Syntax: Building sentences

Module 1 [LE321L Introduction to linguistics], JNU Monsoon Sem. 2020

Yangchen Roy

Infinite systems are unlearnable

Also, sentences can be infinitely long

- The cat chased the mouse

- The cat chased the mouse that ate the cheese.

- The cat chased the mouse that ate the cheese that came from the cow.

- The cat chased the mouse that ate the cheese that came from the cow that grazed in the field.

In principle, this can go on forever.

- The space in the brain is finite, so it cannot be that we store infinitely long sentences in our brains in a dictionary format. Rather, sentences are composed of discrete units that are combined by rules. This system of rules is finite.

Sentences are rule-based

- In the previous Module *Morphology*, you would have studied that **words are the building blocks of sentences.**

1) **ate a pot of soup yesterday spicy Meena**

2) **fire saw the the roof man on**

3) **the well tall was tree beside planted the by my mother**

There is more to sentences than just words. In any languge, when words are put together to build sentences, there are rules that are automatically followed.

The part of grammar that represents a speaker's knowledge of sentences and their structures is called syntax

Let's unjumble the words here to make a sentence:
ate a pot yesterday spicy of soup Meena

- [soup]

  [[spicy [soup]]

  [pot [of [spicy [soup]]]]

  [a [pot [of [spicy [soup]]]]]

- [ate]

- [yesterday]

- [Meena]

[ate]

[ate [a pot of spicy soup]]

[Meena [ate a pot of spicy soup]]

[Yesterday [Meena [ate a pot of spicy soup]]]

or

[[Meena [ate [a pot of spicy soup]]] yesterday]

What we just did is **systematically** group the words into larger meaningful units, called **"constituents"**, and put these constituents together to build **larger constituents**, and finally put these constituents together to get the final unit **– the sentence. What is this system?**

We used an operation called "Merge"

- Merge is a simple operation. It basically says, take two units A and B and put them together to give you a larger unit C. Now take C (which, of course is made up of A and B), and put it together with another element D and so on.

Fun fact! Merge is also the operation you use to add numbers

Let's unjumble the words here to make a sentence:
the read poem meeting at Sophia the

- [poem]

[the [poem]]

- [meeting]

[the [meeting]]

[at [the [meeting]]]

- [Sophia]

- [read]

[read]

[read [the poem]]

[read [the poem] [at the meeting]]

[Sophia [read [the poem][at the meeting]]]

Is Merge enough?

- Let's say we have the group of words [Meena], [a pot of soup], [ate]

- We know that Meena is the person who did the eating and what was eaten was soup. We want to build the appropriate sentence.

- We take [ate] and [a pot of soup] and use Merge to give us the bigger unit. **What are the possible combinations we can get?**

Word Order

- [ate] [a pot of soup]

- [a pot of soup] [ate]

- Merge generates both these options!

- But because we are merging groups of words of the English language, the rule system of English will tell us that (1) is the correct way of building the sentence, not (2).

- The rule system of English will again tell us that [Meena] [ate a pot of soup] is the correct order after merging [Meena] and [ate a pot of soup] and NOT [ate a pot of soup] [Meena]

- What we just did is fix the order of the units after merging the two units. **Word Order of sentences matters.** And **languages differ in the Word Orders they have**. What order of words for the above sentence does your language have

- Let is look at the word order of English

(1) * The dog the cat chased

(2) The dog chased the cat

Now look at the following sentence pair. They have the same set of words, but mean different things

(3) He says what he means (Everything he says is something that he means)

(4) He means what he says (If he means something he says it, but he also says things he doesn't mean)

Also see

(5) My dog chased your cat

(6) Your cat chased my dog

The rules of syntax specify grammatical relations of sentences, like subject (S) and object (O). These relations help the speaker convey important information to the listener: who is doing what to whom. This information is important if we want to know the meaning of a sentence.

- English and Mandarin are Subject—Verb—Object (SVO) languages

(1) Seema bought flowers

    S   **V**   **O**

(2) Zhāngsān   shōudǎole   [yifēng   xìn]
    Zhangsan   receive   one   letter
    S   **V**   **[O**   **]**

    'Zhongsan received a letter'

(3) Japanese is a SOV language

    Johnga   tegamio   yonda

    John   letter   read

    S   **O**   **V**

    'John read the letter.'

Irish is a **VSO** language, **Nias** is **VOS** and **Hixkaryana** is **OVS**

Can you figure out the word-order of your language?

The verb decides

(1) Transitive verbs

- Look at the following sentences

1) *The boy found.

2) *The boy found quickly.

3) *The boy found in the house.

4) The boy found the ball.

A native speaker of English will predict that sentence 1) to 4) are ungrammatical. This is because the rules of syntax specify that when you have a verb like found you need to follow this word with 'what was found', something like 'the ball'

Verbs like found, that **necessarily/obligatorily need** an object like 'the ball' to complete their meaning are called transitive verbs.

More transitive verbs

want

* Sarah wants

* Sarah wants tomorrow

* Sarah wants badly

Sarah wants a new pen

Sarah wants to go to the market

see

*Meena saw

*Meena saw yesterday

*Meena saw closely

Meena saw an owl

Meena saw that I was chopping carrots

Sarah wants a new dress badly, Sarah wants a new dress tomorrow, Meena saw an owl yesterday and *Meena saw an owl closely* are also possible sentences. What does this tell you about the obligatoriness of words like *tomorrow, badly, yesterday* and *closely?*

(4) Intransitive verbs

1) ***Disa slept the baby.**

2) **Disa slept soundly**

3) **Disa slept**

sleep is different from *found*, *see* and *want*. It cannot have an object like *the baby*, *a new pen* or *an owl*. Other verbs like *sleep* are *go*, *run*, *cry* etc.

While transitive verbs obligatorily need an object to complete their meaning, intransitive verbs can never need for an object to complete their meaning.

An easy test to identify transitive verbs and separate them from intransitive verbs is to **ask a 'What'-question.**

What does Sarah want?　　　What did Meena see?　　　*What did Disa sleep

(3) Ditransitive verbs

- Some verbs obligatorily need two objects to complete their meaning. These are called ditransitives.

1) Nitin gave **a sack of rice** to **Sidhra**

2) Sidhra sent **a letter** to **her mother**

If a sentence is a play or film the verb is its director. It is the one that decides whether it needs objects and if so what kinds of objects and how many. This is called "verb subcategorization"

The minimum ingredient needed to make a sentence is a verb.

Go!

Fire! (as in shoot)

16

Types of sentences

- **simple sentence**

(1)   Mani saw a rabbit

- has only one complete clause

A clause is a unit that has (at least) one verb. A sentence is a clause; in fact it is a complete/independent clause. An independent clause or sentence is grammatical (makes sense).

- **compound sentence – two sentences joined together by a conjunction**

- has two independent clauses

(2) Mani saw a rabbit and Jenny saw a peacock

(3) Mani saw a rabbit but Jenny saw a peacock

- **complex sentence– a sentence with an almost-sentence inside it**

- has two or more clauses but only one complete/independent clause. The rest will be dependent clauses.

(4) [Jenny said [that Mani saw a rabbit]]

that Mani saw is rabbit is the dependent clause here. It is not grammatical on its own, and needs the help of another clause (an independent one) to be . The same is the case with *Mani to come home*, below:

(5) [Jenny asked [Mani to come home]]

Other sentences

- **negated sentences**

(6) Mani did not see a peacock

- **questions – polar questions & wh-questions**

(7) Did Mani see a peacock?

(8) What did Jenny see?; Where did Mani see the rabbit?; How big was the rabbit?
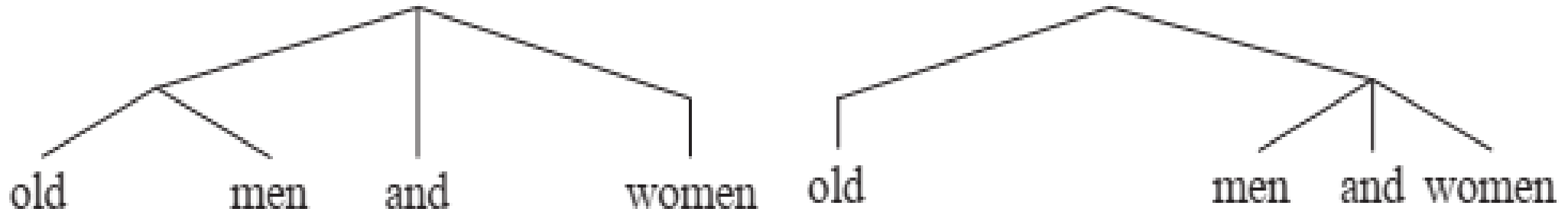
Sentence structure is hierarchical

structural ambiguity as a clue

1) The captain ordered the old men and women off the sinking ship

   [old men] and [women] the adjective old only modifies men

   [old [men and women]] the adjective old modifies both *men* and *women*

1) **Seema hit the man with the umbrella**

2) **I discovered an old English poem**

3) **We will oil your sewing machine and adjust tension in your home for $10.00.**

4) **Enraged cow injures farmer with axe**

5) **Hospitals are sued by seven foot doctors**

What do these sentences have in common?

They are all ambiguous i.e. they have two meanings or "readings". Each reading is correct in a separate context. The two readings that we get is because of the different syntactic structures associated with each reading. **Structural ambiguity** is a good way to know that a sentence has two separate syntactic structures.

Structural ambiguity is not the same as lexical ambiguity: this will make you smart.
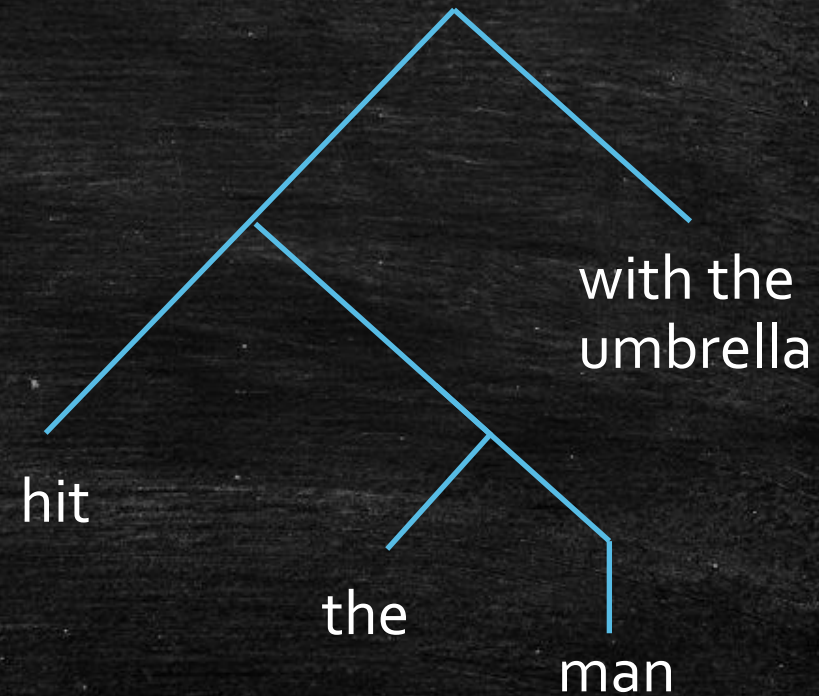
Seema hit the man with the umbrella

hit

the

man   with the
umbrella

hit

the

man

with the
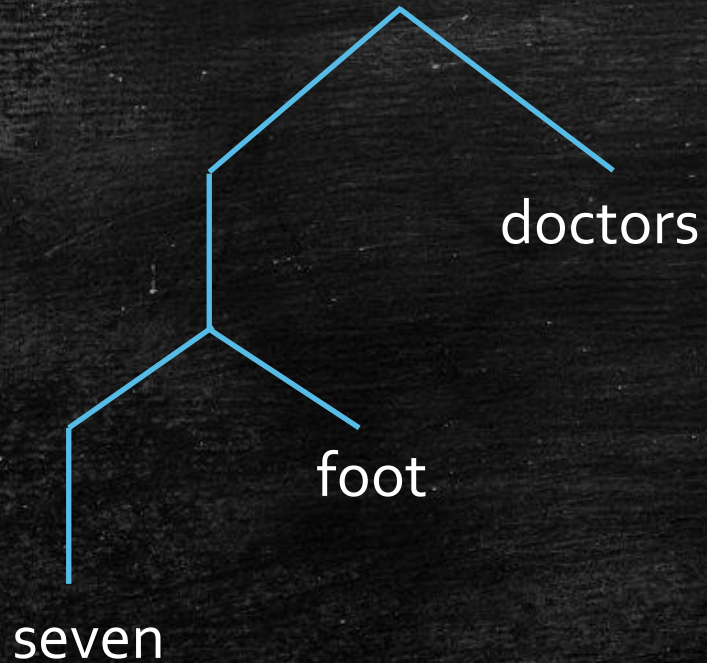umbrella
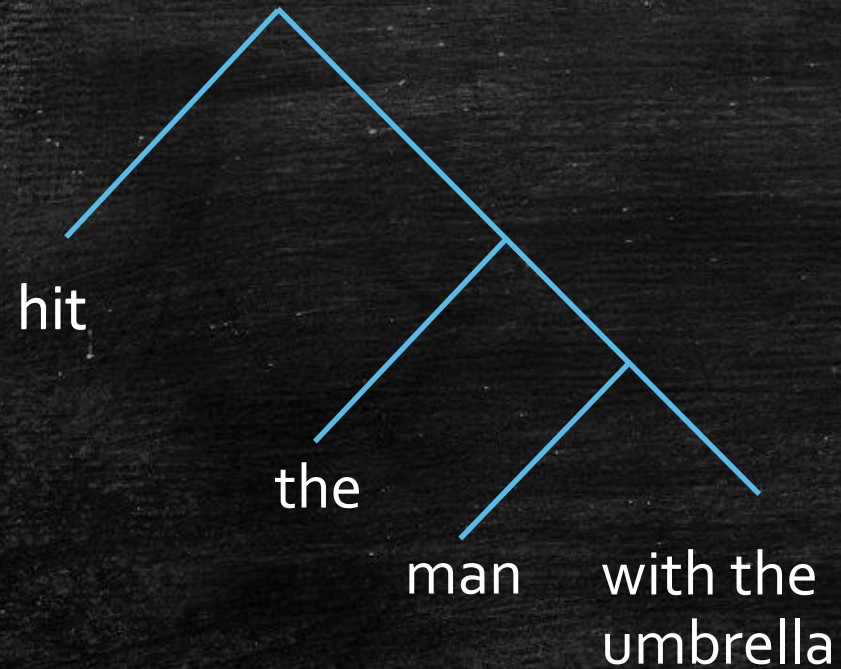
**the man who was hit had
an umbrella with him**

**the hitting was done
with an umbrella**

Hospitals are sued by seven foot doctors

doctors who are each seven feet tall

doctors

foot

seven

foot doctors seven in total

foot doctors

seven

23

What Grammaticality is not based on

- A person's ability to make grammaticality judgments does not depend on having heard the sentence before. You may never have heard or read the sentence "**Enormous crickets in pink socks danced at the prom",** but your syntactic knowledge tells you that it is grammatical.

- Grammaticality also does not depend on the truth of sentences. If it did, lying would be impossible.

What Grammaticality is not based on

- grammaticality and meaningfulness are not the same thing. The sentences in yellow have weird meanings, but they are all grammatical. The ones maked with a '*' sound like gibberish, they are ungrammatical.

1) **Colorless green ideas sleep furiously**

2) ***Furiously sleep ideas green colorless**

3) **A verb crumpled the milk**

4) ***Milk the crumpled verb a**

5) **A pot of soup ate Meena yesterday**

6) *** ate a pot yesterday spicy of soup Meena**

7) **The poem read Sophia at the meeting**

8) *** the read poem meeting at Sophia the**

- If we were to write a template that describes the structure of an English sentence i.e. a template that gives us the correct word order of English, we may get something like the following:

- Det – N – V – Det – N

A determiner (articles like a/an/the) followed by a noun, which is followed by a verb which again is followed by a Det and then an N. This can actually describe many sentences in English:

(1) The child found a puppy

(2) The professor wrote a book

(3) That runner won the race

This makes it seem that sentences are a linear string of words, where the words belong to a particular grammatical category.

But sentences, we have just seen, are not flat structures. Words in a sentence have a hierarchical organization.

Tree diagrams

- [[the] [child] [[found] [[a] [puppy]]]

- This hierarchical organization is best represented using what is called a tree-diagram, an upside down tree with a root encompassing the entire sentence on top.

- The tree conveys the same information as the nested brackets.

- What are the parts of found a puppy?

- [found] [a puppy] and then, further, [a] and [puppy]
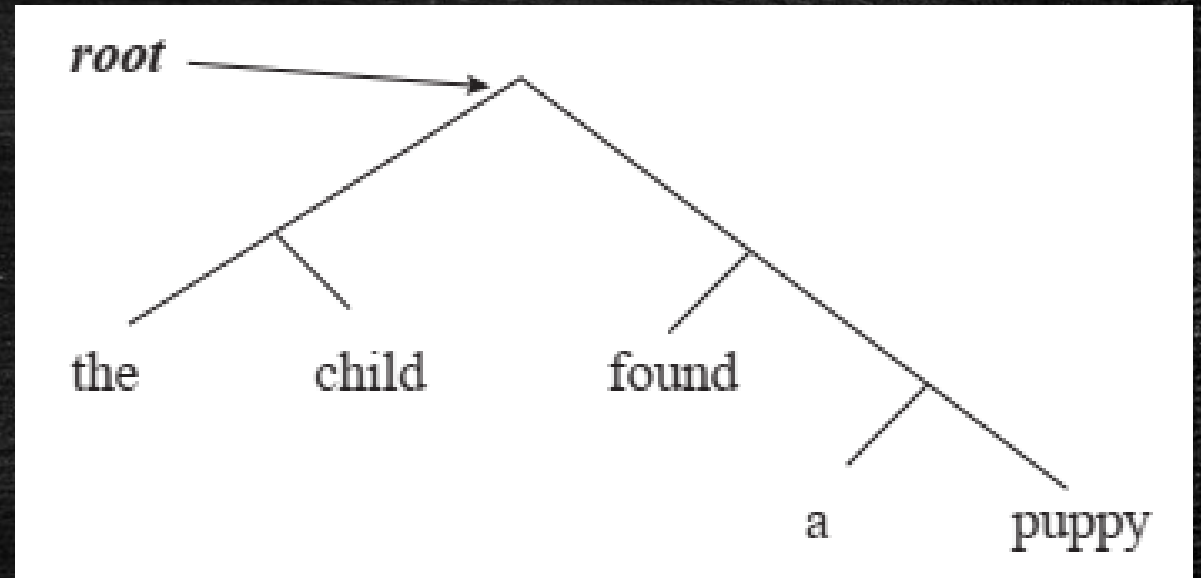
Constituency and constituency tests

the phrase **ran up the hill** behaves differently from **ran up the bill**, even though the two phrases are superficially similar.

Jack and Jill ran [up the hill].

*Jack and Jill ran the hill up.

[Up the hill] ran Jack and Jill.

Jack and Jill ran up the bill.

Jack and Jill ran the bill up.

*Up the bill ran Jack and Jill.

These sentences act differently because they have different syntactic structure associated with them.

In *ran up the hill*, *up the hill* forms **a unit**. But *up the bill* is **not a unit** in *ran up the bill* (rather, *ran up* forms a unit here).

29

- The natural grouping or parts of a sentence are called constituents.

- A sentence itself is a constituent

- "Every sentence in a language is associated with one or more constituent structures. If a sentence has more than one constituent structure, it is ambiguous, and each tree will correspond to one of the possible meanings, like the sentences on Slide 21.

- In *ran up the hill*, *up the hill* is a constituent. But *up the bill* is not a constituent. In *ran up the bill*, rather, *ran up* forms a constituent. When words form a unit of this sort we put them inside "[ ]".

- The way constituents are revealed is via what are called "constituency tests"

(1) the "stand-alone" test

- If a group of words can stand alone, they form a constituent. For example, the set of words that can be used to answer a question is a constituent.

- Let's look at the sentence **The child found a puppy** again. The answer to the question "What did you find?" can be

a puppy

*found a

- Let's look at **Jack and Jill ran up the hill**

Where did Jack and Jill go?

up the hill

*up the

(2) The pronoun replacement test

- Pronouns can substitute for natural groups.

- Let's say you want to test if *a puppy* in the sentence *The child found a puppy* is a constituent

- Where did the child find a puppy?

*the child found **him** in the park*

What has him replaced?

(3) The do-replacement test

- Suppose we want to test if *found a puppy* is a constituent?

Do can take the place of an entire predicate, which is a constituent

John found a puppy and Bill did too

- ***Did*** has replaced ***found a puppy*** here

*Jack and Jill **went up the hill**, and Molly **did** too*

*Rita **saw a sparrow** and Mike **did** too*

(4) The "move-as-a-unit" test

- If a group of words can be moved, they form a constituent

- Let's test the whether *a puppy* *and a child* *are* a constituent in *the child found a puppy*.

*It was **a puppy** that the child found*

*It was **a child** who found the puppy*

***A puppy** was found by the child*

But is *found a* a constituent?

Syntactic categories

Take the sentence *The child found a puppy*, once more.

- Now let's consider the following groups of words:
- the policeman
- ships
- angry
- Hamid
- The woman in the green shoes
- slowly

1) The child found *a puppy*

2) The child found **the policeman**

3) The child found **ships**

4) *The child found **angry**

5) The child found **Hamid**

6) The child found **the woman in the green shoes**

7) *The child found **slowly**

Notice that we can replace *a puppy* with the group of words in yellow but not with those in red.
A family of expressions that can substitute for one another without loss of grammaticality is called a syntactic category.

Let's try another

- **The boy watching the movie** saw an elephant

- **Megha** saw an elephant

- **My grandfather** saw an elephant

- **You** saw an elephant

- **The tallest boy in the classroom** saw an elephant

- **He** saw an elephant

What syntactic category do the words in yellow belong to?

They belong to the category called Noun Phrases (NPs).

NPs may function as subjects or as an object in a sentence. The noun is considered the "head" of the NP. Therefore, the minimum element that is required to form a NP is a noun (or a pronoun).

1) The woman with blue hair *saw a snake on the lawn*

2) The woman with blue hair **fell down**

3) The woman with blue hair **went to the movies last night**

4) The woman with blue hair **slept**

5) The woman with blue hair **gave a snake to Sarah**

But

1) *The woman with blue hair **happy**

2) *The woman with blue hair **fast**

3) *The woman with blue hair **happily**

The groups of words in yellow belong to the same syntactic category, called **Verb Phrases (VPs)**. The main element or "head" of a VP is always a verb. The minimum element required to form a verb phrase, as the name suggests, is a verb.

Other Phrases

Adjectival Phrases

1) Salma saw a **red and yellow** butterfly

2) Salma saw a **pretty** butterfly

3) Salma saw an **extremely big** butterfly

4) Samla saw a **tiny** butterfly

5) Salma saw a **happy-looking** butterfly

Adverb Phrases

1) The man walked home **fast**

2) The man walked home **very fast**

3) The man walked home **very very fast**

4) The man walked home **yesterday**

5) The man walked home **extremely quickly**

Prepositional Phrases

1) The boy **from Jaipur** came to visit the head of the department

2) The boy **in the red kurta** came to visit the head of the department

3) The boy **with the woman** came to visit the head of the department

Caution!

- Maria eats **an ice-cream**

- Maria eats **fast**

Here it **may look like** *an ice-cream* and *fast* are the same syntactic category since they appear to replace one another. But *an ice-cream* is a noun phrase and *fast* is an adverb phrase.

Also see

John walks **to the store**

John walks **very slowly**

to the store **and** *very slowly***, are, again, not the same syntactic category. What categories are they?**

Do not use replacement blindly!

Phrases can have phrases inside them

- [NP the woman in the green shoes]

[NP the woman]

[PP in the green shoes]

    [NP the green shoes]

[NP [NP the woman] [PP in [NP the green shoes]]]

Phrases can have phrases inside them

- saw a snake on the lawn

- The tallest boy in the classroom

- extremely quickly

Every word inside the phrase belongs to another kind of syntactic category, called lexical category

[NP [NP the woman] [PP in [NP the green shoes]]]

lexical categories of the words in the noun phrase the woman in the green shoes:

the – determiner (D)

woman – noun (N)

in – preposition (P)

the – determiner (D)

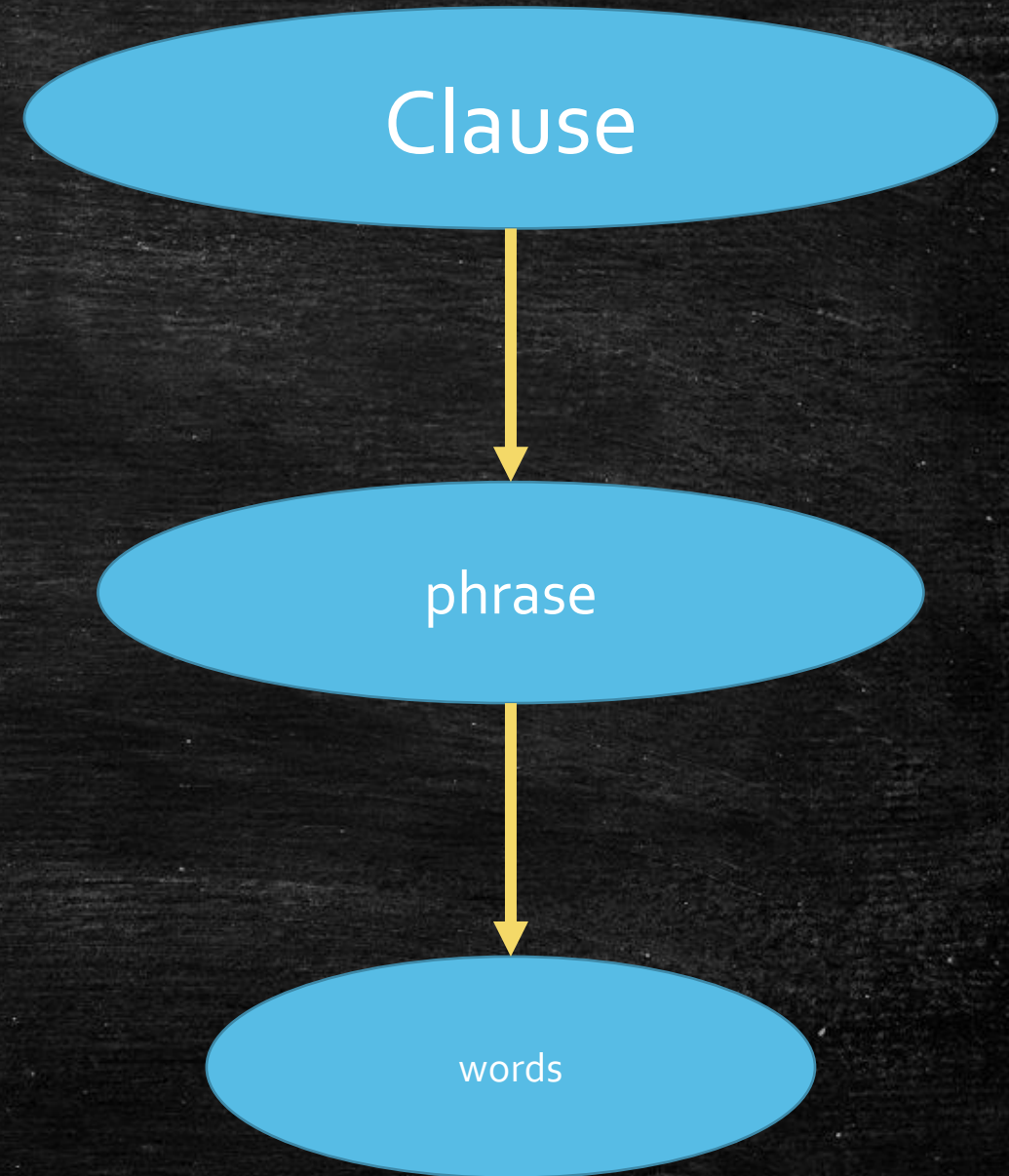green – adjective (Adj)

shoes – noun (N)

"Syntactic categories"

Phrasal categories

NP
VP
PP
AdjP
AdvP

Lexical categories

Noun (N)
Verb (V)
Preposition (P)
Adjective (Adj)
Adverb (Adv)
Determiner (D)

Clause

phrase

words

44

Identifying lexical categories or "parts of speech"

The problem of traditional definitions

In school or college, you would have learnt that **a noun is a "person, place or thing"** and **a verb is "an action, state, or state of being"**

Such definitions are based on semantic criteria i.e. meaning

1) **The destruction of the city bothered the Mongols**

Destruction is not a "person place or thing", it is an action. By the semantic criteria definition we have above, *destruction* should be a verb. But *destruction* is a noun

2) **Sincerity is an important quality**

Sincerity is an attribute, a property associated with adjectives

3) **The police are worried about the assassination of the president**

Assassination is an action, a property associated with verbs

4) **Madurai is a great place to live**

Madurai is a location, a property associated with prepositions/post-positions

Yet these are all nouns!

So, semantics or meaning is not a reliable criteria for determining the lexical category of words.

More problems

1) Gabrielle's **mother** is an axe-murderer. (N)

2) Anteaters **mother** attractive offspring. (V)

3) Wendy's **mother** country is Iceland. (Adj)

Words can change their lexical category depending upon where it appears in a sentence

Even more problems

- The situation in other languages:

Wita-ngku   ka   maliki wajilipinyi.

small-SUBJ   AUX   dog   chase.PRES

"The small (one) is chasing the dog."

Is *wita* an adjective or a noun?

- some words just don't have meaning

Akshay said that lexical categories are easy to study

what is the meaning of *that*?

The subconscious knowledge of lexical categories of your first language

The yinkish dripner blorked quastofically into the nindin with the pidibs.

blorked - verb

yinkish - adjective

dripner - noun

nindin - noun

quastofically - adverb

pidibs - noun

Substituting the word

Suppose you know that *king* is a noun, and want to find out the lexical category o puppy.

Make a sentence with the word *puppy*, and then substitute *puppy* with the word *king*. If the sentence now formed is grammatical, it suggests that *king* is also a noun.

1)   The **puppy** loved pooris

2)   The **king** loved pooris

If we substitute *puppy* with a word and do not get a syntactically well-formed sentence, it means that *puppy* and that word are not the same lexical category.

3)   *The *green* loved pooris (puppy is not an adjective)

4)   *The *in* loved pooris (puppy is not a preposition)

5)   *The *sing* loved pooris (puppy is not a verb)

6)   *The *a* loved pooris (puppy is not a determiner)

The distribution of a word

But what is the problem with substitution?

It's not 100% reliable. It can give you false positives

1) Seema eats quickly

2) Seema eats meat

Both these sentences are grammatical, but *quickly* and *meat* are not the same category.

You have to be sure about the category of the word you are using to do the substitution test.

We also have to look at the distribution of a word

The distribution of a word refers to the linguistic elements in and around the word i.e. **its "environment"**. We examine both the **morphological distribution** and the **syntactic distribution** of a word to determine its lexical category.

The morphological distribution of words

The morphological distribution of a word refers to the kinds of affixes (prefixes and suffixes) and other morphology that appear on the word. So we look at the **derivational and inflectional morphemes** in a word.

Derivational morphemes

We look for **affixes that make words out of other words**. We call these affixes **derivational morphemes**. These morphemes usually result in a particular lexical category.

Let's take the word *distribute*, which is a verb. We can add the **suffix –(t)ion** and we get the noun *distribution*

Words ending in *–(t)ion* are nouns

This is an example of morphological distribution.

Let's look at the suffix *–al*, which creates adjectives.

- Let's take the noun ***distribution*** and add the suffix *–al* to it.

- We get the adjective ***distributional***

Let's look at the suffix *–ment*, which creates nouns. Let's take the verb ***disappoint*** and add the suffix *–ment* to it

- We get the noun **disappointment**

If you are unsure of the category of *disappoint*, adding *–ment* and getting a well-formed word tells you that *disappoint* is a verb.

If you are unsure of the category of *disappointment*, you can see the suffix *–ment* inside the word and know that it is a noun.

Inflectional morphemes

- While **derivational affixes make a word into a particular category**; by contrast **inflectional morphemes** don't make a word into a particular category, but instead **only attach to certain categories**.

- Let's take **–est**

- It only attaches to words that are **already adjective**

1) **biggest**

2) **fastest**

3) **happiest**

4) ***doggest**

5) ***womanest**

The syntactic distribution of words

- It is about what kinds of words appear near the word.

Syntactic distribution of nouns:

- A word next to or near to a determiner will be a noun. Nouns also appear next to adjectives, and may appear next to prepositions. Nouns can also be the subjects and objects of sentences.

- **the girl** walking down the street stopped for a moment

- I saw **the** most beautiful **sunset** last evening

- I saw the most **beautiful sunset** last evening

- Sarah saw a snake **in school**

Syntactic distribution of verbs:

▪ Verbs may appear next to auxiliaries and modals.

auxiliaries are helping verbs like *have/has/had, am/are/were/was*

The girl **was** **eating** a melon

The girl **had** **eaten** a melon last week

modals are words like *can, could, shall, should, will, would*

I **can** **eat** melons

▪ They may appear next to the infinitive marker *to*

I like **to** **eat** melon

Syntactic distribution of adjectives:

- They appear beside nouns

Anil cooked a **tasty** **biryani**

- They can follow the *be* verb

The biryani **was** **tasty**

- They can appear beside adverbs like *very*

The biryani was **very** **tasty**

The syntactic distribution of adverbs

- They cannot appear before a noun

*happily woman saw a hen

*quickly fox ate a hen

- They can appear almost everywhere else

Sadly, the injured man passed away

The man ran fast

He died, sadly

Why don't you quickly cross the bridge?

- They can be modified by the adverb *very* and other *adverbs*:

very quickly, very well, almost tomorrow, extremely near

Compulsory readings (Available on GDrive)

- Pages 77 to 88 (in "Syntax: The sentence patterns of language") from *Fromkin*

- Pages 35 to 43 (from Chapter 2) of Carnie's *Syntax: A generative introduction*

World Atlas of Language Structures Online
https://wals.info/chapter/81

Watch Caroline Heycock on Constituency at https://www.youtube.com/watch?v=bvc9erKrLUQ and https://www.youtube.com/watch?v=mO4WQB1BIq0

References and optional readings/video resources