# DAS
## Semantics

# Theories of meaning and the meaning of words

Francis Bond

**Department of Asian Studies**
**Palacký University**
https://fcbond.github.io/
bond@ieee.org

Lecture 2

# The meanings of words

# Words carry different meanings: *leave*

10070    *Nothing was <u>left</u> save a few acres of ground , and the two-hundred-year-old house , which is itself crushed under a heavy mortgage .*

10079    *The money which my mother had <u>left</u> was enough for all our wants , and there seemed to be no obstacle to our happiness . ”*

10085    *He had no friends at all save the wandering gipsies , and he would give these vagabonds <u>leave</u> to encamp upon the few acres of bramble- covered land which represent the family estate , and would accept in return the hospitality of their tents , wandering away with them sometimes for weeks on end .*

10107    *She <u>left</u> her room , therefore , and came into mine , where she sat for some time , chatting about her approaching wedding .*

10108    *At eleven o'clock she rose to <u>leave</u> me , but she paused at the door and looked back.*

10439 *" The rest you will <u>leave</u> in our hands . "*

10449 *And now , Miss Stoner , we must <u>leave</u> you for if Dr. Roylott returned and saw us our journey would be in vain .*

10526 *Then he turned down the lamp , and we were <u>left</u> in darkness .*

How many different meanings?

# How can we represent the differences?

➤ Definitions

➤ Translations/paraphrases

➤ Semantic Relations

➤ Components

➤ Vector Spaces

# Semantic Representations of Words

➤ Divide meaning into

➤ **reference**: the relation to the world/mental space
➤ **sense**: the rest of the meaning
  ∗ **denotation** the part that distinguishes the meaning from other meanings
  ∗ **connotation** cultural or emotional associations

➤ Introduce CONCEPTS                                              (meaning as font-change)

➤ How can we represent concepts?
➤ How do we learn them?
  ∗ Typically children start off by **underextending** or **overextending** concepts

➤ Example: *That dog*

➤ reference — the animal over there
➤ sense — canine quadruped domesticated by man
➤ connotation — faithful, friendly (or dirty)

# Definitional Semantics

➤ Standard lexicographic approach to lexical semantics:

      **semantics** = *the study of language meaning*
      **tailor** = *a person whose occupation is making and altering garments*

➤ Definitions are conventionally made up of;

   ➤ genus: what class the lexical item belongs to
   ➤ differentiae: what attributes distinguish it from other members of that class

➤ Often hard to understand if you don't already know the meaning!

# Definitional Semantics: pros and cons

➢ Pros:

  ➢ familiarity (we are taught to use dictionaries)

➢ Cons:

  ➢ subjectivity in sense granularity (splitters vs. lumpers) and definition specificity
  ➢ circularity in definitions
  ➢ consistency, reproducibility, …
  ➢ often focus on diachronic (historical) rather than synchronic (current) semantics

# Entries for *leave*

---

**02015598-v**  (72) V1, V2 *get out, go out, leave, exit* "move out of or depart from"

**02356230-v**  (8) V3 *leave, entrust* "put into the care or protection of someone"

**02009433-v**  (149) V1 *leave, go away, go forth* "go away from a place"

**02229055-v**  (7) V3 *leave, will, bequeath* "leave or give by will after one's death"

**02729414-v**  (56) V2 *leave* "act or be so as to become in a specified state"

**02730135-v**  (5) V2 *leave* "have left or have as a remainder"

**06690114-n**  (1) *leave* "permission to do something"

Not to be confused with *left hand* and *the leaves fell*, ….

---

# Paraphrases and translation

➤ Saying the same thing in different words

  ➤ Same language = **paraphrase**
  ➤ Different language = **translation**

➤ We showed some paraphrases in the entries given above

➤ If you speak another language, then you can use that to disambiguate may things.

  ➤ *leave, entrust* = 預ける *azukeru*
  ➤ *get out, go out, leave, exit* = 去る *saru*
  ➤ *leave, will, bequeath* = 遺す *nokosu*

➤ Can you explain the ambiguity in *The money which my mother had left was enough for all our wants*?

# Paraphrase cues

➢ ***that is to say***

    ➢ *I met him that night, and he called next day to ask if we had got home all safe, and after that we met him - that is to say, Mr Holmes, I met him twice for walks, …* IDEN

    ➢ *In three days, that is to say on Monday next …* FINA

➢ ***in other words***

    ➢ *…there is a possibility that these initials are those of the second person who was present - in other words, of the murderer. …* BLAC

    ➢ *…when they closed their League offices that was a sign that they cared no longer about Mr. Jabez Wilson's presence; in other words, that they had completed their tunnel.* REDH

# Lexical Semantics

➢ Lexical semantics is concerned with the identification and representation of the semantics of lexical items

➢ If we are to identify the semantics of lexical items, we have to be prepared for the eventuality of a given word having multiple interpretations

    ➢ **Polysemy**: having multiple meanings
    ➢ **Monosemy**: having only one meaning

➢ **Homonyms** are words with two unrelated meanings:

    ➢ **homographs**: same spelling
       *bow* vs *bow*; *keep* vs *keep*
    ➢ **homophones**: same pronunciation
       *right* vs *write*; *keep* vs *keep*

# Distinguishing Polysemes

➢ **Antagonism**: can the word be used in a sentence with multiple <u>competing</u> interpretations that are incompatible?

*Kim can't bear children*

    ➢ Cannot have children
    ➢ Doesn't like children

➢ **Zeugma**: can the word be used in a context where multiple <u>competing</u> interpretations are simultaneously evoked?

*Kim and her visa expired*

    ➢ died
    ➢ ran out

*Hitmen were quite expensive, so she decided to take out a loan and her husband.*

➢ **Paraphrase/Translation**: Is there more than one (clearly different) way to paraphrase/translate the word.

# Necessary and Sufficient Conditions

➢ Can we define words in terms of **conditions**?

   ➢ *zebra*
      ∗ quadruped
      ∗ animal                              (redundant)
      ∗ black and white striped
      ∗ herbivore

➢ These are **intrinsic**, **generic** properties

➢ Can we use words even if we don't know their properties?

   ➢ *Kway Teow*

➢ We seem to be ok with fairly vague definitions

   ➢ What is a *dog-cart*?
   ➢ What is a *swamp adder*?

# Words/Concepts are related in many ways

We can also look at words (or more properly senses) in terms of their relations to other words.

➢ **Hyponymy/Hypernymy**

➢ **Synonymy**

➢ **Antonymy** (Opposites)

➢ **Meronymy**

  ➢ **Member-Collection**
  ➢ **Portion-Mass**
  ➢ **Element-Substance**

➢ **Domain** (lexical field)

# Hypernymy and Hyponymy

➤ **Hyponymy**: X is a hyponym of Y iff $f(X)$ entails $f(Y)$ but $f(Y)$ does not entail $f(X)$ (for all or most $f$):

*Kim has a pet dog $\vDash$ Kim has a pet animal*
*Kim has a pet animal $\nvDash$ Kim has a pet dog*

N.B. complications with universal quantifiers and negation:

*Kim likes all animals $\vDash$ Kim likes all dogs*
*Kim likes all dogs $\nvDash$ Kim likes all animals*

➤ **Hypernymy**: Y is a hypernym of X iff X is a hyponym of Y

➤ Can a word have multiple hypernyms?

(1)   $tank_1 \subset military\_vehicle_1; \subset tracked\_vehicle_1; \subset armored\_vehicle_1;$
        $? \subset weapon_1$

# What is entailment

**Entailment** ($\vDash$): A sentence $p$ entails a sentence $q$ when the truth of the first ($p$) guarantees the truth of the second ($q$), and the falsity of the second ($q$) guarantees the falsity of the first ($p$).

# Properties of hypernymy/hyponymy

➤ Asymmetric; applies at the sense level

➤ applies only to lexical items of the same word class

➤ Transitive: $dog_1 \subset mammal_1 \subset animal_1$

➤ Not all nodes are lexicalized; can be multiple

| neutral (Hyper) | male | female | child |
|---|---|---|---|
| *sheep* | *ram* | *ewe* | *lamb* |
| *cow* | *bull* | *cow* | *calf* |
| *goose* | *gander* | *goose* | *gosling* |
| *horse* | *stallion* | *mare* | *foal:colt/filly* |
| *dog* | *dog* | *bitch* | *puppy* |
| *snake* | *snake* | *snake* | *snake* |

➤ Can you do this for *pig*, *cat* or *chicken*?

➤ Can you give an example of this in another language?

# Language Change

➢ The meanings of words change over time

➢ *guitar* — "a stringed instrument usually having six strings": originally these were all made of wood

➢ We then get *electric guitar* — "a guitar with a built-in pickup or pickups which convert string vibrations into electrical signals for amplification"

➢ To refer to non-electric guitars we get a new coining *acoustic guitar* – "a guitar that does not require electrical amplification": which used to just be guitar

　➢ *guitar* is now a hypernym of them both and can refer to either

➤ Sometimes this practice becomes politically charged, although linguistically it is unremarkable

  ➤ **woman** "an adult female person"
  ➤ **trans woman** "a person who identifies as a woman but was assigned male at birth"
  ➤ **cis woman** "a person who identifies as a woman and was assigned female at birth"

➤ Can you give other examples of this in English or other languages?

# Synonymy

➢ **Propositional synonymy**: X is a propositional synonym of Y if

➢ (i) X and Y are syntactically identical,
➢ (ii) substitution of Y for X in a declarative sentence doesn't change its truth conditions

e.g., *violin* and *fiddle*

➢ Why propositional synonymy is over-restrictive:

➢ syntactic identity (cf. *eat* and *devour*)
➢ collocations (cf. *cemetery* and *graveyard*)
➢ gradability (cf. *sofa*/*settee* vs. *boundary/frontier*)

# Near Synonymy

➤ Synonyms are substitutable in **some/most** rather than **all** contexts

➤ Synonymy via semantics: synonyms share "common traits" or attributional overlap, walking the fine line between "necessary resemblances" and "permissible differences":

   *grain* vs. *granule*; *green* vs. *purple*; *alsation* vs. *spaniel*

➤ Permissible differentiation via **clarification**:

   *Here is a grain, or granule, of the substance.*
   *\* The cover is green, {or, that is to say} purple.*

   and **contrast**:

   *Here is a grain or, more exactly, granule*
   *\* He likes alsations, or more exactly, spaniels*

# Properties of synonymy

➤ Symmetric

➤ traditionally applies only to lexical items of the same word class but pairs like **can** vs **be able to** suggest otherwise

➤ applied at the sense level?

➤ $\approx$ converse of polysemy

# Antonymy (opposites)

➢ **Simple antonyms**: the negative of one implies the positive of the other.

      (2)   *dead/alive*

      (3)   *pass/fail*

➢ **Gradable Antonyms**: points along a scale

      (4)   *boiling/hot/warm/tepid/cool/cold/freezing*

      (5)   *like this class/fascinating/interesting/dull/boring*

➢ **Reverses**: reverse the direction of a motion

      (6)   *ascend/descend*

      (7)   *up/down; right/left*

➤ **Converses**: the same act from different points of view

(8)    *above/below; right/left*

(9)    *employer/employee*

(Slightly non-standard usage)

➤ **Taxonomic Sisters**: children of the same (grand)parent

(10)    *Monday/Tuesday/…/Sunday*
         in WordNet: ***day of the week*** $\supset$ ***weekday***, ***weekend***
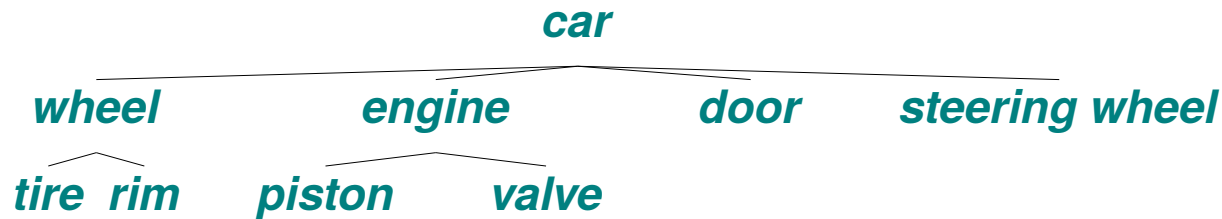
(11)    *LMS/English/Chinese/…*
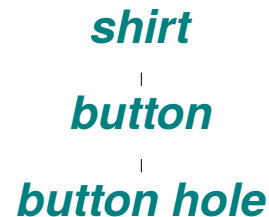         Context dependent

# Meronymy

➢ **Meronomy** refers to the part-whole relation

  ➢ **meronym** is the part
  ➢ **holonym** is the whole

```
                        car
         _____|_____
        |            |           |         |
     wheel        engine        door   steering wheel
      _|_          _|_
     |   |        |   |
   tire rim    piston valve
```

➢ It is not always transitive        *shirt*

                                       |
                                    *button*

                                       |
                                  *button hole*

  But we don't normally say that a ***button hole*** is part of a ***shirt***.

# Member-Collection

➤ The relation between a collection and one of the units that makes it up

    (12)    *tree–forest*

    (13)    *sheep–flock*

    (14)    *fish–school*

    (15)    *book–library*

    (16)    *member–band*

    (17)    *musician–orchestra*

    (18)    *student–class*

# Portion-Mass

➤ The relation between a mass noun and a typical unit of measurement

(19)  *drop–liquid*

(20)  *grain–sand/salt/truth*

(21)  *sheet/ream–paper*

(22)  *lump–coal (or just about anything)*

(23)  *strand–hair*

(24)  *rasher–bacon*

➤ Similar to classifiers in many ways, e.g. in Malay

(25)  *ekor* "tail"–*animal*

(26)  *orang* "human"–*person*

# Domain (lexical field)

The domain in which a word is typically used with this meaning.

(27)  **driver**$_1$ — the operator of a motor vehicle

(28)  **driver**$_2$ — someone who drives animals that pull a vehicle

(29)  **driver**$_3$ — a golfer who hits the golf ball with a driver [GOLF]

(30)  **driver**$_4$ — ($\simeq$ device driver) a program that determines how a computer will communicate with a peripheral device [COMPUTER SCIENCE]

(31)  **driver**$_5$ — ($\simeq$ number one wood) a golf club (a wood) with a near vertical face that is used for hitting long shots from the tee [GOLF]

Some GOLF terms: **approach**$_9$, **approach shot**$_1$, **golf course**$_1$, **links course**$_1$, **wedge**$_5$, **tee**$_1$, **scratch**$_9$, **putt**$_1$, **slice**$_1$, **hook**$_1$

# And More

➤ There are many, many more lexical relations advocated by various theories including:

  ➤ Troponymy/hypernymy (cf. *walk* vs. *lollop*) "way of doing something"
  ➤ Entailment (cf. *snore* vs. *sleep*) "if you do one thing, you must be doing the other"
  ➤ Operator (cf. *question* vs. *ask*) "the thing you do by doing something"
  ➤ Magnifier (cf. *wound* vs. *badly*) "intensifier, diminisher"
  ➤ Usage (cf. *strong-willed* vs. *pig-headed* "stubborn")
    *pig-headed* is **pejorative**

# Wordnet

# WordNet

➤ WordNet is an open-source electronic lexical database of English, developed at Princeton University

    http://wordnet.princeton.edu/

➤ Made up of four separate semantic nets, for each of nouns, verbs, adjectives and adverbs

➤ WordNets exist for many languages, my group has worked on:

  ➤ Japanese
  ➤ Bahasa Malay/Indonesian
  ➤ Chinese (Mandarin and Cantonese)
  ➤ The shared open multi-lingual wordnet (150+ languages)
    https://omwn.org/
  ➤ Kristang
  ➤ Myanmar
  ➤ Czech

Miller (1998); Fellbaum (1998)

# Wordnet Structure

➤ Lexical items are categorised into ~115K (and counting) glossed **synsets** (= synonym sets)

```
1. enrichment -- (act of making fuller or more
   meaningful or rewarding)
2. enrichment -- (a gift that significantly increases
   the recipient's wealth)
```

➤ Lexical relations at either the synset level or sense (= combination of lexical item and synset) level

➤ Strongly lexicalist (orginally):

➤ synsets only where words exist
➤ but many multiword expressions ($\approx 50\%$)

# Psycholinguistic Foundations of WordNet

➢  Strong foundation on hypo/hypernymy (lexical inheritance) based on

  ➢  response times to sentences such as:
      *a canary {can sing/fly,has skin}*
      *a bird {can sing/fly,has skin}*
      *an animal {can sing/fly,has skin}*
  ➢  analysis of anaphora:
      *I gave Kim a novel but the {book,?product,...} bored her*
      *Kim got a new car. It has shiny {wheels,?wheel nuts,...}*
  ➢  selectional restrictions

➢  Is now often used to calculate **semantic similarity**

  ➢  The shorter the path between two synsets the more similar they are
  ➢  Or the shorter the path to the nearest shared hypernym, …

# Word Meaning as a Graph
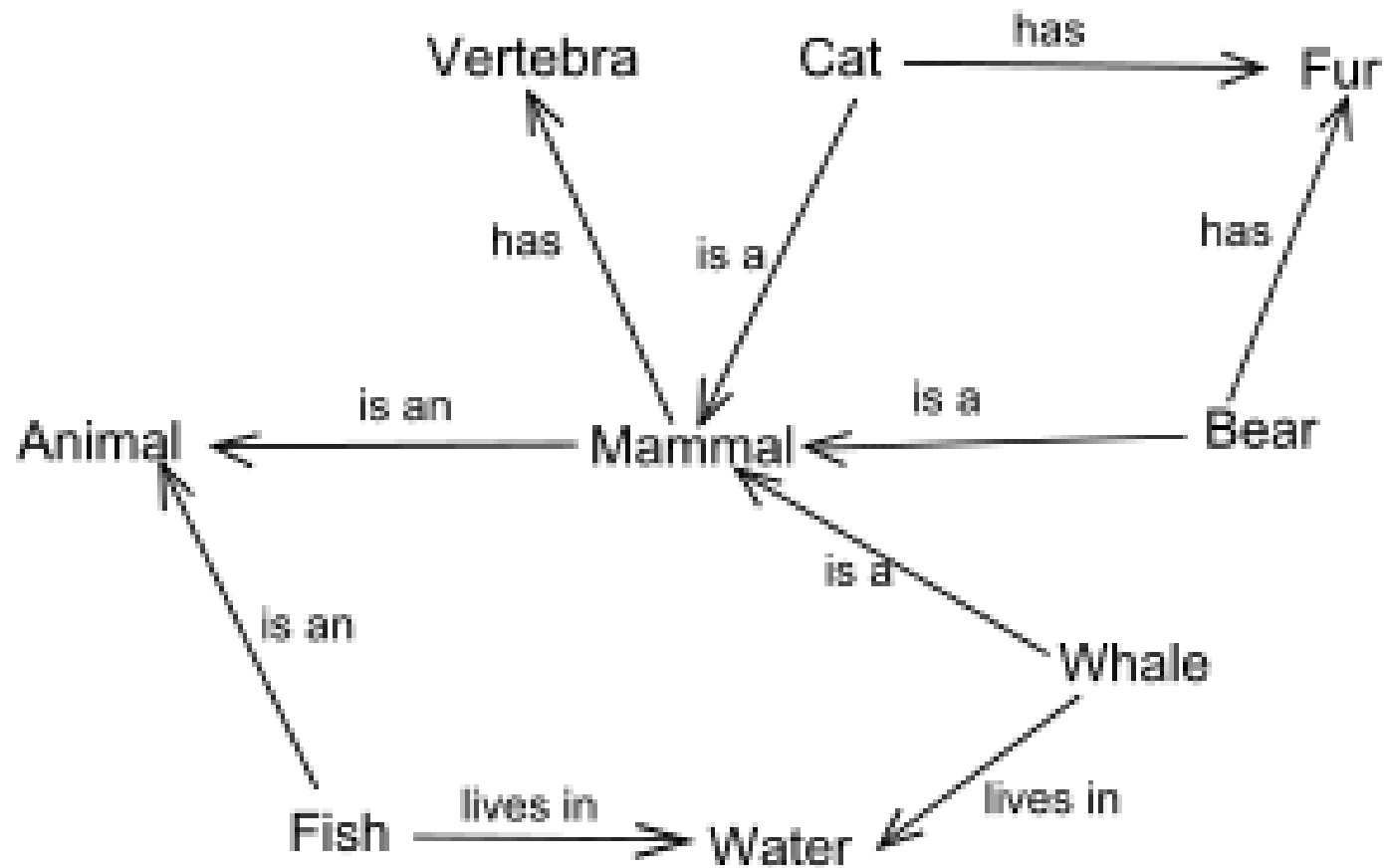


➤ You need a very big graph to capture all meanings

# Wordnet in this course

➢ We will use wordnet to test our skills in determining word meaning

    ➢ tag a short text from this year's story or stories
    ➢ discuss differences with other annotators

➢ As well as a source of examples and inspiration

# Where is the meaning?

# 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:

(32)   *I patted the dog.*
(33)   *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.

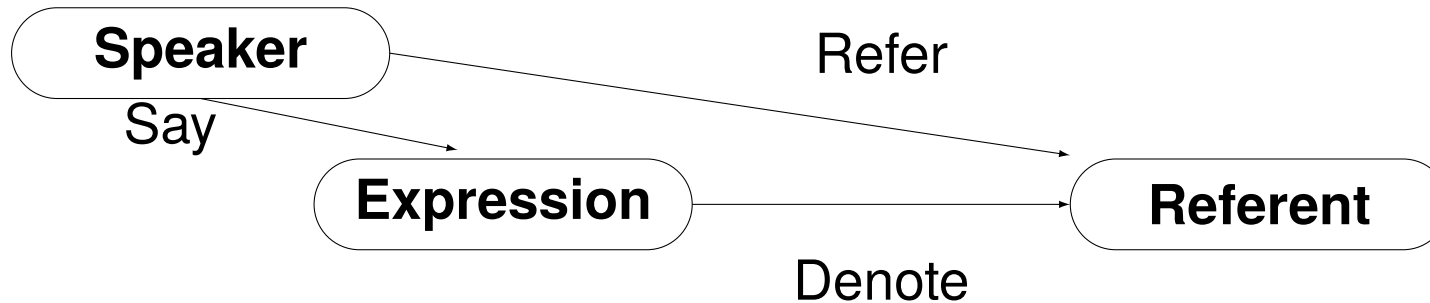But we can represent the same reality in different ways:

(34)   *Ich habe Hunger* "I have hunger"
(35)   *I am hungry*

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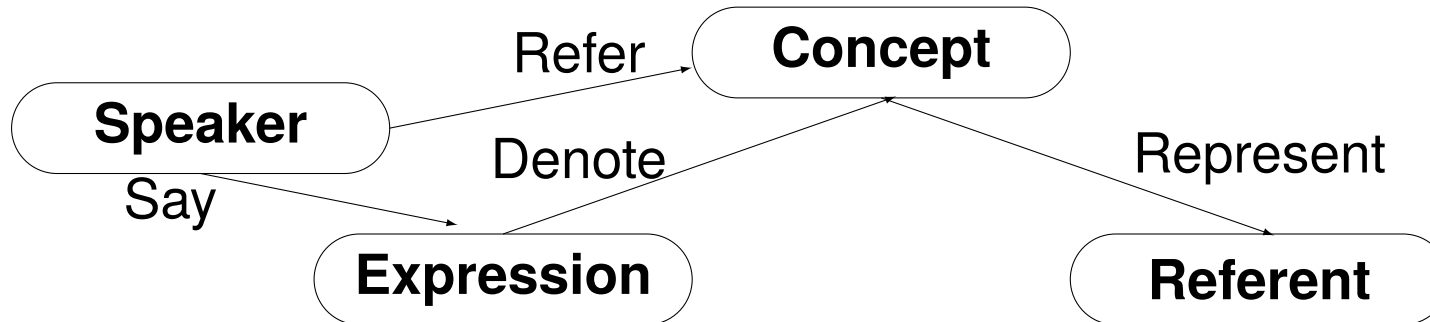
# Referential View



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**)

> (36)  *cat*, *ano kiiro kaban* "that yellow bag"
> (37)  *London Bridge*, *Xiao Ming*

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

> (38)  *maybe, if, is, but*

➢ Not all nominals refer

> (39)  *That is an ugly dog*
> (40)  *If only I had a dog*

➢ 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, …

# Deixis

# What is Deixis

➤ any linguistic element whose interpretation necessarily makes reference to properties of the extra-linguistic context in which it occurs is **deictic**

**Person** relative to the speaker and addressee; *you, me, them*
**Spatial Location** demonstratives; *this, that, over there, here*
**Temporal Location** tense; *yesterday, today, tomorrow*
**Social Status** relative to the social position: *professor, you, uncle, boy*

➤ **Discourse deixis**: referring to a linguistic expression or chunk of discourse

More than 90% of the declarative sentences people utter are indexical in that they involve implicit references to the speaker, addressee, time and/or place of utterance in expressions like first and second person pronouns, demonstratives, tenses, and adverbs like *here*, *now*, *yesterday* (Bar-Hillel, 1954, p366).

# Spatial Deixis

➤ Two way systems (English, …)

| | | | |
|---|---|---|---|
| **proximal** | *this* | *here* | close to the speaker |
| **distal** | *that* | *there* | far to the speaker |

➤ Three (four) way systems (Japanese, …)

| | Gloss | THING | PLACE |
|---|---|---|---|
| **proximal** | close to speaker | *kore* "this" | *koko* "here" |
| **medial** | close to addressee | *sore* "that" | *soko* "there" |
| **distal** | far from both | *are* " 'tother" | *asoko* "over there" |
| **Q** | interrogative | *dore* "what" | *doko* "where" |

➤ Can you do English TIME?

➤ Can you do this in another language?

# More Spatial Deixis

➤ Often lexicalized:

  ➤ *go, come, foreign, home, local, indigenous, national language*

➤ Can lead to **discourse**/**textual deixis**

  (41)   *Here we begin explaining textual deixis*

➤ Often also used for time

  (42)   *This year we are trying a new kind of assignment*

➢ Spatial expressions extend to possession in many languages

(43) *NICT-ga    Kyoto-ni    aru*
NICT-ɴᴏᴍ Kyoto-ʟᴏᴄ be

NICT is in Kyoto

(44) *watashi-ni musuko-ga aru*
I-ʟᴏᴄ        son-ɴᴏᴍ    be

I have a son (lit. a son is in me)

# Person Deixis

➢ Minimally a three way division

| First Person | Speaker | *I* |
| Second Person | Addressee | *you* |
| Third Person | Other | *he/she/it* |

➢ Often combined with

  ➢ **gender**: *he/she/it*
  ➢ **number**: *I/we*, *'anta* "you:m", *'antumaa* "you:dual", *'antum* "you:m:pl" (Arabic)
  ➢ **inclusion**: *núy* "we including you", *níi* "we excluding you" (Zayse)
  ➢ **honorification**: *kimi* "you:inferior", *anata* "you:equal", don't use pronouns for superiors: *sensei* "teacher", …(Japanese)

# Social Deixis

In European languages, a two-way choice in 2nd person pronominal reference is known as the T/V distinction, based on the French forms for "you".

➤ T/V distinctions in European languages

|         | Familiar 2sg | Polite 2sg |
|---------|--------------|------------|
| French  | *tu*         | *vous*     |
| German  | *du*         | *Sie*      |
| Spanish | *tú*         | *usted*    |
| Czech   | *ty*         | *vy*       |

➤ Shift from asymmetric use showing **power** (superior uses *du*; inferior uses *vous*) to symmetric use showing **solidarity** (strangers use *vous*; intimates use *du*): typically the socially superior person must invite the socially inferior person to use the familiar form

# Social Deixis can be marked on other words

(45)  *Tanaka-san-ga kudasaimashita*          [addressee and subject hon.]

Tanaka gave it to me (and I honor him and you)

(46)  *Tanaka-san-ga kudasatta*                [subject honorification]

Tanaka gave it to me (and I honor him)

(47)  *Tanaka-kun-ga kuremashita*            [addressee honorification]

Tanaka gave it to me (and I honor you)

(48)  *Tanaka-kun-ga kureta*                      [no honorification]

Tanaka gave it to me (implies I am higher status than him)

➤ Find examples where someone addresses Sherlock as *Holmes* and compare then to examples where he is addressed as *Mr Holmes*: what is the difference?

➤ Find examples in *Válka s Mloky* where *ty* and *vy* are used: what is the difference?

# Conclusions

➤ We learned a little about word meaning, reference and deixis

# Acknowledgments and References

➤ Definitions from WordNet: `http://wordnet.princeton.edu/`

➤ Images from

   ➤ the Open Clip Art Library: `http://openclipart.org/`
   ➤ Steven Bird, Ewan Klein, and Edward Loper (2009) *Natural Language Processing with Python*, O'Reilly Media
     `www.nltk.org/book`

➤ Video: Dead parrot sketch by Monty Python

**\***

References

Yehoshua Bar-Hillel. 1954. Indexical expressions. *Mind*, 63:359–379.

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

George Miller. 1998. Foreword. In Fellbaum (1998), pages xv–xxii.