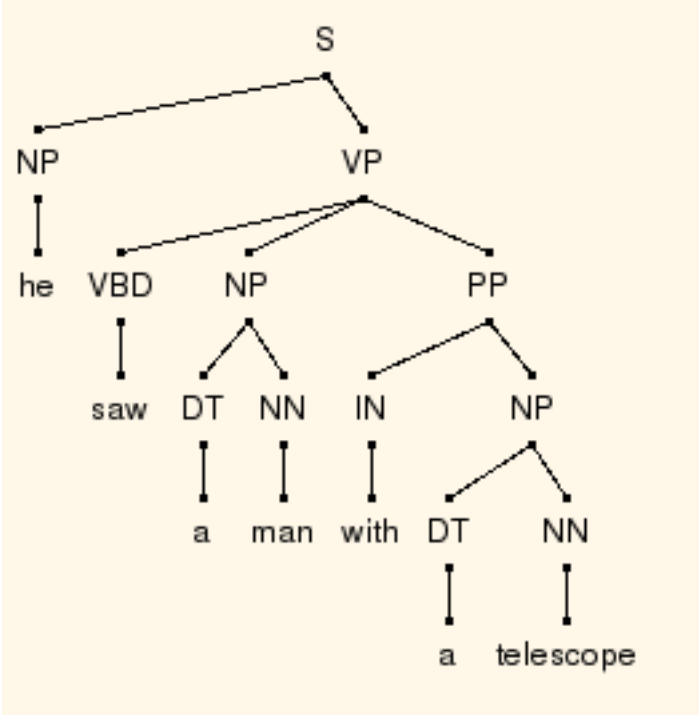
# CFG and Nuts-and-Bolts Parsing

Computational Linguistics

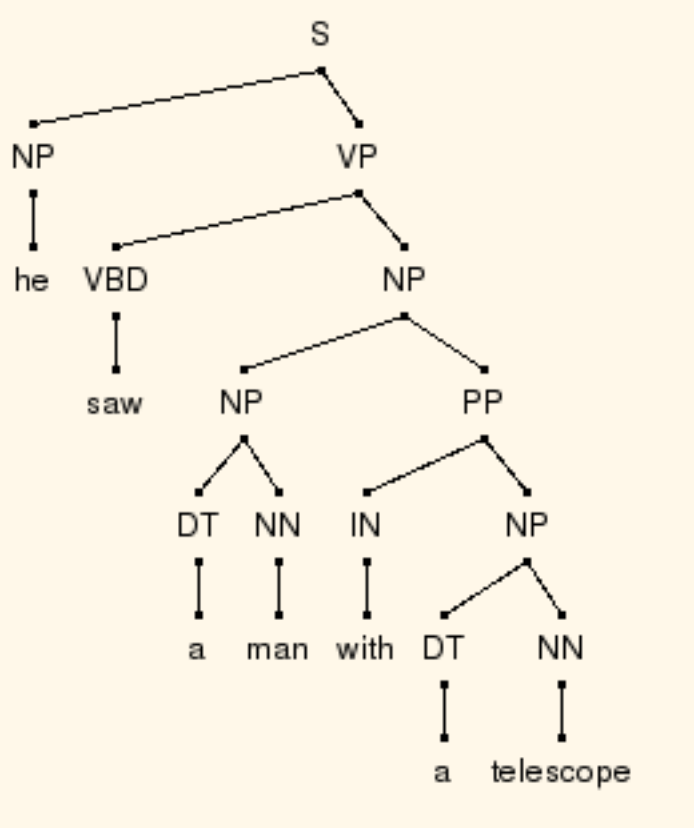
Spring 2023




He saw a man with a telescope.

VP reading:

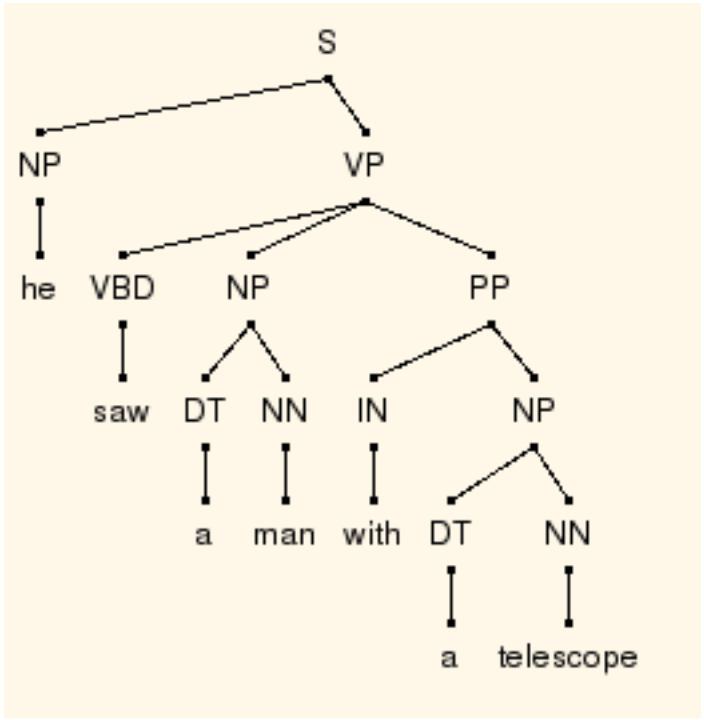
Using a telescope, he saw a man.

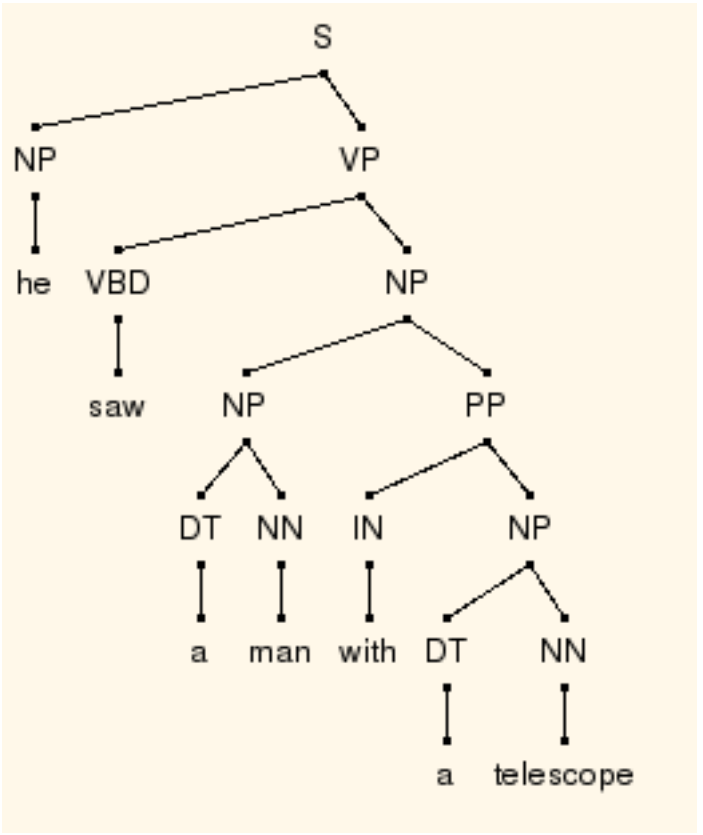



He saw a man with a telescope.

NP reading:

He saw a man who had a telescope.




[a man with a telescope] was seen.

[what man with a telescope] did he see?

[a man with a telescope] he saw.

He saw [ a man with a telescope]

[a man with a telescope] was seen.

[what man with a telescope] did he see?

[a man with a telescope] he saw.

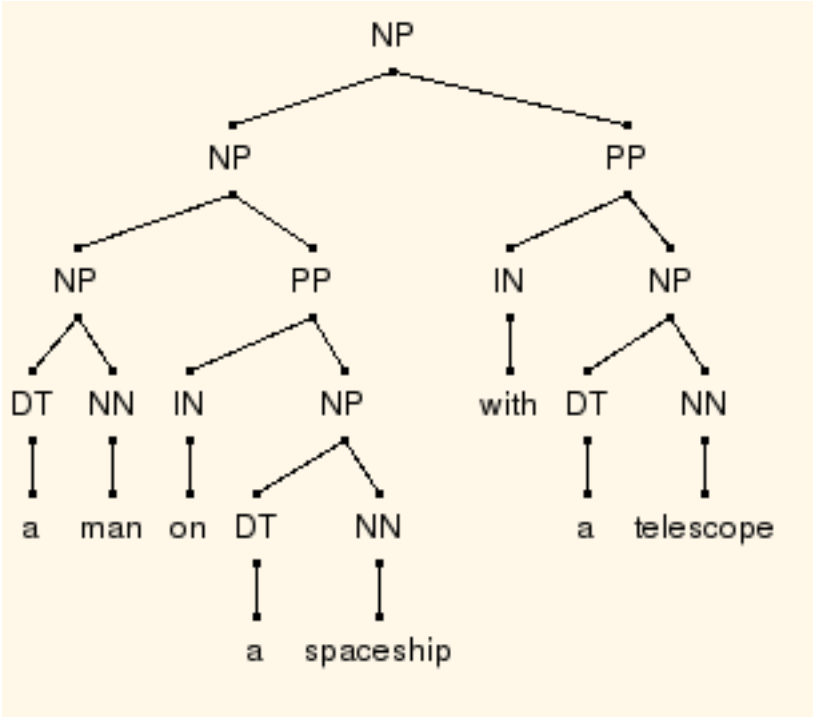
Movement diagnostics: when phrases "move", they do so as units, and transform the semantics in as systematic way.

This is reason to think that in the "NP reading", the PP forms a unit with *man*.

a man was seen with a telescope
what man did he see with a telescope?
a man he saw with a telescope
(… but not a woman)

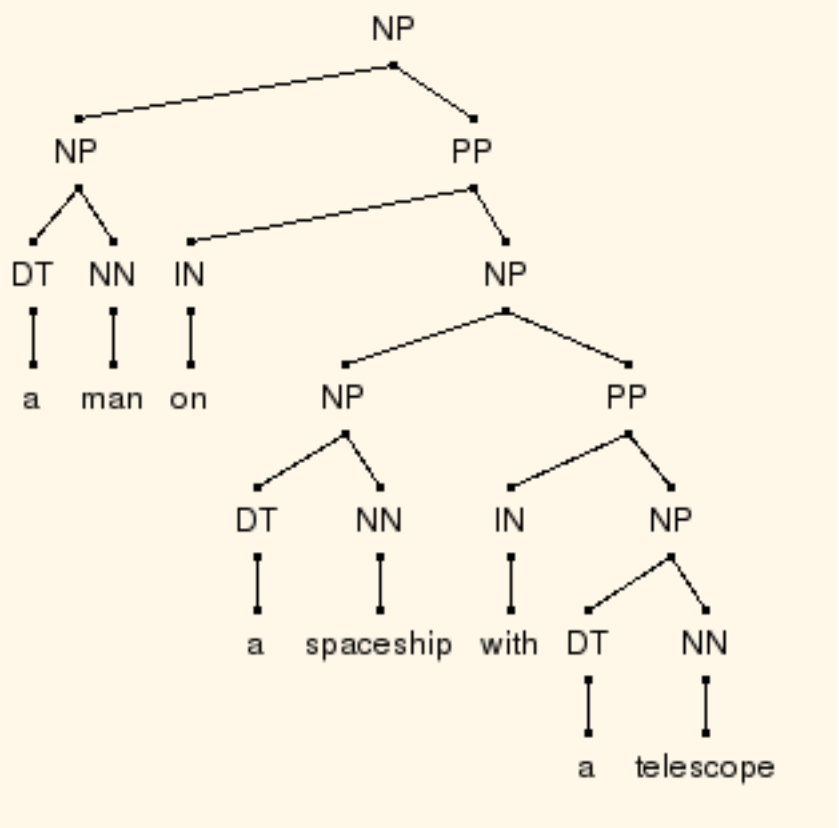
All are intuitively ambiguous.
This could be because PPs can rightward.

the man was seen with a telescope
the man with a telescope was seen




a man on a spaceship with a telescope

a man on a spaceship who has a telescope




a man on a spaceship with a telescope

a man on a spaceship that has a telescope

# Simplified tree language with nouns and PP

## Plural nouns

covers NPL
doors NPL
books NPL
tables NPL
titles NPL
carpets NPL
floors NPL
windows NPL

# Prepositions

by P

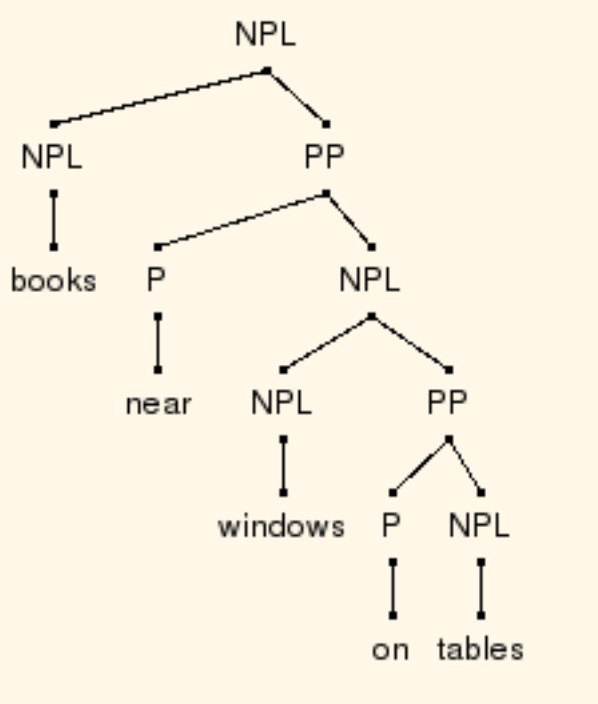
near P

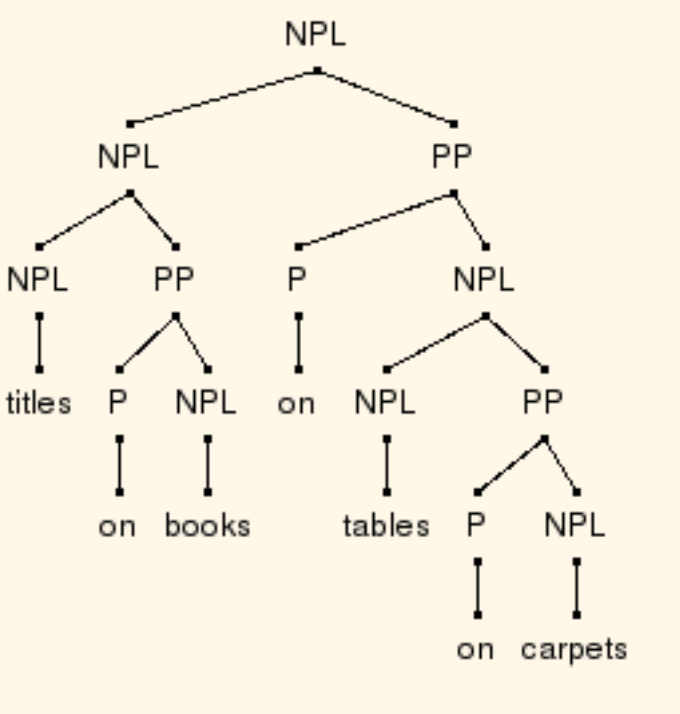
on P

with P

under P

# Tree shapes


NPL
NPL
PP
books
P
NPL
near
NPL
PP
windows
P
NPL
on
tables


NPL
NPL
PP
NPL
PP
P
NPL
titles
P
NPL
on
NPL
PP
on
books
tables
P
NPL
on
carpets

# Productivity

Speakers can use many phrases in this family, of seemingly unbounded size.

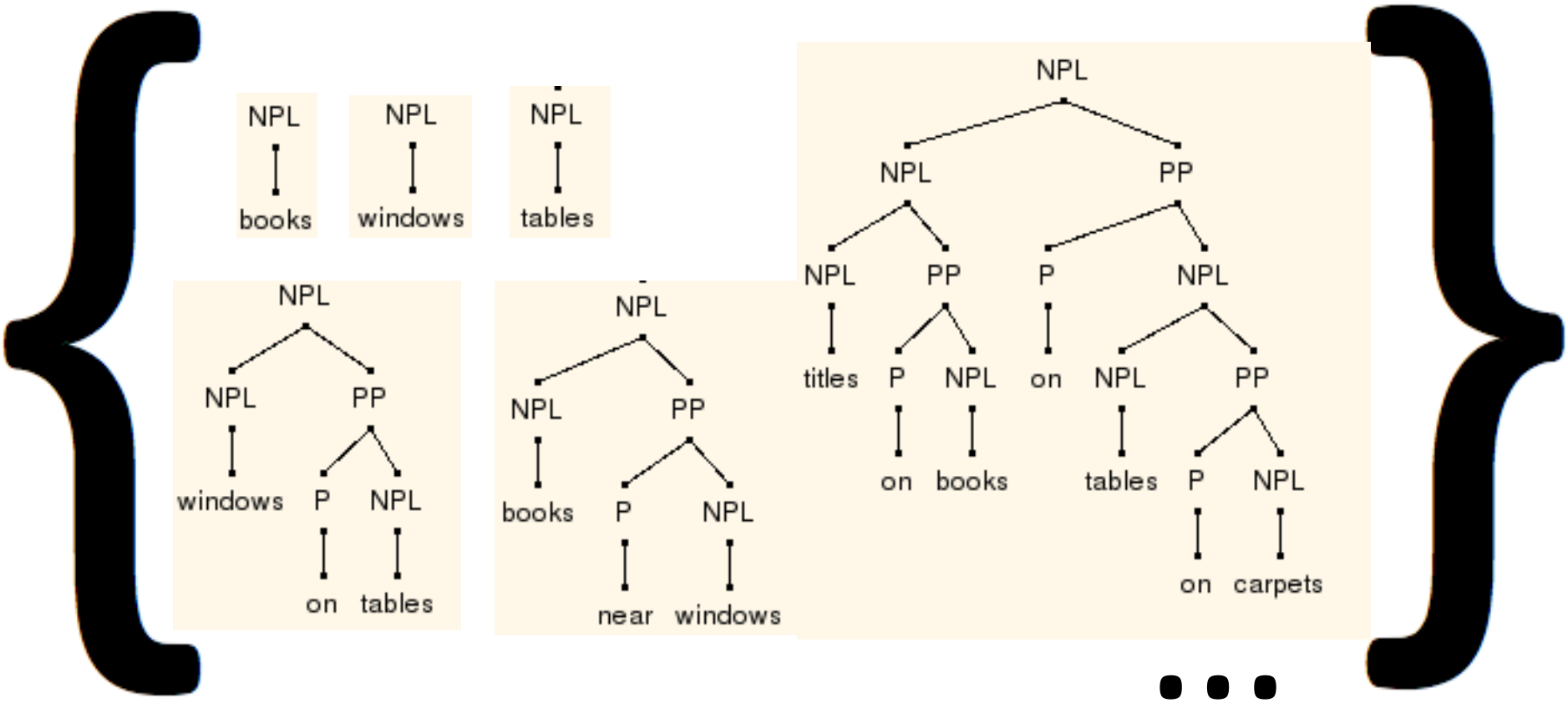
titles of books on tables near windows in gardens by rivers in kingdoms

# Ambiguity

Individual phrases have multiple readings, which (by hypothesis) correlate with tree shape.

# Large tree language

P =



# Grammar for tree language

Context free grammar

Feature constraint grammar -- soon

Minimalist movement grammar

- later in term

-

# Context free grammar for P

Set of “rules” or “productions” or “local trees” constructed from these vocabularies:

Terminals
Σ={covers, doors, books, tables, titles, carpets, floors, windows, by, near, on, with, under}

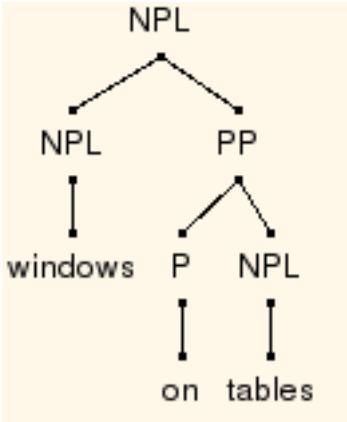
Non-terminals
N = {NPL, P, PP}

# Context free grammar for P

NPL➔NPL PP
PP ➔P NPL
NPL➔covers
NPL➔doors
NPL➔books
NPL➔tables
NPL➔titles
NPL➔carpets
NPL➔floors
NPL➔windows

P➔by
P➔near
P➔on
P➔with
P➔under

# Tree licensing

t= 


is licensed as an NPL because the *local tree* at each non-terminal vertex is a production of the grammar, and the root label t(ε)=NPL.

ε NPL➔NPL PP
0 NPL➔windows
1 PP ➔P NPL
10 P➔on
11 NPL➔tables

NPL→NPL PP

PP →P NPL

NPL→covers

NPL→doors

NPL→books

NPL→tables

NPL→titles

NPL→carpets

NPL→floors

NPL→windows

P→by

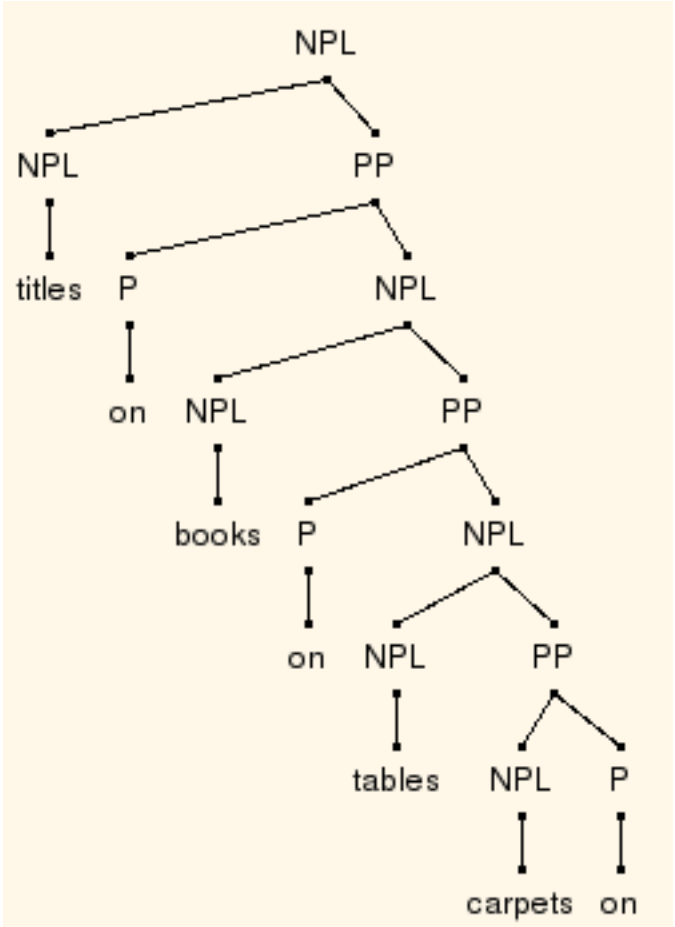
P→near

P→on

P→with

P→under

u=




is *not* licensed because 11111 is a non-terminal vertex in u, and the local tree in u at 11111, namely PP➔NPL P, is not a production of the grammar.

# Parsing

Parsing: given a terminal string and a grammar, find all the trees that have the given terminal string and are licensed by the grammar, with the start symbol of the grammar at the root.

The grammar includes a start symbol.

*Continue in Notebook*