- Next lectures
  - Word sense disambiguation
    - Background from linguistics
      - Lexical semantics
    - On-line resources
    - Computational approaches

# Semantic analysis

- Assigning meanings to utterances/sentences
- **Compositional semantics**: we can derive the meaning of the whole sentence from the meanings of its parts.
  - Marseille ate a green apple.
- Relies on knowing:
  - the meaning of individual words
  - how the meanings of individual words combine to form the meaning of groups of words
  - how it all fits in with syntactic analysis

# Caveats

- Problems with a compositional approach… examples?

# Introduction to lexical semantics

- Lexical semantics is the study of
  - the systematic meaning-related connections among words and
  - the internal meaning-related structure of each word
- Lexeme
  - an individual entry in the lexicon
  - a pairing of a particular orthographic and phonological form with some form of symbolic meaning representation
- Sense: the lexeme's meaning component
- Lexicon: a finite list of lexemes

# Dictionary entries

- right *adj.* located nearer the right hand esp. being on the right when facing the same direction as the observer.
- left *adj.* located nearer to this side of the body than the right.

---

- red *n.* the color of blood or a ruby.
- blood *n.* the red liquid that circulates in the heart, arteries and veins of animals.

# Lexical semantic relations

- Homonyms: *words that have the same orthographic and phonological form and unrelated meanings*
  - Instead, a ***bank***$^{1}$ can hold the investments in a custodial account in the client's name.
  - But as agriculture burgeons on the east ***bank***$^{2}$, the river will shrink even more.
- Homophones: distinct lexemes with a shared pronunciation
  - E.g. *would* and *wood*, *see* and *sea.*
- Homographs: identical orthographic forms, different pronunciations, and unrelated meanings
  - The expert angler from Dora, Mo., was fly-casting for **bass** rather than the traditional trout.
  - The curtain rises to the sound of angry dogs baying and ominous **bass** chords sounding.

# Why do these distinctions matter?

- One type or another is more likely to affect specific NLP applications.
  - Spelling correction?
  - Speech recognition?
  - Text-to-speech?

# Lexical semantic relations: polysemy

- Polysemy: the phenomenon of multiple *related* meanings within a single lexeme
  - Example: While some ***banks*** furnish blood only to hospitals, others are much less restrictive.
  - New sense, e.g. bank[3]?
  - Polysemy allows us to associate a lexeme with a set of related senses.
- Distinguishing homonymy from polysemy is not always easy. Decision is based on:
  - Etymology: history of the lexemes in question
  - Intuition of native speakers

# Synonymy

- Lexemes with the same meaning
- Invoke the notion of **substitutability**
  - Two lexemes will be considered synonyms if they can be substituted for one another in a sentence without changing the meaning or acceptability of the sentence
    - How *big* is that plane?
    - Would I be flying on a *large* or small plane?
    - Miss Nelson, for instance, became a kind of *big* sister to Mrs. Van Tassel's son, Benjamin.
    - We frustrate 'em and frustrate 'em, and pretty soon they make a *big* mistake.
    - Also issues of **register**
      - Social factors that surround the use of possible synonyms, e.g. politeness, group status.

# Word sense disambiguation

- Given a *fixed* set of senses associated with a lexical item, determine which of them applies to a particular instance of the lexical item in running text
- Two fundamental approaches
  - WSD occurs during semantic analysis as a side-effect of the elimination of ill-formed semantic representations
  - Stand-alone approach
    - » WSD is performed independent of, and prior to, compositional semantic analysis
    - » Makes minimal assumptions about what information will be available from other NLP processes
    - » Applicable in large-scale practical applications

- Next lectures
  - Word sense disambiguation
    - Background from linguistics
      - Lexical semantics
    - On-line resources
    - Computational approaches