

Detecting Meaning with Sherlock Holmes* †

Theories of meaning and the meaning of words

Francis Bond

Division of Linguistics and Multilingual Studies

<http://www3.ntu.edu.sg/home/fcbond/>

bond@ieee.org

Lecture 2

Location: LT4

*Creative Commons Attribution License: you are free to share and adapt as long as you give appropriate credit and add no additional restrictions: <https://creativecommons.org/licenses/by/4.0/>.

The Adventure of the Word's Meaning

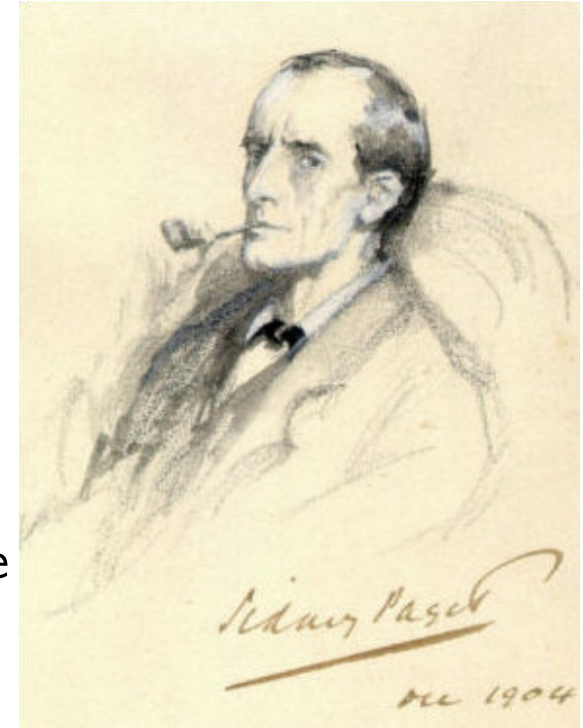
- Revision
 - Introduction to Sherlock Holmes
 - Introduction to Semantics
- How to represent meaning
- Referential theories
- Deixis
- Concepts
- Derivation

Revision: Sherlock Holmes



Why Sherlock Holmes?

- Enjoyable, accessible, popular stories (and adaptations)
- Sherlock Holmes
 - A London-based consulting detective
 - who solves many cases with a quirky personality
 - and immense powers of observation and raticionation
 - and his faithful friend: John Watson
- A fictional character, invented by Sir Arthur Conan Doyle
 - in 60 stories, published between 1887 and 1927
 - Mainly narrated by Watson
- Some stories are annotated with word senses as part of our research in semantics
 - The NTU — Multilingual Corpus (NTU-MC)
 - We will use and experience this during the course — learn by doing



Revision:

Introduction to Semantics

What is Semantics

- Very broadly, semantics is the study of meaning
 - Word meaning
 - Sentence meaning

- Layers of Linguistic Analysis
 1. Phonetics & Phonology
 2. Morphology
 3. Syntax
 4. Semantics
 5. Pragmatics
 6. Stylistics

Meaning in the larger context

- **Semiotics** is the study of interpreting symbols, or **signification**
 - We refer to the **signified**
 - Using a **signifier** Saussure
- We use language to talk about language, which can get messy. So we try to use certain words with very specific technical senses.
 - **technical term** ← remember me!
 - *word* “gloss” or *utterance*: example of a word/expression being used
 - *lexeme*: the abstraction of word in the lexicon
 - **predicate**: the abstraction of the meaning in the lexicon/formal semantic system
 - **CONCEPT**: the meaning in a person’s mind

Utterances, Sentences and Propositions

- **utterance**: an actual instance of saying (or writing or ...) something
- **sentence**: an abstraction, the type of what was said
 - (1) *Caesar invades Gaul*
- **proposition**: a further abstraction, normally ignoring some non-literal meaning
 - (2) `invade(Caesar, Gaul)`
- **interpretation**: our mental representation (linked to our existing knowledge)

The meanings of words

Words carry different meanings

- 10070 *Nothing was left 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 left 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 leave 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 left 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 leave me , but she paused at the door and looked back.*
- 10439 *” The rest you will leave in our hands . ”*
- 10449 *And now , Miss Stoner , we must leave 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 left in darkness .*

How many?



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 many 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

The polysemy of a word can be tested by a variety of means, including:

- **Antagonism**: can the word be used in a sentence with multiple competing 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 competing 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
- * black and white striped
- * herbivore

(redundant)

➤ 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 \models Kim has a pet animal

Kim has a pet animal $\not\models$ Kim has a pet dog

N.B. complications with universal quantifiers and negation:

Kim likes all animals \models Kim likes all dogs

Kim likes all dogs $\not\models$ Kim likes all animals

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

- Can a word have multiple hypernyms?

(3) *tank*₁ \subset *military_vehicle*₁; \subset *tracked_vehicle*₁; \subset *armored_vehicle*₁;
? \subset *weapon*₁

What is entailment

Entailment (\models): 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
<i>sheep</i>	<i>ram</i>	<i>ewe</i>	<i>lamb</i>
<i>cow</i>	<i>bull</i>	<u><i>cow</i></u>	<i>calf</i>
<i>goose</i>	<i>gander</i>	<u><i>goose</i></u>	<i>gosling</i>
<i>horse</i>	<i>stallion</i>	<i>mare</i>	<i>foal:colt/filly</i>
<i>dog</i>	<u><i>dog</i></u>	<i>bitch</i>	<i>puppy</i>
<i>snake</i>	<u><i>snake</i></u>	<u><i>snake</i></u>	<u><i>snake</i></u>

- Can you do this for *pig*, *cat* or *chicken*?
- Can you give an example of this in another language?



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.

(4) *dead/alive*

(5) *pass/fail*

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

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

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

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

(8) *ascend/descend*

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

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

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

(11) *employer/employee*

(Slightly non-standard usage)

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

(12) *Monday/Tuesday/.../Sunday*

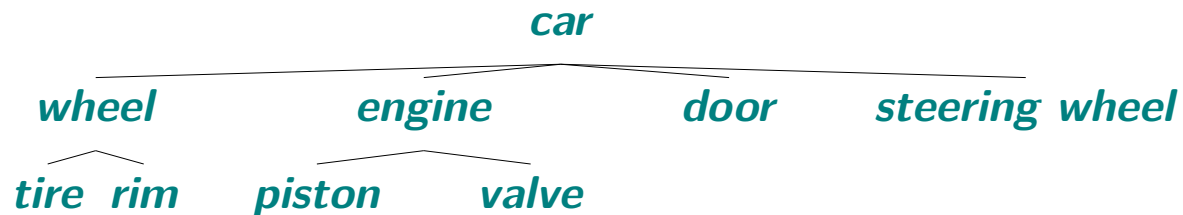
in WordNet: *day of the week* \supset *weekday, weekend*

(13) *LMS/English/Chinese/...*

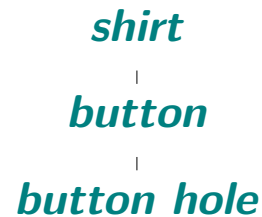
Context dependent

Meronymy

- **Meronymy** refers to the part-whole relation
 - **meronym** is the part
 - **holonym** is the whole



- It is not always transitive



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

(14) *tree-forest*

(15) *sheep-flock*

(16) *fish-school*

(17) *book-library*

(18) *member-band*

(19) *musician-orchestra*

(20) *student-class*

Portion-Mass

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

(21) *drop-liquid*

(22) *grain-sand/salt/truth*

(23) *sheet/ream-paper*

(24) *lump-coal (or just about anything)*

(25) *strand-hair*

(26) *rasher-bacon*

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

(27) *ekor* “tail”–*animal*

(28) *orang* “human”–*person*

Domain (lexical field)

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

- (29) *driver*₁ — the operator of a motor vehicle
- (30) *driver*₂ — someone who drives animals that pull a vehicle
- (31) *driver*₃ — a golfer who hits the golf ball with a driver [GOLF]
- (32) *driver*₄ — (\simeq device driver) a program that determines how a computer will communicate with a peripheral device [COMPUTER SCIENCE]
- (33) *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: approach₉, approach shot₁, golf course₁, links course₁, wedge₅, tee₁, scratch₉, putt₁, slice₁, hook₁

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”)
strong-willed 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, at LMS we work on:

- Japanese
- Bahasa Malay/Indonesian
- Chinese
- The shared open multi-lingual wordnet (150+ languages)

<http://compling.hss.ntu.edu.sg/omw/>

Wordnet Structure

- Lexical items are categorised into $\sim 115\text{K}$ (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)

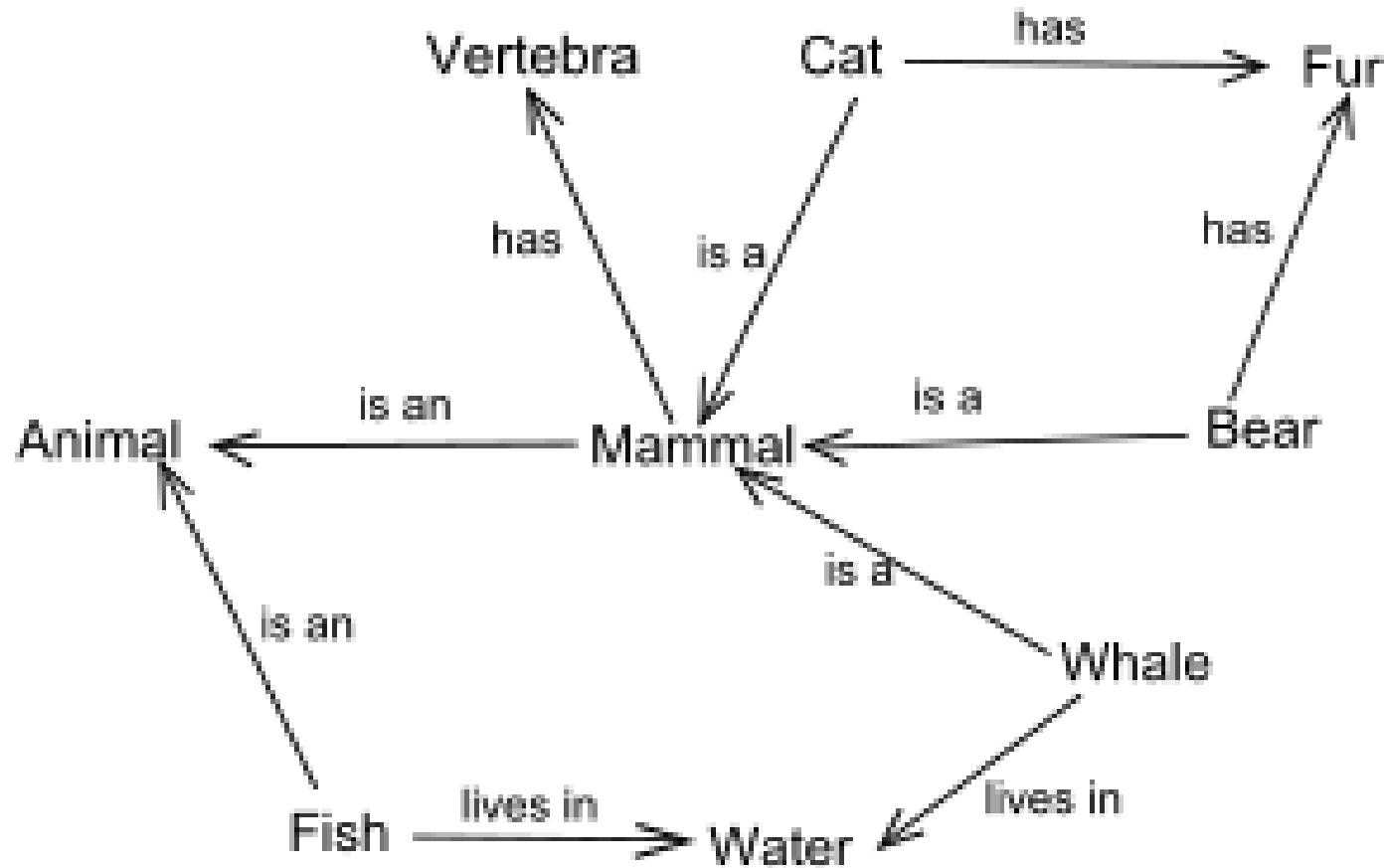
- Lexical relations at either the synset level or sense (= combination of lexical item and synset) level

- Strongly lexicalist (originally):
 - 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 *The Red-Headed League*
 - 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:

(34) *I patted the dog.*

(35) *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.

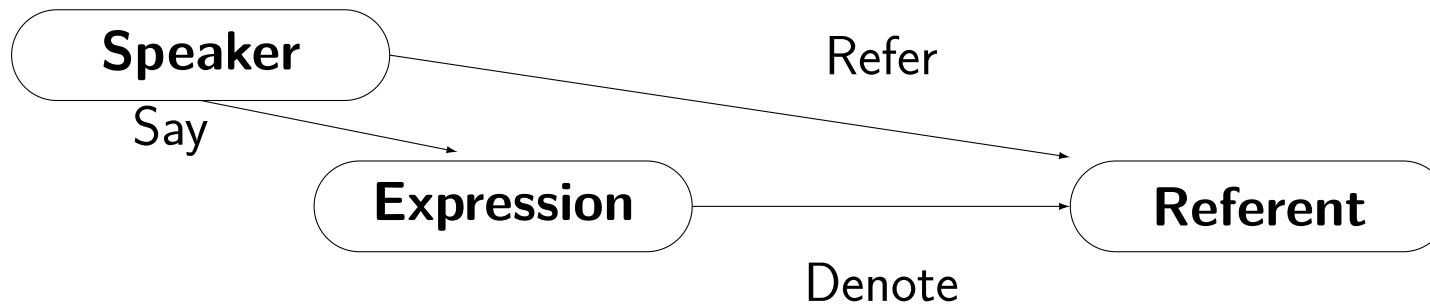
But we can represent the same reality in different ways:

(36) *Ich habe Hunger* “I have hunger”

(37) *I am hungry*

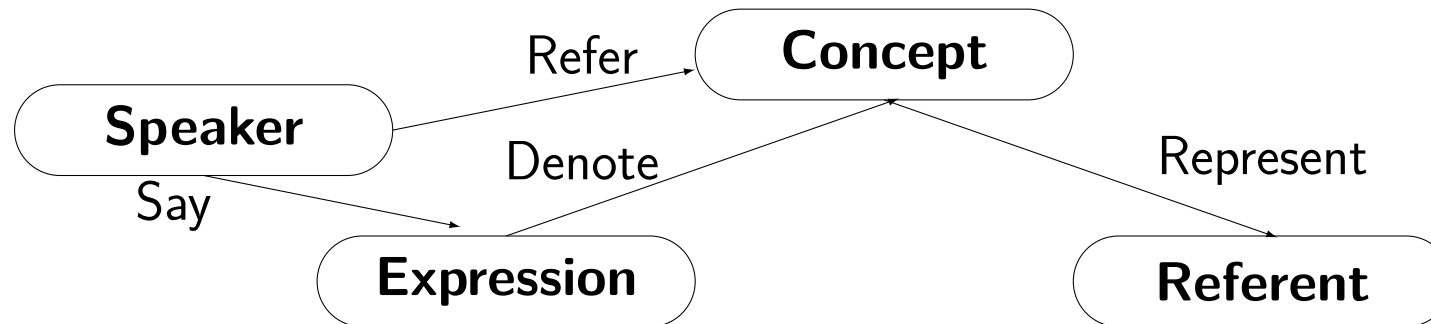
Representational theories are interested in how we represent reality, and how our representations are influenced by conceptual structures conventionalized in language.

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**)

(38) *cat, ano kiiro kaban* “that yellow bag”

(39) *London Bridge, Xiao Ming*

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

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

- Not all nominals refer

(41) *That is an ugly dog*

(42) *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

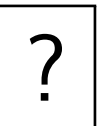
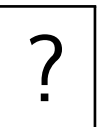
- Two way systems (English, ...)

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	<i>kore</i> “this”	<i>koko</i> “here”
medial	close to addressee	<i>sore</i> “that”	<i>soko</i> “there”
distal	far from both	<i>are</i> “’tother”	<i>asoko</i> “over there”
Q	interrogative	<i>dore</i> “what”	<i>doko</i> “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**

(43) *Here we begin explaining textual deixis*

➤ Often also used for time

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

➤ Spatial expressions extend to possession in many languages

(45) *NICT-ga Kyoto-ni aru*

NICT-NOM Kyoto-LOC be

NICT is in Kyoto

(46) *watashi-ni musuko-ga aru*

I-LOC son-NOM be

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

Person Deixis

- Minimally a three way division

First Person	Speaker	<i>I</i>
Second Person	Addressee	<i>you</i>
Third Person	Other	<i>he/she/it</i>

- 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

- 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

- (47) *Tanaka-san-ga kudasaimashita* [addressee and subject hon.]
Tanaka gave it to me (and I honor him and you)
- (48) *Tanaka-san-ga kudasatta* [subject honorification]
Tanaka gave it to me (and I honor him)
- (49) *Tanaka-kun-ga kuremashita* [addressee honorification]
Tanaka gave it to me (and I honor you)
- (50) *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 them to examples where he is addressed as *Mr Holmes*: what is the difference?



Types of Deixis

(a) Gestural; (b) Symbolic; (c) Non-deictic uses (Levinson 1983:66):

- (51) a. *You, you, but not you, are dismissed*
b. *What did you say?*
c. *You can never tell what they want nowadays*
- (52) a. *This finger hurts*
b. *This city stinks*
c. *I met this weird guy the other day*
- (53) a. *Push, not now, but now*
b. *Let's go now rather than tomorrow*
c. *Now, that is not what I said*
- (54) a. *Not that one, idiot, that one*
b. *That's a beautiful view*
c. *Oh, I did this and that*

Non-standard usage of deixis

- (55) *You take your screwdriver, right, and screw her home*
- (56) *Are we ready for our medicine now, Dr Smith?*
- (57) *We now turn to a discussion of globalisation in Chapter Three*
- (58) *When you're hot you're hot*
- (59) *Sometimes you wonder about the quality of the political leadership*
- (60) *She's a beauty all right* [said of a car]

Concepts

Prototypes

- Concepts are organized in groups around a **prototype**
- These have typical members (remembered as **exemplars**)
 - What is typical FURNITURE?
 - What is a typical BIRD?
- prototypes have **characteristic features**
 - has feathers
 - warbles
 - flies
 - lays eggs
- This work was pioneered by Eleanor Rosch (1973, 1975) (very readable)

Relations between Concepts

- Concepts are linked in many ways
- Most common relationship is **hypernymy**: DOG is-a ANIMAL
- Typically subordinate terms inherit properties from superordinate terms
Birds fly so *Sparrows fly*
- Larger units of knowledge, such as **frames** are similar

Basic Level Categories

- Some categories (concepts) seem to be more psychologically basic than others
 - Pictures of objects are categorized faster at the basic level
 - Basic level names used more often in free-naming tasks
 - Children learn them earlier
 - Basic-level names are more common in adult discourse
 - Basic-level categories are common in different cultures
 - Basic level names tend to be short
 - Basic-level names tend to be common in compound nouns

➤ superordinate	basic	subordinate
<i>vehicle</i>	<i>bus</i>	<i>school bus</i>
<i>jewelry</i>	<i>necklace</i>	<i>pearl necklace</i>
<i>animal</i>	<i>dog</i>	<i>poodle</i>

-
- Basic level categories are a decomposition of the world into maximally informative categories.
 - BLCs maximize the number of attributes shared by members of the category
 - BLCs minimize the number of attributes shared with other categories
 - It can be hard to agree on what is the Basic Level: whereas dog as a basic category is a species, bird or fish are at a higher level, etc.
 - Similarly, the notion of frequency is very closely tied to the basic level, but not exactly the same.

What is a word?

Defining word

➤ How many words are there in the following?

(61) *He who laughs last laughs longest.*

(62) *If he is right and I am wrong, are we both in trouble?*

(63) *I'm gonna go to the station-master.*

(64) *Sorry to knock you up, Mr Holmes.*

(65) 他们结婚了 *ta1men jie2hun1 le* “they got married” (他们结了婚)

➤ **Tokens:** Individual instances of a class

➤ **Types:** The class as a whole

-
- Why do we need a definition for *word*?
 - Psychological reality: People can divide language into words
 - Phonological contours: People pronounce words as unit
 - Orthographic practice: Many language put spaces between words
 - * Some put them between phrases (Korean)
 - * Some words include spaces *New York*, *ad hoc*

Bloomfield's grammatical definition

A word, then, is a free form, which does not consist entirely of (two or more) lesser free forms; in brief, a word is a *minimum free form*.

(Bloomfield 1984: p178)

In practice, the definition is somewhat task specific: it may make more sense to talk of **orthographic words**, **semantic words** or **predicates**,

Derivational Relations

Diathesis Alternations

➤ **Causative/inchoative alternation:**

*Kim broke the window ↔ The window broke
also *the window is broken* (state)*

➤ **Middle construction alternation:**

Kim cut the bread ↔ The bread cut easily

➤ **Conative alternation:**

Kim hit the door ↔ Kim hit at the door

➤ **Body-part possessor ascension alternation:**

Kim cut Sandy's arm ↔ Kim cut Sandy on the arm

Diathesis Alternations and Verb Classes

- A verb's (in)compatibility with different alternations is a strong predictor of its lexical semantics:

	<i>break</i>	<i>cut</i>	<i>hit</i>	<i>touch</i>
Causative	YES	NO	NO	NO
Middle	YES	YES	NO	NO
Conative	NO	YES	YES	NO
Body-part	NO	YES	YES	YES

break = {*break, chip, crack, crash, crush, ...*}

cut = {*chip, clip, cut, hack, hew, saw, ...*}

hit = {*bang, bash, batter, beat, bump, ...*}

touch = {*caress, graze, kiss, lick, nudge, ...*}

- **Corollary:** we can predict the syntax of novel words we are given the semantic class for
- The principal weakness of syntax-based verb classification is that there are often subtle divergences in semantics between synonyms (cf. *eat* vs. *devour* vs. *gobble*)

Agentitive Nouns

- The entity who/which performs the action of the verb
verb + *-er, -or, -ant*

(66) *murderer, commentator, whaler, director, computer*

(67) ?? *undertaker, cooker, footballer, crofter*

- Should *murderer* be listed separately from *murder* in the dictionary? Why or why not?



- Also the undergoer: **verb** + *-ee*: *employee*

- *-er|-or* has equivalents in most languages:
Japanese: 員, 者, 人, 機, ...

What if we had fewer words?

He was a fine creature, this man of the old English soil, simple, straight and gentle, with his great, earnest, blue eyes and broad, comely face. His love for his wife and his trust in her shone in his features. Holmes had listened to his story with the utmost attention, and now he sat for some time in silent thought. DANC

- Can we get the same message (denotation and connotation) with a smaller vocabulary? <http://xkcd.com/simplewriter/>
- It is hard!
- That is why we have so many words
- and why some writers are better than others.

Conclusions

- We learned about how to talk about meaning

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

Synonyms for a *dead Parrot*

be dead, be demised, be deceased, pass on, be no more, cease to be, expire, go to meet one's maker, be a stiff, be bereft of life, rest in peace, push up the daisies, one's metabolic processes are now history, be off the twig, kicked the bucket, shuffle off this mortal coil, ring down the curtain, join the choir invisible, be an ex-parrot

From the “Dead Parrot Sketch”, also known as the “Pet Shop Sketch” or “Parrot Sketch”, originally in *Monty Python's Flying Circus*, first performed in the eighth episode of the show's first series, “Full Frontal Nudity” (7 December 1969).



References

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

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

Beth Levin. 1993. *English Verb Classes and Alternations*. University of Chicago Press, Chicago, London.

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