HAL
archives-ouvertes.fr

Exploiting naive vs expert discourse annotations: an experiment using lexical cohesion to predict Elaboration / Entity-Elaboration confusions

Clémentine Adam, Marianne Vergez-Couret

▶ **To cite this version:**

HAL Id: hal-00986433

https://hal.archives-ouvertes.fr/hal-00986433

Submitted on 2 May 2014

Exploiting naive *vs* expert discourse annotations: an experiment using lexical cohesion to predict Elaboration / Entity-Elaboration confusions

Clémentine Adam
CLLE & University of Toulouse
`adam@univ-tlse2.fr`

Marianne Vergez-Couret
LPL, Aix-Marseille University
and Hong-Kong Polytechnic University
`marianne.vergez@gmail.com`

Abstract

This paper brings a contribution to the field of discourse annotation of corpora. Using ANN-ODIS, a french corpus annotated with discourse relations by naive and expert annotators, we focus on two of them, Elaboration and Entity-Elaboration. These two very frequent relations are (a) often confused by naive annotators (b) difficult to detect automatically as their signalling is poorly studied. We propose to use lexical cohesion to differentiate between them, and show that Elaboration is more cohesive than Entity-Elaboration. We then integrate lexical cohesion cues in a classification experiment, obtaining highly satisfying results.

1 Introduction

This paper brings a contribution to the field of corpus annotation at the discourse level. Discourse structure is based on coherence links existing between discourse units. These links can be captured using the notion of discourse relations (Mann and Thompson, 1987; Asher and Lascarides, 2003). Handling and detecting elements of discourse structure is very challenging for Natural Language Processing. Applications such as natural language generation (Bateman and Zock, 2005), automatic summarization (Marcu, 2000), among others, could take advantage of discourse level information detection. In the current state of research, providing reliably annotated corpora at the discourse level is really groundbreaking and opens new possibilities of investigation in discourse studies.

The ANNODIS project (Péry-Woodley et al., 2009; Afantenos et al., 2012) will provide the scientific community with access to such a corpus for French (see section 2). It is the first ressource in French annotated with discourse relations. Similar corpora have already been developped for English, including the Penn Discourse TreeBank (Prasad et al., 2007), the RST Tree Bank (Carlson et al., 2001) or the Discor corpus (Reese et al., 2007). But ANN-ODIS has distinct characteristics. For our concern, the main difference is the two-level annotation: first a pre-annotation done by naive annotators (called "naive annotation") and then a revised annotation done by expert annotators (called "expert annotation"). This allows investigation on the whole process of annotation.

In this paper, we focus on Elaboration and Entity-Elaboration, the two most frequent and frequently confused relations (see section 3). We propose a new approach based on lexical cohesion cues to differentiate between these relations, and show its reliability using expert annotation (see section 4). We integrate this approach in a machine learning experiment and highlight the improvement it brings (see section 5). We show how the obtained classifier can be used to automatically improve naive annotation or to reduce the experts' workload.

2 The ANNODIS corpus

The ANNODIS corpus[1], enriched with annotations at the discourse level, considers two different approaches of discourse organization, a top-down ap-

---

[1] `http://w3.erss.univ-tlse2.fr/annodis`

proach[2] and a bottom-up approach. Here, we focus only on the bottom-up approach which aims at constructing the structure of discourse, starting from elementary discourse units (EDUs) and recursively building more complex discourse units (CDUs) via discourse relations (EDUs contain at least one eventuality, most of the time only one). At the end, a hierarchical structure is defined for the whole text. This part of the corpus is composed of newspaper articles (from *Est Républicain*) and extracts from Wikipedia articles.

The specifications in the annotation manual were adapted from the SDRT model, a semantic approach of discourse structure (Asher and Lascarides, 2003), but were also inspired by other discourse models such as the RST framework (Mann and Thompson, 1987), the Linguistic Discourse Model (Polanyi, 1988), the graphbank model (Wolf and Gibson, 2005) etc. The discourse relations linking discourse units described in the manual are a set of relations[3] largely inspired from discourse relations common to most discourse theories mentioned above.

The ANNODIS corpus was annotated using a three-step process: it contains preliminary, naive and expert annotations of discourse relations. During the first preliminary phase, 50 documents were annotated by 2 postgraduate students in language sciences. The key purpose of this initial phase was to assist the drafting of the annotation manual which was used afterwards for the second main phase of annotation. 86 different texts were then doubly annotated with the help of the aforementioned manual by 3 postgraduate students in language sciences. The naive annotation was performed in order to discover cognitive principles of discourse organization (Afantenos et al., 2012). The double annotation allowed evaluating the inter-annotators agreement. The Kappa score (Cohen, 1960) on common attached discourse units for the full set of relations is 0.4, which indicates a moderate to weak agreement

and reveals the difficulty of the discourse annotation task. The expert annotation was performed as a third phase. 42 texts randomly selected from naive annotation were then reviewed and corrected by expert annotators. 44 texts from naive annotations remain to be reviewed and corrected.

This paper will focus on one of the frequent mistakes concerning two close relations: Elaboration and Entity-Elaboration (hereafter E-Elaboration) in the naive annotation and their correction in the expert annotation.

3 On Elaboration and Entity-Elaboration

The distinction between an elaboration of a state or an event (Elaboration) and an elaboration of an entity (E-Elaboration) is common in discourse theories. But the status of E-Elaboration as a discourse relation is not obvious and divides the scientific community. In the RST framework (Mann and Thompson, 1987), distinction points exist between Elaboration and E-Elaboration but both are regrouped in a single discourse relation. Knott (1996) considers discourse markers as the basis to motivate a set of coherence relations. Therefore Knott et al. (2001) reject E-Elaboration as a discourse relation for two reasons. The first is absence of obvious discourse markers. The second is that the E-Elaboration relation does not relate two propositions, as discourse relations usually do. Conversely, Fabricius-Hansen and Behrens (2001) introduce separate relations (called E[ventuality] Elaboration and I[ndividual] Elaboration). Prévot et al. (2009) note the need to introduce this relation in order to avoid confusions in annotation, arguing that keeping all the embedded segments in one discourse segment smudges the discourse contribution of the including segment. In ANNODIS, the choice was made to consider two different relations for annotation.

3.1 Elaboration and Entity-Elaboration in ANNODIS

For each relation, the annotation manual gives an informal description, several illustrations and additional information on the possible confusions between the described relation and other discourse relations. Here are the descriptions of Elaboration and

---

[2]The top-down approach focuses on the selective annotation of multi-level discourse structure such as enumerative structures (Ho-Dac et al., 2009).

[3]Alternation, Attribution, Background, Comment, Continuation, Contrast, Elaboration, Entity-Elaboration, Explanation, Flashback, Frame, Goal, Narration, Parallel, Result, Temporal Location. Fusion is also used when expert annotation disagreed with segmentation. Fusion(1,2) means that segments 1 and 2 are considered one segment.

E-Elaboration in the annotation manual of ANN-ODIS:

The Elaboration relation relates two propositions only if the second proposition describes a sub-state or sub-event of the state or event described in the first proposition. Elaboration also includes exemplification, reformulation and paraphrase cases.

The E-Elaboration relation relates two segments for which the second one specifies a property of one of the involved entities in the first segment. This property can be important (e.g. identificatory) or marginal.

Example (1) illustrates both relations. Each segment corresponding to one EDU is numbered. Segments sharing a same rhetorical role in the discourse must be joined into complex segments.

(1)　　[La Lausitz, [une région pauvre de l'est de l'Allemagne,]$_1$ [réputée pour ses mines de charbon à ciel ouvert,]$_2$ a été le théâtre d'une première mondiale, mardi 9 septembre.]$_3$ [Le groupe suédois Vattenfall a inauguré, dans la petite ville de Spremberg, une centrale électrique à charbon expérimentale]$_4$ [qui met en œuvre toute la chaîne des techniques de captage et de stockage du carbone]$_5$
[Lausitz, [a poor region in east Germany,]$_1$ [famous for its open air coal mines,]$_2$ was the scene of a world first, on Tuesday September 9th.]$_3$ [The swedish group Vattenfall inaugurated, in the small town of Spremberg, an experimental coal power plant]$_4$ [involving the complete carbon capture and storage chain.]$_5$

The expert annotation for this mini-discourse is given below :

E-Elaboration (3,[1-2])
Elaboration (3,4)
E-Elaboration (4,5)

Complex segment [1-2] is embedded in segment 3 and is given properties of the entity "La Lausitz". It is therefore attached to this segment by Entity-Elaboration. Segment 4 describes the event "to inaugurate a power plant" which is a reformulation of "to be the scene of a world first" and is attached to segment 3 with Elaboration. Finally, segment 5 gives a property of the entity "a power plant" in segment 4 and is attached to it via E-Elaboration.

The annotation manual also discusses possible confusions between Elaboration and E-Elaboration (and conversely). The discussion mostly highlights how the distinction between state and event could help to avoid confusion. It also reminds the reader of the major distinction between the two relations, e.g. Elaboration gives details on a state or an event while E-Elaboration gives details on an entity.

Despite these precautions, the naive annotators are often prone to error when confronted with these two relations.

3.2 Quantitative analysis in ANNODIS

Elaboration and E-Elaboration are the more frequent relations in the ANNODIS corpus, both in the naive annotation with 50% of the annotated relations and in the expert annotation with 35% of the annotated relations. The low inter-agreement for these relations in the naive annotation indicates that the relations are not well-understood. This hypothesis is reinforced by overestimation of annotated Elaboration and E-Elaboration: in 60% of the cases, an agreement between two naive annotators does not ensure that the annotation is correct (Vergez-Couret, 2010).

Note that when experts review and correct naive annotation, most of the corrections involve wrong annotations of Elaboration and E-Elaboration. Table 1 presents the expert annotation for each Elaborations and E-Elaborations annotated by the naives.

		Naive		
		Elab	E-Elab	Total
Expert	Elab	302	70	372
	E-Elab	158	216	374
	Total	460	286	746
Expert	Fusion	81	57	
	Continuation	70	32	
	Background	32	18	
	Other	150	59	

Table 1: Expert annotations for E-Elaborations and Elaborations in naive annotation

This table shows that confusions between Elaboration and E-Elaboration are the most important compared to confusions with other discourse relations. Elaboration is mistaken for E-Elaboration (hereafter Elaboration → E-Elaboration) and more importantly E-Elaboration is mistaken for Elabo-

ration (here after noted E-Elaboration → Elaboration). This paper only focuses on these two relations for methodological reasons: this choice allows first to give careful considerations to the linguistic features involved in the two relations (see section 3.3) and also to highlight and evaluate the improvements brought by using new kinds of linguistic cues (see section 4).

3.3 Linguistic features of Elaboration and Entity-Elaboration

Annotating Elaboration and E-Elaboration, manually or automatically, is very challenging since no prototypical marker exists for the two relations (Knott, 1996, among others). Some possible markers given in the ANNODIS manual (*à savoir, c'est-à-dire, notamment*, etc.) are not discriminatory for one of the two relations, and they are relatively rare.

One could think of other possible linguistic features of Elaboration and E-Elaboration. Prévot et al. (2009) underline possible linguistic realisations of E-Elaboration such as relative clauses and appositions (nominal and adjectival appositions, brackets...). Adam and Vergez-Couret (2010) point out that French gerund clauses may express serveral discourse relations including Elaboration but not E-Elaboration. Even if these syntactic features are not discriminatory with respect to all discourse relations (for instance gerund clauses and appositions may express Explanation or Background), we will see in section 4 if these syntactic features allow to distinguish Elaboration and E-Elaboration.

But more importantly, we would like to focus on one of the major distinctions between the two relations, e.g. Elaboration provides details on a state or an event while E-Elaboration provides detail on an entity, and how to highlight this distinction. The hypothesis we are testing is that this distinction results in differences concerning the lexical cohesion between the two segments. Cohesion includes all the links holding a text together as a whole, including reference, ellipsis and lexical cohesion. Lexical cohesion encompasses relations such as synonymy, hyperonymy, lexical similarity, etc. Our hypothesis is that Elaboration involves more lexical cohesion links since it relates two propositions and its interpretation involves information given by lexical semantics and world knowledge (Asher and Las-

carides, 2003). Adam and Vergez-Couret (2010) show that the use of lexical cohesion cues reliably detect gerund clauses which are Elaborations. In contrast, E-Elaboration only relates a proposition to an entity. In example (2), where Elaboration relates [17-19] to the target segment 16, it is indeed possible to highlight lexical cohesion links playing a role in Elaboration.

(2) [Un soir, il faisait un temps horrible,]$_{16}$ [les éclairs se croisaient,]$_{17}$ [le tonnerre grondait,]$_{18}$ [la pluie tombait à torrent.]$_{19}$
[One night, the weather was horrible,]$_{16}$ [flashes of lightning were crossing,]$_{17}$ [thunder growled,]$_{18}$ [rain fell heavily.]$_{19}$

In this case, cohesion lexical links between "temps" (weather) in 16 and "éclair" (flash of lightning), "tonnerre" (thunder) and "pluie" (rain) in [17-19] play a role in the interpretation of Elaboration.

On the other hand, E-Elaboration does not provide details about the whole proposition in the target segment, but provides details on an entity of this segment. Lexical cohesion links are not expected in this case.

(3) [Pourquoi a-t-on abattu Paul Mariani, [cinquante-cinq ans]$_4$, [attaché au cabinet de M. François Doubin,]$_5$?]$_6$
[Why was Paul Mariani, [fifty-five]$_4$, [personal assistant to M. François Doubin,]$_5$ gunned down?]$_6$

In example (3), the age and the profession of Paul Mariani is not lexically linked to the fact that he was gunned down.

In the next section, we discuss how to highlight lexical cohesion links in order to differenciate Elaboration and E-Elaboration.

4 Differentiating between Elaboration and Entity-Elaboration using lexical cohesion

4.1 Preamble

The interplay of lexical cohesion and discourse structure is an often studied but still not fully understood issue (Barzilay, 2008; Berzlánovich et al., 2008). Lexical cohesion cues are typically used in diverse automated approaches of discourse, but as these cues are used among others, their impact is not precisely evaluated. We aim at demonstrating

that lexical cohesion cues can be successfully applied to differentiation between Elaboration and E-Elaboration.

Adam and Morlane-Hondère (2009) propose to use a distributional semantic model (Baroni and Lenci, 2010) in order to detect lexical cohesion. Adam and Vergez-Couret (2010) use the lexical links identified by this method in a practical experiment of Elaboration detection. They show that the use of distributional neighbors in combination with an ambiguous marker of Elaboration (the gerund clause) very reliably detects some cases of Elaboration. This result confirms that Elaboration implies lexical cohesion, and that a distributional semantic model is a good lexical resource for identifying lexical cohesion links in texts.

As an extension to those studies, we want to use lexical cohesion cues to help differentiating between Elaboration and E-Elaboration. We first present how distributional neighbors can be used to estimate lexical cohesion between two text segments (section 4.2). Then, we compare the lexical cohesion of Elaboration and E-Elaboration and show that Elaboration is significatively more cohesive than E-Elaboration (section 4.3).

4.2 Methods: How to evaluate the strength of lexical cohesion between two segments

In order to evaluate the strength of lexical cohesion between two text segments S_a and S_b, we proceed in two steps. First, the two segments are annotated with part-of-speech and lemma information using the TreeTagger (Schmid, 1994). Then, all the lexical proximity links between the two segments are annotated. To detect these links, we use a lexical proximity measure based on the distributional analysis of the french Wikipedia (Bourigault, 2002). Internal links in a segment are not considered.

The number of lexical links N_ℓ can be directly interpreted as a cohesion cue. But this cue is skewed since this number is correlated to the segment's size (longer segments have more items to be linked). To reduce this skew, we built a score where the number of lexical links is normalized. Calling N_a the number of neighbours (linked or not) in the first segment (S_a) and N_b the number of neighbours in the second

segment (S_b), our normalized score Sc is defined as:

$$Sc = \frac{N_\ell}{\sqrt{N_a \cdot N_b}}$$

4.3 Application to Elaboration and E-Elaboration relations in ANNODIS

From the ANNODIS corpus, we extracted all the Elaboration and E-Elaboration relations according to the expert annotation. Then, we projected the neighbourhood links as described in section 4.2. The results are given in the Table 2.

	Elab.	E-elab.
Number of cases	625	527
Average segment length	54.61	27.84
Average # of proj. links N_ℓ	**5.99**	1.39
Average cohesion score Sc	**0.61**	0.32

Table 2: Comparison between Elaboration and E-Elaboration lexical cohesion

Table 2 shows that Elaborations contain much more lexical links than E-Elaborations (4 to 5 times more). This can partially be explained by the length of Elaboration segments : Elaborations are typically 2 times longer than E-Elaborations. From an application point of view, the skew on N_ℓ is not a problem. Using N_ℓ as a cue is then equivalent to combining two cues: the higher lexical cohesion of Elaboration relation and the fact than Elaborations are longer than E-Elaborations. From a theoretical point of view, we expect to observe that Elaboration is more lexically cohesive than E-Elaboration even for the normalized score Sc. Data in Table 2 confirms this expectation. This first result is interesting in itself, as it provides an experimental validation based on a corpus for the theoretical descriptions of Elaboration and E-Elaboration (Asher and Lascarides, 2003; Prévot et al., 2009).

Based on this result, we propose to use lexical cohesion cues to improve ANNODIS annotations, by predicting the errors of the annotators. In the next section (5) we present an experiment set up in order to reach this goal.

5 Predicting the confusions between Elaboration and E-Elaboration: implementation

In section 3, we highlighted that Elaboration and E-Elaboration are the relations that are most frequently mistaken in the naive annotation of ANNODIS corpus. However, as shown in section 4, Elaboration and E-Elaboration can be distinguished using their lexical cohesion, which can be evaluated by using distributional neighbours. In this section, we present a machine learning experiment aiming at automatically classifying Elaboration and E-Elaboration using lexical cohesion cues, among other features.

5.1 Experiment methodology

From the ANNODIS corpus, we extracted all Elaboration and E-Elaboration relations according to the naive annotation. We restricted this subset to relations having an Elaboration or E-Elaboration annotation in the expert annotation. Indeed, we only defined cues for these two relations; considering other relations would require specifying markers for them. Then, for each $< S_a, S_b >$ couple, we computed the attributes listed in Table 3.

Att.	Description	Values
N_ℓ	see section 4.2	$N_\ell \in \mathbb{N}$
Sc	see section 4.2	$Sc \in \mathbb{R}^+$
rel	S_b is a relative clause	boolean
app	S_b is a nom. / adj. apposition	boolean
ger	S_b is a gerund clause	boolean
bra	S_b is in brackets	boolean
emb	S_b is an embedded segment	boolean
w_{Sa}	# of words in S_b	$w_{S1} \in \mathbb{N}$
w_{Sb}	# of words in S_b	$w_{S2} \in \mathbb{N}$
w_{tot}	$w_{Sa} + w_{Sb}$	$w_{tot} \in \mathbb{N}$
s_{Sa}	# of segments in S_a	$s_{S1} \in \mathbb{N}$
s_{Sb}	# of segments in S_b	$s_{S2} \in \mathbb{N}$
s_{tot}	$s_{Sa} + s_{Sb}$	$s_{tot} \in \mathbb{N}$

Table 3: Attributes computed

Thus, we considered:

- lexical cohesion cues described in section 4.2 (N_ℓ and Sc);

- linguistic features presented in section 3.3 (rel, app, ger and bra): these features were detected using patterns based on the part-of-speech annotation of the segments;

- structural features regarding the two segments: is S_b embedded in S_a? (emb) How many words are there in the two segments? (w_{Sa}, w_{Sb} and w_{tot}) Are they simple segments or complex segments? (s_{Sa}, s_{Sb} and s_{tot}).

We then processed the data produced using the machine learning software Weka (Hall et al., 2009). More specifically, we used Weka's implementation of the Random Forest classifier (Breiman, 2001). In the following sections, we present our results (section 5.2) and discuss the way they could be exploited in an annotation campaign (section 5.3).

5.2 Classification results

Table 4 shows again the results for naive annotation when compared to the annotation provided by experts. The accuracy is satisfying at 69.4%, but closer examination reveals that a large set of E-Elaboration are mistakenly classified as Elaboration by the naive annotators. Using the classifier introduced in sec-

	elab	e-elab	← Naive annot.
elab	302	70	
e-elab	158	216	
↑Expert annot.			Accuracy : 69.4%

Table 4: Confusion matrix for naive annotation

tion 5.1, we performed a classification experiment on this data set, considering the naive annotation as an additional unreliable cue. Results from this experiment, using 10-fold cross-validation, are presented in Table 5. The accuracy increases to 75.7% and both E-Elaboration→Elaboration and Elaboration→E-Elaboration confusions are significantly reduced. This 6.3% improvement on the naive annotation is highly satisfying.

	elab	e-elab	← Naive-aided
elab	306	66	auto. annot.
e-elab	115	259	
↑Expert annot.			Accuracy : 75.7%

Table 5: Confusion matrix for naive-aided automatic annotation

27

In order to evaluate the impact of the different attributes used in the classifier (see Table 3), we repeated the classification experiment, using a single attributes category at a time. The results are summarized in Table 6. Structural attributes bring only a

Attributes used	Accuracy
Naive annotation	69.4%
Naive + lexical cohesion cues	72.3% (+2.9%)
Naive + linguistic cues	71.7% (+2.3%)
Naive + structural cues	69.7% (+0.3%)
All	75.7% (+6.3%)

Table 6: Impact of the different attributes categories

0.3% gain. As expected, lexical cohesion cues bring a noticeable improvement (+2.9%). Moreover, this improvement is stronger than the one brought by all linguistic features combined (+2.3%). This confirms the importance of lexical cohesion to differentiate between Elaboration and E-Elaboration. The synergy between the attributes categories is highlighted by the gain brought by the combination of all attributes, significantly higher than the sum of individual gains.

5.3 Exploiting our classifier's results in an annotation campaign

In the context of an iterative annotation campaign such as ANNODIS, an automatic classifier could hold different roles: (a) providing a first annotation, *i.e.* replacing the naive annotation (b) improving the naive annotation, *i.e.* replacing the expert annotation (c) helping the expert annotation, with an intermediate process between naive and expert annotation.

Role (a) is irrelevant to the present study. Indeed, the automatic annotations experiments were performed only on cases identified by naive annotators as Elaboration or E-Elaboration. In its current form, the automatic annotation system developed can only be used as a processing step following the required naive annotation (in the ANNODIS context, naive annotation is the only one available for 44 texts, see section 2). As demonstrated by the results of section 5.2, our system can directly be used to improve the naive annotation (b): a significant amount of confusions between the frequent relations Elaboration and E-Elaboration can be corrected (from 69.4% to 75.7% accuracy).

Finally, we show below how our classifier can be exploited to help expert annotation (c). This last proposal is relevant to workload reduction for the experts annotators, which are still required here (contrary to proposal (b)) . We have seen (Table 4) that naive annotators are not very reliable for E-Elaboration identification, so that in practice this classification should always be reviewed. However, presenting all naive E-Elaboration results to the expert introduces a significant overhead. Automatic classification can be used to isolate the most critical cases, allowing to reduce this overhead by presenting only those cases to the expert.

Table 8 illustrates the expected performance for such a system. From 286 relations classified as E-Elaboration by the naive annotators, 159 are automatically validated as E-Elaboration and not presented to the experts. Aiming for an error rate below 10%, we used the cost matrix presented in Table 7. Thus, only 8.2% of the accepted annotations are er-

0	10
1	0

Table 7: Cost matrix

roneous. The experts are then presented with the 127 cases that the automated classifier identified as possible Elaborations. For the data on which Table 8 is based, this represents a $159/286 = 55.6\%$ workload reduction for expert annotators.

	elab	e-elab	← automatic annot.
elab	57	13	(naive annot=e-elab)
e-elab	70	146	
↑Expert	127	159	

second look   accepted annot.
by expert   (error : 8.2%)

Table 8: Confusion matrix for naive e-elab second-look setup

Going further, our system could also be used to suggest improvements to the annotation manual, by highlighting the causes for frequent mistakes and by allowing an analysis of the reliability of the different cues taken in consideration (or not) by the annotators

6 Conclusion

In this paper, we used ANNODIS, a french corpus annotated with discourse relations, which provides the results of two annotation steps, to study two particular discourse relations: Elaboration and E-Elaboration. These two very frequent relations are (a) often erroneously interchanged by annotators (b) difficult to detect automatically as their signalling is poorly studied. We considered these relations from a lexical cohesion viewpoint.

We introduced a method to evaluate the lexical cohesion between two segments, using distributional neighbors. This approach allowed us to confirm that Elaboration is more cohesive than E-Elaboration. We therefore integrated lexical cohesion cues in a machine learning system, employed in a classification experiment with promising results.

These results bring improvements that could be used to facilitate future annotation campaigns. Going further, this study is especially interesting because (a) it fully exploits two levels of annotation, which is very rare; (b) it enhances the linguistic description of the considered relations, based on attested data; (c) it validates our approach based on lexical cohesion detection.

References

C. Adam and F. Morlane-Hondère. 2009. Détection de la cohésion lexicale par voisinage distributionnel : application à la segmentation thématique. In *RECITAL'09*, Senlis, France, juin 2009.

C. Adam and M. Vergez-Couret. 2010. Signalling elaboration: Combining gerund clauses with lexical cues. In *Multidisciplinay Approaches to Discourse (MAD 2010)*, Moissac, France, 17-20 March.

S. Afantenos, N. Asher, F. Benamara, M. Bras, C. Fabre, M. Ho-dac, A. Le Draoulec, P. Muller, M.-P. Péry-Woodley, L. Prévot, J. Rebeyrolle, L. Tanguy, M. Vergez-Couret, and L. Vieu. 2012. An empirical resource for discovering cognitive principles of discourse organisation: the annodis corpus. In *Eighth Language Resources and Evaluation Conference (LREC 2012)*, Istanbul (Turkey), 21-22 May.

N. Asher and A. Lascarides. 2003. *Logics of conversation*. Cambridge:CUP.

M. Baroni and A. Lenci. 2010. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36(4):673–721.

M. Barzilay, R. & Lapata. 2008. Modeling local coherence: An entity-based approach. *Computational Linguistics*, 34(1):1–34.

J. Bateman and M. Zock, 2005. *The Oxford Handbook of Computational Linguistics*, chapter Natural Language Generation, pages 284–304. Oxford Universtity Press, New York.

I. Berzlánovich, M. Egg, and G. Redeker. 2008. Coherence structure and lexical cohesion in expository and persuasive texts. In *Proceedings of the Workshop Constraints in Discourse III*, pages 19–26, Potsdam, Germany.

D. Bourigault. 2002. UPERY : un outil d'analyse distributionnelle étendue pour la construction d'ontologies à partir de corpus. In *Actes de la 9ème confrence sur le Traitement Automatique de la Langue Naturelle*, pages 75–84, Nancy.

L. Breiman. 2001. Random forests. *Machine Learning*, 45(1):5–32.

L. Carlson, D. Marcu, and M. E. Okurowski. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the 2nd SigDial Workshop on Discourse and Dialogue*. version papier + version lectronique.

J. Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20 (1):37–46.

C. Fabricius-Hansen and B. Behrens. 2001. Elaboration and related discourse relations viewed from an interlingual perspective. In *Proceedings from Third Workshop on Text Structure*, Austin, Texas.

M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. 2009. The weka data mining software: An update. *SIGKDD Explorations*, 11(1).

L.-M. Ho-Dac, C. Fabre, M.-P. Péry-Woodley, and J. Rebeyrolle. 2009. A top-down approach to discourse-level annotation. In *Corpus linguistic Conference*, Liverpool, UK, 20-23 juillet.

A. Knott, J. Oberlander, M. O'Donnell, and C. Mellish. 2001. Beyond elaboration : the interaction of relations and focus in coherent text. In T. Sanders, J. Schilperoord, and W. Spooren, editors, *Text representation : linguistic and psycholinguistic aspects*, pages 181–196. Amsterdam : Benjamins.

A. Knott. 1996. *A data-driven methodology for motivate a set of coherence relations*. Ph.D. thesis, University of Edinburgh.

W. C. Mann and S. A. Thompson. 1987. Rhetorical structure theory : a theory of text organisation. Technical report, Technical report ISI/RS-87-190, Information Sciences Intitute.

D. Marcu. 2000. The rhetorical parsing of unrestricted texts : a surface-based approach. *Computational Linguistics*, 26(3):395–448.

M.-P. Péry-Woodley, N. Asher, P. Enjalbert, F. Benamara, M. Bras, C. Fabre, S. Ferrari, L.-M. Ho-Dac, A. Le Draoulec, Y. Mathet, P. Muller, L. Prévot, J. Rebeyrolle, M. Vergez-Couret, L. Vieu, and A. Widlöcher. 2009. Annodis : une approche outillée de l'annotation de structures discursives. In *Actes de la conférence TALN*, Senlis, France.

L. Polanyi. 1988. A formal model of the structure of discourse. *Journal of Pragmatics*, 12:601–638.

R. Prasad, E. Miltsakaki, N. Dinesh, A. Lee, A. Joshi, L. Robaldo, and B-L. Webber. 2007. The penn discourse treebank 2.0. annotation manual. Technical report, IRCS, Institute for Research in Cognitive Science, University of Pennsylvania.

L. Prévot, L. Vieu, and N. Asher. 2009. Une formalisation plus précise pour une annotation moins confuse: la relation d'elaboration d'entité. *Journal of French Language Studies*, 19(2):207–228.

B. Reese, P. Denis, N. Asher, J. Baldridge, and J. Hunter. 2007. Reference manual for the analysis and annotation of rhetorical structure. Technical report, University of Texas at Austin.

H. Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the Conference on New Methods in Language Processing*, Manchester, UK.

M. Vergez-Couret. 2010. *Etude en corpus des réalisations linguistiques de la relation d'Elaboration*. Ph.D. thesis, Université de Toulouse.

F. Wolf and E. Gibson. 2005. Representing discourse coherence: A corpus-based study. *Computational Linguistics*, 31(2):249–287, June.