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

Marine Damiani
MoDyCo, UMR 7114, Université Paris Ouest
marinedamiani@gmail.com

Delphine Battistelli
STIH, EA 4509, Université Paris Sorbonne
delphine.battistelli@paris-sorbonne.fr

Abstract

In this paper we present the development of a corpus of French newswire texts annotated with enunciative and modal commitment information. The annotation scheme we propose is based on the detection of predicative cues - referring to an enunciative and/or modal variation - and their scope at a sentence level. We describe how we have improved our annotation guideline by using the evaluation (in terms of precision, recall and F-Measure) of a first round of annotation produced by two expert annotators and by our automatic annotation system.

1 Introduction

This paper concerns the design of a reference corpus that can be used to evaluate an automatic annotation system of enunciative and modal commitment in newswire texts in French. This complex linguistic phenomenon refers to the fact that a situation can be presented as certain, or only possible/probable, by an enunciator who can be the author of the text but who can also be another enunciator (explicitly named or not) from whom the author reports some content that he has heard, read, imagined, etc. Different kinds of linguistic cues are involved. In addition to the need to identify and semantically classify these cues, one has to deal with the question of their scope. This question is even more complex as many cues can be present together in a sentence, thus complexifying the interpretation of the interaction of different scopes (see Example 1.).

1. M. Arabi **a exprimé**[cue1] [**le souhait**[cue2] [d'aider la Syrie à surmonter cette phase]scope2]scope1] // [Mr. Arabi **expressed**[cue1] [**a desire**[cue2] [to help Syria overcome this phase.]scope2]scope1]

Another major difficulty concerns the fact that evidential and modal characteristics are very similar (see for example a noun like *desire*). Our work addresses the question of annotating these cues and their semantic scope. Unlike most other approaches, we have chosen not to treat these two kinds of characteristics separately, since both are implicated in what is called enunciative commitment. We will focus here on our practice for the development of a reference corpus.

After a brief presentation of the theoretical background (section 2), we describe which kinds of linguistic cues are considered and what kind of semantic scopes are then encountered (section 3). Our annotation procedure aims to delimit textual segments that are semantically impacted by the presence of enunciative and modal cues. In this light, we will focus only on what we will describe below as *predicative cues*. Then we will explain how we have improved our annotation guideline by using the evaluation of a first round of annotation produced for the same task by two expert annotators and by our automatic annotation system (section 4).

2 The phenomenon of enunciative and modal commitment

In the field of linguistics, the notion of modality can be considered from an enunciative perspective (see Bally, 1932 Benveniste, 1966; Culioli, 1973). From this perspective, which is the one we adopt here, the construction of an utterance (or a text) has to take into account certain language operations such as predication or operations of commitment, the expression of which leaves a certain number of surface linguistic traces (or cues). The enunciator's degree of commitment to a predicative content is marked in the utterance by different kinds of linguistic traces. In other words, it can be said that in discourse the enunciator expresses different degrees of commitment to the truth of the propositional content.

Very close to this issue is thus the long tradition of tracking veridicality in discourse. Whether – in the most recent work - under the term of "factuality degrees of events" (Sauri and Pustejovsky, 2012), "event veridicality" (De Marneffe et al., 2012), "detection of uncertainty" (CoNLL-2010 Shared Task) or "attributions" and

"private states" (Wilson and Wiebe, 2005), this notion refers to the relationship between language and reader commitment. In our approach, we do not attempt to access the notion of veridicality directly but rather via the organization of the text into different textual segments that have different enunciative and modal validation contexts. However, the cues we have to take into account to achieve this goal are mostly the same as in veridicality studies (modal verbs, reported speech verbs, verb of propositional attitude, hedging adverbs, and so on). Moreover, beyond traditional lexical cues, we also include in our work other cues such as morphological inflection (e.g. inflection of the French conditional tense), syntactic constructions such as subordinate clauses of condition or prepositional constructions (e.g. *according to X, at first sight*…). Furthermore, we have to take into account the fact that a lot of cues are embedded (as seen in Example 1 with *express* and a *desire*). If we want to interpret the enunciative and modal context of the textual segment *to help Syria overcome this phase*, we have to consider the fact that it is embedded in the segment *a desire to help Syria overcome this phase*. From this point of view our work is related to Kilicoglu (2012) who studied "embedding predications". Thus, we do not only consider the type of cues we find in text but also the way they interact. This methodology also enables us to consider cues that play a role at a discursive level. This question of discursive markers is discussed in (Charolles et al., 2005).

Although modality markers in French - in their close relationship with the markers of evidentiality - have been systematically described (see for example Gosselin, 2010; Le Querler, 2004) there is still no reference corpus proposing the annotation of enunciative and modal characteristics as a discursive delimitation task and this is the goal we seek to achieve. This problem of identifying modal cues related to a scope was initially researched in biomedical texts (Vincze et al., 2008). This applicative task made it possible to renew the linguistic approach to modality by adopting a more concrete approach, focusing first on the variety of cues that can be identified in a text. This perspective also enables the issue of the influence of textual genre on modality markers to be addressed.

In the next section, we present the way we propose to annotate this enunciative and modal commitment variation in text in terms of cues and scopes.

3 Annotating enunciative and modal commitment in term of cues and scope

Our annotation goal is to define in which enunciative and modal context a propositional content occurs. Observation of the cues in our corpus showed that there are two kinds of cues: predicative cues that lead to the opening of a new textual segment (this kind of cue has the syntactic property of governing another textual segment, e.g. cue1 in Example 2.) and what we called modifier cues (mainly adverbs and some adjectives, e.g. cue2 in Example 2.). The identification of predicative cues (and their scope) leads to split the text into different textual segments and then the identification of modifier cues influence the validation context of the textual segment previously identified.

2. Paul **veut**^{cue1} *sûrement*^{cue2} que [Mary vienne.]$_{scope}$ // Paul *certainly*^{cue1} **wants**^{cue2} [Mary to come]$_{scope}$.

The annotation task we present here consists in annotating these *predicative cues* (that lead to modify the level of enunciative and/or modal commitment of a textual segment) and their *scope*. The scope of a predicative cue corresponds to the textual segment impacted by the variation in the level of enunciative and/or modal commitment. Table 1 presents the four classes of predicative cues that we consider and for each of them gives some examples of the syntactic components that can be under the scope of the cue.

Cues	Scope
Verbs	**Direct and/or indirect object**
Reporting verb, modal verbs	*Paul **promet**^{cue} qu'[il viendra]$_{scope}$ / Paul **promises**^{cue} that [he will come]$_{scope}$* *Paul **veut**^{cue}[venir]$_{scope}$ / Paul **wants**^{cue} [to come]$_{scope}$*
Nouns	**Noun complements, relative clause**
Predicative nouns	*C'est **son souhait**^{cue} [d'être impliqué]$_{scope}$ / It is **his wish**^{cue} [to be involved]$_{scope}$*
Morphological	**All the verb complements**
Future, conditional	*John **viendra**^{cue} [plus tard]$_{scope}$ / John **will**^{cue} [come later]$_{scope}$*
Syntactic	**Main clause**
Subordinate clauses of condition	*[Mary refuse de donner son approbation]$_{scope}$ **à moins que Paul accepte**^{cue} / [Mary refuses to give her approval]$_{scope}$ **unless Paul accepts**^{cue}*
Prepositional construction	***D'apres Paul**^{cue}, [Mary va venir]$_{scope}$ / **According to Paul**^{cue}, [Mary is coming]$_{scope}$*

Table 1: Cues and associated scopes

As can be seen, depending on the type of predicative cue, the syntactic dependents we consider in the scope vary. This description of what we consider as a predicative cue and how to delimit the corresponding scope is reported in the first version of an annotation guideline. In order to refine our descriptions and measure their relevance on the corpus, the following section presents the inter-annotator agreement between two expert annotators and the first results of the automatic system for the same annotation task. This evaluation process should lead to the production of a more precise guideline that can reveal fine discursive shades and also stimulate reflection on how best to deal with syntactic and semantic information in the automatic annotation system.

4 Annotation and evaluation process

Our final goal is to develop an automatic annotation system that produces the annotation of enunciative and modal cues and their scope in newswire texts. In this light, we have to build a guideline of our annotation aim and a reference corpus that can be used to evaluate the system.
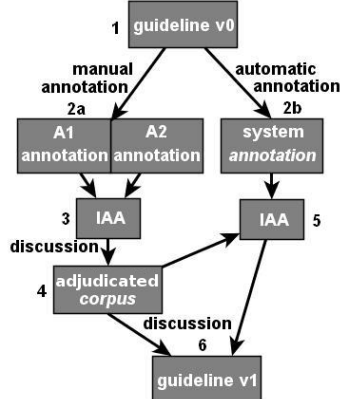

Figure 1.Workflow of guideline improvement

Figure 1 illustrates the steps in the workflow applied to improve our annotation guideline. For this purpose, two annotators (henceforth A1 and A2), both of them experts in linguistics, worked together to build a guideline and then the reference corpus[1]. First of all, the two annotators defined the annotation goals together (see step 1 in Figure 1). Then they annotated separately a corpus of 20 newswire texts (see step 2a in Figure 3). This corpus contains 256 predicative cues and their associated scopes (see Table 2).

# Sent	Total	Verbs	Nouns	Morpho	Syntactic
199	256	210	4	11	31

Table 2: Corpus statistics

---

[1] Our annotation process is based on Morante and Daelemans (2012).

This manual annotation task was carried out using the Glozz Annotation Tool (Widlöcher and Mathet, 2012) that relies on the URS (Unit-Relation-Schema) meta model and produces an xml output. The model permits to annotate textual units that can be embedded or not (in our case the predicative cues and their scope) and relations (for us, the opening relation links the predicative cue to its scope).

After this first annotation round, inter-annotator agreement was calculated (see table 3). The results show that the agreement between the two annotators is high for the cues but not very good for the scopes. By comparing the two sets of annotations in detail, we observed in our corpus that some textual segments can be either included or excluded from the scope depending on the annotator's interpretation. Example (3) shows the scope annotation proposed by annotator A1. As we can see, the textual segment *qui a débuté lundi* is included in the scope by this annotator but it is excluded in the annotation proposed by A2. In this particular case, we consider that both interpretations are acceptable since we cannot say for sure if this segment is presented from the viewpoint of the journalist or from the viewpoint of the source *un de ses avocats*. The same phenomenon is often observed with temporal adverbials that cannot be interpreted unambiguously as being a part of the scope or not. In these two kinds of cases the annotator needs to use the context and his linguistic background to decide. This raises the issue - already mentioned in Farkas et al. (2010) – as to whether it is advisable to set a strict boundary for the scope.

We propose to address this issue by evaluating the scope annotation both strictly and more flexibly. In the flexible interpretation we distinguish the segments that are detected with an exact match boundary from those that are detected with different boundaries but that are still correct in the interpretation (as in example 3).

3. [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.

To measure the distinction of using strict or flexible boundaries for scope, we propose to distinguish the scope evaluation (for strict scope boundaries) from the weighted scope evaluation (for flexible boundaries).

Flexible boundaries are calculated with a 0.5 factor as follows:

$$Weighted\ Precision = \frac{SB + 0.5 \times FB}{Ref}$$

$$Weighted\ Recall = \frac{SB + 0.5 \times FB}{Rel}$$

- *SB (strict boundaries)*: the number of entities with a strict scope boundary
- *FB (flexible boundaries)*: the number of entities with a flexible scope boundary
- *Ref*: the number of reference entities (*i.e.* ideally identified)
- *Rel*: the number of relevant entities (*i.e.* correctly identified)

The distinction between the evaluation of scope and weighted scope revealed that in a significant number of cases (in this experimentation about 10 %) the two annotators disagreed in their annotation but that both interpretations were correct. This observation helped us to rethink our annotation goals and based on the result of inter-annotator agreement, the two annotators produced a common adjudicated version of their annotation[2] (step 4 in Figure 1). This new annotated version is the result of a reflection on the two annotators' disagreements and considers the context to delimit scope boundaries.

Adjudicated /System	precision	recall	F1
Cues	0.85	0.86	0.86
Scopes	0.79	0.72	0.76
Weighted Scopes	0.84	0.77	0.80

SB	FB	Rel	Ref
185	22	256	234

Table 3: IAA: the annotations of annotator A1 are evaluated against the annotations of annotator A2

Adjudicated /System	precision	recall	F1
Cues	0.83	0.85	0.84
Scopes	0.52	0.59	0.55
Weighted Scopes	0.67	0.76	0.71

SB	FB	Rel	Ref
59	33	100	113

Table 4: System evaluation: annotations from the system are evaluated against the adjudicated version

In a second step, we evaluated the first annotation version of our automatic system (step 2b in Figure 1) on a subset of the corpus against the annotation of the adjudicated version (see table 4). The subset corpus contains 100 cues and their associated scopes. Our automatic annotation system is based on the analysis dependency syntactic parser combined with scope detection rules (see Battistelli and Damiani, 2013). The results

of this evaluation show that the detection of cues is good, as with the manual annotation, while the scope detection is not as good. This can be explained partly by the fact that the syntactic parser analysis produces some analysis errors (tagging or parsing errors, wrong syntactic attachment especially with coordinating conjunctions). Moreover, this evaluation shows that with an automatic system, distinguishing strict and flexible boundaries can highlight the results in another way. Indeed, if we look at the scope evaluation, the F-measure is not really satisfactory. If we take into account only this measure, it could be concluded that our system is not efficient. However, with the measure of weighted scope we see that while in many cases the scope did not match exactly with the reference corpus, it was not wrong either. This phenomenon of scope boundaries that are not easily decidable represents 10% of disagreement in the IAA (ie 22 cases) and 30% in the system evaluation (ie 33 cases), and has to be taken into account to improve the guideline and the annotation system. This first annotation experiment on a small corpus helped us to define new annotation goals that must be integrated both in the new version of the guideline (step 6 in Figure 1) and in the automatic annotation system.

5   Conclusion

In this paper, we have focused on a methodology to produce a reference corpus proposing the annotation of enunciative and modal commitment information as a discursive delimitation task. The annotation scheme we propose is based on the detection of predicative cues and their scopes. The results of the evaluation presented here show that the most challenging task is not to find the predicative cues but to delimit their scopes and beyond this delimitation question to define how to assess whether a scope is correct or not. Next step of our work is to launch a larger annotation campaign involving more human annotators and a bigger corpus. In this second step, our model will integrate modifier cues such as hedging adverbs that modify the semantic value of the textual segments that have been first delimited and introduce discursive cues that can impact more than a single sentence At last, in order to make our work available for the community our guideline and reference corpus will soon be available on *Chronolines* project website[3].

---

[2] This adjudicated version is available for consultation: http://vmoaxc.1fichier.com/predicative_cue_scope.zip

[3] http://www.chronolines.fr/

References

Bally, C. 1932. *Linguistique générale et Linguistique française*. Paris : Leroux, 2éd. (1944), Berne.

Battistelli, D. and Damiani, M. 2013. *Analyzing modal and enunciative discursive heterogeneity: how to combine semantic resources and a syntactic parser analysis*. In IWCS 2013 Workshop: WAMM, Potsdam.

Benveniste, E. 1966. *Problèmes de linguistique générale, 1*, Paris : Gallimard.

Charolles, M., Le Draoulec, A., Péry-Woodley, M. P. and Sarda, L. 2005. *Temporal and spatial dimensions of discourse organisation*. Journal of French Language Studies, 15(2), 115.

Culioli, A. 1973. *Sur quelques contradictions en linguistique. Communications*, 20(1), 83-91.

De Marneffe, M. C., Manning, C. D. and Potts, C. 2012. *Did it happen? The pragmatic complexity of veridicality assessment*. Computational Linguistics 38(2):301-333.

Farkas R., Vincze V., Móra G, Csirik J. and Szarvas G. 2010. *The CoNLL 2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text.* In Proceedings of the 2010 Conference on Computational Natural Language Learning.

Kilocoglu, H.H. 2012. *Embedding predications*. PhD Dissertation, Concordia University, Montreal.

Morante, R. and Daelemans, W. 2012. *ConanDoyle-neg: Annotation of negation in Conan Doyle stories*. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC).

Sauri, R. and Pustejovsky, J. 2012. *Are You Sure That This Happened? Assessing the Factuality Degree of Events in Text*. Computational Linguistics, 38(2):261-299.

Widlöcher A. and Mathet Y. 2012. *The Glozz Platform: A Corpus Annotation and Mining Tool*. In Proceedings of the 2012 ACM symposium on Document engineering, 171-180.

Wilson, T. and Wiebe, J. 2005. *Annotating Attributions and Private States*. In ACL 2005 Workshop: Frontiers in Corpus Annotation II: Pie in the Sky, 53-60.