Analyzing modal and enunciative discursive heterogeneity: how to combine semantic resources and a syntactic parser analysis

Delphine Battistelli
Université Paris Sorbonne,
STIH, EA 4509
Paris, F-75006, France
`delphine.battistelli@paris-sorbonne.fr`

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

Abstract

This paper introduces our methodology for annotating variations in enunciative and modal commitment in a text. We first present the theoretical background of the study which puts the emphasis on the close interaction between time, aspect, modality and evidentiality (TAME) categories (and also markers). We then present our semantic resources which encompass not only lexical items, but also morphological inflections and syntactic constructions. We finally describe the first step of our global natural language processing (NLP) workflow which uses a syntactic analysis parser.

1 Introduction

Our paper concerns the design and development of lexical and grammatical resources for French in order to annotated textual segments in texts with regard to their modal and enunciative characteristics. The present study is part of a broader project (named *ChronoLines*[1]) which deals with the generation of innovative interfaces to display information according to temporal criteria from newswire texts in French provided by the Agence France Presse (AFP). The main goal of this project is closely related to applications such as the construction of timelines from texts (e.g. Alonso and al. 2009); its originality, however, compared to current timelines, is that we aim to take explicitly into account the problem of different levels of temporal referencing. These levels are associated with the different strategies to manage enunciative and modal expression which can be identified within the texts. Along similar lines to the annotation of "attributions" and "private states" (Wilson and Wiebe, 2005) or the calculation of "factuality degrees of events" (Sauri and Pustejovsky, 2012), we aim to take into account the fact that, independently of their calendar anchoring, situations can be presented as certain, fully accomplished, 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. ... It should be noted that around 90% of newswire text sentences contain at least one clue of an epistemic modal and/or enunciative shift.

In section 2, we present the theoretical background of the study. Section 3 details our methodology for detecting variations in enunciative and modal commitment in a text (via the detection of certain lexical and grammatical cues, including certain kinds of syntactic constructions) and then identifying embedded textual segments which correspond to these shifts/variations (thanks to a syntactic parser analysis). Section 4 describes the first step of our global natural language processing (NLP) methodology.

2 Theoretical background: *modality meaning* and TAME categories

The notion of modality, which is closely linked to the notion of evidentiality[2], has been studied from

---

[2] This notion is implied when a new source of information is introduced in the discourse (reported, hearsay, inferred,…).

many different perspectives: logic, philosophy and linguistics - see for example (Bybee and al., 1994; Palmer, 2001; Nuyts, 2006). In the field of linguistics, modality can be considered from an enunciative perspective - see (Bally, 1935; 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, discourse operations and operations of commitment3, the expression of which leaves a certain number of surface linguistic traces (or clues). The enunciator's degree of commitment to a predicative content introduces variations in aspect, time and modality and also what is called enunciative modalities, all marked in the utterance by traces that the enunciator leaves in the discourse. In the case of direct speech, these traces can manifest themselves in the introductory portion of the direct reported speech in different forms: they may indicate the enunciator's stance towards what is reported, describe the speaker's attitude towards what is being said (in general) or towards what the speaker himself is saying; or refer to the relationship between the speaker and the enunciator, etc. The linguistic act of reported speech permits the enunciator to indicate his/her commitment to what is said or written by another source. Thus, what is emphasized in an enunciative perspective is the intrinsic presence of the enunciator in every discourse production. This presence is mainly manifested via time, aspect, modal and/or evidential traces. These four kinds of linguistic traces refer to four kinds of intrinsically linked semantic categories (see the well-known acronym TAME). When only the categories of modality and evidentiality are considered, it can be said that the enunciator expresses different degrees of commitment to the truth of the propositional content.

At an analytical and descriptive level, which is essential in annotation tasks, and which is, moreover, sensitive to a specific language, several questions arise:

- what kinds of linguistic markers (even if they are most of the time ambiguous) can be considered as prototypical for each of the TAME categories (let's see for example "modal verbs" (*as must*, …), "hedging adverbs" (*as probably*), "reported speech verbs" (*as to say or to tell*), …)?

- how can markers that encode semantic instructions belonging to several TAME categories be handled? For example, certain tenses can play a role at all TAME levels - see for example (de Haan 1999, Hassler, 2010)

- how can the four (cognitivo-linguistic) operations that TAME categories refer to be ordered - see (Dik, 1997)?

- how can a textual perspective of TAME categories be taken into account, that is to say how can these categories be explored beyond the syntactic level of utterances? This problem has been very rarely addressed in theoretical linguistics and in NLP approaches – see for example (Battistelli, 2011)

- how should one deal with the problem of the interaction of different epistemic modality markers, at a sentence level – see (Sauri et Pustejovsky, 2012) - and at a textual level? Note that this question directly leads to the question of possibly different levels of predication. Let's take the case of sentence level. One sentence frequently expresses more than one situation (predicative content) and each of the situations can be qualified with a different degree of certainty;

- how can one deal with the problem, closely linked to the previous one, of the scope of modality markers, and therefore of the length of "modalised" textual segments?

In order to analyze the highly complex interaction between the categories of modality and evidentiality, our methodology first focuses on detecting these "modalised" textual segments and on analyzing their hierarchical organization. This is what we choose to present in this paper.

From our enunciative perspective, this leads us to take a look at the text in its enunciative and/or

modal variations. Indeed, throughout a text (even within a single sentence), enunciative and modal values tend to vary. For example, encountering a reporting verb such as *dire/to say*, *répondre/to answer*, *annonce/to announce*, introduces a variation in the enunciative value but no variation in the modal value; an adverb such as sûrement/ probably, *probablement/likely*, introduces a variation in the modal value; a verb such as *prétendre/to claim*, *croire/to believe*, *imaginer/to imagine*, introduces a variation in both the modal and the enunciative value.

In the next section, we detail the methodology developed to analyze these variations.

3 Methodology for analyzing discursive heterogeneity by detecting E_M segments

Any discourse (and by extension even a discourse comprising a single sentence) is necessarily presented from the viewpoint of a human cognizer (in our case, the journalist who writes the newswire). Thus, any discourse always has a default source who is its author. Moreover, as explained in the previous section, we consider modality from an enunciative perspective. Any newswire can then be considered as a textual segment having "default" enunciative (= 'author') and modal (= 'true', based on the Gricean maxim of quality which testifies that sources are trustworthy) values. Most of the time, in a text (even within a single sentence), textual segments which have different enunciative and modal values can be identified. This is due to the complex mechanisms of interaction between semantic clues that we outlined in section 2.

Our methodology consists in focusing on the tracking on semantic clues which have to be taken into account in order to identify different enunciative and modal textual segments. This kind of approach to modal meaning focuses on discursive heterogeneity and also makes it possible to deal with the interaction between Modality and other related linguistic categories (Time, Aspect and Evidentiality). Furthermore, we believe that developing such an approach could - beyond the applicative interest in information retrieval (cf. *ChronoLines* already mentioned) - help to achieve a better understanding of this complex TAME

interaction at a discursive level, i.e. at a level where pertinent textual units are not only sentences or utterances. In the remainder of the article, we will give just examples of sentences, but we wish to emphasize that our methodology also enables a discourse level analysis.

We propose to set up a NLP workflow (see Fig. 1) that automates the annotation process of the text into textual segments that have enunciative and modal features. These textual segments will be called from now on E_M segments. In this workflow, we distinguish two main steps:

Step 1: concerns the E_M splitter module which produces annotation of a text as a succession or imbrication of E_M segments. It uses semantic clues (organized in E_M semantic resources database) that lead to the opening of new E_M segments; it also uses a syntactic parser to calculate the length of an E_M segment at the sentence level;

Step 1bis: aims at linking different sentences in a single E_M segment if they denote a homogeneous discourse unit;

Step 2: the E_M value assigner determines the values of E_M segments. Semantic clues are again used insofar as some of them have an influence only on the enunciative value of an E_M segment, or only on its modal value, or on both its enunciative and modal value.
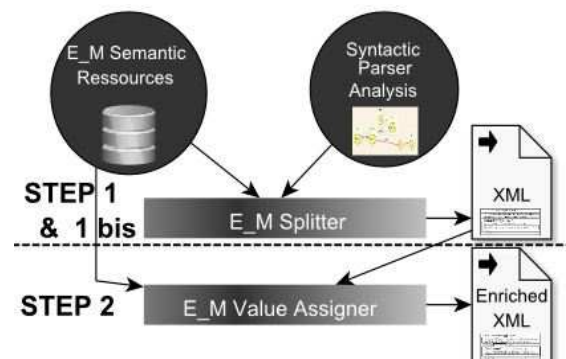

Figure 1. NLP workflow for analyzing enunciative and modal discursive heterogeneity

Figure 2 shows as embedded boxes the E_M segments (from 'E_M_1_dft' to 'E_M_4') that we want the final system to produce for the sentence in example 1. The involved semantic clues (from c=1 to 3) are highlighted.

1. *La Syrie a nié jeudi avoir fourni au Hezbollah libanais des missiles Scud capables d'atteindre l'ensemble du territoire israélien, accusant l'Etat*

hébreu de vouloir, avec de telles accusations, faire monter la tension au Proche-Orient. (*Syria denied on Thursday having supplied the Lebanese Hezbollah with Scud missiles capable of reaching the whole territory of Israel, accusing the Hebrew State of seeking, through such charges, to heighten tension in the Middle East.*)

'E_M_1_dft' is the "default" segment. Over the text, every occurrence of a linguistic clue "opens" a new E_M segment. We will see later how we identify the length of each E_M segment depending on the different types of linguistic clues. For now, we can say that clue 1 *a nié/denied* opens segment 'E_M_2' and that clue 2 "accuser/to accuse" opens segment 'E_M_3'. Finally, clue3 *vouloir/seeking* opens segment 'E_M_4'. As can be seen, an E_M segment can follow another E_M segment (for example 'E_M_3' follows 'E_M_2') or be embedded in another E_M segment (for example 'E_M_4' is embedded in 'E_M _3').
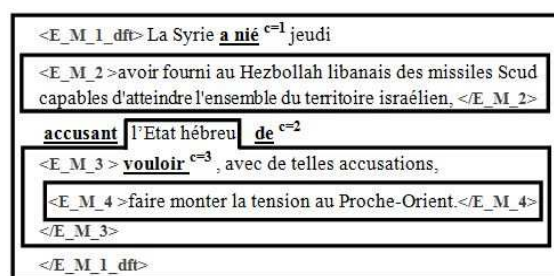

Figure 2. Example of an E_M splitter output[4]

In the next section, we detail Step1 and then focus on the building of semantic resources and the relevance of using a syntactic parser to identify E_M segments. Step1bis and Step2 will be dealt with in other papers.

4 Using semantic resources and a syntactic parser analysis to detect E_M segments

The identification of an E_M segment in the text is, as we have just seen, founded on the tracking of semantic clues that open a new segment. In other words, every variation in the enunciative and/or modal meaning introduces a new E_M segment.

Section 4.1 describes the organization of semantic resources and illustrates them with some examples. Then in section 4.2 we explain how we use the syntactic analysis produced by the parser.

4.1 Semantic resources

Semantic clues that are able to open a new E_M segment can be (see Fig. 3): lexical items (verbs of propositional attitude, speech verbs, nouns, adverbs, etc.), morphological inflections (for example tense inflections like in French *conditionnel* and *imparfait* tenses[5]), or specific syntactic constructions (subordinate clauses of condition, prepositional constructions). All kinds of them are used during Step1 in order to detect and split E_M segments. Note that these resources are also organized in the database at a deeper level in order to be used during Step2 to calculate the precise E_M segments values.
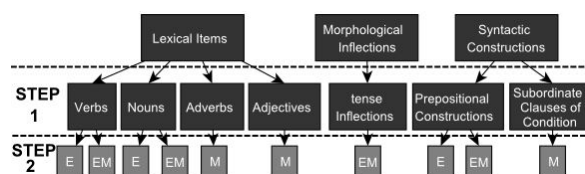

Figure 3. Organization of E_M semantic resources: at a surface level and at a deeper level

Lexical items

Lexical clues can be split into two groups: predicative clues and modifier clues.
- *Predicative clues: verbs and nouns*

A predicative clue introduces a new E_M segment which includes all the syntactic dependents of the predicate. These two categories of predicative clues involve a semantic variation: at the enunciative level (see example (2), where the clue *dire/says* introduces a new enunciative source for the E_M segment *Jean viendra/Jean will come*); or at both the enunciative and modal levels (see example (3) where the clue *pense/thinks* introduces both a new source and a modal variation). Example (4) illustrates the case of a sequence of several clues: clue1 *a exprimé/expresses* opens segment "E_M1" and influences only the modal validation context. Inside this segment, clue2 *le souhait/a*

---

[4] Note that the purpose of this representation as interlocking boxes is to illustrate our methodology; the actual annotation is in XML.

[5] See respectively (Kronning, 2002) and (De Mulder and Vetters, 2002).

desire impacts on both the enunciative and the modal validation context of segment "E_M2".

2. [Paul **dit**^clue que [Jean viendra]_E_M]_E_M_dft. ([Paul **says**^clue that [Jean will come] _E_M]_E_M_dft)

3. [Paul **pense**^clue que [Jean viendra]_E_M]_E_M_dft. ([Paul **thinks**^clue that [Jean will come] _E_M]_E_M_dft)

4. [M. Arabi **a exprimé**^clue1 [**le souhait**^clue2 [d'aider la Syrie à surmonter cette phase]_E_M]_E_M]_E_M_dft ([Mr. Arabi **expressed**^clue1 [**a desire**^clue2 [to help Syria overcome this difficult.]_E_M2]_E_M]_E_M_dft 1)

- *Modifier clues: adverbs and adjectives*

Unlike predicative clues, modifier clues do not open a new E_M segment, but will modify the value of the current E_M segment during Step2. Sentence (5) is an example of this phenomenon: the clue *peut-être/may* directly modifies the predicate *viendra/come*. The modal value of the whole segment is then impacted.

5. [Jean viendra **peut-être**^clue]_E_M_dft. ([Jean **may** ^clue come]_E_M_dft.)

An adverb can also modify a verbal predicate which itself introduces an E_M segment. In this case, the adverb will not modify the value of the current E_M segment but the value of the E_M segment opened by the verbal predicate. Example (6) illustrates this case: clue2 *apparemment/apparently* impacts on clue1 *dit/says*. The value of the E_M segment will change at the enunciative level because of clue1, and at the modal level because of clue2.

6. [Paul **dit**^clue1 **apparemment**^clue2 que [Jean viendra]_E_M]_E_M_dft . ([Paul **says**^clue 1 **apparently**^clue2 that [Jean will come]_E_M]_E_M_dft)

- *Building lexical resources*

The development of these lexical resources is based on the most frequent lexical items that occur in press AFP newswires. To find these most frequent items, we used a corpus of 20,000 texts (from the years 2010 and 2011). We then generated the frequency of the lemmatized corpus to find the most frequent lexical items. Our goal is to build resources that can cover at least 80% of the corpus. This coverage is reached with the 320 most frequent verbs. From this list of 320 verbs, we manually sorted 140 verbs which introduce a

modal and/or an enunciative variation. We used the same technique to build the list of 15 adverbs and the list of 10 adjectives. For predicative nouns, we had to use another method because the number of nouns was too high and we also wanted to keep only predicative nouns. To solve this issue, we used Verbaction [6] which is a lexicon of nouns morphologically related to verbs. For each of our 140 verbs, we searched whether a related noun existed in Verbaction (this task has been automated). We obtained a list of nouns which can potentially be clues but the results then had to be manually filtered as some items were irrelevant for our work. For example Verbaction gives two predicative nouns coming from the verb *accepter/to accept*: *acceptation/acceptance* and *acception/acceptation*, and we decided to keep only the first noun (*acceptation/acceptance*). After this filtering process, the list of predicative nouns comprised 80 items. Let's mention we are currently still exploring additional resources, in particular by integrating syntactic and semantic information about verbs from (Hadouche and Lapalme, 2010).

Morphological inflections

Several morphological inflections in French form another type of clue implied in the opening of a new E_M segment; for example, the French conditional tense, which has the morphological inflection (*-rais, -rais,-rait,-rions,-riez,-raient*). This tense is prototypical of French journalistic practice. It introduces uncertainty about the source of the information and/or about the epistemic modal status of the information.

In example (7), the verb form *aurait annoncé/would have announced*[7] is composed of two clues. The clue *annoncé/announced* introduces a new enunciative source 'Jean' (distinct from the journalist source). The other clue is the morphological inflection "-rait". This morpheme is interpreted as a trace of the journalist's lack of commitment: the hearer/reader interprets this trace as an ambiguity marker concerning the true source of information (is it 'Jean' or someone else who

---

[6] http://redac.univtlse2.fr/lexicons/verbaction.html
[7] The verb form 'would have + verb' is the literal translation of the French pattern. A correct translation would be *Jean is said to have announced*.

announced…?) and/or the modal status of the E_M segment *le départ de Paul/Paul's departure*.

The ambiguity about the origin of the uncertainty (which can be enunciative and/or modal) implied by the use of the French *conditionnel* tense comes across particularly clearly when we look at the possible translations of the sentence in English. The translation we have chosen (*Jean is said to have announced Paul's departure*) shows that the ambiguity concerns the source and not the epistemic value of *le départ de Paul/Paul's departure*. However, this interpretation really depends on the context and it is often difficult in French to decide which of the two interpretations is intended. We will therefore simply consider that the French *conditionnel* tense leads to the opening of a new E_M segment, with indetermination concerning the origin of the journalist's lack of commitment.

7. [Jean **aurait**^{clue1} **annoncé**^{clue2} [le départ de Paul.]$_{E\_M}$]$_{E\_M\_dft}$ ([Jean **is said**^{clue1} **to have announced**^{clue2} [Paul's departure.]$_{E\_M}$]$_{E\_M\_dft}$)

Syntactic constructions

The third class of clues contains syntactic constructions which are able to open E_M segments.

- *Subordinate clauses of condition*

A subordinate clause of condition indicates in which conditions the propositional content of the main clause is realized. The main clause is therefore considered as an E_M segment and the subordinate clause as a semantic (modal) clue. In example (8), the subordinate clause *si Paul n'accepte pas/if Paul does not accept* acts as a clue opening the E_M segment associated to the main clause. Example (9) shows a similar case with a subordinate clause introduced by the conjunction *à condition de/on condition that*.

8. [[Marie a refusé de donner son accord]$_{E\_M}$ **si Paul n'accepte_pas**^{clue1}]$_{E\_M\_dft}$·([[Mary refused to give her approval]$_{E\_M}$ **if Paul does not accept**^{clue1}]$_{E\_M\_dft}$)

9. [[Marie acceptera]$_{E\_M}$ **à condition que Paul vienne**^{clue1}]$_{E\_M\_dft}$ ([[Marie will accept]$_{E\_M}$ **on condition that Paul comes** ^{clue1}]$_{E\_M\_dft}$

- *The prepositional constructions*

A prepositional construction such as selon X /according to X can also be considered as a semantic (enunciative) clue. Placed or not at the beginning of a sentence, this kind of expression introduces a new source and thus opens a new E_M segment, as illustrated in example (10). Another propositional construction such as à première vue/at the first sight can be considered as a semantic (enunciative and/or modal) clue.

10. [**Selon Paul**^{clue1}, [Jules vient]$_{E\_M}$]$_{E\_M\_dft}$ ([**According to Paul**^{clue1}, [Jules is coming]$_{E\_M}$)]$_{E\_M\_dft}$

11. [**A première vue**^{clue1}, [Marie a raison]$_{E\_M}$]$_{E\_M\_dft}$ ([**At first sight**, [Marie is right]$_{E\_M}$]$_{E\_M\_dft}$)

4.2 Using syntactic parser analysis to detect E_M segments boundaries

A semantic (modal and/or enunciative) clue is linked to an E_M segment: either the semantic clue opens a new E_M segment or it modifies the (modal and/or enunciative) value of the current E_M segment. In this section, we address the issue of defining the boundaries of E_M segments, using a syntactic parser. We use a large-coverage syntactic parser for French, FRMG (FRench MetaGrammar) (De La Clergerie et al., 2009). The main syntactic contexts in which semantic clues can occur are as follows:

- clausal complements (*Paul **dit** que.../Paul **says** that…*)

- adverbial clause modifiers (*Paul viendra **s'il ne pleut pas**/Paul will come **if it is not raining***)

- constructions with subject inversion (*"Marie va venir", **a dit Paul**/"Marie will come", **said Paul***)

- direct and indirect objects (*Paul **a demandé** l'intervention de la police/Paul **asked** the police to intervene*)

- verb modifiers (*Il est venu **mardi**/He came on **Tuesday***).

- noun complements (*Ceci est **le souhait de ces pays d'être impliqués**/This is **the wish of these countries to be involved***)

- relative clauses (*Il a exprimé **l'espoir que la guerre finisse**/He expressed **the hope that the war would end***)

From the observation of all these kinds of contexts, we have developed the general rule RULE_G1 which makes it possible to delimit an E_M segment: if a new E_M segment has been opened due to the presence of one predicative clue, then all the complements of this predicative clue except its modifiers 8 (e.g. temporal modifiers, purpose clause modifiers, etc.) are part of this new E_M segment. Furthermore, since any text (taken as a whole) is considered as an E_M segment having "default" enunciative and modal features, the text is always associated to E_M_1_dft and this constitutes the second general rule RULE_G2.

To illustrate the application of these two general rules, let's return to the example of sentence (1) given above. Figure 2 illustrates both the syntactic relations produced by the parser (shown by arrows) and the semantic clues implied in the analysis (shown in dotted lines). The dotted boxes delimit the E_M segments we want the system to detect.

In the case of figure 2, the text comprises a single sentence, and is thus associated to E_M_1_dft (RULE_G2). The semantic clue1 *a nié/denied* marks the opening of a new segment named E_M_2. The length of this segment depends on syntactic information coming from the parser. According to RULE_G1, this segment is composed of the clausal complement *avoir fourni au Hezbollah libanais des missiles Scud capables d'atteindre l'ensemble du territoire israélien / having supplied the Lebanese Hezbollah with Scud missiles capable of reaching the whole territory of Israel*, and does not include the temporal modifier *jeudi/Thursday* nor the modifier clause *accusant l'Etat hébreu (…)/accusing the Hebrew State of (…)*. Those two components remain in the segment E_M_1_dft.

Clue2 *accusant de/accusing of* marks the opening of a new segment E_M_3 which is composed of its complements (RULE_G1): the direct objet of the verb *l'Etat hébreu/the Hebrew state* and the clausal complement *vouloir avec de telles accusations faire monter la tension au proche orient/seeking*

through such charges to heighten tension in the Middle East. Clue 3 is the verb *vouloir/seeking*. This clue marks the opening of a new segment E_M_4 which is, according to RULE_G1, composed of the direct object of the verb *vouloir/seek*, i.e. *faire monter la tension au Proche-Orient/to heighten tension in the Middle East*, but excludes the modifier *avec de telles accusations/through such charges*.

The E_M segment splitting system we have started to develop (named E_M splitter) takes into account the two general rules described above. We are currently working on E_M value assigner module (see Fig. 1) which is dedicated to assign values to the E_M segments. This module uses the deeper level of semantic resources organization (see Fig. 3), that is to say the distinction between intrinsic modal and/or enunciative meaning of clues. It uses also a heritage mechanism able to apprehend segment embedding (and thus the interaction between several clues).
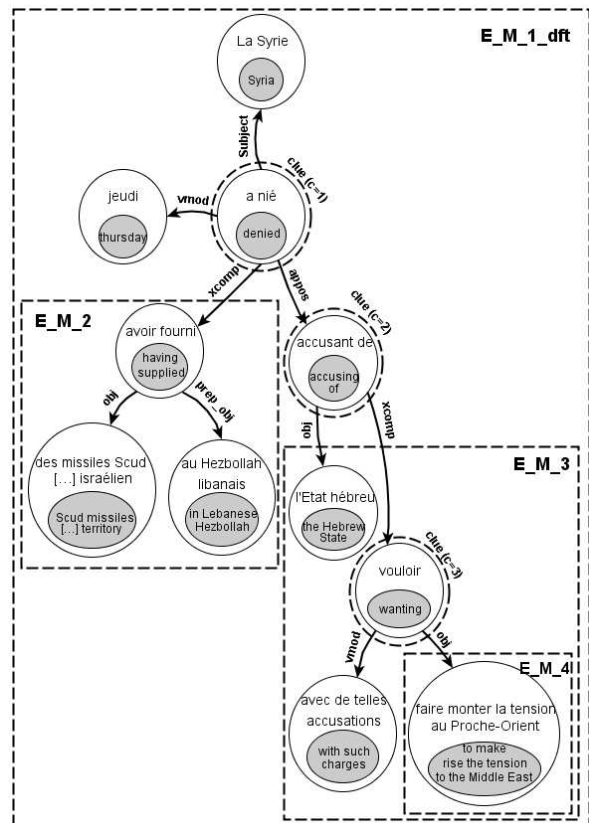

Figure 4. Using syntactic information to split E_M segments

---

8 We decided to consider that, by default, information coming from modifiers has to be allocated to the writer and thus constitutes a form of textual background.

5   Conclusion

In this paper, we have focused on the different kinds of French resources that need to be involved in the annotation of textual segments according to - what is called in our enunciative theoretical background terminology - their enunciative and modal characteristics (hence the notion, and notation, of E_M segment). These two types of characteristics are involved both:

- in an information retrieval purpose (when dealing with - or calculating - the factual status of situations described in a text, as for example in certain textual entailment applications, or when dealing with attributions and polarity, as for example in opinion or sentiment analysis applications);

- and in a theoretical semantic purpose (when addressing the difficult matter of the shifting boundaries – either lexical or grammatical, and variously marked in different languages - between the notions of evidentiality and epistemic modality).

Thus, in the methodology that we have proposed to explore modal meaning and its annotation, the precise E_M characteristics are deliberately not addressed in the first step, but only in the second step (cf. Step 2 in Fig. 1), when the task of delimiting textual segments has already been approached (cf. Step 1 in Fig. 1). As we have seen, the present paper focuses on the problem of splitting a text into textual segments, and it presents how we envisage combining E_M semantic resources and a dependency syntactic analysis parser results to achieve segmentation.

We would like to point out that the way we build these semantic resources is quite original especially in an NLP perspective: firstly, because these resources encompass not only lexical items, but also morphological inflections and syntactic constructions (see section 4.1); secondly, because they aim to reflect some important theoretical investigations about the close interaction between TAME categories.

References

Alonso, O., Gertz, M. & Baeza-Yates, R. 2009. Clustering and exploring search results using timeline constructions. In Proceedings of the 18th ACM conference on Information and knowledge management (pp. 97-106).

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

Battistelli, D. 2011. Linguistique et recherche d'information. Hermès Publishing. Paris

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

Bybee, J. L., Perkins, R., & Pagliuca, W. 1994. The evolution of grammar: tense, aspect and modality in the language of the world. University of Chicago Press.

De La Clergerie E., Sagot B., Nicolas L. et Guénot M. 2009. FRMG: évolutions d'un analyseur syntaxique TAG du français. In Journée ATALA "Quels analyseurs syntaxiques pour le français ?", Paris, France

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

De Brabanter, P., & Dendale, P. 2008. Commitment: The term and the notions. Belgian Journal of Linguistics , 22, 1-14.

De Haan, F. 1999. Evidentiality and epistemic modality. Setting boundaries. Southwest Journal of Linguistics, 18, 83-101.

De Mulder, W., & Vetters, C. 2002. The French imparfait, determiners and grounding. In F. Brisard (Éd.), Grounding: The Epistemic Footing of Deixis and Reference (pp. 113-149). Berlin/NY: Mouton de Gruyter.

Dik, S. C. 1997. The Theory of Functional Grammar. Part 1: The structure of clause (éd. Second). Berlin: Mouton de Gruyter.

Hadouche, F., & Lapalme, G. 2010. Une version électronique du LVF comparée avec d'autres ressources lexicales. Langages, 10(179-180), 193-220.

Hassler, G. 2010. Epistemic modality and evidentiality. Dans G. Diewald, & E. Smirnova (Éds.), Linguistic Realization of Evidentiality in European Languages (pp. 223-248). Berlin/NY: Mouton de Gruyter.

Nuyts, J. 2006. Modality: Overview and linguistic issues. In W. J. Frawley (Éd.), The Expression of Modality. Berlin: De Gruyter Mouton.

Palmer, F. 2001. Mood and Modality. Cambridge University Press.

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

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