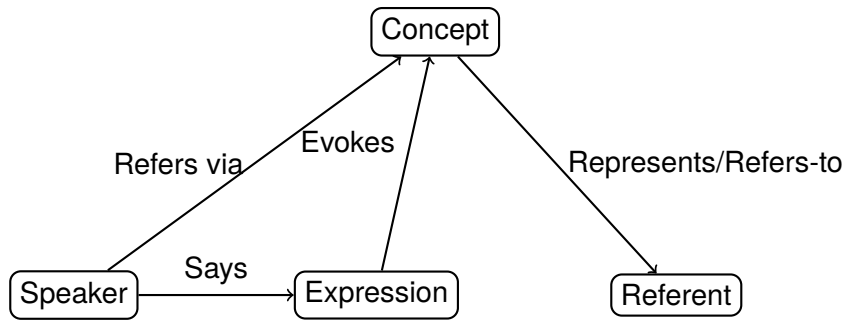
The **referential view** is focused on direct relationships between expressions (words, sentences) and things in the world (realist view).

Representational View

The **representational view** is focused on how relationships between expressions (words, sentences) and things in the world are mediated by the mind (cognitive linguistics).

This gives a more complex, but richer model.

Referring vs Non-Referring

- **Referring expressions** are expressions that identify entities in the world (typically **nominals**)

  (10) *cat*, *ano kiiro kaban* "that yellow bag"

  (11) *London Bridge*, *Xiao Ming*

- **Non-referring expressions** don't have referential properties

  (12) *maybe, if, is, but*

- Not all nominals refer

  (13) *That is <u>an ugly dog</u>*

  (14) *If only I had <u>a dog</u>*

- And, of course, all this is made more confusing if we model the fictional world and our interpretation of it as separate from the characters' interpretations, …

Further Reading

- Introduction What does it mean to mean?
  - ▶ Saeed: § 2

Roadmap

Faculty
of Arts

Close Reading

- Reading (and often re-reading) a text to uncover multiple aspects of meaning that lead you to understand a text better
- Looking at what the text actually says, as well as the inferences you make from reading it
- After a close reading you should be able to support your conclusions with specific examples from the text
- You can consider many aspects of the text, such as
  - ▶ The Title
  - ▶ Word Choice
  - ▶ The Tone and Style
  - ▶ Discerning Patterns
  - ▶ Point of View and Characterization
  - ▶ Symbolism

Word Choice and Diction

- What word(s) stand out? Why? (typically vivid words, unusual choices, or a contrast to what a reader expects)
- How do particular words get us to look at characters or events in a particular way? Do they evoke an emotion?
- Did the author use nonstandard language or words in another language? Why? What is the effect?
- Are there any words that could have more than one meaning? Why might the author have played with language in this way?
- Do some words have extra connotations?

Faculty
of Arts

Word Sense Disambiguation

- Knowing what individual words mean is the first step towards understanding
- For the assignment, we will try to identify the **sense** of words
    - ▶ We use Wordnet (Fellbaum, 1998) as the sense inventory because it contains semantic relations as well as definitions and it is accessible: there are good interfaces to it
    - ▶ For every word we chose the most appropriate sense in wordnet or write a comment if we think there isn't one
    - ▶ Once we have identified a sense, it is then easy to look at synonyms and other closely related words
    - ▶ For Czech we use the Czech wordnet (Pala and Smrž, 2004)
    - ▶ Both wordnets have been extended as part of the the Natural Text Understanding — Multilingual Corpus (**ntu-mc**)

WSD: A challenging task! I

This is difficult for many reasons

- Meaning boundaries are not clear: the sense distinctions impose a structure on something that is actually fuzzy
- Dictionaries are imperfect
  - ▶ senses may be missing
  - ▶ senses may be too fine-grained
- Processing a text by computer is difficult
  - ▶ The computer may have misinterpreted
    - the part-of-speech
      Does that go "Female deer which go" "Is it the case that is goes?"
      The speckled band "the band that is speckled" "the band that someone speckled"
    - Or the sentence boundaries
    - Or the words boundaries
- People use language idiosyncratically
  - ▶ extending meanings metaphorically
  - ▶ sometimes so strangely that we might even say wrongly

WSD: not impossible

- Typically people agree around 72.5% of the time (Snyder and Palmer, 2004).
    - ▶ Verbs are hardest (67.8%), then nouns (74.9%) and adjectives (78.5%)
    - ▶ Disagreements tend to cluster around a relatively small group of difficult words.
    - ▶ For example *national*
        - In six out of seven instances one annotator chose "limited to or in the interests of a particular nation" and the other annotator chose "concerned with or applicable to or belonging to an entire nation or country"
        - They are hard to distinguish!

Roadmap

Faculty
of Arts

Administrivia

Coordinator Francis <u>Bond</u> `<bond@ieee.org, francis.bond@upol.cz>`

Course details are all online

Extra Credit

- If you submit a correction that gets accepted for one of the resources we use then it shows good mastery of the material
  - ▶ you can get 1-5% extra credit (depending on the size/difficulty) Mark $n \propto 10^{n-1}$ lines of code/documentation
  - ▶ You can't go over 100%
- A correction can involve
  - ▶ fixing an error in transcription or annotation
    - spelling error
    - wrong sense
    - error in the dictionary
  - ▶ making the documentation easy to read
  - ▶ pointing out an error in a translation / finding a new translation
  - ▶ coming up with a good example of a phenomenon

Student Responsibilities

By remaining in this class, the student agrees to:

1. Make a genuine effort to learn and engage.
2. Read messages and participate.
3. Do assignments on time.
4. Attend regularly.
5. Seek help early, not last minute.
6. Treat peers respectfully.

Faculty of Arts

Attendance

1. You are expected to attend all classes.
2. Be on time - lateness is disruptive to your own and others' learning.
3. Valid reasons for missing class include the following:
   1. A medical emergency (including mental health emergencies)
   2. A family emergency (death, birth, natural disaster, etc).
4. There will be significant material covered in class that is not in your readings. You cannot expect to do well without coming to class.
5. If you miss a class, it is your responsibility to get the notes, any handouts you missed, schedule changes, etc. from a classmate.

Faculty of Arts

Remediation and Academic Integrity

1. No late work will be accepted, except in the case of a documented excuse.
2. For planned, justified, absences on class days or days on which assignments are due, advance notice must be provided.
3. Cheating will not be tolerated. Violations, including plagiarism, will be seriously dealt with, and could result in **a failing grade for the entire course**.
4. Refer to the University Honour Code
5. As always, use your common sense and conscience.

The winning strategy

- Read the textbook before class (and after again, if necessary)
- Work together: make study groups
- Tasks: Discuss as much as you want, annotate your own answers
- Tutorials: Discuss as much as you want, write your own answers
- Treat LLMs as though they were a human — you can discuss the answer with them, but they should not give you the entire answer
- Ask questions … early and often!

References I

Christine Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Karel Pala and Pavel Smrž. 2004. Building Czech wordnet. *Romanian Journal of Information Science*, 7:79–88.

Benjamin Snyder and Martha Palmer. 2004. The English all-words task. In *Proceedings of Senseval-3*, pages 41–44. ACL, Barcelona.

Glossary of Key Terms (English–Czech) I

English	Čestina
Q	otázka (Q)
analysis	analýza
autonomous	autonomní
collocation	kolokace
communication	komunikace
composition	kompozice
compositional	kompoziční
concordance	konkordance
connotation	konotace
context	kontext
corpus	korpus
deictic	deiktický
deixis	deixe

Glossary of Key Terms (English–Czech) II

denotation	denotace
discourse	diskurz
distal	distální
expression	výraz
gender	(gramatický) rod
honorification	honorifika
idiom	úsloví
inclusion	inkluze
lexical	lexikální
meaning	význam
medial	mediální
mental lexicon	mentální lexikon
metaphor	metafora
nominals	nominální výrazy
non-referring expressions	nereferenční výrazy
number	číslo

Glossary of Key Terms (English–Czech) III

person	(gramatická) osoba
phrase	slovní spojení
power	moc
pragmatics	pragmatika
proximal	proximální
refer	odkazovat
reference	reference
referential	referenční
referential view	referenční pohled
referring expressions	referenční výrazy
representational	reprezentační
representational view	reprezentační pohled
semantics	sémantika
sense	význam (smysl)
sentence meaning	význam věty
sentiment	sentiment

Glossary of Key Terms (English–Czech) IV

social status	společenský status
solidarity	solidarita
spatial location	prostorová lokalizace
temporal location	časová lokalizace
textual deixis	textová deixe
utterance	výpověď
word meaning (sense)	slovní význam