

Corpus arborés et parsing

Cours 8

Santiago Herrera
s.herrera@parisnanterre.fr

Correction simple du DM

1. Explorez les relatives en français sur Grew-match et décrivez ce que vous avez trouvé.
 - a. Quelles sont les parties du discours des gouverneurs des relatives ?
<https://universal.grew.fr/?custom=655712ecc613b>
 - b. Quelles sont les fonctions syntaxiques (relations de dépendance) du pronom relatif dans les relatives ? Montrez aussi leur distribution.
<https://universal.grew.fr/?custom=6557142daacc8>
 - c. Combien de sujets d'une relative sont en position post-verbale ?
<https://universal.grew.fr/?custom=655713912ea3f>
 - d. Tous les pronoms relatifs sont-ils dans une relative ?
<https://universal.grew.fr/?custom=6557146e423e9>

Correction simple du DM

2. Trouvez les relations de dépendance non projectives

- a. Écrivez des requêtes Grew-match pour trouver ces relations sans utiliser la requête `global { is_not_projective }`

Au moins deux requêtes...

- b. Quelles sont les constructions syntaxiques non projectives en français ? Décrivez vos résultats avec des exemples.

Des objets pronominaux qui dépendent d'un participe passé, comparatives, des pronoms relatifs dans des relatives, etc.

Correction simple du DM

3. Transformez, avec des règles Grew de réécriture, le corpus de trois phrases "corpus_DM1_CAetP.conllu" depuis le schéma d'annotation UD vers le schéma d'annotation SUD (les deux versions sont disponibles sur Grew-match).

Nous l'avons vu en classe. La semaine prochaine, je vous montrerai les résultats.

Annotation

- Guidelines

Exemple en SUD : <https://surfacesyntacticud.github.io/guidelines/u/relations/subj/>

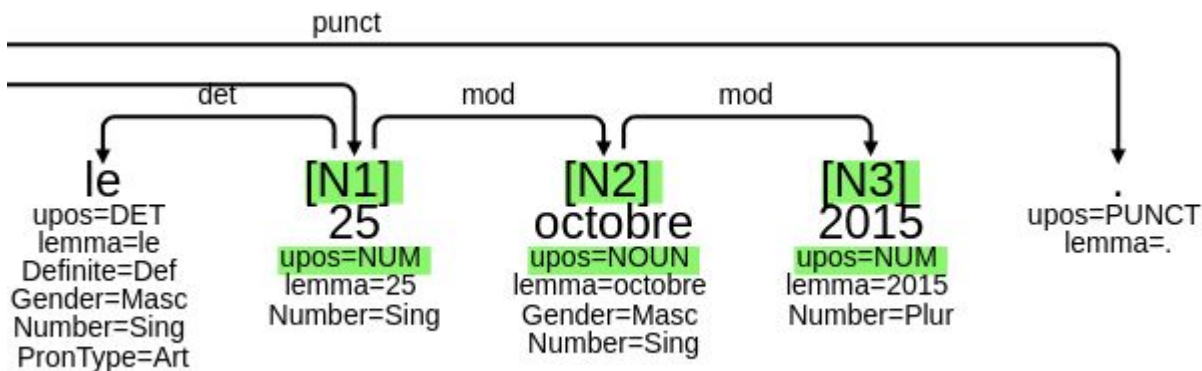
Exemple en UD : <https://universaldependencies.org/guidelines.html>

Corpus GUM : <https://wiki.gucorpling.org/gum/dependencies>

- Différents objectifs, différente granularité, des organisations différentes

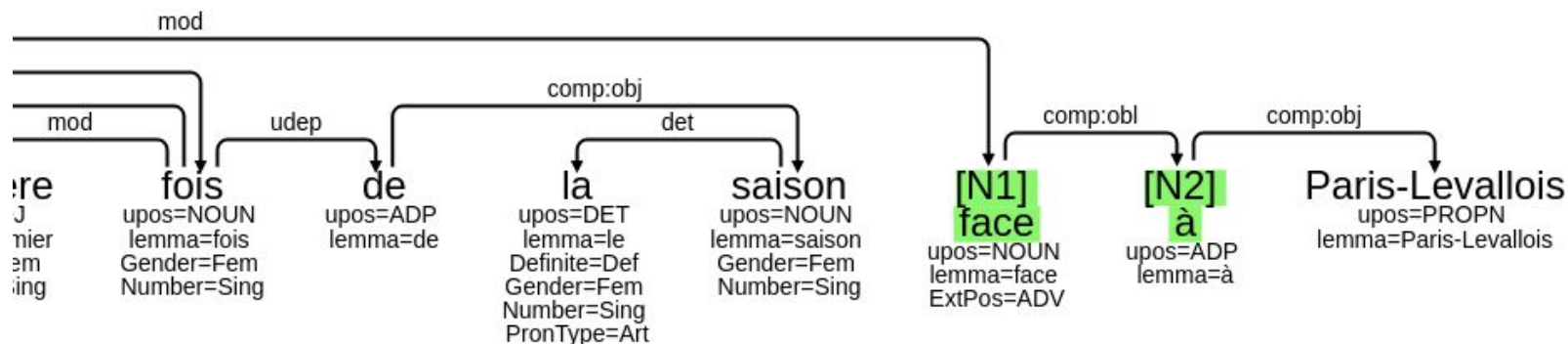
Dates

Link : <https://universal.grew.fr/?custom=65572485d00af>



ExtPos en SUD

Rouen s'impose pour la première fois de la saison **face** **à** Paris-Levallois (91-89).

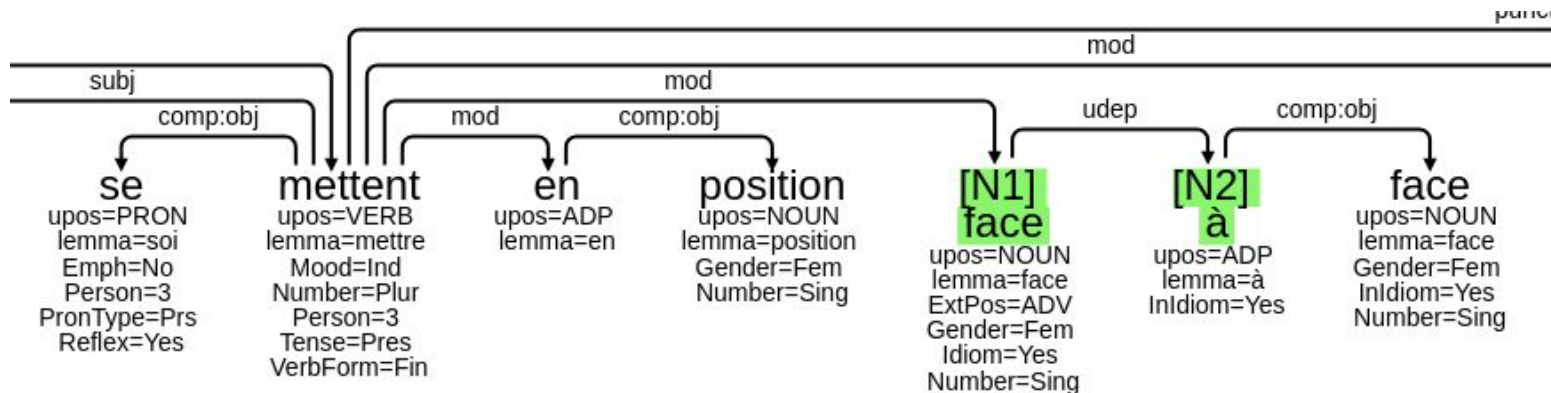


Idiom et InIdiom

Phrasème, locution, expression idiomatique, expression figée

Ex. Se retourner dans sa tombe / tiré par les cheveux

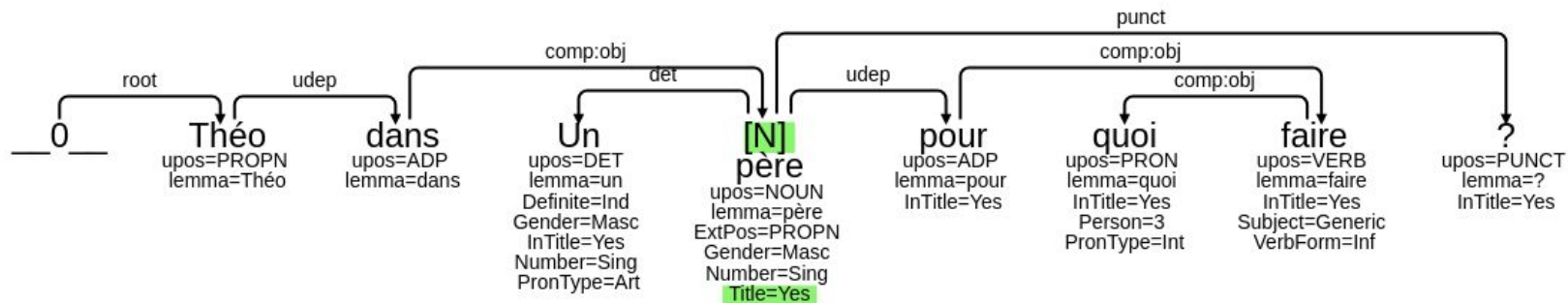
Les deux armées se mettent en position **face à face**, à portée de canon.



Autres exemples plus subtils : “en vigueur”, “tout à fait”

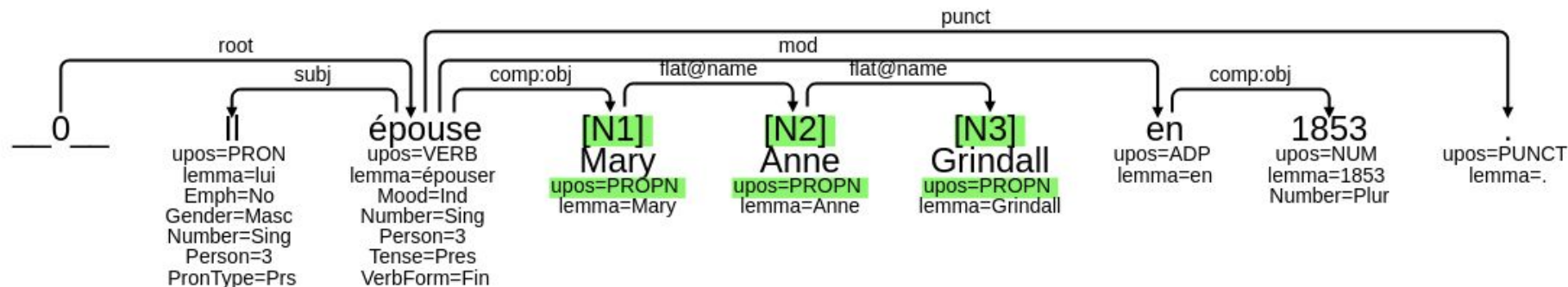
Title et InTitle

Théo dans Un père pour quoi faire ?



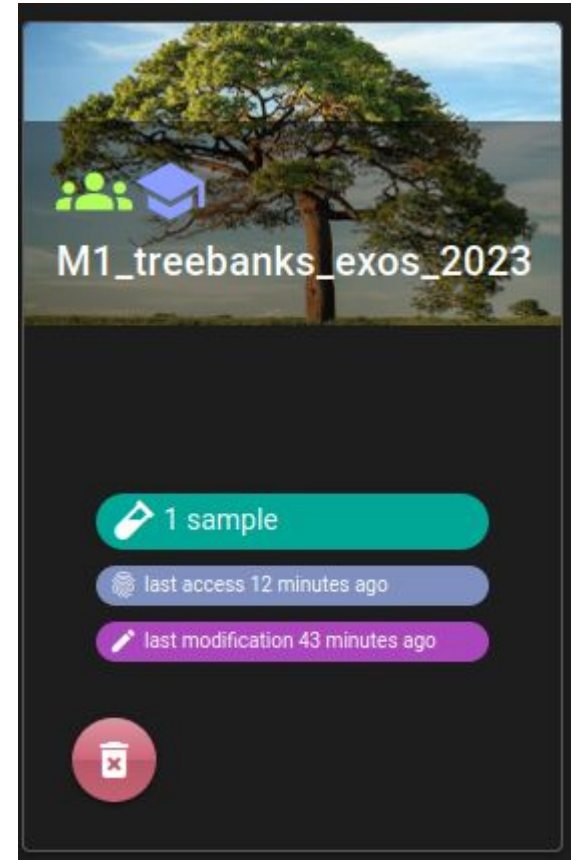
Les noms propres

Il épouse Mary Anne Grindall en 1853.

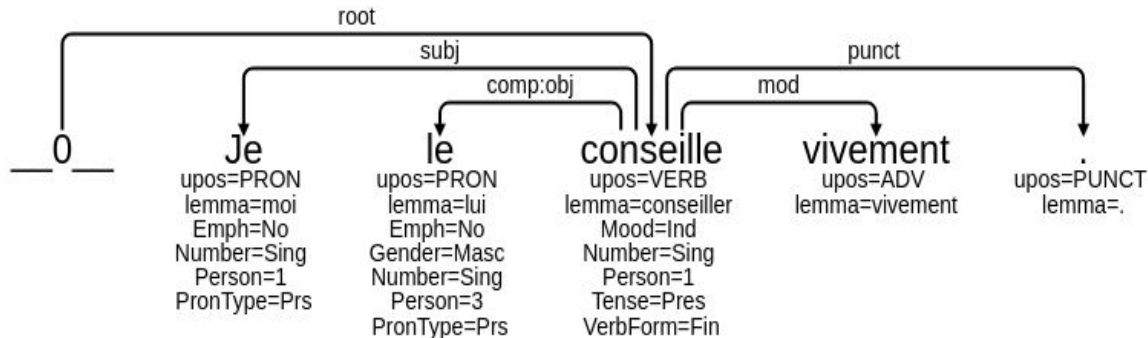


Annotation en ligne : Arborator

- <https://arboratorgrew.elizia.net>
- Projet = M1_treebanks_exos_2023
- **1a, 1b, 1c.trees**



Fichier conllu



```

# global.columns = ID FORM LEMMA UPOS XPOS FEATS HEAD DEPREL DEPS MISC
# sent_id = fr-ud-train_06412
# text = Je le conseille vivement.
1  Je  moi  PRON      _      Emph=No|Number=Sing|Person=1|PronType=Prs  3  subj  _      wordform=je
2  le  lui  PRON      _      Emph=No|Gender=Masc|Number=Sing|Person=3|PronType=Prs  3  comp:obj  _      _
3  conseille  conseiller  VERB      _      Mood=Ind|Number=Sing|Person=1|Tense=Pres|VerbForm=Fin  0  root  _      _
4  vivement  vivement  ADV      _      _      3  mod  _      SpaceAfter=No
5  .      .      PUNCT     _      _      3  punct  _      _
  
```

Notebook

- Télécharger le zip
- Décompresser le fichier dans un dossier

```
mkdir dossier  
cd dossier  
python3 -m venv .venv  
source .venv/bin/active  
pip install -U -r requirements.txt  
jupyter notebook
```

pour sortir de l'environnement virtuel : *deactivate*

Windows

Windows venv activation

To activate your venv on Windows, you need to run a script called `myenv`, the command would be:

```
# In cmd.exe
venv\Scripts\activate.bat
# In PowerShell
venv\Scripts\Activate.ps1
```