

Syntax

What is syntax?

- The word syntax comes from Ancient Greek: "coordination", which consists of syn, "together," and táxis, "an ordering".
- Syntax is a subfield of linguistics that studies how expressions can combine to form larger expressions, i.e. how words and phrases make sentences.
- The words and phrases are linguistic expressions, pieces of languages with a certain form, a certain meaning, and some syntactic properties.

Syntax

- Any speaker of any human language can produce and understand an infinite number of possible sentences
- The kindhearted boy had many girlfriends.
The kindhearted, intelligent boy had many girlfriends.
The kindhearted, intelligent, handsome boy had many girlfriends.
- Thus, we can't possibly have a mental dictionary of all the possible sentences
- Rather, we have the rules for forming sentences stored in our brains
- **Syntax** is the part of grammar that pertains to a speaker's knowledge of sentences and their structures

What the Syntax Rules Do

- The rules of syntax combine words into phrases and phrases into sentences
- They also specify the correct word order for a language
 - For example, English is a Subject-Verb-Object (SVO) language
 - *The President nominated a new Supreme Court justice*
 - **President the new Supreme justice Court a nominated*
- They also describe the relationship between the meaning of a group of words and the arrangement of the words
 - *I mean what I say vs. I say what I mean*

Syntax

- The rules of the syntax also specify the **grammatical relations** of a sentence, such as **subject** and **direct object**. In other words, they provide information about who is doing what to whom.
- Syntax rules specify constraints on sentences based on the verb of the sentence
- *Zack believes Robert to be a gentleman*
- **Zack believes to be a gentleman*
Zack tries to be a gentleman
**Zack tries Robert to be a gentleman*

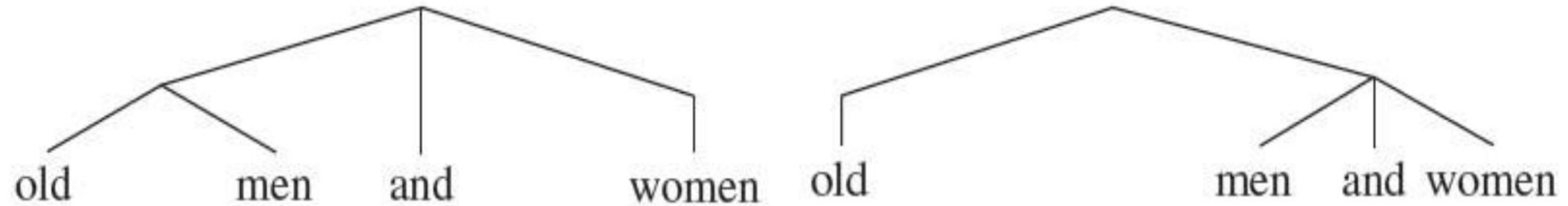
What the Syntax Rules Do

- Syntax rules also tell us how words form groups and are hierarchically ordered in a sentence

“The captain ordered the old men and women off the ship”

- This sentence has two possible meanings:
 - 1. The captain ordered the old men and the old women off the ship
 - 2. The captain ordered the old men and the women of any age off the ship
- The meanings depend on how the words in the sentence are grouped (specifically, to which words is the adjective ‘old’ applied?)
 - 1. The captain ordered the [old [men and women]] off the ship
 - 2. The captain ordered the [old men] and [women] off the ship

- These groupings can be shown hierarchically in a tree



- These trees reveal the structural ambiguity in the phrase “old men and women”
 - Each structure corresponds to a different meaning

What Grammaticality Is Not Based On

- People can judge grammaticality without ever having heard the sentence before

“Enormous crickets in pink socks danced at the prom.”

- Grammaticality is not based on meaningfulness

“Colorless green ideas sleep furiously.”

“A verb crumpled the milk.”

*‘Twas brillig, and the slithy toves
Did gyre and gimble in the wabe*

- Grammaticality is not based on truthfulness

Sentence Structure

- We could say that the sentence

“The child found the puppy” is based on the template

Det—N—V—Det—N

- But this would imply that sentences are just strings of words without internal structure
- This sentence can actually be separated into several groups:
 - [the child] [found a puppy]
 - [the child] [found [a puppy]]
 - [[the] [child]] [[found] [[a] [puppy]]]

Constituents and Constituency Tests

- **Constituents** are the natural groupings in a sentence
- Tests for constituency include:
 - 1. “stand alone test”: if a group of words can stand alone, they form a constituent
 - A: “What did you find?”
 - B: “A puppy.”
 - 2. “replacement by a pronoun”: pronouns can replace constituents
 - A: “Where did you find a puppy?”
 - B: “I found him in the park.”
 - 3. “move as a unit” test: If a group of words can be moved together, they are a constituent
 - A: “The child found a puppy.” → “A puppy was found by the child.”

Constituents and Constituency Tests

- Experimental evidence shows that people perceive sentences in groupings corresponding to constituents
- Every sentence has at least one constituent structure
 - If a sentence has more than one constituent structure, then it is ambiguous and each constituent structure corresponds to a different meaning

Constituents and Constituency Tests

- Experimental evidence shows that people perceive sentences in groupings corresponding to constituents
- Every sentence has at least one constituent structure
 - If a sentence has more than one constituent structure, then it is ambiguous and each constituent structure corresponds to a different meaning

Syntactic Categories

- A **syntactic category** is a family of expressions that can substitute for one another without loss of grammaticality

The child found a puppy.

A police officer found a puppy.

Your neighbor found a puppy.

The child **found a puppy**.

The child **ate the cake**.

The child **slept**.

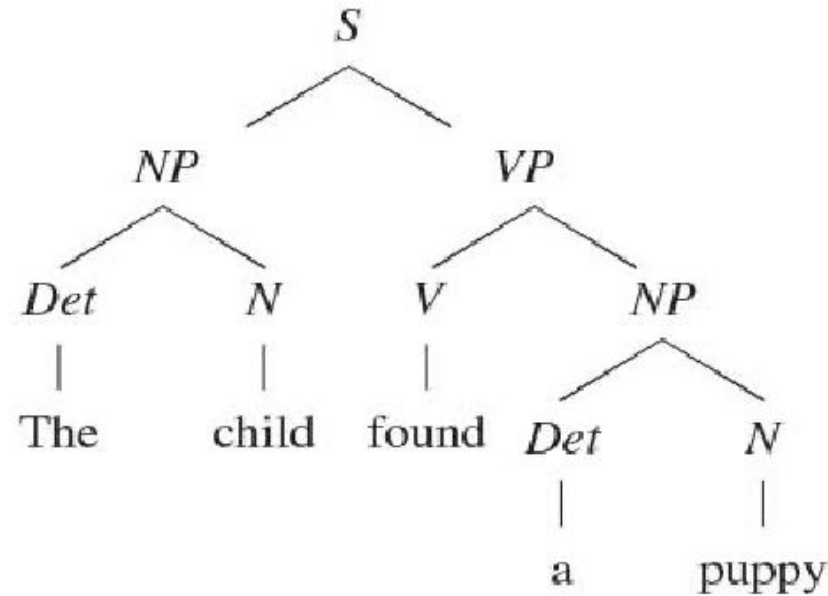
- All the underlined groups constitute a syntactic category known as a **noun phrase (NP)**
 - NPs may be a subject or an object of a sentence, may contain a determiner, proper name, pronoun, or may be a noun alone
- All the bolded groups constitute a syntactic category known as a **verb phrase (VP)**
 - VPs must always contain a verb but may also contain other constituents such as a noun phrase or a **prepositional phrase (PP)**

Syntactic Categories

- Phrasal categories: NP, VP, PP, AdjP, AdvP
- Lexical categories:
 - Noun: *puppy, girl, soup, happiness, pillow*
 - Verb: *find, run, sleep, realize, see, want*
 - Preposition: *up, down, across, into, from, with*
 - Adjective: *red, big, candid, lucky, large*
 - Adverb: *again, carefully, luckily, very, fairly*
- Functional categories:
 - Auxiliary: verbs such as *have*, and *be*, and modals such as *may, can, will, shall, must*
 - Determiners: *the, a, this, that, those, each, every*

Phrase Structure Trees and Rules

- A **phrase structure (PS) tree** (or **constituent structure tree**) is a tree diagram with syntactic category information.



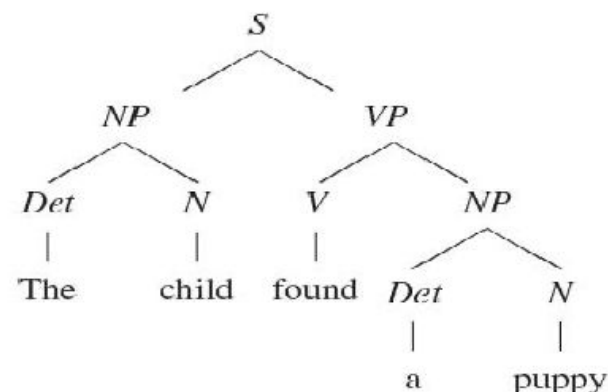
Phrase Structure Trees and Rules

- In a PS tree, every higher node **dominates** all the categories beneath it

- S dominates everything

- A node **immediately dominates** the categories directly below it

- The VP immediately dominates the V and the NP



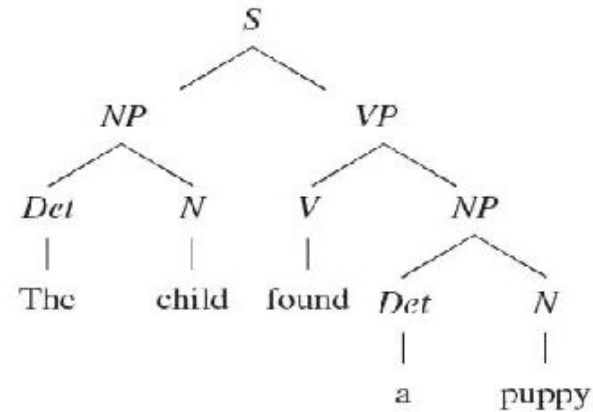
- **Sisters** are categories that are immediately dominated by the same node

- The V and the NP are sisters

Phrase Structure Trees and Rules

- Phrase structure rules specify the well-formed structures of a sentence
 - A tree must match the phrase structure rules to be grammatical
- This tree is formed using the following rules:

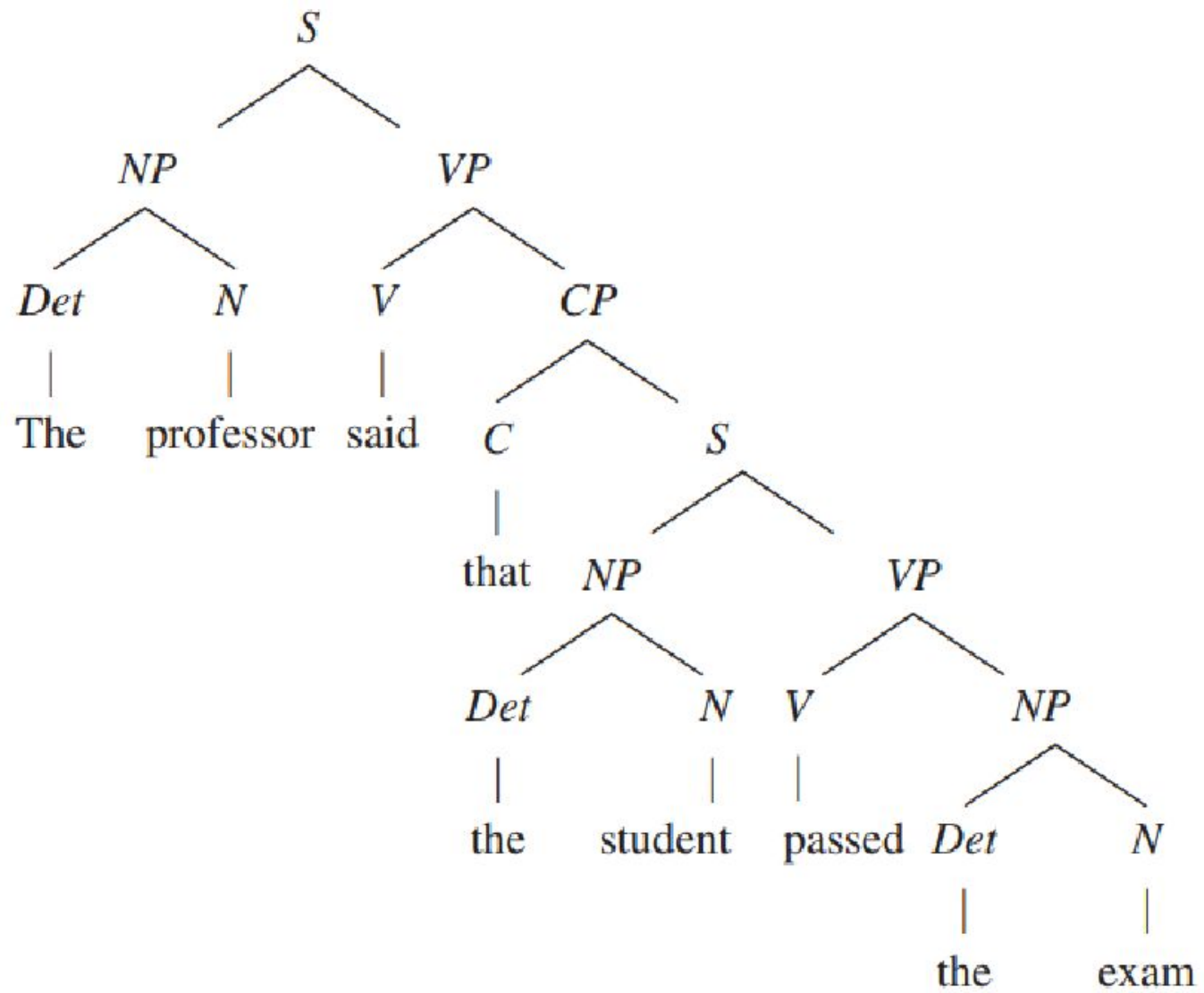
$S \rightarrow NP VP$
 $NP \rightarrow Det N$
 $VP \rightarrow V NP$



Phrase Structure Trees and Rules

- But, a VP could also contain:
 - A verb only: *The woman laughed.*
 - A PP: *The woman laughed in the garden.*
 - A CP: *The man said that the woman laughed.*
- We therefore have to account for these possible sentences in our phrase structure rules and need the following rules so far:

1. S → NP VP
2. NP → Det N
3. VP → V NP
4. VP → V
5. VP → V PP
6. PP → P NP
7. VP → V CP
8. CP → C S



Phrase Structure Trees and Rules

- Phrase structure rules are used as a guide for building trees
- To build a tree you expand every phrasal category until only the lexical categories remain
- By following the guidelines in the phrase structure rules, we can generate all the possible grammatical sentences in a language
 - Any tree that violates the phrase structure rules will represent an ungrammatical sentence

- <https://www.youtube.com/watch?v=B1r1grQiLdk>
- <https://www.youtube.com/watch?v=n1zpnN-6pZQ>