Formal Language vs Natural Language

Syntax Rules

Designed by people for specific applications

- Tokens
- Structure

- Programming languages
- Math notations
- Chemistry

$$3 + 3 = 6$$

$$3 = + 6\$$$

Evolved naturally

Parsing – a process of determining structures
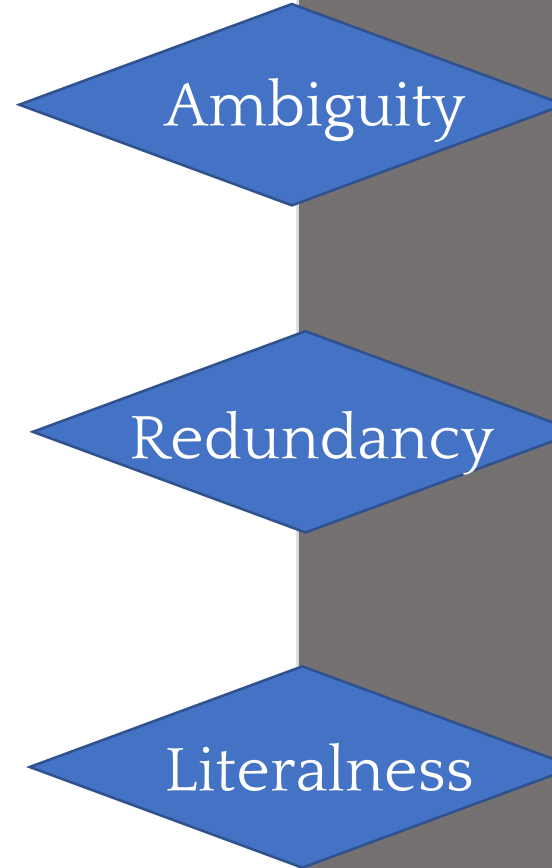
Semantics – a process of determine their meaning

Language Structure

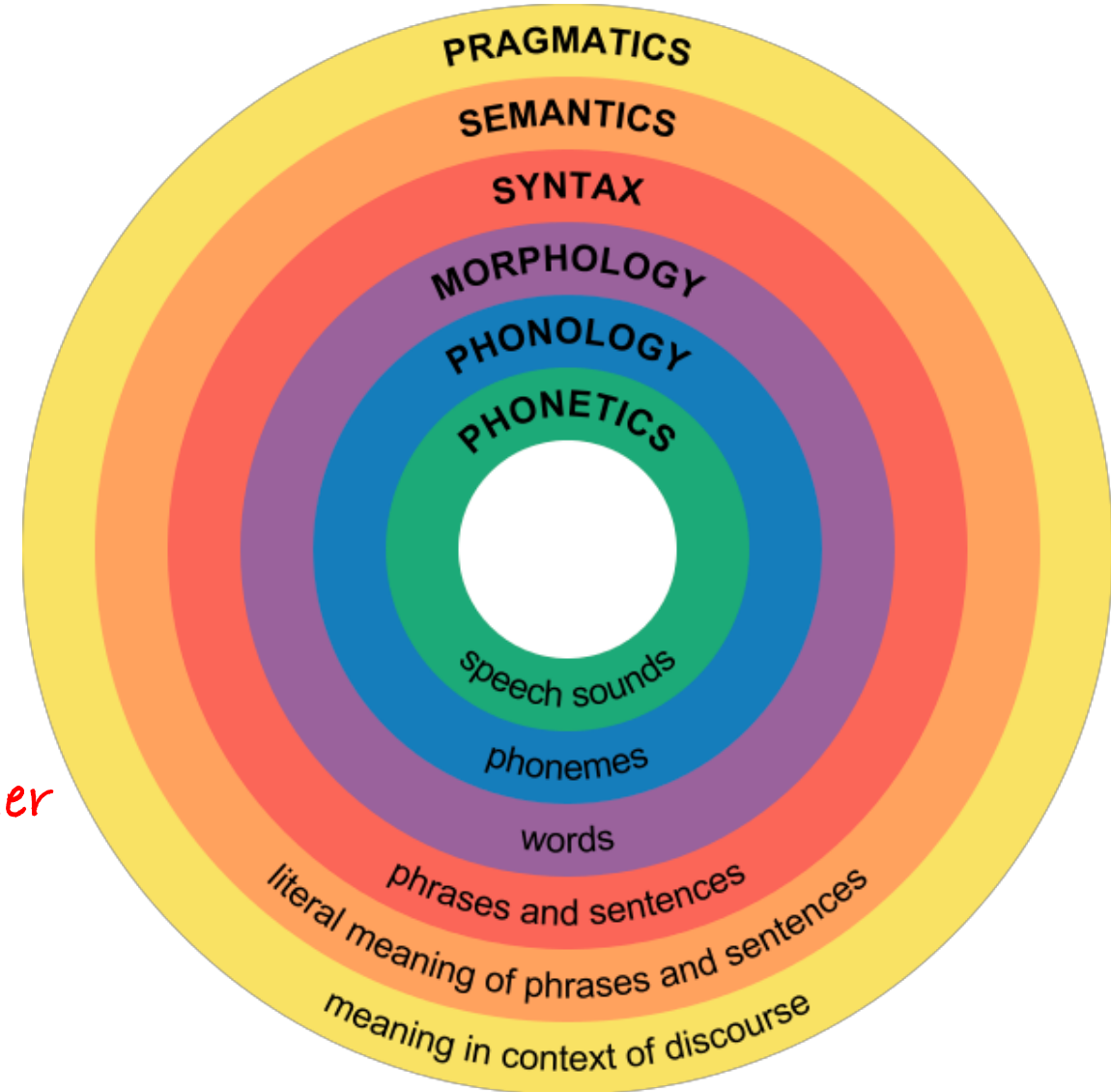**Phonetics** – the study of speech sounds in their physical aspects

Phonology – the study of speech sounds in their cognitive aspects

Morphology – the study of the formation of words
dog dogs walk walked

Syntax – the study of the formation of sentences
She walked the dog versus The dog walked her

Semantics – the study of meaning

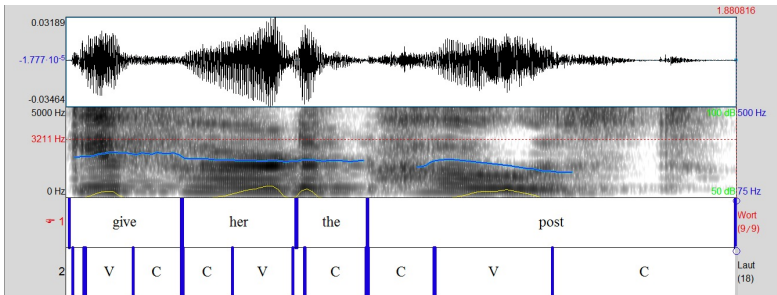**Pragmatics** – the study of language use

PRAGMATICS
SEMANTICS
SYNTAX
MORPHOLOGY
PHONOLOGY
PHONETICS

speech sounds
phonemes
words
phrases and sentences
literal meaning of phrases and sentences
meaning in context of discourse

Language Units

Phonemes

- the basic units of phonology
- the smallest unit of sound that may cause a change of word meaning
- does not have meaning by itself

Morphemes

- the basic units of morphology
- the smallest meaningful unit of language

prefix root word suffix

uncomfort**able**
irregular**ly**
disorganis~~ed~~
unconfident**ly**
disrespect**ful ly**

Lexemes

- the set of inflected forms taken by a single word
- they are included under one entry in the dictionary

run
running
ran

a alamy stock photo

Language Morphology

Noun
Words depicting object or entity that could be living or non-living:
fox, dog, book

Verb
Words describing certain actions, states, or occurrences:
running, jumping, read, and write

Adjective
Words describing or qualifying other words, typically nouns and noun phrases:
beautiful flower

Adverb
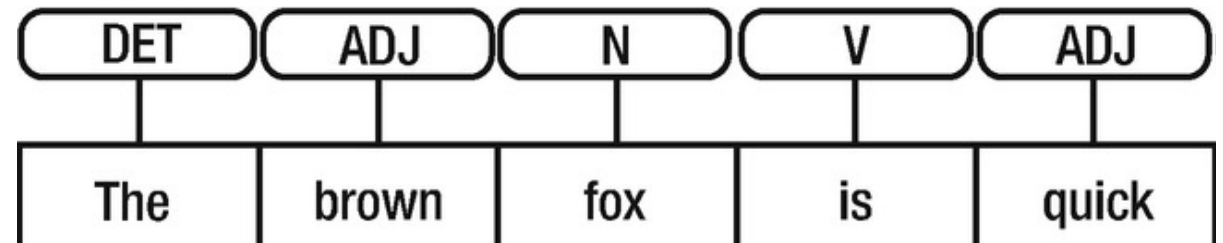Modifiers for other words including nouns, adjectives, verbs, or other adverbs:

very beautiful flower

Part of Speech Tags (POS): N V ADJ ADV

POS Tags

Tag	Description
CC	Coordinating conjunction
CD	Cardinal number
DT	Determiner
EX	Existential *there*
FW	Foreign word
IN	Preposition or subordinating conjunction
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative
LS	List item marker
MD	Modal
NN	Noun, singular or mass
NNS	Noun, plural
NNP	Proper noun, singular
NNPS	Proper noun, plural
PDT	Predeterminer

- POS is a grammatical tagging or word–category disambiguation
- POS tags are used to determine major syntactic categories

https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

Language Pragmatics
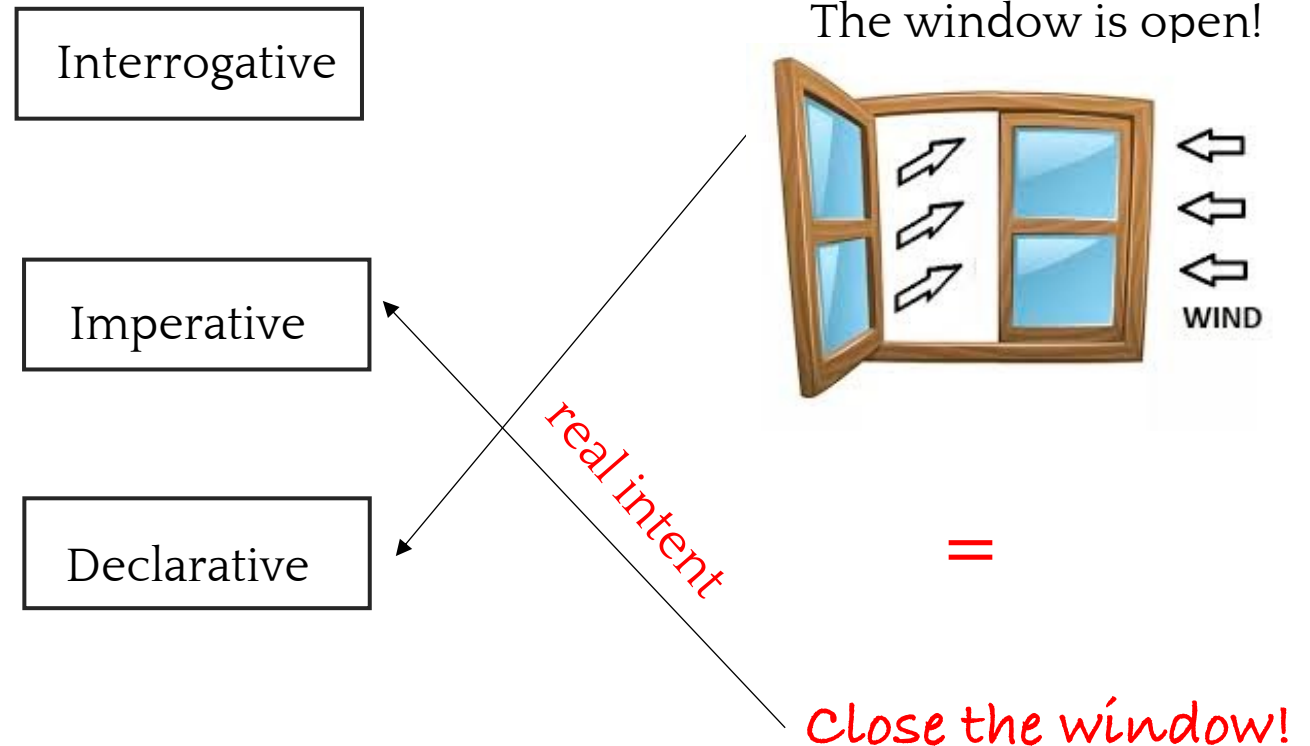
Direct Speech Act

Indirect Speech Act

the study of the context-dependent aspects of meaning

Greeting: "How are things going?"
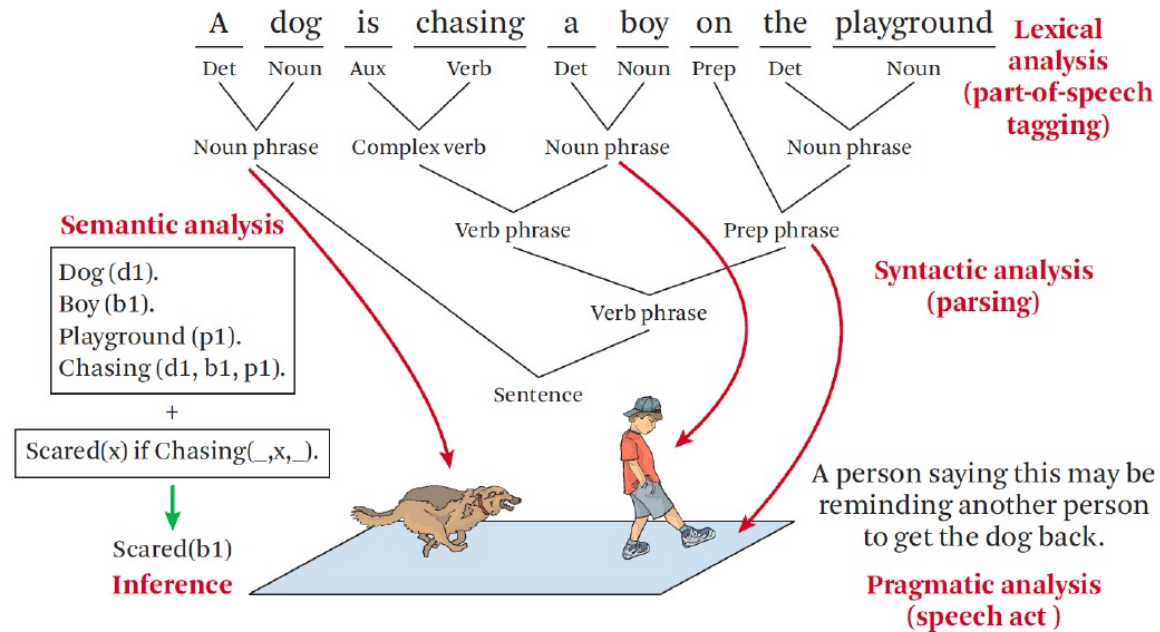**Request:** "Could you pass me the mashed potatoes, please?"
Complaint: "I've already been waiting three weeks for the computer. Please deliver my computer!"
Invitation: "We're having some people over Saturday evening and wanted to invite you to join us."

Interrogative

Imperative

Declarative

Sentence Type

The window is open!

WIND

real intent

=

Close the window!

Summary: NLP Tasks

A	dog	is	chasing	a	boy	on	the	playground
Det	Noun	Aux	Verb	Det	Noun	Prep	Det	Noun

Lexical analysis (part-of-speech tagging)

Noun phrase · Complex verb · Noun phrase · Noun phrase

Semantic analysis

Dog (d1).
Boy (b1).
Playground (p1).
Chasing (d1, b1, p1).

+

Scared(x) if Chasing(_,x,_).

↓

Scared(b1)
Inference

Verb phrase · Prep phrase

Syntactic analysis (parsing)

Verb phrase

Sentence

A person saying this may be reminding another person to get the dog back.

Pragmatic analysis (speech act)

1. **Lexical analysis** - meaningful units in a language
2. **Syntactic analysis** - how words are related in a sentence
3. **Semantic analysis** - the meaning of a sentence
4. **Pragmatic analysis** - meaning in context (speech acts)
5. **Discourse analysis** - connection between sentences

Zhai (2016)