Enunciative and modal variations in newswire texts in French: From guideline to automatic annotation

Marine DAMIANI

Université Paris Ouest – MoDyCo, UMR 7114
Nanterre, F-92000, France
marinedamiani@gmail.com

Delphine BATTISTELLI

Université Paris Sorbonne - STIH, EA 4509
Paris, F-75006, France
delphine.battistelli@paris-sorbonne.fr
1. Goals

- Although modality markers in French have been systematically described there is still no reference corpus proposing the annotation of enunciative and modal characteristics as a discursive delimitation task.

- Developing a corpus of French newswire texts annotated with enunciative and modal commitment information.

- Improving the annotation guideline by using the evaluation of a first round of annotation produced by two expert annotators and by our automatic annotation system.
Two particularity of the annotation are that:

- a predicative cue can be embedded in the scope of another predicative cue
- the predicative cue is not part of the scope but is a trigger of a semantic scope
3. Workflow of guideline improvement

1. guideline v0

2a. A1 annotation
2b. automatic annotation

3. IAA

4. adjudicated corpus

5. IAA

6. discussion

The 7th Linguistic Annotation Workshop & Interoperability with Discourse
ACL Workshop, Sofia, August 8-9, 2013
4. Disagreement example

[Le procès devant un tribunal militaire d'un blogueur égyptien arrêté pour avoir critiqué l'armée, qui a débuté lundi, a été ajourné à dimanche]scope, a indiqué cue mardi un de ses avocats.

[The trial before a military court of an Egyptian blogger arrested for criticizing the army, which began on Monday, has been postponed to Sunday]scope, said cue one of his lawyers on Tuesday.
5. Perspectives

Next steps are:

- Launching a larger annotation campaign involving more human annotators and a bigger corpus
- Integrating modifier cues such as hedging adverbs
- Introduce discursive cues impacting more than a single sentence