

M1 master ISIL

Traitement automatique des langues

Plan

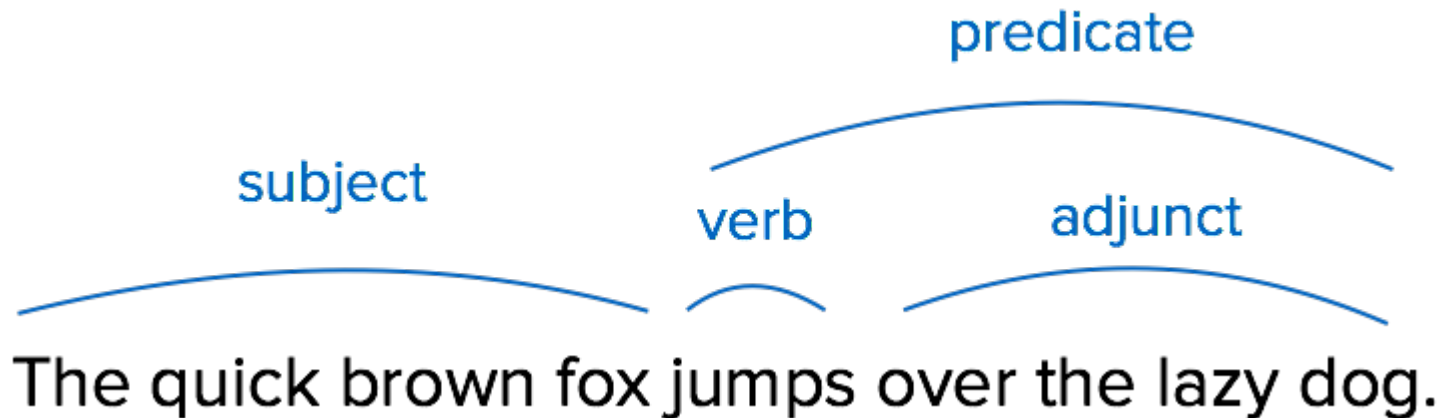
- Introduction
- Applications
- Niveau lexical
- Niveau syntaxique
- Niveau sémantique
- Niveau pragmatique

Prochain TP

- Installer NLTK sous python avant de venir
- Un deuxième test est programmé après 2 Cours

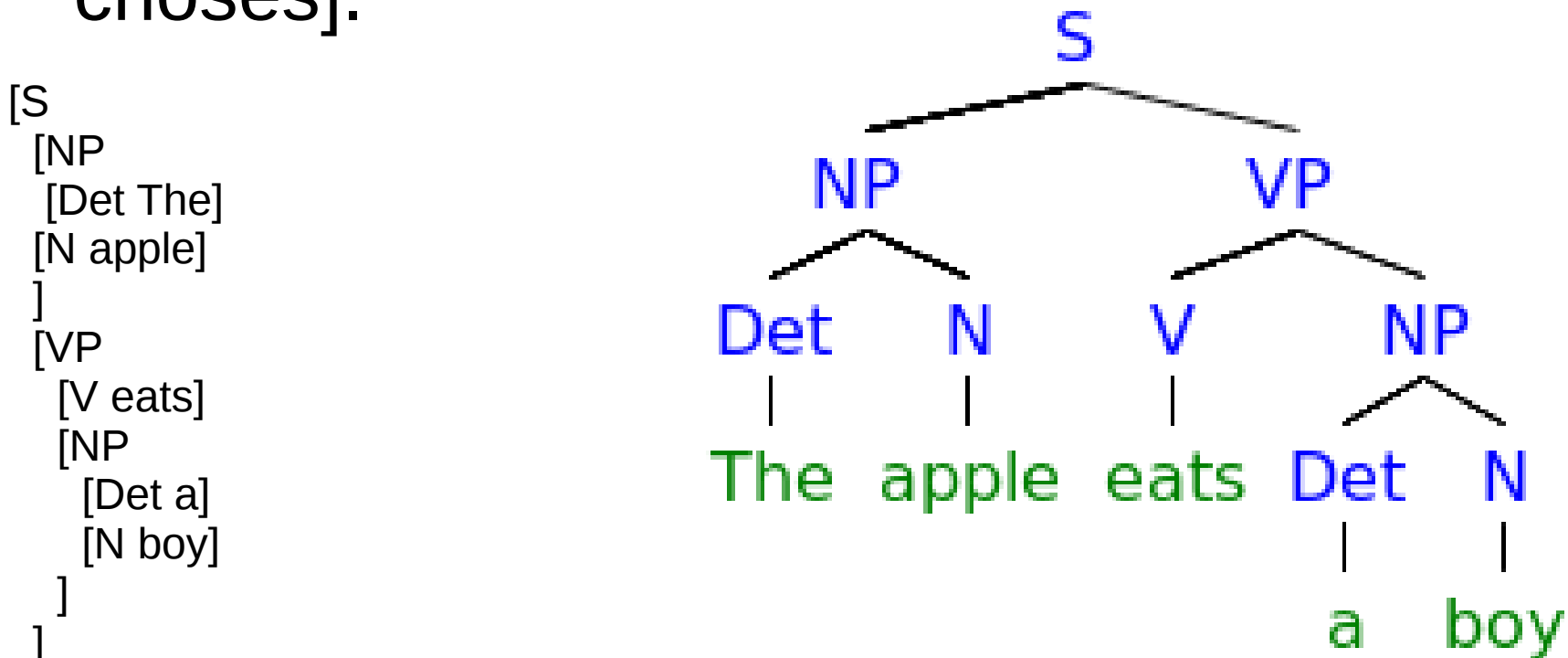
Syntaxe

- La syntaxe est, à l'origine, la branche de la linguistique qui étudie la façon dont :
- les mots **se combinent** pour former des phrases ou des énoncés dans une langue.



Syntaxe

- On distingue la syntaxe, qui concerne les expressions [les mots],
- de la **sémantique**, qui concerne ce qui est visé par les expressions [le sens, la signification/les choses].



Analyse Syntaxique

- L'analyse syntaxique consiste à mettre en évidence la structure d'un texte,
- généralement une phrase écrite dans une langue naturelle, mais on utilise également cette terminologie pour l'analyse d'un programme informatique.

Analyse Syntaxique

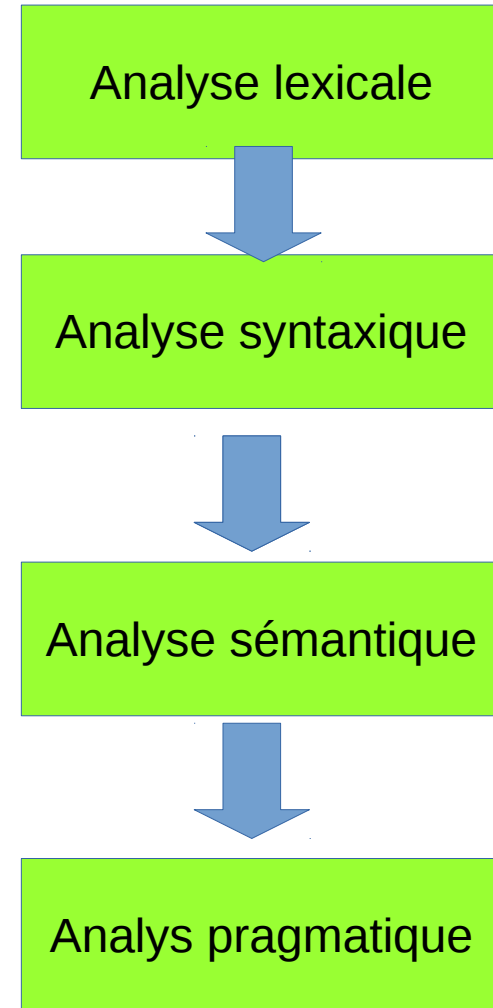
- L'analyseur syntaxique (parser, en anglais) est le programme informatique qui réalise cette tâche.
- Cette opération suppose une formalisation du texte, qui est vu le plus souvent comme un élément d'un langage formel,
- défini par un ensemble de **règles de syntaxe** formant une **grammaire formelle**.

Analyse Syntaxique

- La structure révélée par l'analyse donne alors précisément la façon dont les règles de syntaxe **sont combinées** dans le texte.
- Cette structure est souvent une **hiérarchie de syntagmes**, représentable par un **arbre syntaxique** dont les nœuds peuvent être décorés (dotés d'informations complémentaires).

Analyse syntaxique

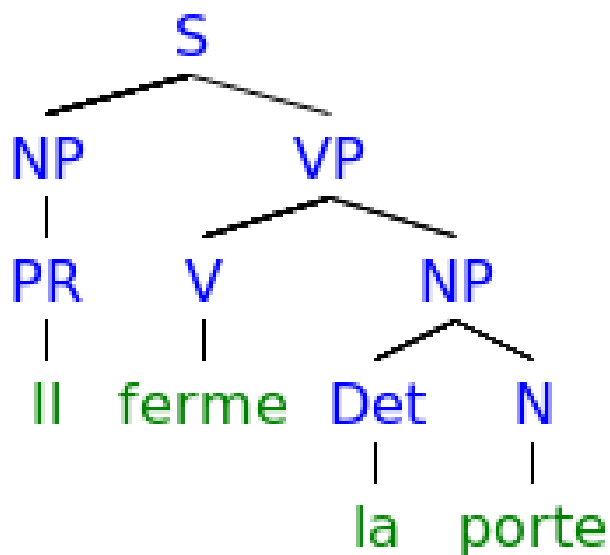
- L'analyse syntaxique fait habituellement suite à une analyse lexicale qui découpe le texte en un flux de lexèmes,
- et sert à son tour de préalable à une analyse sémantique.



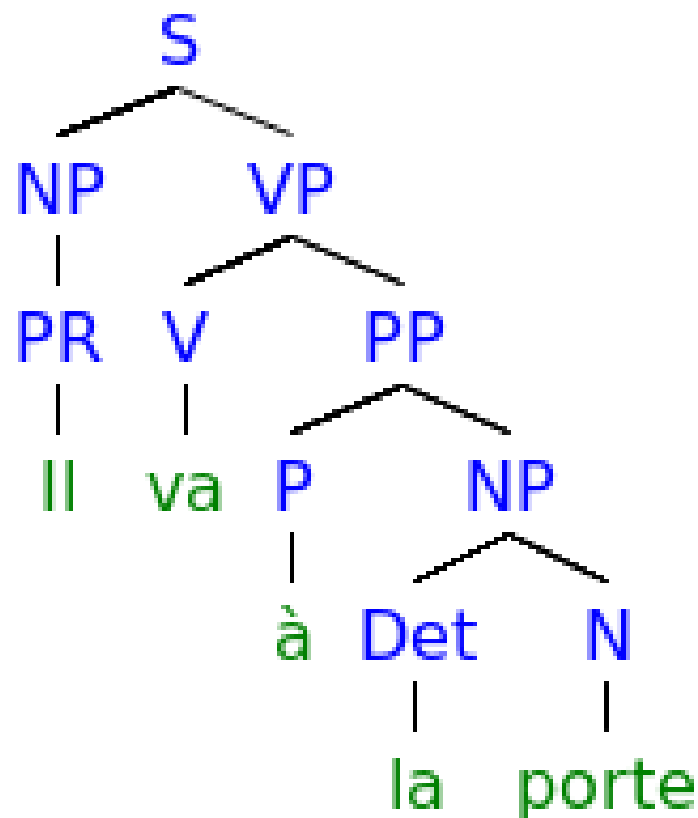
Étiquetage morpho-syntaxique

- Assigne chaque mot d'un texte à sa catégorie grammaticale. Par exemple, le mot ferme peut être un verbe dans
 - « il ferme la porte »
 - et un nom dans « il va à la ferme ».

Étiquetage morpho-syntaxique



[S [NP [PR Il]] [VP [V ferme]
[NP [Det la] [N porte]]]



[S [NP [PR Il]] [VP [V va]
[PP [P à] [NP [Det la] [N porte]]]

Étiquetage morpho-syntaxique

- Étiquetage morpho-syntaxique de chacun des mots d'un texte, comme dans un arbre syntaxique.
- Certaines phrases ambiguës peuvent être interprétées de plusieurs manières différentes, comme :
- je regarde l'homme avec les jumelles,

Étiquetage morpho-syntactique

- je regarde **l'homme** avec les jumelles



je regarde l'homme en utilisant des jumelles

- je regarde **l'homme avec les jumelles**

– « je regarde l'homme qui a des jumelles »

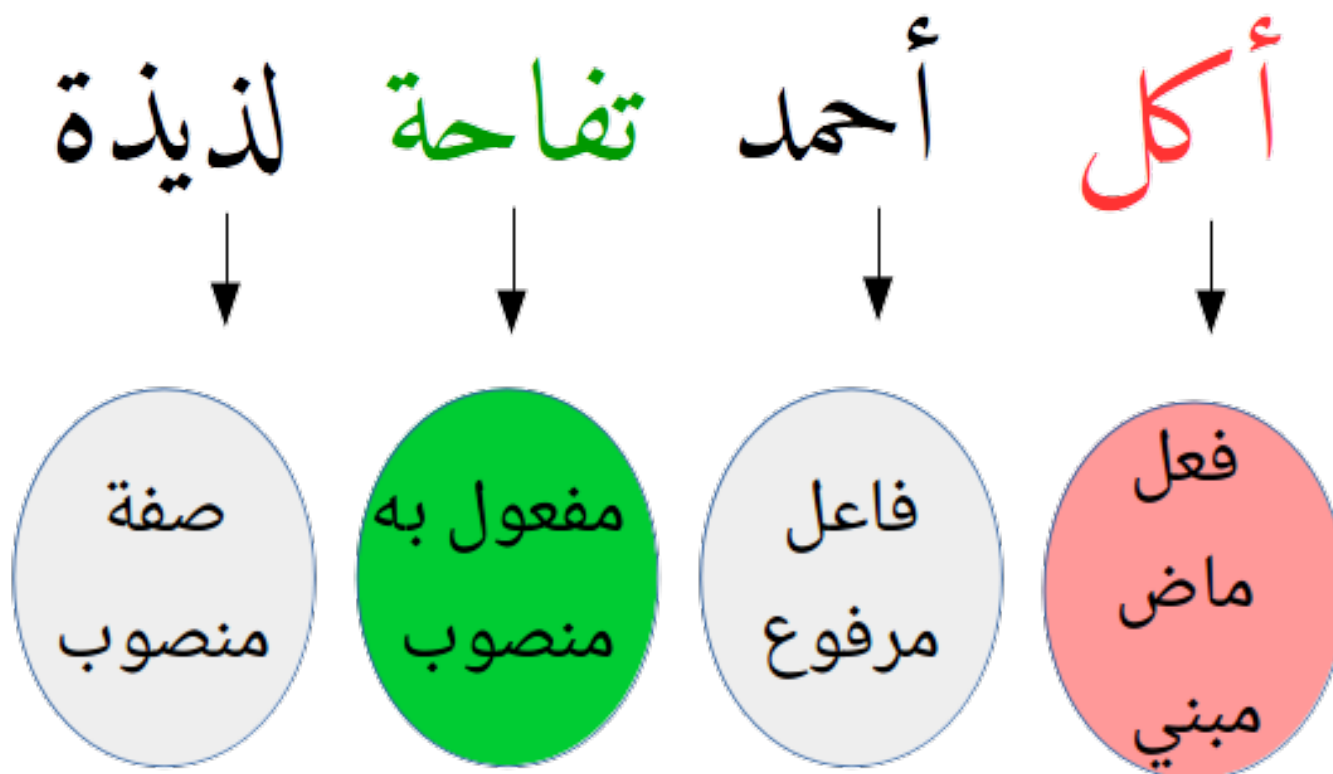


- je regarde **l'homme** avec les **jumelles**

– « je regarde l'homme qui est accompagné de sœurs jumelles ».



Étiquetage morpho-syntaxique



Ordre de mots entre langues

The word order of a 被(bèi) sentence

Learn more at litaochinese.com

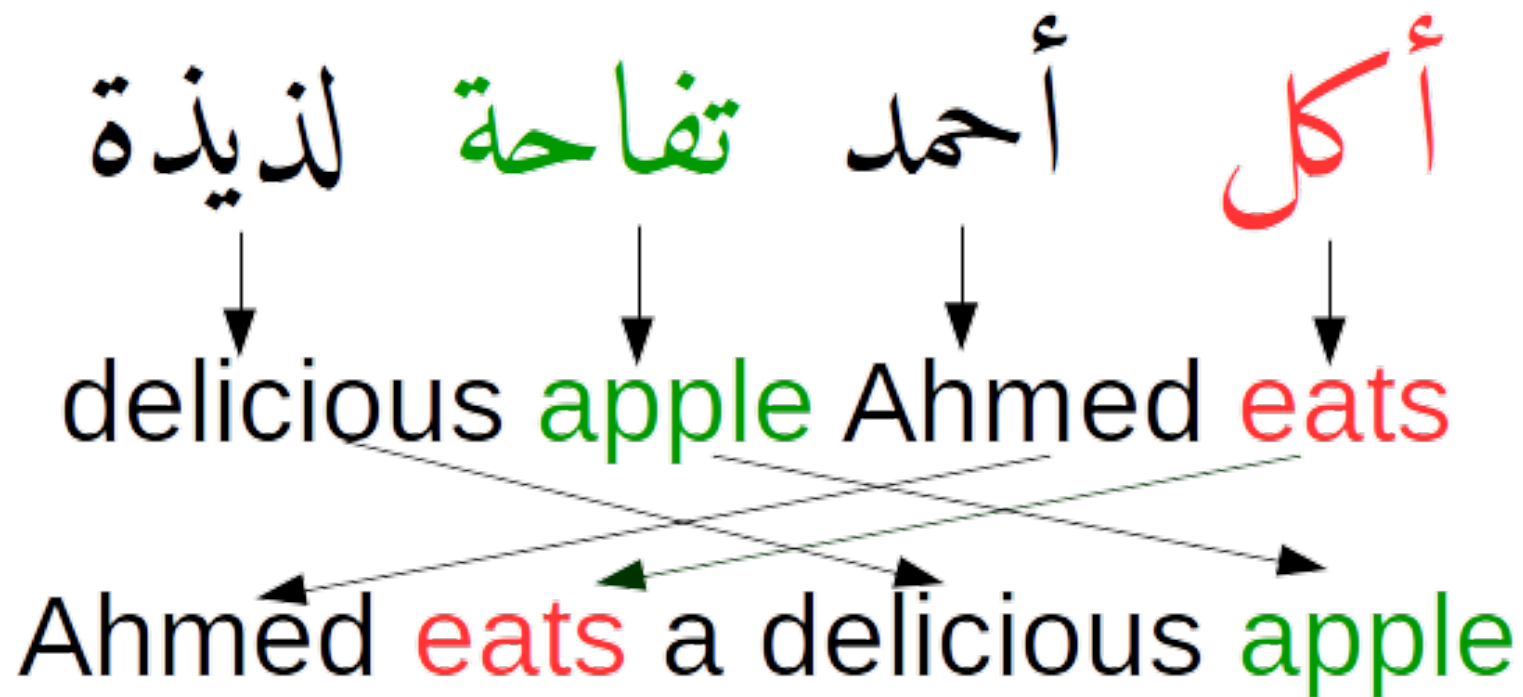


The word order of a 把(bǎ) sentence with 在(zài)

Learn more at litaochinese.com



Application de traduction



Accord

Le soleil

La lune

Ils mangent

إلى الجزائريين

الطالبات المجتهدات

الطلابان المجتهدان

Erreur grammaticales

حافظوا على المدينة نظيفة وجميلة

Enlever l'ambiguïté

John saw **old** [**men and women**]

John saw [**old men**] and women

إنما يخشى الله من عباده العلماءُ

إنَّ الله بريء من المشركين ورسولهُ

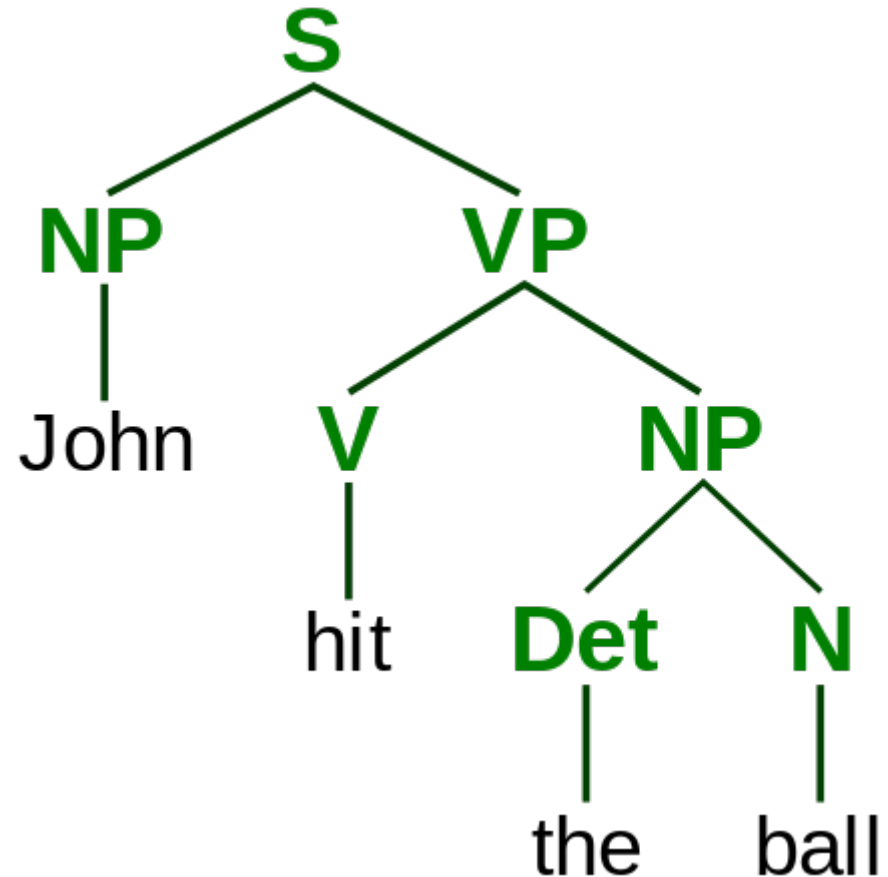
Enlever l'ambiguïté

أكل محمد تفاحة جميلة

أكل محمد تفاحة جميلة

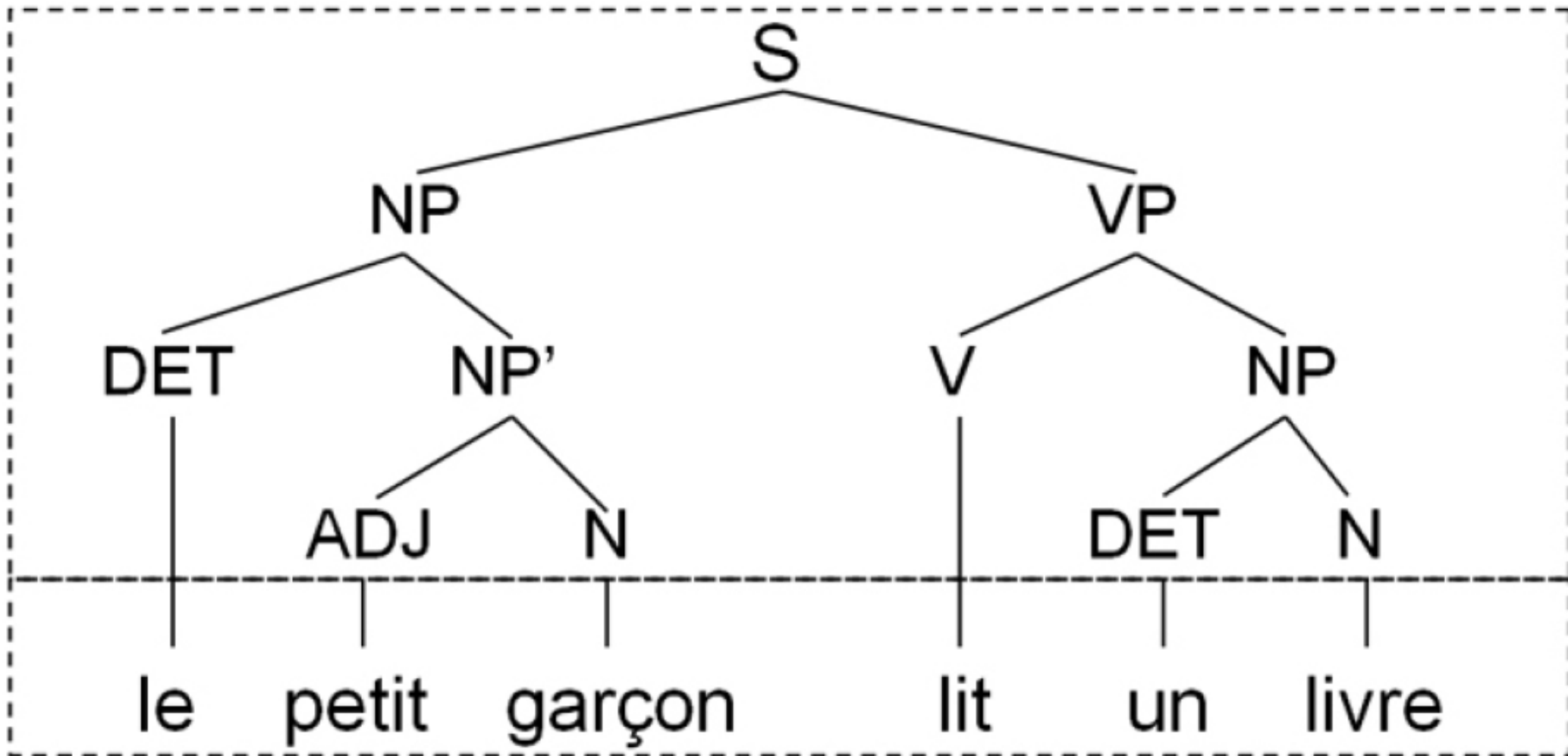
Grammaire

- English
- $S \Rightarrow NP VP$
- $NP \Rightarrow Det N$
- $VP \Rightarrow V NP$
- $N \Rightarrow John \mid Ball$
- $V \Rightarrow hit$
- $Det \Rightarrow the$



Exercice

- Donner la grammaire qui généré l'arbre suivant

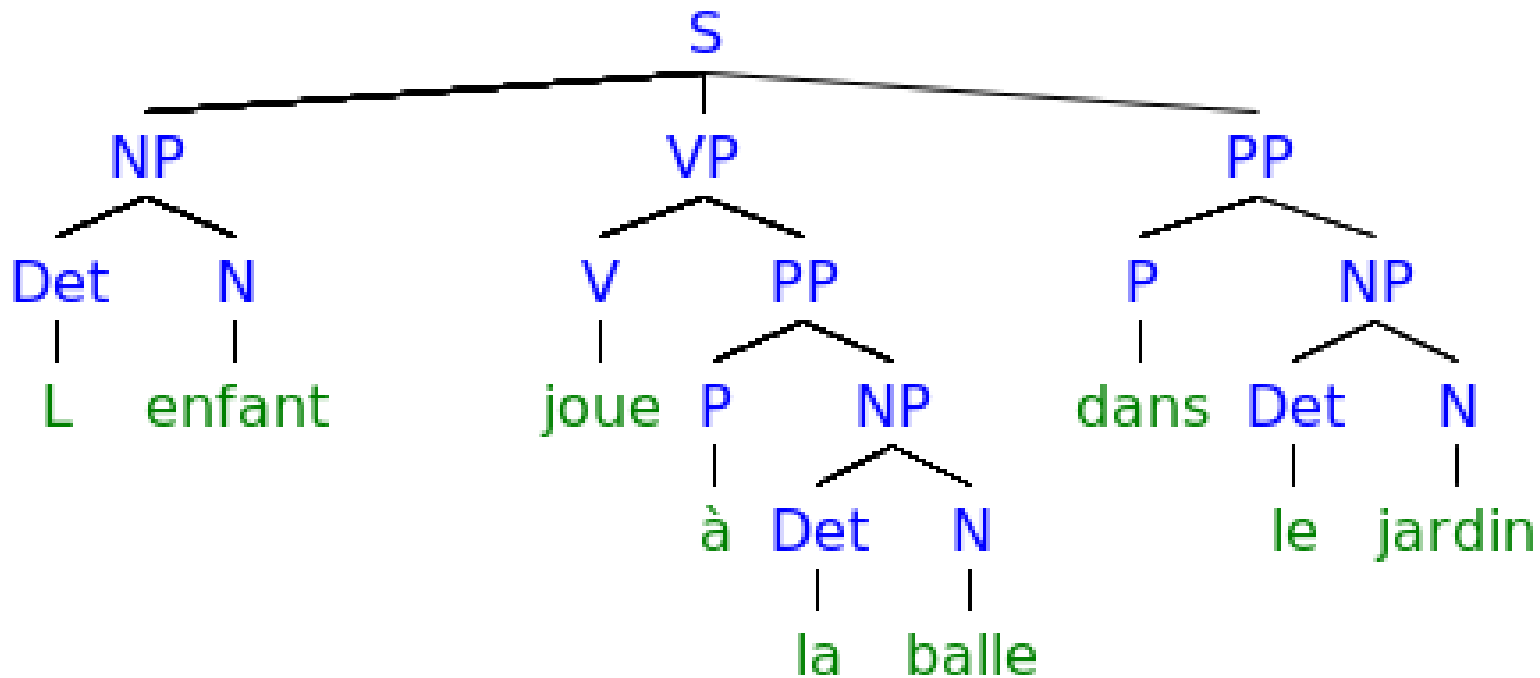


Exercice

- Tracer l'arbre syntaxique pour la phrase suivante :
- L'enfant joue à la balle dans le jardin

Exercice

- [S [NP [Det L] [N enfant]] [VP [V joue] [PP [P à] [NP [Det la] [N balle]]]] [PP [P dans] [NP [Det le] [N jardin]]]]



Exercice

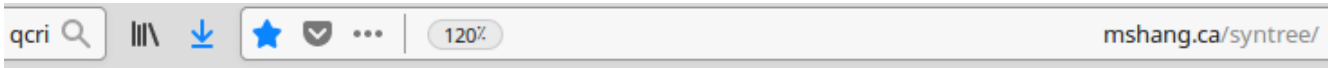
- Tracer l'arbre syntaxique pour les phrases suivantes :

• يسرق الثعلب الماكر تفاحا لذيذا من البستان

• الشمس طالعة

• في البيت فأر

Sites pour générer l'arbre syntaxique



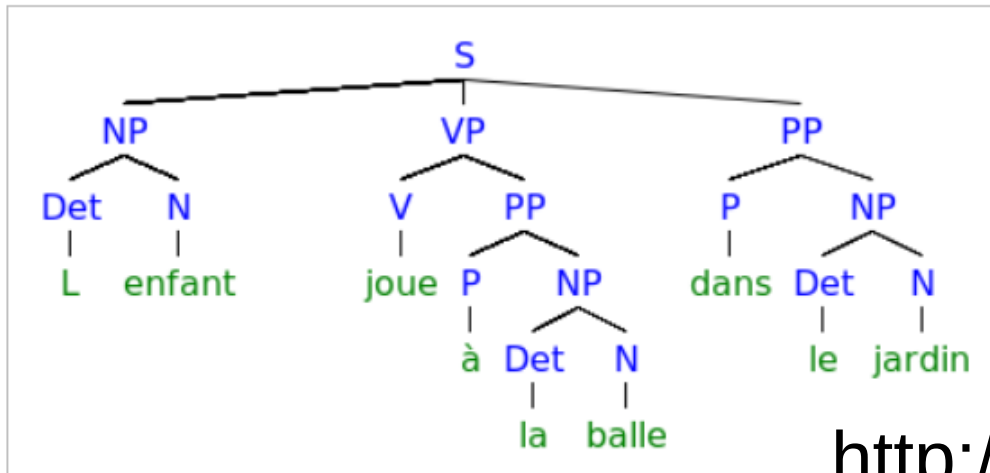
Syntax Tree Generator

```
[S [NP [Det L] [N enfant]] [VP [V joue] [PP [P à] [NP [Det la] [N balle]]] [PP [P dans] [NP [Det le] [N jardin]]]]
```

(C) 2011 by [Miles Shang](#), see [license](#).

Options

Help



<http://mshang.ca/syntree/>

Sites pour générer l'arbre syntaxique

- <http://mshang.ca/syntree/>
-
- <https://yohasebe.com/rsyntaxtree/>

Théorie X-barre

- La théorie X-barre (notée \bar{X})
 - a été proposée par **Noam Chomsky** en 1970 dans le cadre de la linguistique générative
 - et développée ensuite par Ray Jackendoff (1977).
- Elle permet d'**imposer des restrictions** sur la classe des catégories grammaticales possibles.

Théorie X-barre

- La théorie X-barre permet de **simplifier** les grammaires en désignant par **très peu** de symboles les classes de catégories naturelles.

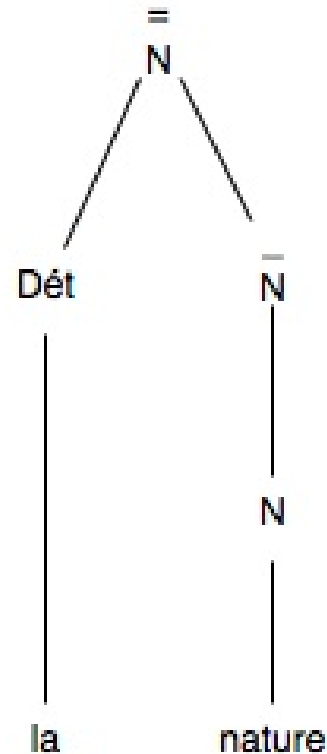
-

Formulation

- Soit la métarègle :
 $XP \rightarrow \text{Spec}_{\bar{X}} + \bar{X}$
- $\bar{X} \rightarrow X - \text{Comp}_x$
- $\bar{X} \rightarrow X - \text{Adjunct}_x$
 - avec $X : \{V, N, A\}$
 - Spéc = spécificateur;
 - Comp = complément
 - adjunct = circonstance

Formulation Arborescente

- Cette métarègle peut être formulée selon une arborescence (plus facile à interpréter). Les catégories spécificateur et complément désignent des types de constituants.
- Ainsi, on a :
- Spéc \bar{N} = Dét.
- Spéc \bar{V} = Aux.
- Spéc \bar{A} = DEG.
- Comp \bar{X} = $\bar{\bar{N}}$.



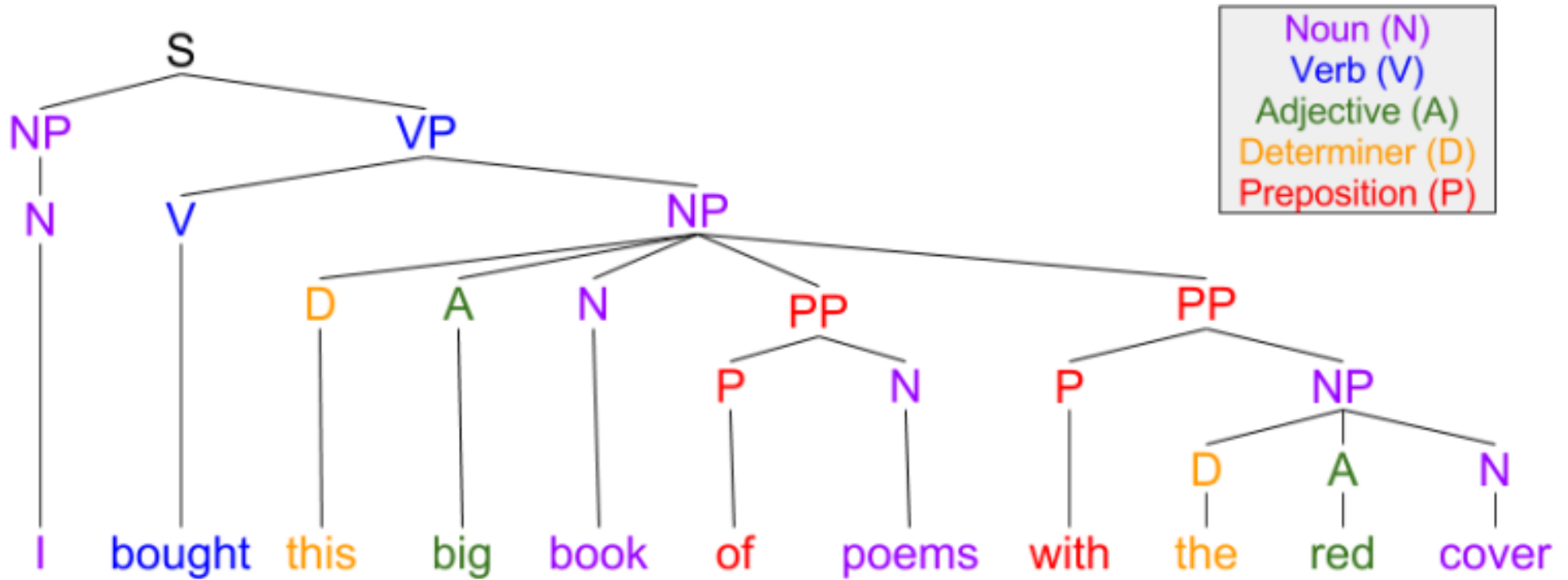
Formulation Arborescente

- Cette métarègle peut être formulée selon une arborescence (plus facile à interpréter). Les catégories spécificateur et complément désignent des types de constituants.
- Ainsi, on a :
- Spéc \bar{N} = Dét.
- Spéc \bar{V} = Aux.
- Spéc \bar{A} = DEG.
- Comp \bar{X} = $\bar{\bar{N}}$.

Exercice

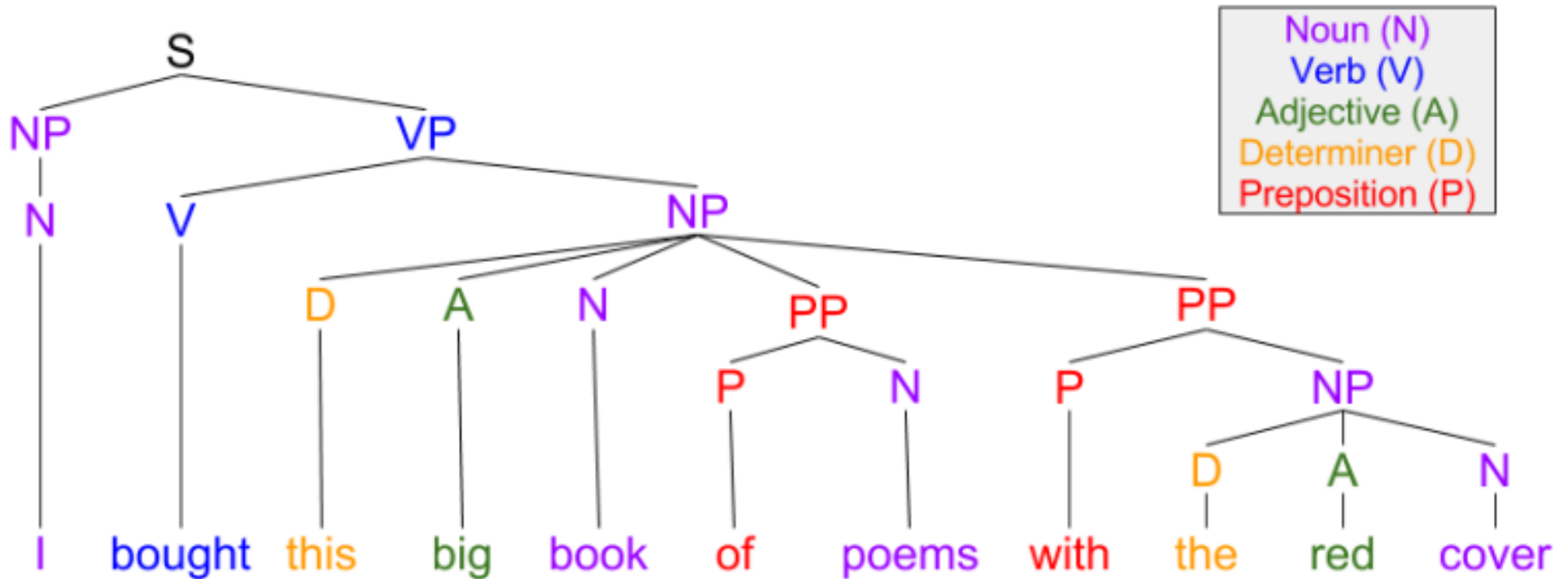
- Tracer l'arbre syntaxique de la phrase suivante
- "I bought this big book of poems with the red cover".

Solution



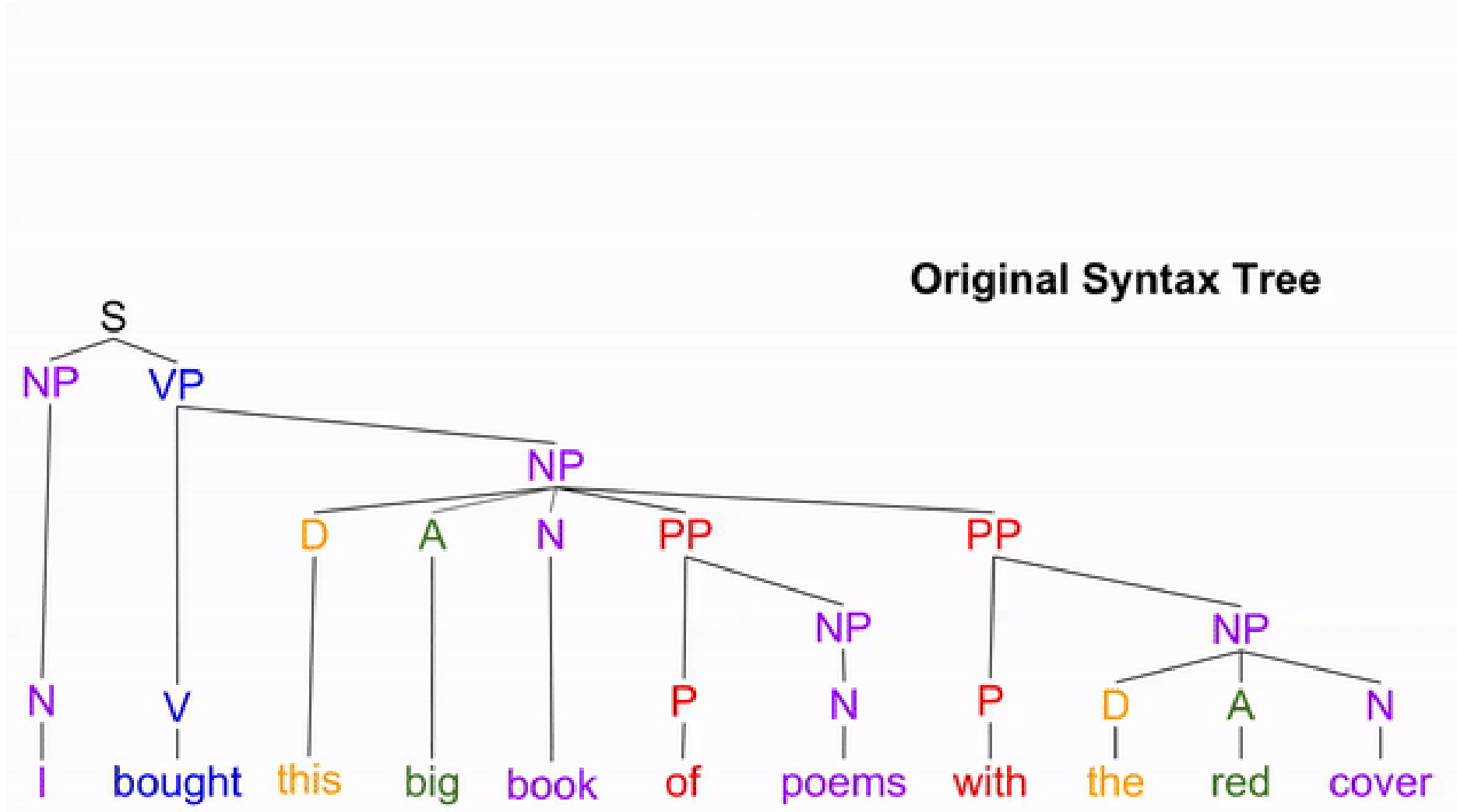
Xbar

- On veut balancer cet arbre comme arbre binaire



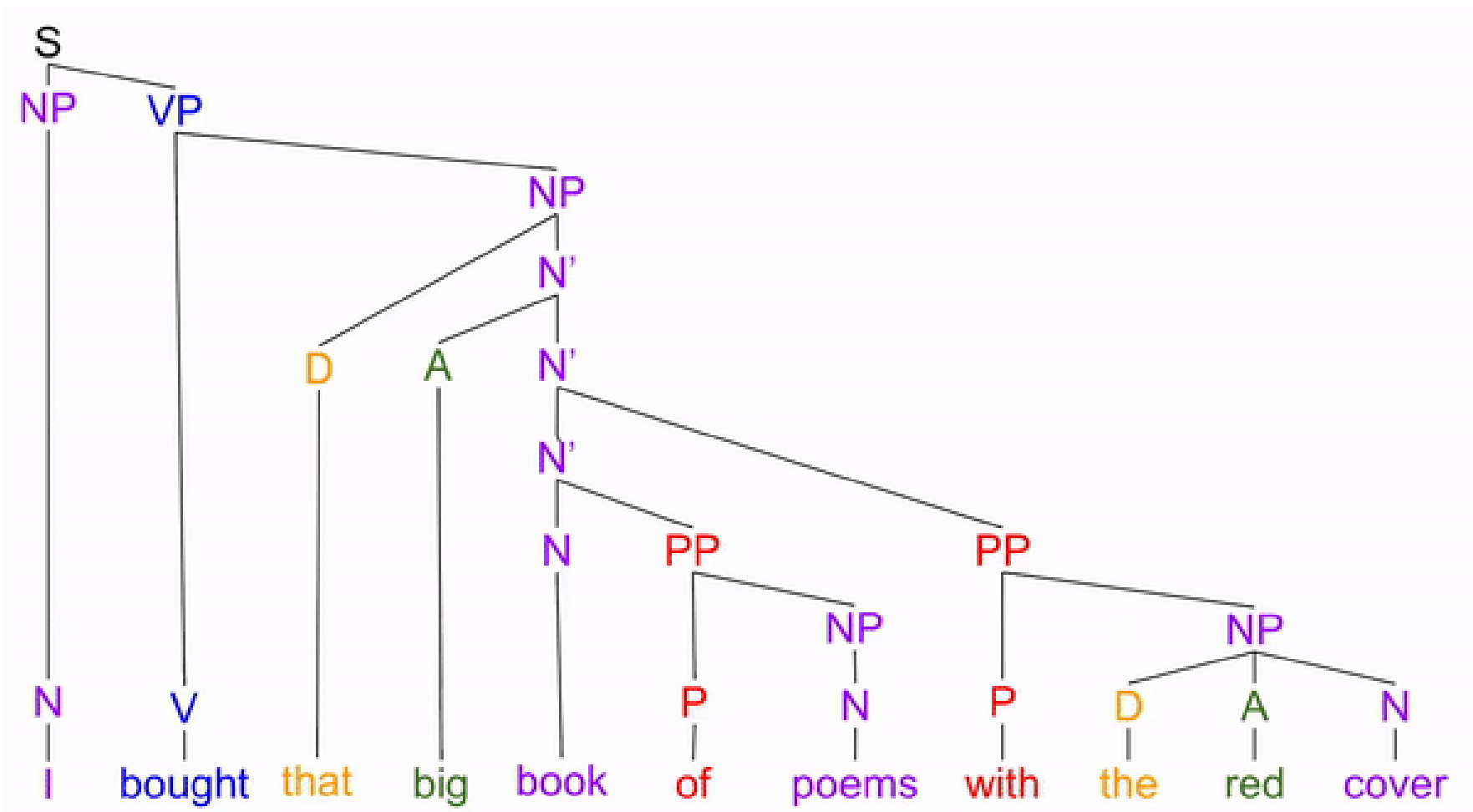
Xbar

- On veut balancer cet arbre comme arbre binaire



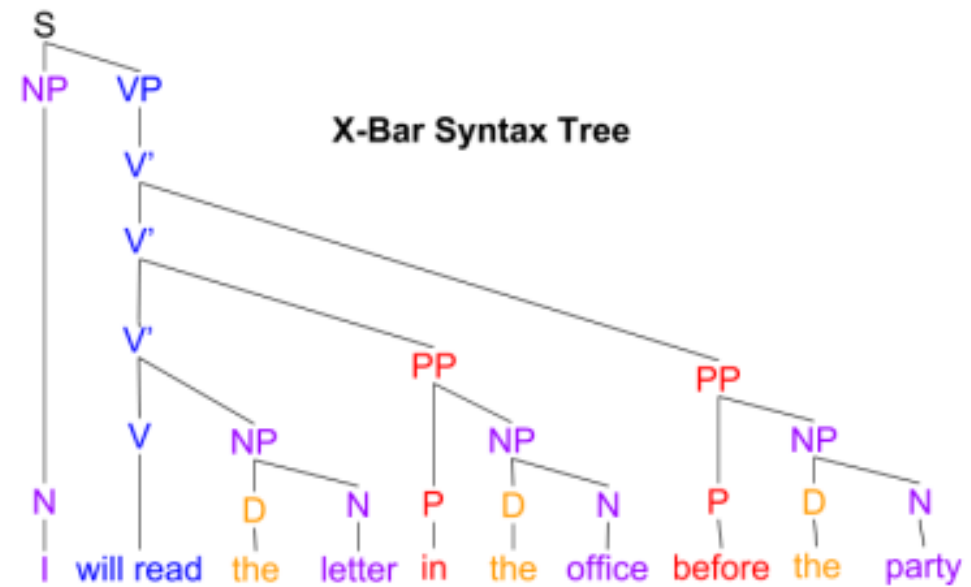
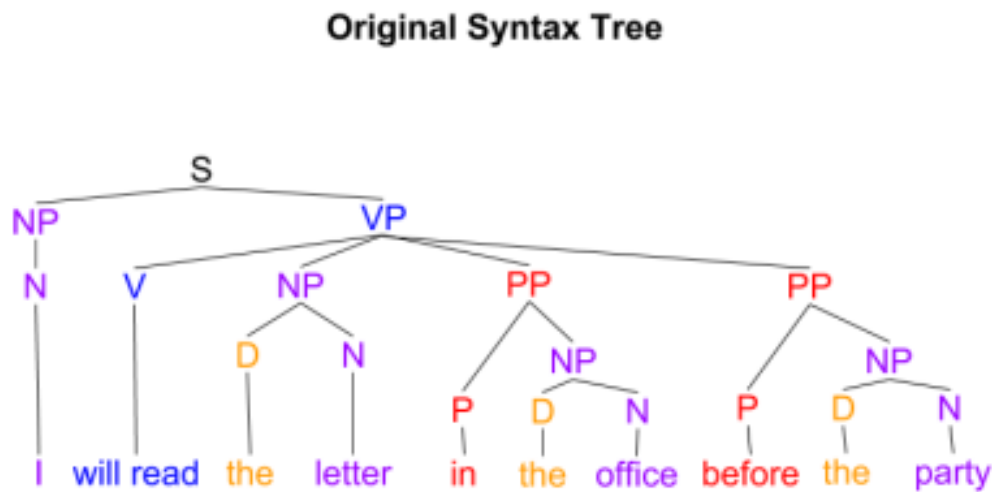
Xbar

- On veut balancer cet arbre comme arbre binaire



Xbar

- On veut balancer cet arbre comme arbre binaire



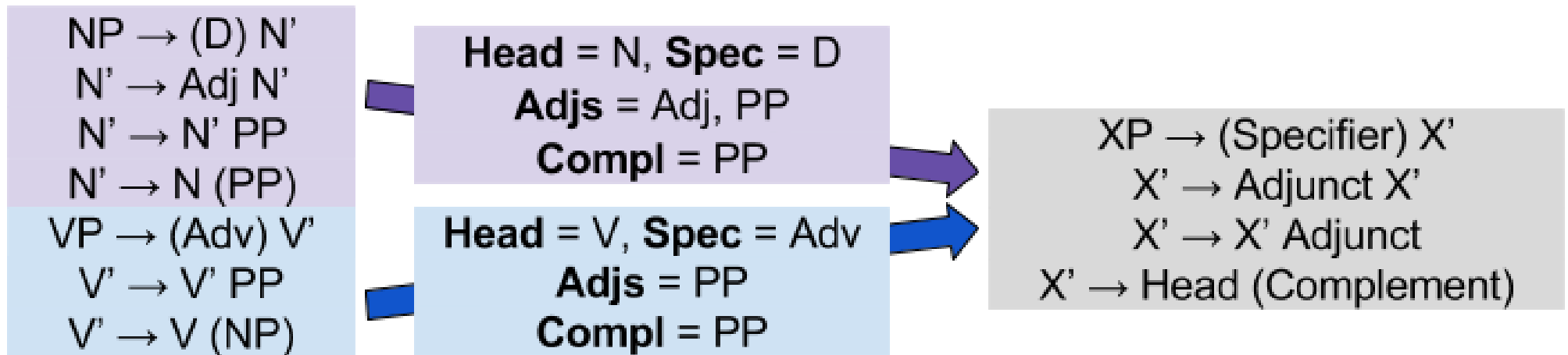
Règles de la grammaire

- 2

Original Rules	X-Bar Rules
$S \rightarrow NP VP$	$S \rightarrow NP VP$
$NP \rightarrow (D) (Adj) N (PP+)$	$NP \rightarrow (D) N'$ $N' \rightarrow Adj N'$ $N' \rightarrow N' PP$ $N' \rightarrow N (PP)$
$VP \rightarrow (Adv) V (NP) (PP+)$	$VP \rightarrow (Adv) V'$ $V' \rightarrow V' PP$ $V' \rightarrow V (NP)$
$PP \rightarrow P (NP)$	$PP \rightarrow P (NP)$

Règles de la grammaire

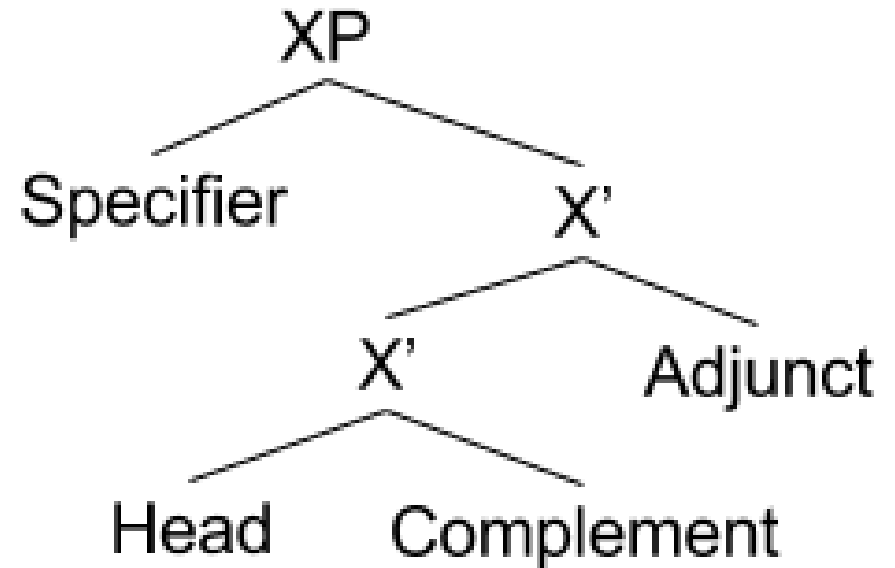
- 2



Règles de la grammaire

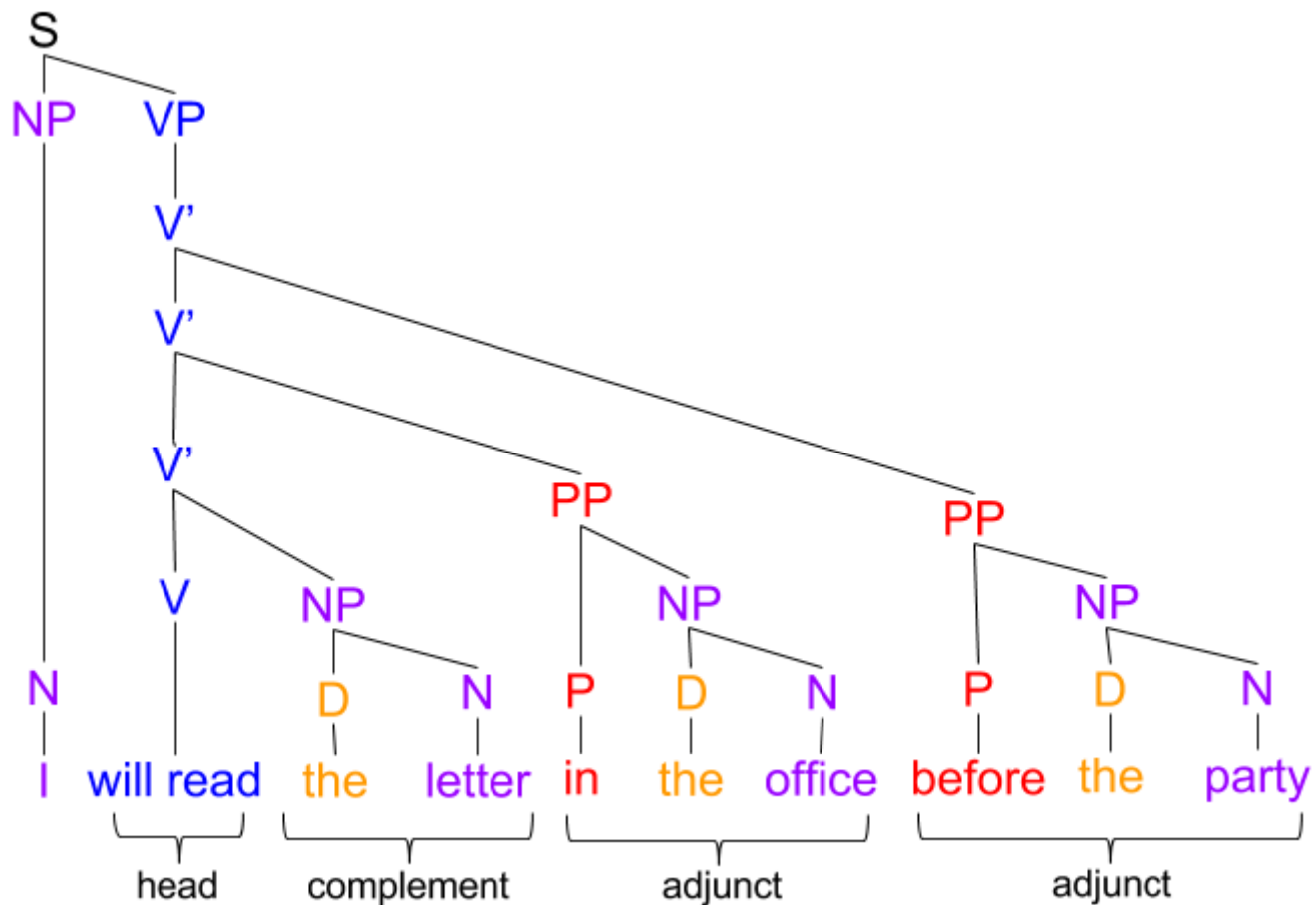
- 2

$XP \rightarrow (\text{Specifier}) X'$
 $X' \rightarrow \text{Adjunct } X'$
 $X' \rightarrow X' \text{ Adjunct}$
 $X' \rightarrow \text{Head (Complement)}$



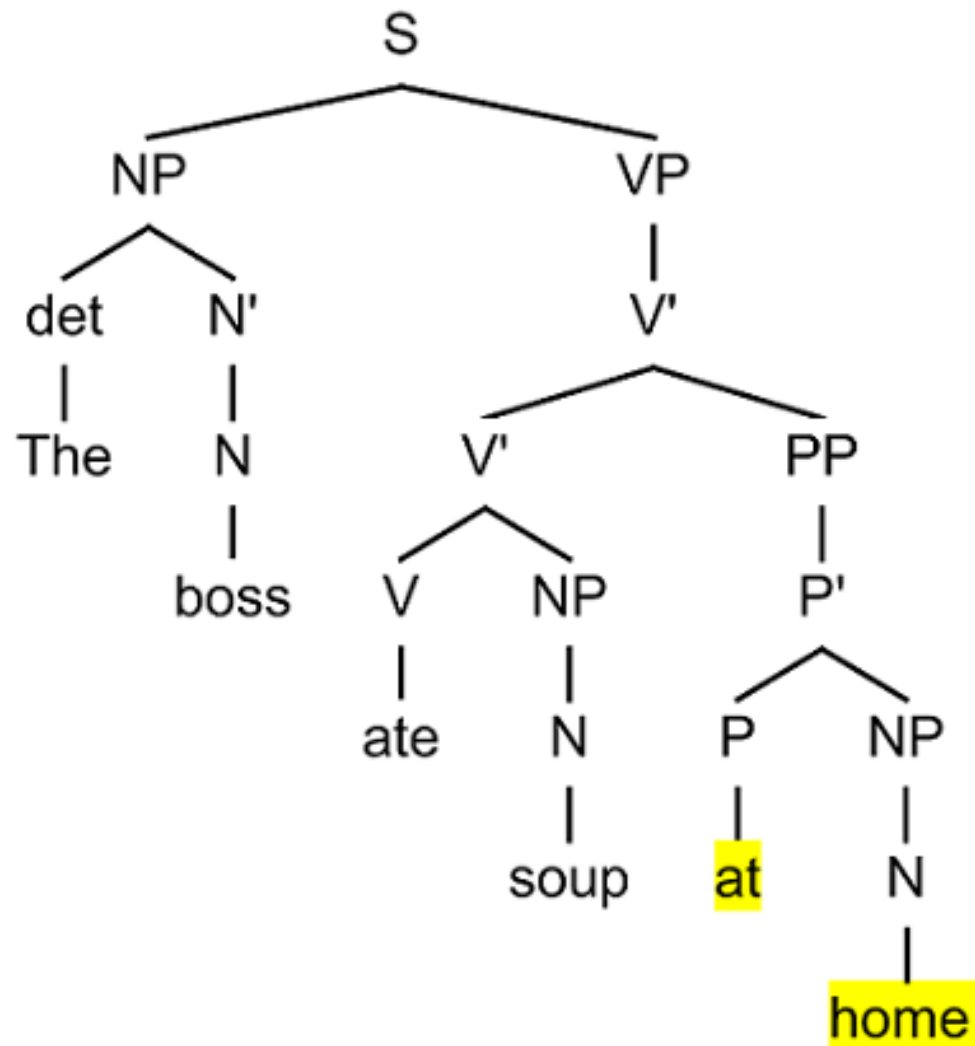
Adjuncts vs Complement

- La circonstance peut être réordonnée librement



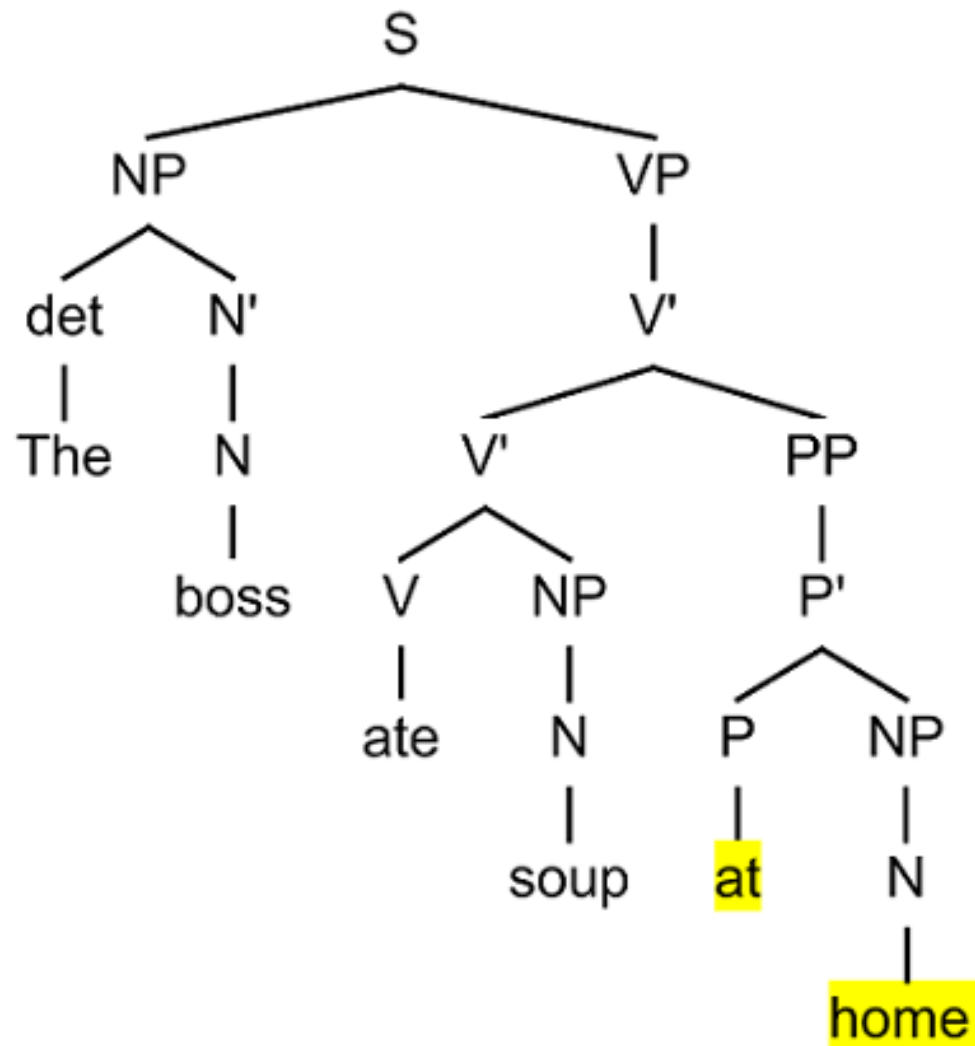
Exercice

- Etudier cet exemple



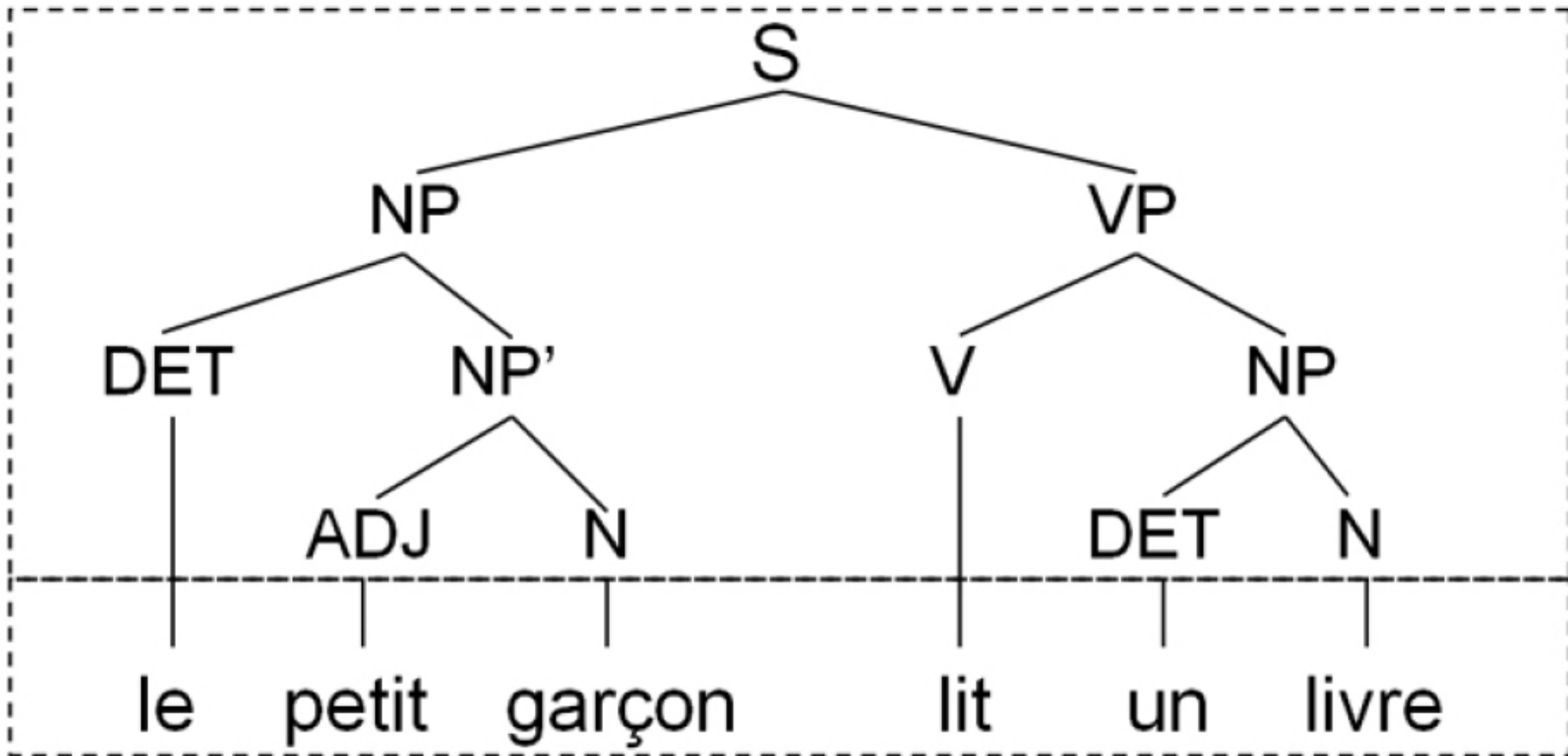
Exercice

- Etudier cet exemple



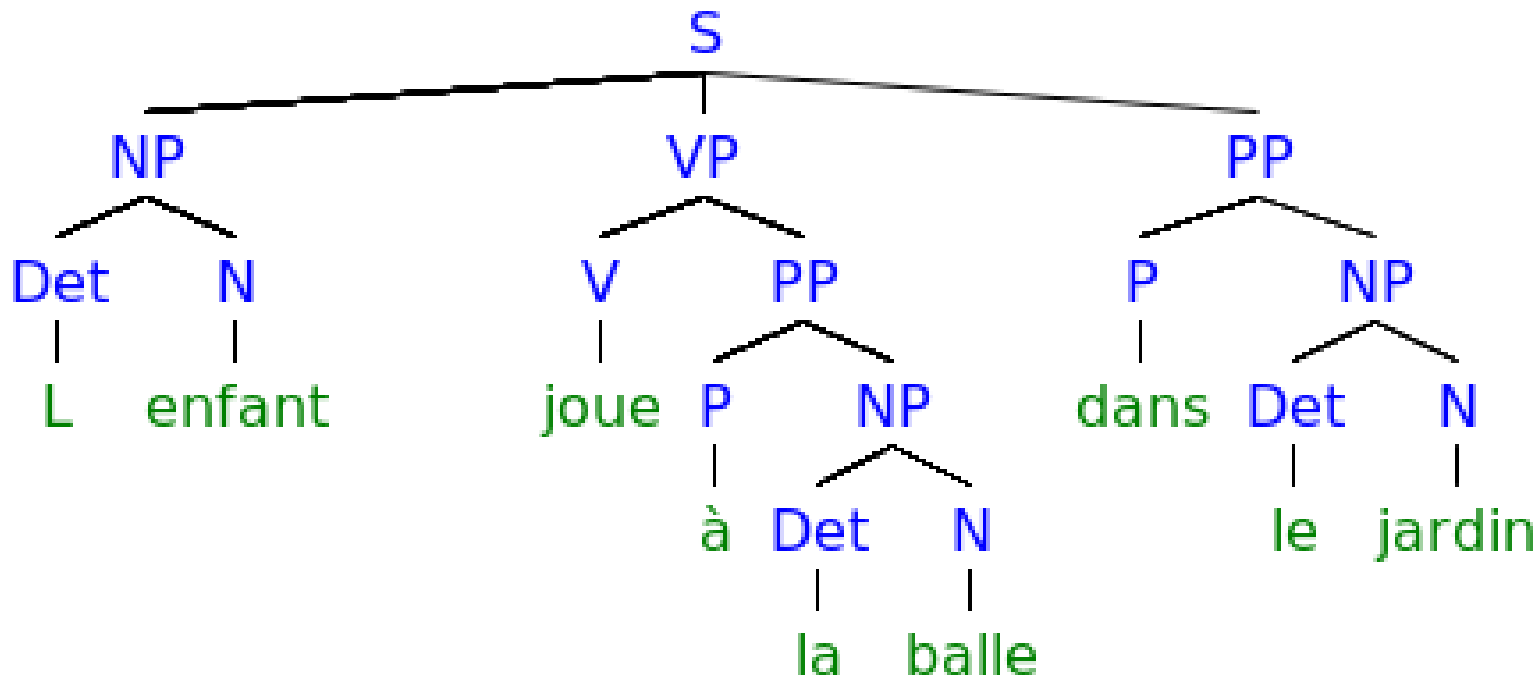
Exercice

- ReTracer l'arbre à l'aide de X-bar



Exercice

- ReTracer l'arbre à l'aide de X-bar



Exercice

- ReTracer l'arbre à l'aide de X-bar

