

Annotation Lexicale : Une Mesure de Complexité Contextuelle

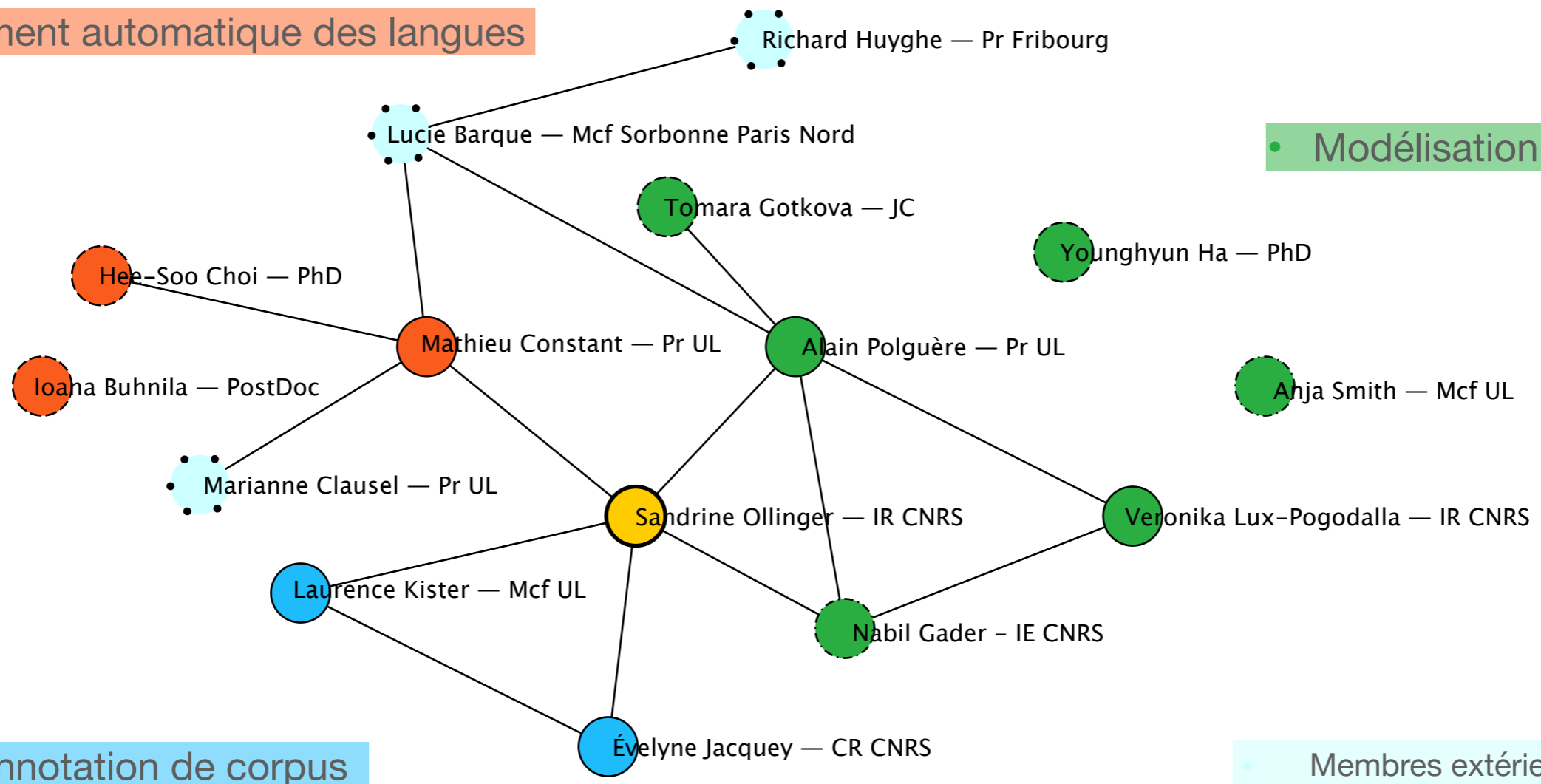


Sandrine Ollinger - 4 juin 2024

Contexte

Groupe de Recherche Polysémie et Ambiguïté Lexicale

• Traitement automatique des langues



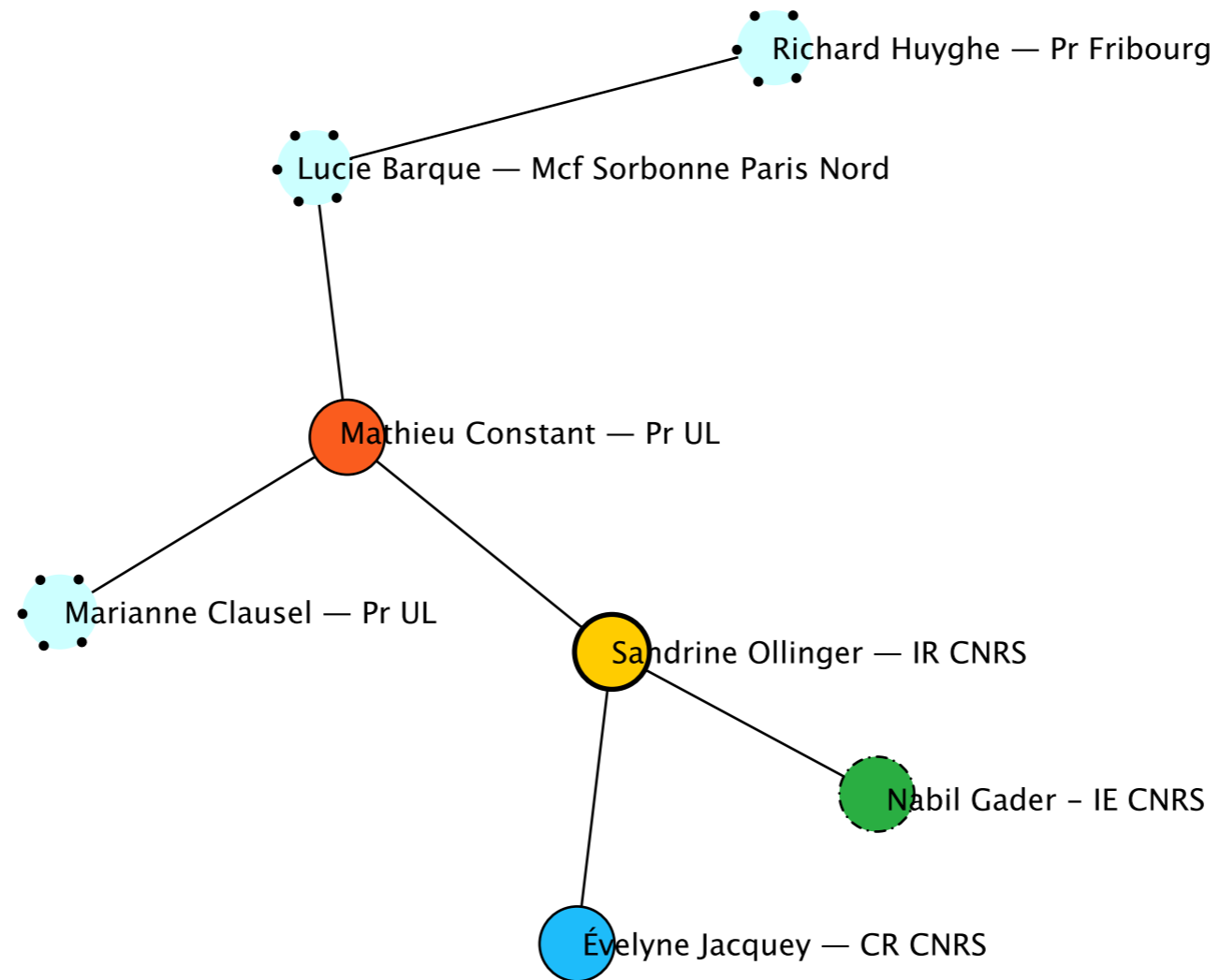
• Modélisation lexicale

• Annotation de corpus

• Membres extérieurs

Projet ALUMCoCo

Financement pôle CLCS – ATILF



Annotation en sens lexicaux

Ce que l'on cherche à faire

Une **tortue** est un reptile tétrapode doté d'une carapace osseuse.

Wikipédia, 19 février 2024

Annotation en sens lexicaux

Ce que l'on cherche à faire

Une **tortue** est un reptile tétrapode doté d'une carapace osseuse.

Wikipédia, 19 février 2024

tortue I

[CG]
nom
fém

[DF]
animal
tortue
=
animal
(petit)
dont le corps est entouré d'une carapace
qui se déplace lentement
terrestre ou aquatique

[FL]
Syn_S ● tortue de mer, ● tortue marine
Gener ● animal_N I.1b ; ● spéc reptiles b
Contr ● lièvre¹ a
Mero ● bec I.2 ; ● carapace

OU

tortue II

[CG]
nom
fém

[DF]
individu qui a un certain comportement
X, qui est une **tortue**

[FL]
Syn_n ● escargot II
A₁ Pred ● lent I.2
Fact₀ ● traîner IV.1

?

Annotation en sens lexicaux

Ce que l'on cherche à faire

tortue l

Une **tortue** est un reptile tétrapode doté d'une carapace osseuse.

Wikipédia, 19 février 2024

● **tortue l**

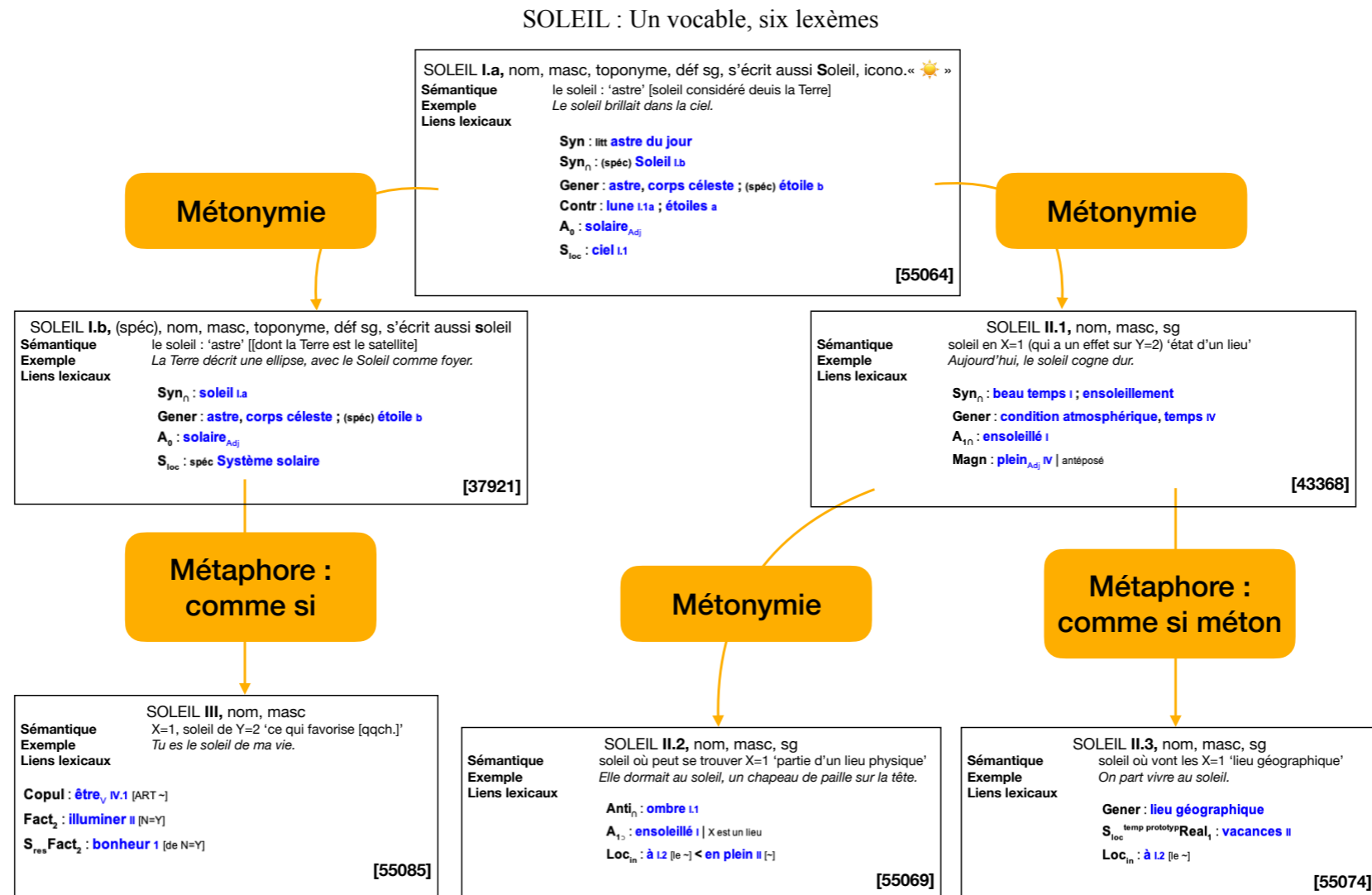
[CG]
nom
fém

[DF]
animal
tortue
=
animal
(petit)
dont le corps est entouré d'une carapace
qui se déplace lentement
terrestre ou aquatique

[FL]
Syn ● tortue de mer, ● tortue marine
Gener ● animal_N l.1b ; ● spéc reptiles b
Contr ● lièvre¹ a
Mero ● bec l.2 ; ● carapace

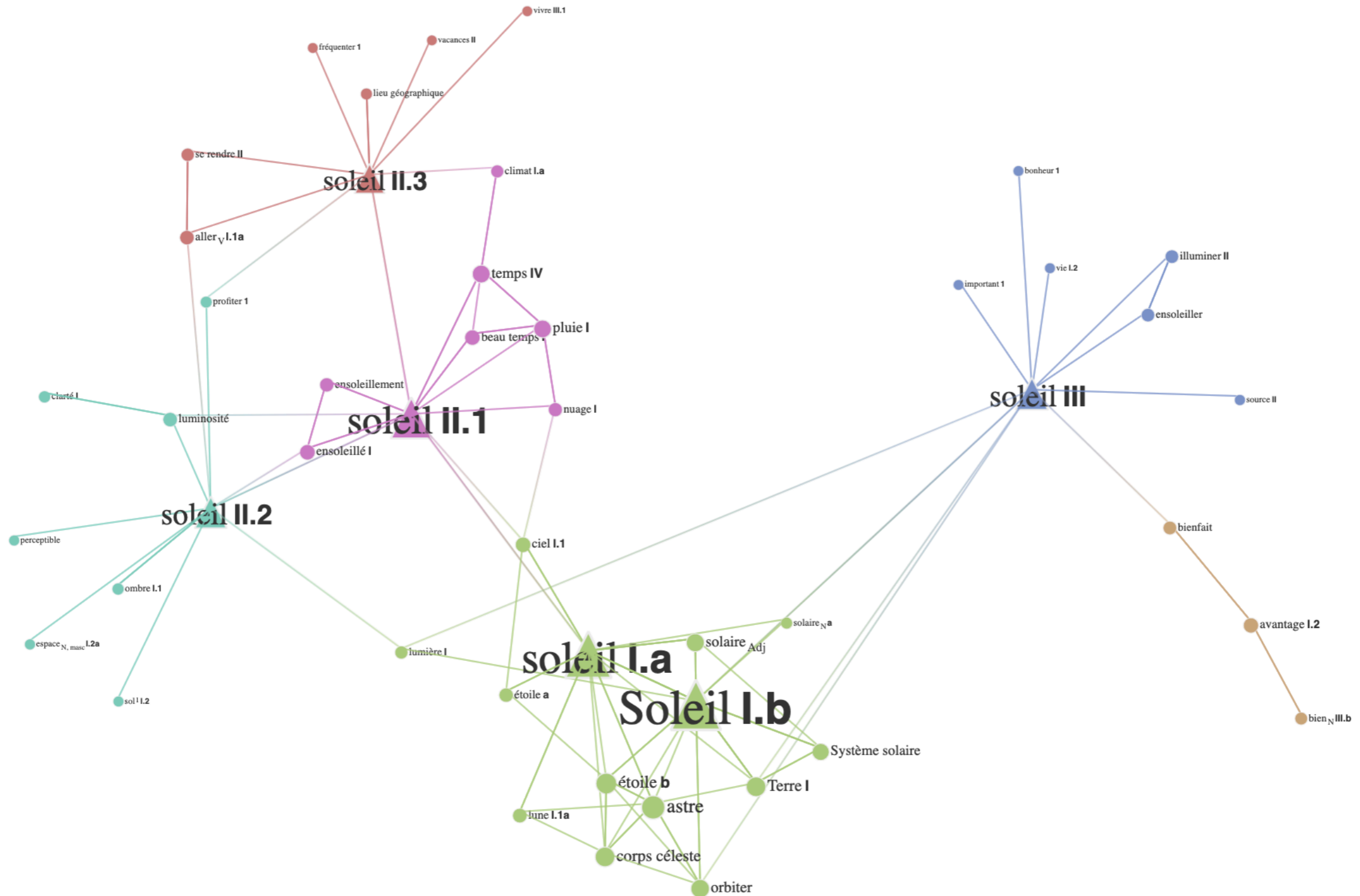
Inventaire de sens

RL-fr — SOLEIL 6 sens



Inventaire de sens

RL-fr — SOLEIL 6 sens



Quelques exemples (1)

Annotation manuelle en sens lexicaux

Quelques exemples (1)

Annotation manuelle en sens lexicaux

soleil II.1

[79] Le soleil fait briller les chromes des vélos à la devanture des magasins de cycles qui se succèdent le long de l'avenue.

soleil I.a

[155] Cet hiver-là fut tout plein de belles journées. [...] ; le soleil brillait souvent dans un grand ciel tout d'une pièce.

Quelques exemples (1)

Annotation manuelle en sens lexicaux

[249] Quand la famille en deuil fait son entrée, le soleil rit dans les assiettes, le vin scintille dans les carafes, le baeckeoffe qui mitonne encore dans la cuisine remplit toute la maison de son fumet.

Quelques exemples (2)

Annotation automatique en sens lexicaux

```
#####
cit31223: ['Elle murmurait mon prénom sans pouvoir s'arrêter\xa0: «\xa0Anouche, Anouche,
ma toute petite Anouche, mon soleil , tu es revenue voir ta vieille nourrice.\xa0»']
source ['ls:fr:ex:31492']

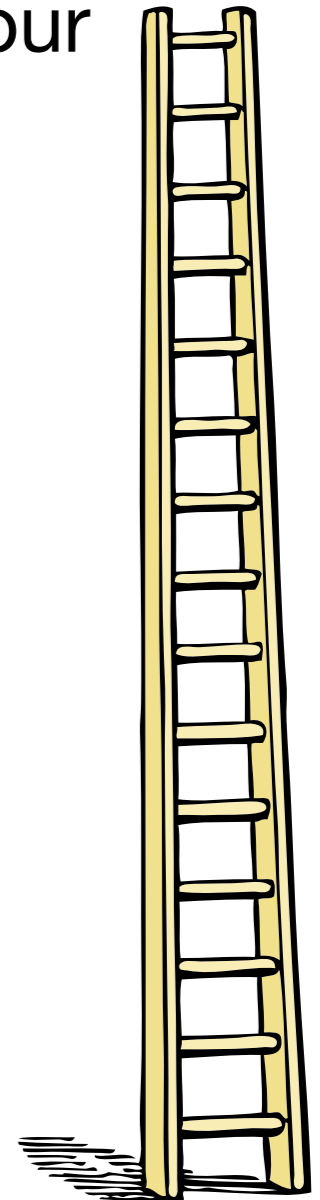
exp_name          BEL-RL-fr.d000.cit31223.t0
instance_id      BEL-RL-fr.d000.cit31223.t0
key              soleil__noun
pos              noun
source           BEL-RL-fr
first_label      __ws_6_soleil__noun__1,ls:fr:node:55085
pred             __ws_3_tarte.N__noun__1,ls:fr:node:56238
masked_pred      __ws_3_soleil__noun__1,ls:fr:node:56238
correct          0/0

candidates: ['ls:fr:node:41918/0.9021100997924805',
'ls:fr:node:38677/0.002905439119786024', 'ls:fr:node:40041/0.0006721526733599603',
'ls:fr:node:26529/0.0006341899861581624', 'ls:fr:node:26663/0.0004211456107441336',
'ls:fr:node:42332/0.0002631803508847952', 'ls:fr:node:32378/0.0002549559576436877',
'ls:fr:node:54878/0.00021936529083177447', 'ls:fr:node:40030/0.00021713468595407903',
'ls:fr:node:33139/0.0001958870270755142']
label_URL: https://spiderlex.atilf.fr/fr/id/55085
pred_URL: https://spiderlex.atilf.fr/fr/id/56238
mpred_URL: https://spiderlex.atilf.fr/fr/id/55064
#####
```

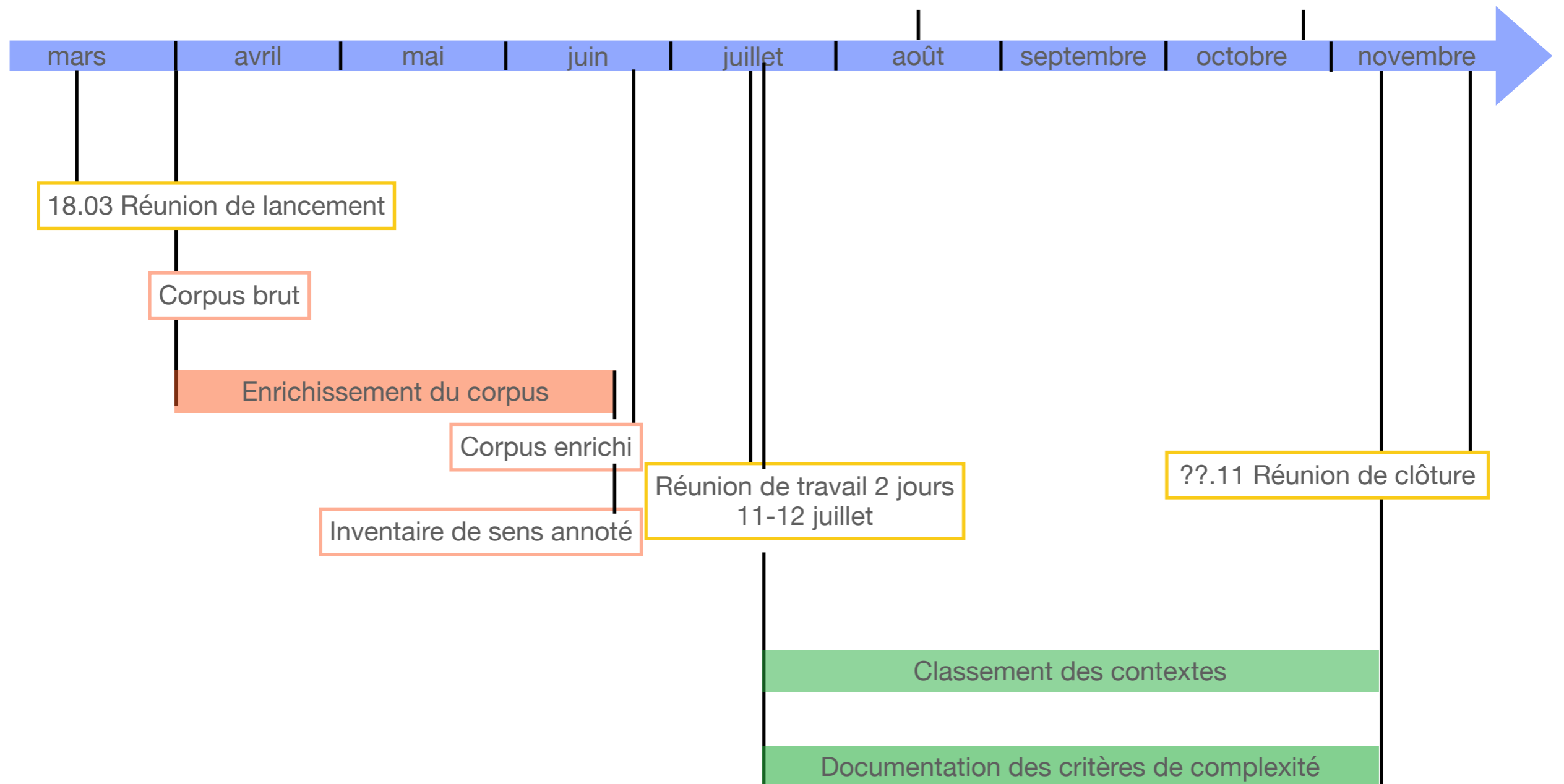
Se concentrer sur une question

Complexité du contexte

- En quoi un contexte est-il plus complexe qu'un autre pour identifier le sens lexical d'une occurrence ?
- Peut-on classer les contextes sur une échelle de complexité ?
- Peut-on mettre au point une mesure de complexité ?
- La complexité contextuelle est-elle la même pour une annotation manuelle et une annotation automatique ?



Les grandes étapes



Merci

Enrichissement du corpus

WSD auto	Is:fr:node :55064								
manuelle	soleil l.a								
	Le	soleil	levant	fait	briller	les	chromes	des	vélos
classes sémantiques		Object					Artefact		Artefact
voisins du RL-fr			l.a						
	le	soleil	levant	faire	briller	le	chrome	de	vélo
syntaxe	D	N	A	V	V	D	N	P+D	N
	det	subj	mod	aux_caus	root	det	obj	dep	obj

Enrichissement du corpus

Le corpus enrichi

format WebAnno pour INCEpTION

```
#FORMAT=WebAnno TSV 3.3
#T_SP=webanno.custom.SensSOLEILauto|value
#T_SP=webanno.custom.SensSOLEILmanuel|value
#T_SP=webanno.custom.ClassesSemantiques|value
#T_SP=webanno.custom.VoisinsRLfr|value
#T_SP=webanno.custom.Lemma|value
#T_SP=webanno.custom.POS|value
#T_SP=webanno.custom.Syntax|value

#Sentence.id=75
#Text=Le soleil levant fait briller les chromes des vélos
1-1 0-4 Le _ _ _ le D det
1-2 5-14 soleil 55064 I.a Object _ soleil N suj
1-3 15-18 levant _ _ _ I.a levant A mod
1-4 19-25 fait _ _ _ _ faire V aux_caus
1-5 26-30 briller _ _ _ _ briller V root
1-6 31-38 les _ _ _ _ le D det
1-7 39-41 chromes _ _ Artefact _ chrome N obj
1-8 41-48 des _ _ _ _ de P+D dep
1-9 49-50 vélos _ _ Artefact _ vélo N obj
```

Échelle de complexité

Méthode à imaginer — exemple T. François (2015) Amesure

