Methods in Lexical Semantics
Word Meanings and Wordnets

Francis **Bond**

Department of Asian Studies,
Palacký University, Olomouc, Czechia

<bond@ieee.org>

LOT Winter School 2026

Faculty
of Arts

Welcome!

- In this course we will learn to analyze word meanings
  - How we can model lexical relations using wordnets
  - How we can compare and contrast languages
  - A deeper look at metaphor and metonymy
  - Distributional approaches to word meaning
  - Lexical resources for low resource languages

Faculty of Arts

Textbook and Readings

- 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)
  - Saeed, John (2015) *Semantics*. 4rd Edition. Wiley-Blackwell.
  - Lyons, John (1977) *Semantics*. Cambridge University Press

Faculty of Arts

Roadmap

1. Introduction to Semantics
2. The meanings of words
3. Lexical Semantics
4. Componential Analysis
5. Wordnet

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?

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

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)

Faculty of Arts

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

Faculty of Arts

Layers of Linguistic Analysis

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

Faculty of Arts

Do people share a common conceptual system?

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

  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?

Faculty of Arts

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*

Faculty of Arts

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

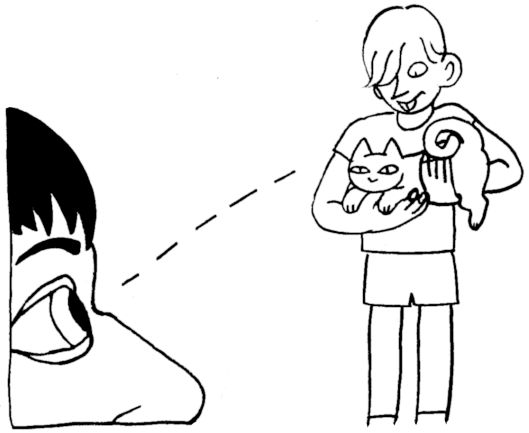
Faculty of Arts

Language is normally under-specified

We get **words**:

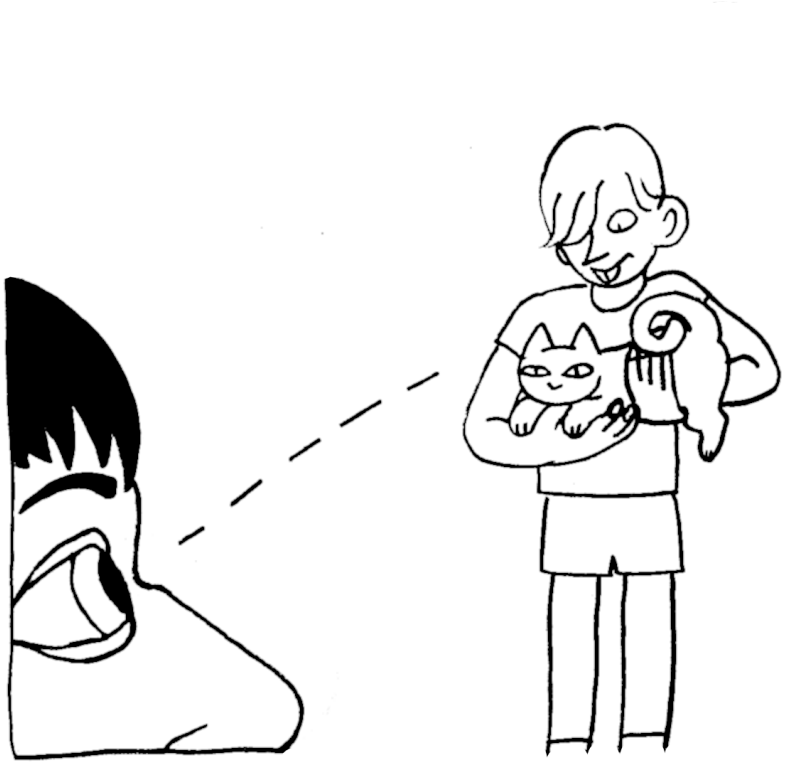
I saw a kid with a cat.

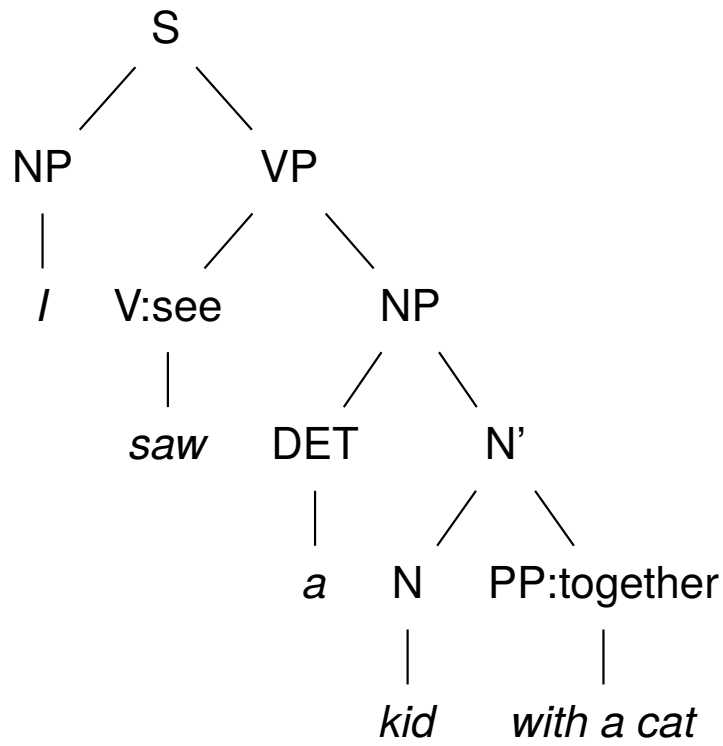
We want **meaning**:



Faculty of Arts

I saw a kid with a cat$_1$



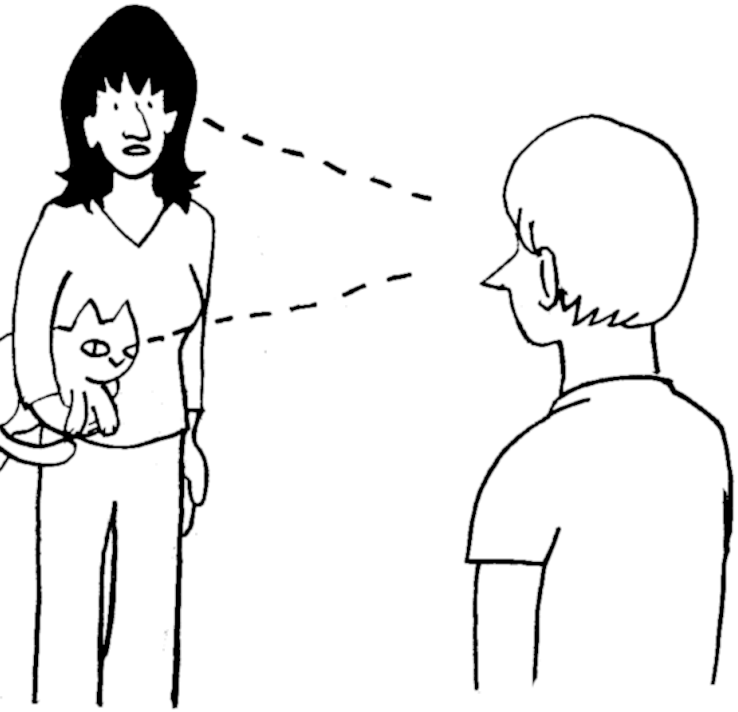


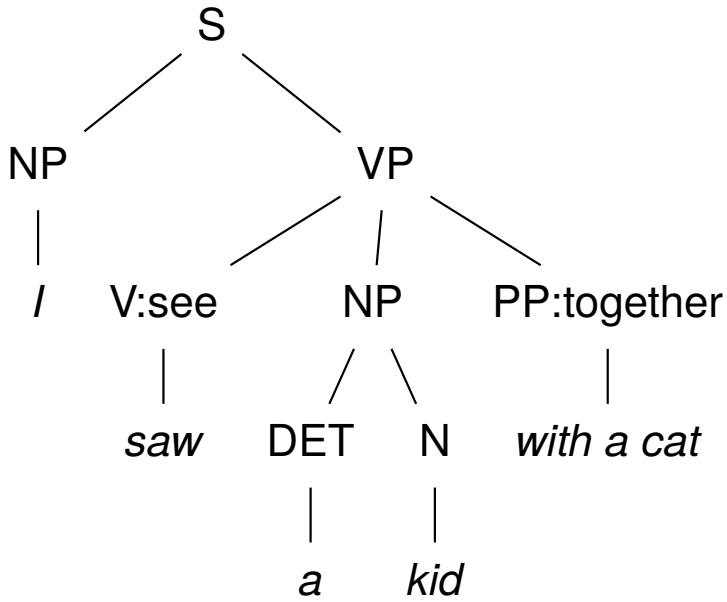

```
see(I, kid: past); with(kid, cat)
```

```
see ⊂ perceive
kid ∼ child
with ⊂ together
```

Faculty of Arts

I saw a kid with a cat$_2$



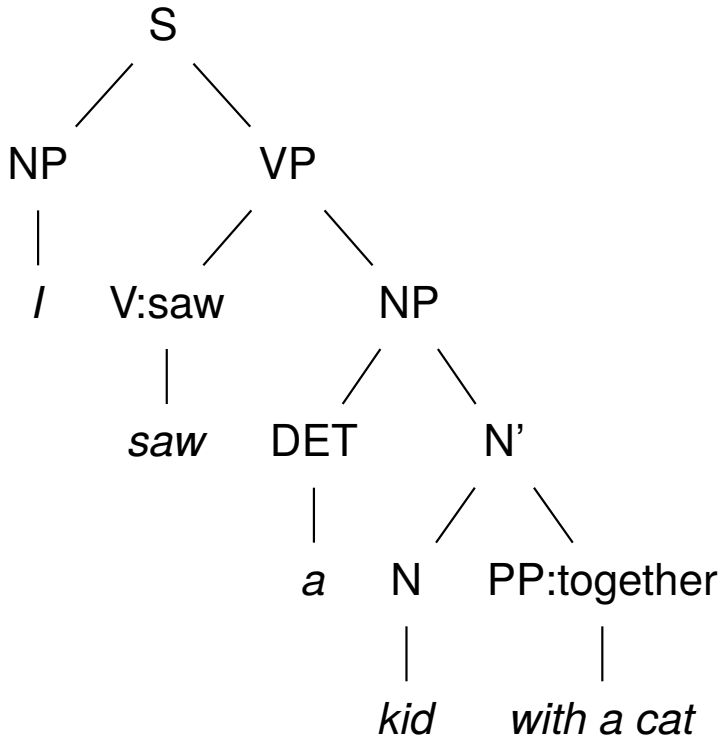



```
see(I, kid: past) with(I, cat)
```

```
see ⊂ perceive
kid ∼ child
with ⊂ together
```

Faculty of Arts

I saw a kid with a cat$_3$

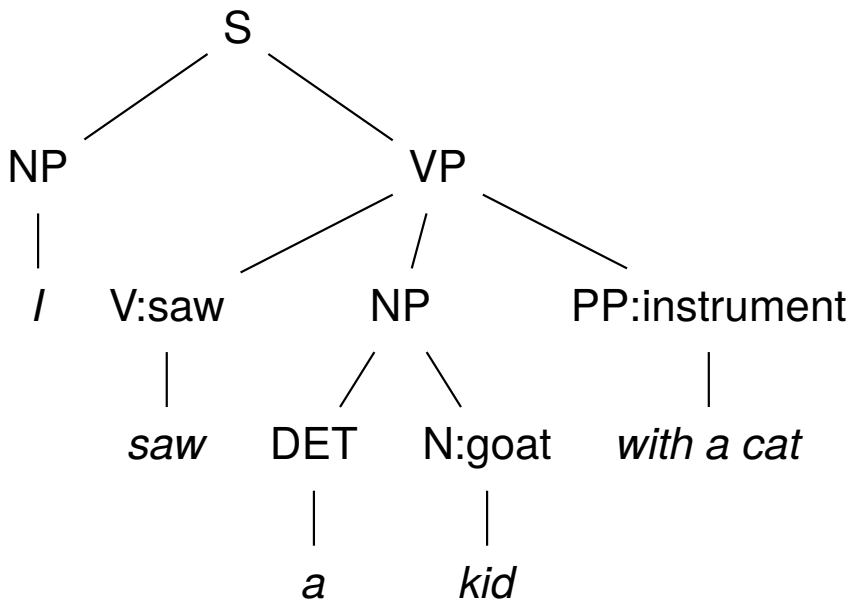



```
saw(I, kid: pres); with(kid,
cat)

saw ⊂ cut
kid ∼ child
with ⊂ together
```

Faculty of Arts

I saw a kid with a cat$_4$

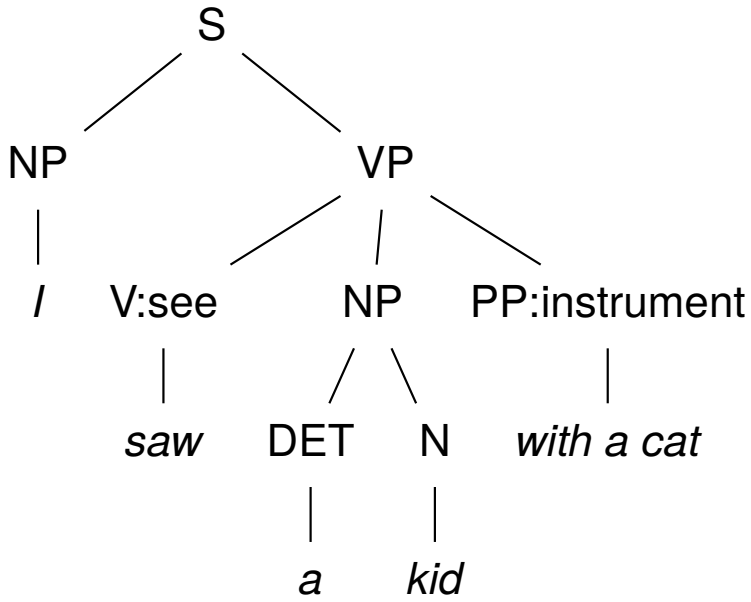



```
saw(I, kid: present) with(I, cat)

saw ⊂ cut
kid ∼ young goat
with ⊂ together
```

Faculty of Arts

I saw a kid with a cat$_5$




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

```
see ⊂ perceive
kid ∼ child
with ⊂ instrumental
```

Faculty of Arts

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 鋸 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 Your turn: try to paraphrase — translate into English
aim to be unambiguous, even if slightly disfluent

Overview

- Word Meaning
  - Definitions
  - Translations/paraphrases
  - Semantic Relations
  - Components
  - Word Embeddings
- Wordnet
- Close Reading and Word Sense Disambiguation

Roadmap

1. Introduction to Semantics
2. The meanings of words
3. Lexical Semantics
4. Componential Analysis
5. Wordnet

Faculty of Arts

Words carry different meanings: ***leave*** I

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 .*

Faculty of Arts

Words carry different meanings: ***leave*** II

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 different meanings??

From the NTU Multilingual Corpus (*Adventure of the Speckled band*, concept lemma = *leave*)

Faculty of Arts

How can we represent the differences?

- Definitions
- Translations/paraphrases
- Semantic Relations
- Components
- Word Embeddings

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`
  - 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)

Faculty of Arts

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!

Faculty of Arts

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

Faculty of Arts

Entries for ***leave*** I

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*, ….

Faculty of Arts

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*?

Faculty of Arts

Roadmap

1. Introduction to Semantics
2. The meanings of words
3. Lexical Semantics
4. Componential Analysis
5. Wordnet

Faculty of Arts

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*

Faculty of Arts

Distinguishing Polysemes

- **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.

Faculty of Arts

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
  - An albino zebra with three legs is still a zebra
- 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 ***grass snake***?
  - 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)

Faculty of Arts

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?

  (6) ***tank***$_1$ $\subset$ ***military_vehicle***$_1$; $\subset$ ***tracked_vehicle***$_1$; $\subset$ ***armored_vehicle***$_1$; ? $\subset$ ***weapon***$_1$

Faculty of Arts

What is **entailment**

Entailment *(⊨): 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 and Auto-hyponyms I

- The meanings of words change over time
  - ***guitar*** — “a stringed instrument usually having six strings”: originally these all used the body to make sound
  - 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
- we can also refer to the prototypical guitar (acoustic) using reduplication
  What kind of guitar do you play? Guitar guitar

Faculty of Arts

Language Change and Auto-hyponyms II

- 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?

Faculty of Arts

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***)

Faculty of Arts

Near Synonymy

- Near 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.*

  ** He likes alsations, in other words, spaniels*

Faculty of Arts

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) I

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

  (7) *dead/alive*

  (8) *pass/fail*

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

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

  (10) *fascinating/interesting/dull/boring*

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

  (11) *ascend/descend*

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

Faculty of Arts

Antonymy (opposites) II

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

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

  (14) *employer/employee*

  (Slightly non-standard usage by Saeed)

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

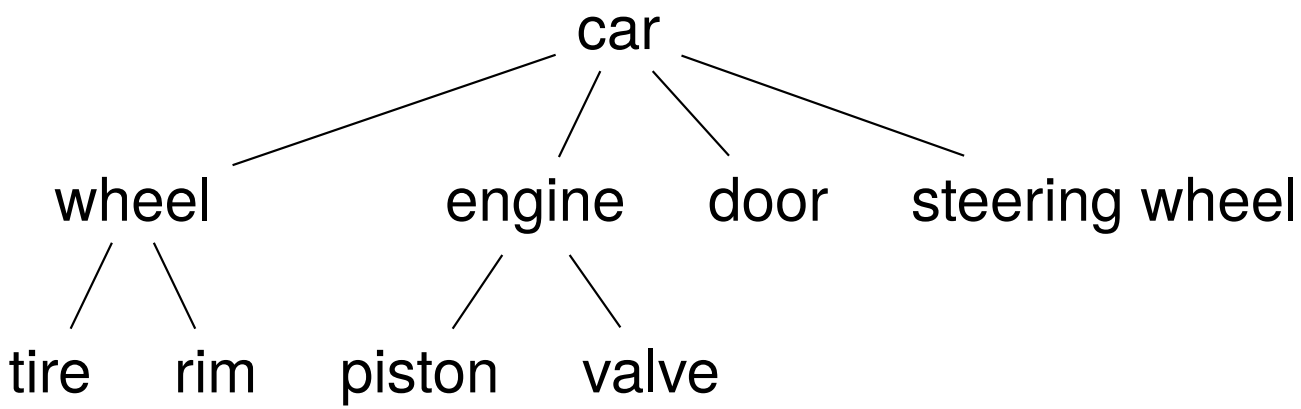
  (15) *Monday/Tuesday/…/Sunday*
  in WordNet: ***day of the week*** ⊃ ***weekday***, ***weekend***

  (16) *LMS/English/Chinese/…*
  Context dependent

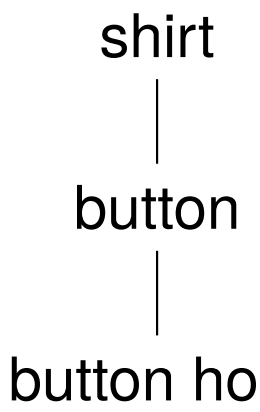
Faculty of Arts

Meronymy

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



- It is not always transitive:



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

  (17) *tree–forest*

  (18) *sheep–flock*

  (19) *fish–school*

  (20) *book–library*

  (21) *member–band*

  (22) *musician–orchestra*

  (23) *student–class*

Faculty of Arts

Portion-Mass

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

  (24) *drop–liquid*

  (25) *grain–sand/salt/truth*

  (26) *sheet/ream–paper*

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

  (28) *strand–hair*

  (29) *rasher–bacon*

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

  (30) *ekor* “tail”–*animal*

  (31) *orang* “human”–*person*

Faculty of Arts

Domain (lexical field)

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

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

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

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

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

(36) ***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***$_1$

Faculty of Arts

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**

Faculty of Arts

Roadmap

1. Introduction to Semantics
2. The meanings of words
3. Lexical Semantics
4. Componential Analysis
5. Wordnet

Faculty of Arts

Break word meaning into its components

- For example:

woman	[FEMALE]	[ADULT]	[HUMAN]	
spinster	[FEMALE]	[ADULT]	[HUMAN]	[UNMARRIED]
man	[MALE]	[ADULT]	[HUMAN]	
bachelor	[MALE]	[ADULT]	[HUMAN]	[UNMARRIED]
wife	[FEMALE]	[ADULT]	[HUMAN]	[MARRIED]
girl	[FEMALE]	[CHILD]	[HUMAN]	
boy	[MALE]	[CHILD]	[HUMAN]	

semantic components/**primitives** shown as [COMPONENT]

- components allow a compact description
- interact with morphology/syntax
- form part of our cognitive architecture

Faculty of Arts

Defining Relations using Components

- **hyponymy**

  A lexical item P is a hyponym of Q if all the components of Q are also in P.

woman	[FEMALE]	[ADULT]	[HUMAN]	
spinster	[FEMALE]	[ADULT]	[HUMAN]	[UNMARRIED]
wife	[FEMALE]	[ADULT]	[HUMAN]	[MARRIED]

  spinster $\subset$ ***woman***; ***wife*** $\subset$ ***woman***

- **incompatibility**

  *A lexical item P is incompatible with Q if they share some components but differ in one or more **contrasting** components*

  spinster $\not\approx$ ***wife***

Binary Features

- We can make things more economical (fewer components):

woman	[+FEMALE]	[+ADULT]	[+HUMAN]	
spinster	[+FEMALE]	[+ADULT]	[+HUMAN]	[–MARRIED]
bachelor	[–FEMALE]	[+ADULT]	[+HUMAN]	[–MARRIED]
wife	[+FEMALE]	[+ADULT]	[+HUMAN]	[+MARRIED]
girl	[+FEMALE]	[-ADULT]	[+HUMAN]	

  - Which should be +? [+FEMALE] or [–MALE]
  - Presumably also [–ELECTRIC], [–CONICAL], …
    Only show **relevant** features
  - **antonyms** differ in only one binary component

Redundancy Rules

- We can add relations between components:

[+HUMAN]	→	[+ANIMATE]	
[+ADULT]	→	[+ANIMATE]	
[+ANIMATE]	→	[+CONCRETE]	
[+MARRIED]	→	[+ADULT]	
[+MARRIED]	→	[+HUMAN]	…

- Which allows us to write:

woman	[+FEMALE]	[+ADULT]	[+HUMAN]	
spinster	[+FEMALE]	[+ADULT]	[+HUMAN]	[–MARRIED]
bachelor	[–FEMALE]	[+ADULT]	[+HUMAN]	[–MARRIED]
wife	[+FEMALE]			[+MARRIED]

Can we say [–MARRIED] → [+HUMAN]?

Faculty of Arts

More Complex Breakdowns

- We can add relations between components:

[+FATHER]	→	[+MALE] [+PARENT]
[+FATHER](x,y)	→	[+MALE](x) [+PARENT](x,y)
[+SON](x,y)	→	[+MALE](x) [+PARENT](y,x)
[+BROTHER](x,y)	→	[+MALE](x) [+PARENT](z,x) [+PARENT](z,y
[+GRANDFATHER](x,y)	→	[+MALE](x) [+PARENT](x,z) [+PARENT](z,y

- Assume [+PARENT](x,y) means "x is the parent of y"
- There are various ways you can formalize such relationships
  - Many parts of language can be formalized in such a way

These are great for many sub-systems of language, but it is hard to make components for everything, …

Faculty of Arts

Word Embeddings

- Represent words as a vector of numbers (instead of a set of components)
- Every word has a unique word embedding (or “vector”), which is just a list of numbers for each word.
- Embeddings start being useful from 50-500 dimensions
  LLMs typically are much larger
- The embedding captures the “meaning” of the word.
- Similar words end up with similar embedding values
- Context based word embeddings give a different vector depending on the context

More later, …

Faculty of Arts

Derivational Relations

- Often words are linked by more or less systematic relations, sometimes morphologically marked
  - ***beauty*/*beautiful*/*beautify***
  - ***cute*/*cuteness***

Faculty of Arts

Agentive Nouns

- An **agentive noun** is a word that is typically derived from another word denoting an action, and that identifies an entity that does that action.
  verb + ***-er, -or, -ant***

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

  (38) ?? ***undertaker, cooker, footballer*** (Saeed also includes these)

- Should ***murderer*** be listed separately from ***murder*** in the dictionary? Why or why not?
- Also **recipient noun**s that show the undergoer: **verb** + ***-ee***: ***employee***, ***trustee***

Faculty of Arts

Agentive Nouns in Other Languages

- Japanese (suffix distinguishes person/machine)
  - 運転する → 運転者 *unten-sha* “driver”
  - 計算する → 計算者 計算機 *keisan-sha/ki* “computer”
  - 研究する → 研究者 研究員 *kenkyuu-sha/in* “researcher”
  - 読む → 読み 読者 *yomite/dokusha* “reader”
- Malay (prefix can convert any part of speech)
  - *bantu (v)* “help” → *pembantu* “assistant/helper”
  - *potong (v)* “cut” → *pemotong* “cutter (human/machine)”
  - *terbang (v)* “fly” → *penerbang* “pilot (not passenger)”

Faculty of Arts

Roadmap

1. Introduction to Semantics
2. The meanings of words
3. Lexical Semantics
4. Componential Analysis
5. Wordnet

Faculty of Arts

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

Faculty of Arts

Wordnet Structure

- Lexical items are categorised into ∼115K (and counting) glossed **synsets** (= synonym sets)
- 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 ($\approx 50\%$)

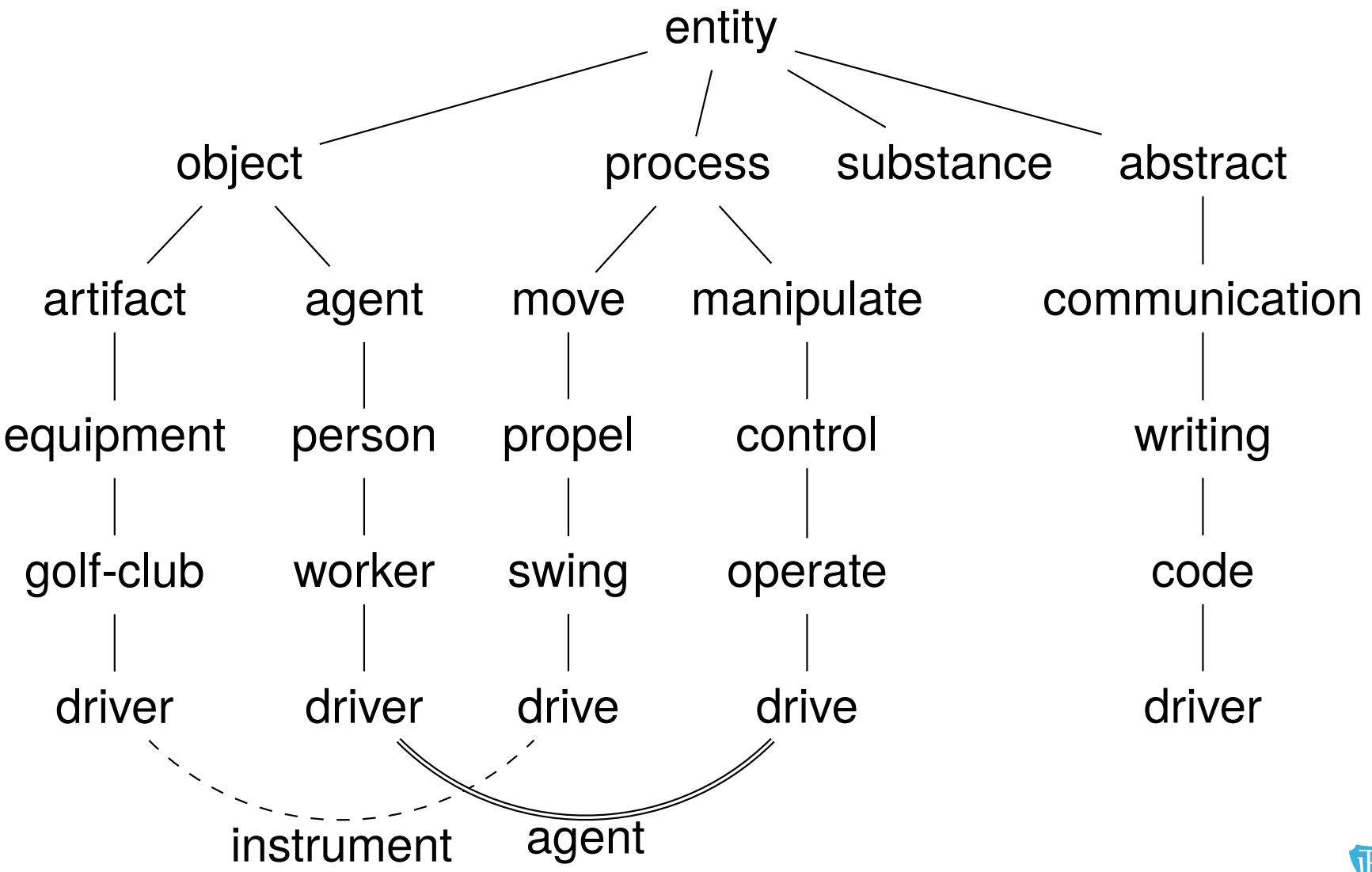
Psycholinguistic Foundations of WordNet

- 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



Faculty of Arts

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
- my students have used wordnets for:
  - Japanese derivational relations (Bond and Wei, 2019)
  - pronoun representation for Japanese, Mandarin and English (Seah and Bond, 2014)
  - exclamatives and classifiers (Mok et al., 2012; Morgado da Costa and Bond, 2016)
  - sentiment analysis (Le et al., 2016; Bond et al., 2019)
  - cross-lingual sense annotation (Bonansinga and Bond, 2016)
  - multilingual crosswords (Tan, 2012)

  ...

Faculty of Arts

References I

Giulia Bonansinga and Francis Bond. 2016. Exploring cross-lingual sense mapping in a multilingual parallel corpus. In *Proceedings of the 8th Global Wordnet Conference (GWC 2016)*, pages 45–49.

Francis Bond, Arkadiusz Janz, and Maciej Piasecki. 2019. A comparison of sense-level sentiment scores. In *Proceedings of the 11th Global Wordnet Conference (GWC 2019)*.

Francis Bond and Ryan Lim Dao Wei. 2019. Generating derivational relations for the Japanese wordnet: The case of agentive nouns. In *2019 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC)*, pages 1–7.

Tuan Anh Le, David Moeljadi, Yasuhide Miura, and Tomoko Ohkuma. 2016. Sentiment analysis for low resource languages: A study on informal Indonesian tweets. In *Proceedings of The 12th Workshop on Asian Language Resources*, page 123–131. Osaka.

Faculty of Arts

References II

Hazel Shuwen Mok, Eshley Huini Gao, and Francis Bond. 2012. Generating numeral classifiers in Chinese and Japanese. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*. Matsue. 211-218.

Luís Morgado da Costa and Francis Bond. 2016. Wow! what a useful extension to wordnet! In *10th International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož.

Yu Jie Seah and Francis Bond. 2014. Annotation of pronouns in a multilingual corpus of Mandarin Chinese, English and Japanese. In *10th Joint ACL - ISO Workshop on Interoperable Semantic Annotation*. Reykjavik.

Jeanette Yi Wen Tan. 2012. *Automatic Generation of Multilingual Crossword Puzzles with WordNet*. Final year project, Linguistics and Multilingual Studies, Nanyang Technological University.

Faculty of Arts