Cogs 532 Spring 2022
Theoretical Linguistics Assignment 01 Apr 15, before class

The most common way to formalize constituency is to define a context free grammar, which is a set of rewrite rules for symbols, where at each step you are allowed to rewrite a single symbol. We will call such grammars phrase structure grammars (PSGs).

Here is a PSG for a microscopic fragment of English:

i. $S \rightarrow NP\ VP$

ii. $NP \rightarrow D\ N$

iii. $VP \rightarrow V\ NP$

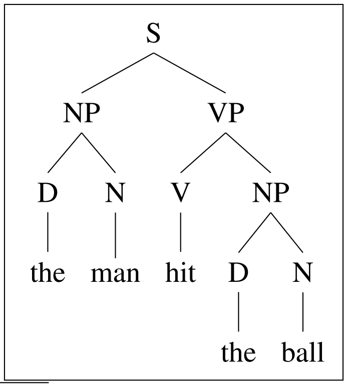
iv. $D \rightarrow the$

v. $N \rightarrow man, ball,$ etc.

vi. $V \rightarrow hit, took,$ etc.

A **leftmost**[1] derivation of the string *the* $+$*man* $+$ *hit* $+$ *the* $+$ *ball*:

Derivation	*Rules*
S	$S \rightarrow NP\ VP$
NP VP	$NP \rightarrow D\ N$
D N VP	$VP \rightarrow V\ NP$
D N V NP	$D \rightarrow the$
the N V NP	$N \rightarrow man$
the man V NP	$V \rightarrow hit$
the man hit NP	$NP \rightarrow D\ N$
the man hit D N	$D \rightarrow the$
the man hit the N	$N \rightarrow ball$
the man hit the ball	



[1] In every step of the derivation, you expand the leftmost nonterminal in the current string. The notion of **rightmost** derivation is defined similarly.

Question 1

Define a PSG that would capture the two constituency trees given for *John saw the man with the telescope* in lecture videos.

Question 2

Find a way to integrate feature specifications into your PSGs such that your grammar generates:

(1) a. John admires her.
b. They admire her.
c. John hopes she will win.
d. John wants her to win.

But cannot generate:

(2) a. *John admire her.
b. *They admires her.
c. *Him admires John.
d. *John hopes her will win.
e. *John wants she to win.

and so on.