

Lecture 02: Sentences and Grammars

Andrei Antonenko

LIN 311: Syntax

August 30, 2018

Outline

- ① Sentences
Strings or Trees?
- ② Phrase-Structure Rules
- ③ Methodology of Syntax
Scientific Method
Competence vs. Performance

Sentences

Sentences: strings of words?

Question: Are sentences linear chains of words? Or is there more structure?

- (1)
- a. The \Rightarrow cat \Rightarrow eats \Rightarrow fish.
 - b. One \Rightarrow camel \Rightarrow hates \Rightarrow goat \Rightarrow cheese.

To reduce the possibilities we might assign transitional probabilities:

- $P(\textit{the} \Rightarrow V) = 0$: Verbs don't occur after *the*
- $P(\textit{eat} \Rightarrow \text{TYPE OF FOOD}) > P(\textit{eat} \Rightarrow \text{TYPE OF FURNITURE})$

Strings of words

Problems with string approach

- Some strings would have very low probabilities, but are grammatical nevertheless.

(2) Colorless green ideas sleep furiously.

- It is impossible to produce long distance dependencies, such as *either ... or* or *if ... then*:

- (3)
- a. **Either** the girl eats ice cream, **or** the girl likes candy.
 - b. **If** the girl eats ice cream, **then** the boy eats burgers.
 - c. ***Either** the girl eats ice cream, **then** the boy eats burgers.
 - d. ***If** the girl eats ice cream, **or** the girl likes candy.

Strings of words

Question rule

- **Yes/no questions** can be answered by “yes” or “no” or “maybe”
 - (4)
 - a. Alex can't eat chocolate covered almonds.
 - b. Can't Alex eat chocolate covered almonds?
- How do we get the question sentence?

Strings of words

Question rule

- (5)
- a. Alex can't eat chocolate covered almonds.
 - b. Can't Alex eat chocolate covered almonds?

Hypothesis #1

Move the 2nd word to the front of the sentence.

Problem

- (6)
- a. The TA can't eat chocolate covered almonds.
 - b. *TA the ___ can't eat chocolate covered almonds?

Strings of words

Question rule

- (7)
- a. Alex can't eat chocolate covered almonds.
 - b. Can't Alex eat chocolate covered almonds?

Hypothesis #2

Move the auxiliary to the front of the sentence.

Problem

Which auxiliary?

- (8)
- a. The TA has been eating chocolate covered almonds.
 - b. *Been the TA has ___ eating chocolate covered almonds?

Strings of words

Question rule

- (9)
- a. Alex can't eat chocolate covered almonds.
 - b. Can't Alex eat chocolate covered almonds?

Hypothesis #3

Move the first auxiliary to the front of the sentence.

Problem

- (10)
- a. The TA who is here can eat chocolate covered almonds.
 - b. *Is the TA who ___ here can eat chocolate covered almonds?
 - c. Can the TA who is here ___ eat chocolate covered almonds?

Strings of words

Question rule

- (11) a. Alex can't eat chocolate covered almonds.
b. Can't Alex eat chocolate covered almonds?

Hypothesis #4

Move the first auxiliary after the main clause subject to the front of the sentence.

It finally works!

- (12) a. [The TA who is here] has been eating chocolate covered almonds.
b. Has [the TA who is here] ___ been eating chocolate covered almonds?

Structure

Yes/no question rule

Move the first auxiliary after **the main clause subject** to the front of the sentence.

- Notice that the rule we came up with refers to *chunks* of the sentence, such as the *subject*.
- It means that sentences are not just linear strings.
- Let's try to build the theory of *chunks*.

Phrase-Structure Rules

Building sentences

Phrase-structure rules

- $X \rightarrow Y Z$ means that X consists of two parts: Y and Z.
 - Every times we see X, we should rewrite it as Y Z.
 - If we have several ways to rewrite X, choose any.
-
- For example, $S \rightarrow N V$ means that S consists of N and V.
 - **Sentence** consists of a **Noun** and a **Verb**.
 - Let's try to come up with rules for English.

Building sentences

Grammar #1

Structural rules

$S \rightarrow N V$

Lexical rules

$N \rightarrow \textit{fish}$, $N \rightarrow \textit{cat}$, $N \rightarrow \textit{dogs}$, ...

$V \rightarrow \textit{dance}$, $V \rightarrow \textit{sing}$, $V \rightarrow \textit{sleep}$, ...

Fish dance.

Cat sleep.

Dogs sing.

What about *Happy fish dance*? We need **Adjectives**

Building sentences

Grammar #2

Structural rules

$S \rightarrow NP V$

$NP \rightarrow Adj N$

Lexical rules

$N \rightarrow fish, N \rightarrow cat, N \rightarrow dogs, \dots$

$V \rightarrow dance, V \rightarrow sing, V \rightarrow sleep, \dots$

$Adj \rightarrow happy, Adj \rightarrow ugly, Adj \rightarrow grumpy, \dots$

Happy fish dance.

Grumpy cat sleep.

Ugly dogs sing.

But what about *Dogs sing* now? We need **optionality** of Adj!

Building sentences

Grammar #3

Structural rules

$S \rightarrow NP V$

$NP \rightarrow (Adj) N$

(Adj) means that Adj is optional

Lexical rules

$N \rightarrow \textit{fish}$, $N \rightarrow \textit{cat}$, $N \rightarrow \textit{dogs}$, ...

$V \rightarrow \textit{dance}$, $V \rightarrow \textit{sing}$, $V \rightarrow \textit{sleep}$, ...

$Adj \rightarrow \textit{happy}$, $Adj \rightarrow \textit{ugly}$, $Adj \rightarrow \textit{grumpy}$, ...

Fish dance. Happy fish dance.

Cat sleep. Grumpy cat sleep.

Ugly dogs sing. Dogs sing.

But what about *Dogs sing songs* or *Grumpy cat eat fish*? We need **objects!**

Building sentences

Grammar #4

Structural rules

$S \rightarrow NP V (NP)$ (NP) means that the second NP is optional
 $NP \rightarrow (Adj) N$ (Adj) means that Adj is optional

Lexical rules

$N \rightarrow \textit{fish}$, $N \rightarrow \textit{cat}$, $N \rightarrow \textit{dogs}$, ...
 $V \rightarrow \textit{dance}$, $V \rightarrow \textit{sing}$, $V \rightarrow \textit{sleep}$, ...
 $Adj \rightarrow \textit{happy}$, $Adj \rightarrow \textit{ugly}$, $Adj \rightarrow \textit{grumpy}$, ...

Fish dance. Happy fish dance.
 Ugly dogs sing. Dogs sing. Dogs sing songs.
 Grumpy cat eat fish. Grumpy cat eat happy fish.

Is there another way to write this grammar?

Building sentences

Grammar #4'

Structural rules

$S \rightarrow NP VP$

NP/VP stand for Noun/Verb Phrase

$NP \rightarrow (Adj) N$

(Adj) means that Adj is optional

$VP \rightarrow V (NP)$

(NP) means that the NP is optional

Lexical rules

$N \rightarrow \textit{fish}, N \rightarrow \textit{cat}, N \rightarrow \textit{dogs}, \dots$

$V \rightarrow \textit{dance}, V \rightarrow \textit{sing}, V \rightarrow \textit{sleep}, \dots$

$Adj \rightarrow \textit{happy}, Adj \rightarrow \textit{ugly}, Adj \rightarrow \textit{grumpy}, \dots$

Fish dance. Happy fish dance.

Ugly dogs sing. Dogs sing. Dogs sing songs.

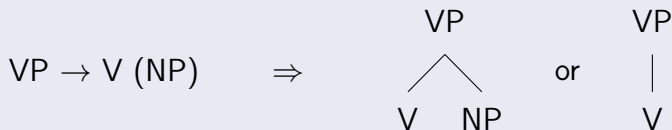
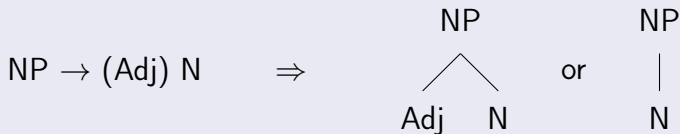
Grumpy cat eat fish. Grumpy cat eat happy fish.

Phrase Structure Rules and Trees

Each phrase structure rule can be diagrammed as a tree:



Phrase Structure Rules and Trees



Phrase Structure Rules and Trees

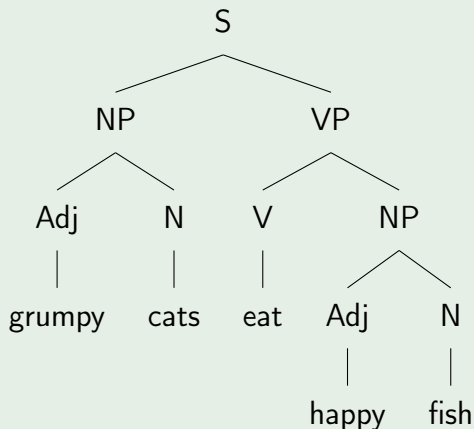
(13) Grumpy cats eat happy fish.

Derivation

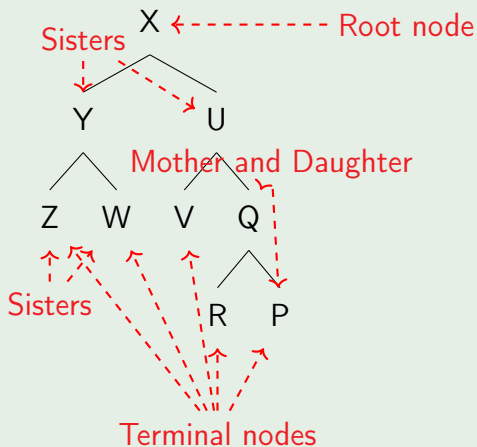
Start	S	Start with symbol S
Step 1	NP VP	$S \rightarrow NP VP$
Step 2	Adj N VP	$NP \rightarrow Adj N$
Step 3	Adj cats VP	$N \rightarrow \textit{cats}$
Step 4	grumpy cats VP	$Adj \rightarrow \textit{grumpy}$
Step 5	grumpy cats V NP	$VP \rightarrow V NP$
Step 6	grumpy cats V Adj N	$NP \rightarrow Adj N$
Step 7	grumpy cats V happy N	$Adj \rightarrow \textit{happy}$
Step 8	grumpy cats eat happy N	$V \rightarrow \textit{eat}$
Step 9	grumpy cats eat happy fish	$N \rightarrow \textit{fish}$

Phrase Structure Rules and Trees

(14) Grumpy cats eat happy fish.



Trees: terminology



Methodology of Syntax

Scientific method

How do linguists study grammar?

- ① Observe the linguistic data.
- ② Formulate grammar as a hypothesis, make sure that the observed data is accounted for.
- ③ Test and modify our grammar in view of the predictions it makes and in view of new data we come across.

Scientific method

Linguistic data

Speakers can judge which strings on words are sentences or not:
well-formedness judgments.

- Well-formed \neq Natural

- (15) a. Colorless green ideas sleep furiously
b. Revolutionary new ideas happen infrequently.

- Well-formed \neq Educated

- (16) a. Who are you going with?
b. Me and my friend just got back from the movies.

Corpora

- **Corpora**: recorded real world speech, newspapers, books, magazines, etc.
- Unfortunately, it is **not enough to look at corpora**:
 - They don't contain **negative information** (such as what sentences are unacceptable);
 - They can **never contain all the sentences of a language**.
- Hypotheses about the grammar often can only be proven wrong by ungrammatical sentences.
 - Hypothesis may predict something, but it may turn out to be ungrammatical.

Competence vs. Performance

- **Performance:** refers to what we actually produce.
- **Competence:** refers to what we know about language.
- The focus of generative grammar is **Competence**.

Infinite number of grammatical sentences

Speakers might not produce all of these sentences ever (**performance**), but we can judge them to be well-formed (**competence**).

- (17)
- The horse behind Pegasus is gray.
 - The horse behind the horse behind Pegasus is gray.
 - The horse behind the horse behind the horse behind Pegasus is gray.
 - ...

Revising Grammars

Grammar #4'

Structural rules

$S \rightarrow NP VP$

$NP \rightarrow (Adj) N$

$VP \rightarrow V (NP)$

Lexical rules

$N \rightarrow \textit{fish}$, $N \rightarrow \textit{cat}$, $N \rightarrow \textit{dogs}$, ...

$V \rightarrow \textit{dance}$, $V \rightarrow \textit{sing}$, $V \rightarrow \textit{sleep}$, $V \rightarrow \textit{hit}$...

$Adj \rightarrow \textit{happy}$, $Adj \rightarrow \textit{ugly}$, $Adj \rightarrow \textit{grumpy}$, ...

Notice that our current grammar generates ungrammatical:

- (18) a. *Dogs sleep cat.
 b. *Grumpy fish hit.

How should this grammar be revised?

Revising Grammars

Lexical rules

We need to incorporate different types of verbs into our grammar – intransitive vs. transitive.

- Note, some verbs can be in both classes!

$V_i \rightarrow \text{sleep}, V_i \rightarrow \text{dance}, V_i \rightarrow \text{sing}, \dots$

$V_t \rightarrow \text{hit}, V_t \rightarrow \text{dance}, V_t \rightarrow \text{sing}$

Structural rules

Our old rule

$VP \rightarrow V (NP)$

should be replaced with a pair of new structural rules:

$VP \rightarrow V_i$

intransitive verbs, no object

$VP \rightarrow V_t NP$

transitive verbs, must have object