

GREW, a tool for annotating corpora and exploiting annotated corpora - <http://grew.fr>

Guy Perrier and Bruno Guillaume, LORIA, Nancy

Why is Graph Rewriting relevant to NLP ?

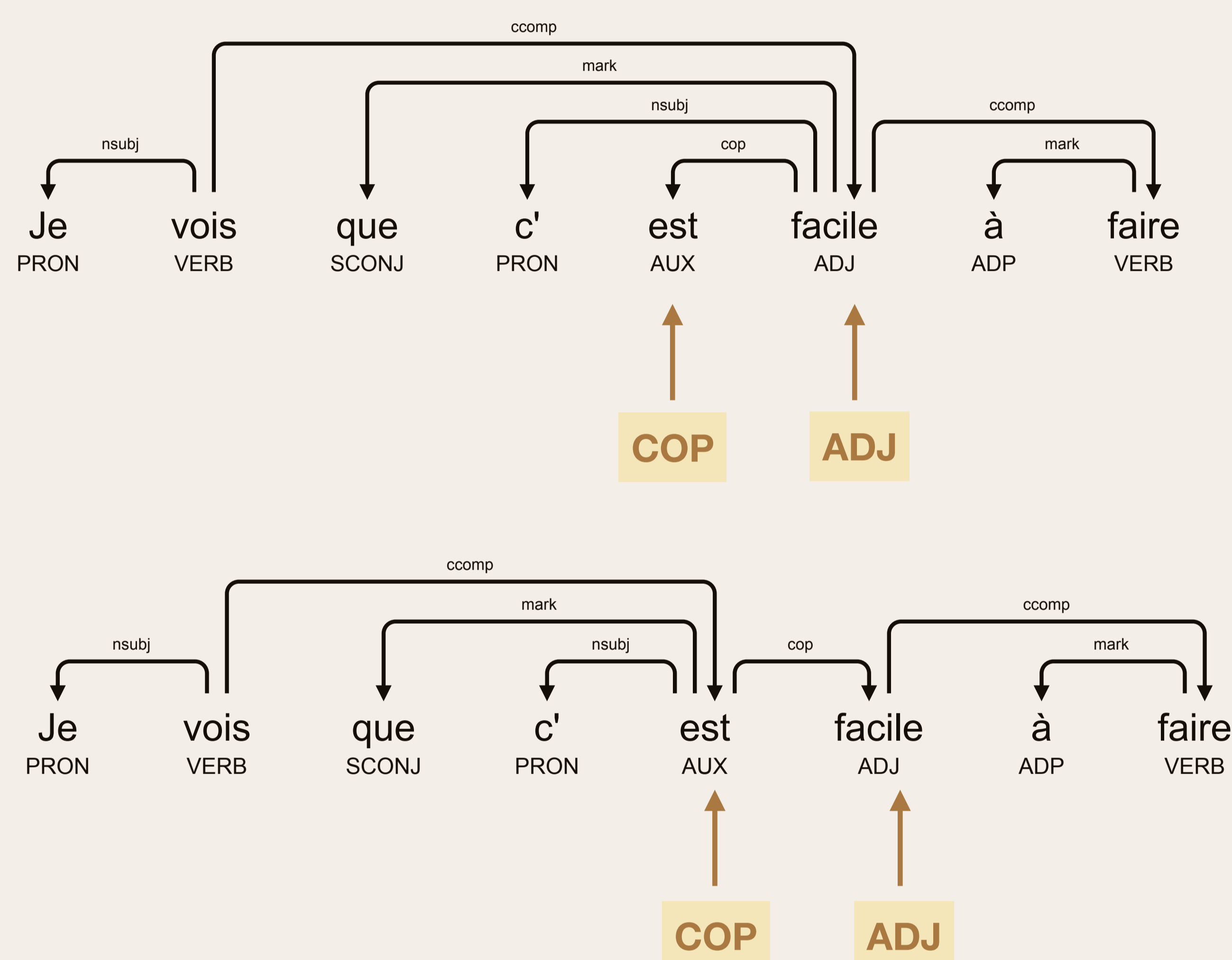
- **Graphs** are a common framework well suited to linguistic representation.
- **Graph rewriting** is a nice computational model for linguistic graphs:
 - ✓ Every **rewriting rule** is able to model a **local linguistic phenomenon**.
 - ✓ Computations are **non deterministic** and can model the **ambiguity** of natural languages.
 - ✓ Applications for **converting and correcting annotations** can be easily implemented.

GREW, an implementation of Graph Rewriting adapted to NLP

- Right hand sides of rules are sequences of **atomic commands**.
- Nodes contain **feature structures** and represent linguistic entities.
- Edges are **labelled** and represent linguistic relations.
- Rules can be **parametrised** by **lexical information**.
- **Strategies** can be used to control the rules applications.

A GREW rule for the Graph Rewriting System from UD to SUD

```
rule rev_cop {
  pattern {
    ADJ[upos=ADJ];
    e: ADJ -[cop]-> COP;
  }
  without {
    ADJ -[case]-> PREP % Ex: je suis en colère
  }
  commands {
    del_edge e;
    shift_in ADJ ==> COP;
    shift_out ADJ =[csubj|nsubj|mark]==> COP;
    add_edge e:COP -> ADJ;
  }
}
```



Graph Matching, a very useful by-product of Graph Rewriting

- The **pattern** & **without** part of rewriting rules can be used to search for patterns in corpora.
- This is useful for **corpora exploration**, **error mining** and **linguistic studies**.
- Available as an online service: <http://match.grew.fr>

A Graph matching example

Searching number disagreement between verbs and their subject in **UD_French-GSD**

All cases of disagreement

```
pattern { M [upos=VERB]; M -[nsubj]-> N; M.Number <> N.Number }
```

When the “avoir” auxiliary is used, the verb does not agree with the subject: *ils ont mangé un couscous*

```
pattern { M [upos=VERB]; M -[nsubj]-> N; M.Number <> N.Number }
without { M -[aux]-> A; A [lemma=avoir] }
```

When the subject is the head of a coordination, the verb does not agree with it but with the whole coordinated expression: *Jean et son père arrivent*.

```
pattern { M [upos=VERB]; M -[nsubj]-> N; M.Number <> N.Number }
without { M -[aux]-> A; A [lemma=avoir] }
without { N -[conj]-> C; }
```