A framework for feature-based description of low level discourse

Laura Alonso Alemany
Ezequiel Andújar Hinojosa and **Robert Sola Salvatierra**
Grial
Departament de Lingüística General
Universitat de Barcelona

Abstract

We present the first stage of the discursive annotation of a corpus in Spanish. In contrast with previous annotation projects, we have applied a feature-based description of discourse relations, and we have focussed in the low level organization of discourse.

This first annotation phase has served to settle the annotation framework and to provide empirical support to theoretical claims. Systematic guidelines have been established for the interpretation of relations between discourse units, by tests and interpretation procedures that characterize the presence of the features characterizing discourse relations.

1 Introduction and Motivation

Annotated corpora have shown to be very useful to build, evaluate or improve linguistic resources and computational analyzers. Big corpora enriched with linguistic information at sentential and sub-sentential level (morphologic, syntactic, semantic) are widely available, but they are scarce and fragmentary at discourse level.

Nevertheless, more and more efforts have been devoted to create discursively annotated corpora in recent years, mainly for English. Various annotation schemata have been developed in order to improve the consistency of annotation and to reduce the annotation cost (Discourse Resource Initiative, 1997; Cooper et al., 1999; Carlson and Marcu, 2001).

In this paper we present a schema for discursive annotation of Spanish written corpora focussed in the low level organization of discourse. Relations between discourse units are described as a conglomerate of features, instead of assigning them atomic labels.

This annotation is specially oriented to improve text summarization and question answering. Therefore, we have focussed in those aspects of discourse that allow to identify coherence relations and relative relevance, two aspects of discourse that are necessary to produce readable texts and to condensate information, respectively. We have disregarded those distinctions that are beyond the capabilities of the state-of-the-art analyzers for Spanish.

We focus in the **low level of discourse** because it consitutes a natural enhancement of the NLP pipeline towards discourse, but also because a reliable and informative representation of low level discourse can be obtained with shallow linguistic information. However, this kind of information comes short to resolve ambiguities at higher levels.

A **feature-based** annotation schema seems to be descriptively more adequate than relation-based approaches, like those based in the Rhetorical Structure Theory (RST) (Mann and Thompson, 1988). In these approaches, relations between discourse units are described by assigning them an atomic label. However, discourse relations can convey various kinds of meaning simultaneously (Moore and Pollack, 1992), which atomic labels fail to distinguish.

In contrast, if relations between discourse units are described compositionally, heterogeneous meanings can be expressed in a distinct, transparent manner. Such a description is arguably more adaptable to different necessities of representation, and also more comparable with other theoretical frameworks. Moreover, the interpretation of discourse relations can be systematized by associating tests and decision procedures to each distinct feature, in order to improve the consistency of annotation.

The rest of the paper is structured as follows. Our theoretical framework is sketched in Section 2, its descriptive adequacy is discussed in Section 3, comparing it with related work. Section 4 describes the schema for corpus annotation based in this framework. The results of an exploratory corpus annotation are discussed in Section 5, and implications are put forward.

`[`$_1$` En este caso, ] [`$_2$` [`$_3$` y ] [`$_4$` gracias al ] excelente trabajo de la antropóloga Silvia Ventosa, ] [`$_5$` autora de "Trabajo y vida de las corseteras de Barcelona", ] [`$_6$` esta leyenda urbana se comprobó ] [`$_7$` que era un calco de una historia ] [`$_8$` que conmocionó a la localidad francesa de Orleans ] [`$_9$` en 1969 ] .`

In this case, and thanks to the excellent work of the anthropologist Silvia Ventosa, author of "Work and life of Barcelona's seamstresses", it was found that this urban legend was a copy of a story that shook the French town of Orleans in 1969.

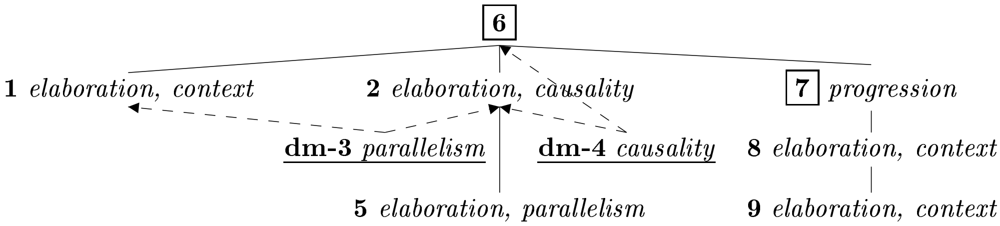

Figure 1: Condensation of a text based on its representation in discourse units and relations. Segments 6 and 7 are chosen as the condensation of the information conveyed by the text; segment 7 introduces a progression, which indicates that it conveys relevant information; segment 6 has a coherence relation with 7, namely, the fact that it is its dominating node in the hierarchical structure of discourse.

2 The low level of discourse

Under the term "*low level of discourse*" we group those phenomena that cannot be explained within a clausal scope (like anaphora, coherence relations, illocutionary force, etc.) but that are realized by features of utterances at sentential level (like syntactic constructions, referential expressions discourse particles, etc.).

Our description of the low level of discourse is focussed on coherence and relative relevance relations between minimal discourse units. This information is useful to produce coherent summaries of texts, as illustrated in Figure 1.

More concretely, we aim to describe extra-argumental relations between sentence constituents[1] that have the status of discourse unit, and also comparable relations between sentences. We believe that these relations are basic linguistic mechanisms that hold for virtually every kind of text, oral or written, and can be described independently of higher-level discourse structures, like genres or domains.

2.1 Minimal discourse units

For space constraints, we are not providing extensive grounding here on the nature of discourse units. Our position is roughly based in the classical Linguistic Discourse Model, as recently synthesized by Polanyi et al. (2004).

Two kinds of minimal discourse units are considered: **discourse segments**, "*syntactic constructions that encode a minimum unit of meaning and/or discourse function interpretable relative to a set of contexts*" (Polanyi et al., 2004) and **discourse markers**, functional units that convey information on the relations between segments. However, discourse segments can function as discourse markers, as in example (1), where the underlined segment elicits a relation between its matrix clause and a previous discourse unit, in this case, a grouping of segments. In our annotation schema allows segments can be characterized with the functional information that is characteristic of discourse markers.

(1) No satisfecho con todo ello, el ciudadano vasco Urquiza desoía a su partido y cada día se tomaba sus potes y vermuts en el "batzoki" del PNV de Durango.
Not happy with all this, the basque citizen Urquiza didn't listen to his party and he had his "potes" and vermouths in the "batzoki" of the PNV in Durango.

Unlike Carlson and Marcu (2001), we consider clausal subjects and objects of verbs and collocations of saying and opinion as discourse segments. It has been shown that their status in discourse is equal to that of a stand-alone segment (Verhagen, 2001), while their matrix clause performs a function comparable to that of lexical discourse markers. In the following example, the construction with a verb of opinion (2.a), conveys the same discursive meaning as the one with a discourse marker (2.b).

[1] We consider subordinated clauses as sentence constituents, because their distributional properties are strongly equivalent to those of a phrase.

feature	discursive effect	dimension
context	provides the setting for a discourse entity *People started demonstrating as soon as the war began.*	matter
parallelism	establishes an equivalence between two elements *A is (mod) B* *Some other governments supported the war, as in Spain.*	matter
causality	elicits a causal relation between two elements *They lost the elections because they manipulated information.*	matter
revision	negates some previous information, explicit or implied *No weapons of mass destruction were found, but Iraq was invaded.*	matter
progression	introduces a new topic or intention *The "Prestige" wandered about for a week, and it finally sunk.*	argumentative
elaboration	continues a presented topic or intention *The "Prestige" wandered about for a week, all along the coast.*	argumentative
symmetric	attachement to a node at the same level in the discourse tree *They lied to voters and so they lost the elections.*	structure
asymmetric	attachment to a node in a different level in the discourse tree *Because they lied to voters, they lost the elections.*	structure

Table 1: Basic components of meaning of discourse relations, organized in three dimensions.

(2) (2.a) [..] estoy convencido de que dentro de poco superará el examen de plancha, [...]
[...] I am sure that she will pass the ironing exam in short, [...]

(2.b) [..] seguramente dentro de poco superará el examen de plancha, [...]
[...] she will surely pass the ironing exam in short, [...]

2.2 Proposed features of the meaning of discourse relations

We have established eight basic features to describe the meaning of discourse relations in various dimensions of discourse meaning simultaneously, summarized in Table 1. These features do not aim to provide a comprehensive description of relations between discourse segments, instead, we have focussed in aspects of discourse useful to identify relevance and coherence relations between discourse segments and treatable by shallow NLP, that is, that are identifiable in text by surface clues, like syntactical constructions, lexical and referential chains and, most of all, discourse markers.

Features are organized in three co-existing dimensions (Alonso et al., 2003):

structural (*coordinating, subordinating*) expresses the relation between a discourse segment and the node where it is attached in a tree-like structure of discourse as that proposed by Webber (1978) or Polanyi (1988).

argumentation (*progression, elaboration*) describes the relation of segments with intentions and/or topics, similar to Grosz and Sidner (1986) intentional/attentional levels

matter (*context, causality, parallelism*) accounts for extra-argumental relations between the propositional content of discourse segments. It is comparable to Halliday and Hasan (1976)'s *internal*, Mann and Thompson (1988)'s *subject-matter*, or Kehler (2002)'s discourse relations.

3 Discussion of features of meaning

3.1 Relation with previous work

A number of components of meaning have been proposed in the literature to describe discourse relations, either explicitly as distinct features or implicitly as part of the meaning of atomic labels. Some of them are strongly equivalent to the ones we propose here, but some others have been ignored, either because they are beyond our shallow NLP capabilities or because we don't believe that they make relevant distinctions for information condensation tasks. We will discuss some of them in this section.

We have not distinguished subtypes of *context*, like *manner*, *instrument*, *location* or the popular family of *temporal* relations, because we believe that relevant distinctions within this range of meaning are strongly dependent on the final application. Basic coherence and relevance assessment seems to be independent of finer-grained distinctions within this meaning.

We believe the distinction between volitional and non-volitional relations (Mann and Thompson, 1988) is not useful for information condensation. Other distinctions within the family of causal relations can be obtained by combination of the proposed features. For example, *purpose* can be expressed as *asymmetric+progression+cause*, *consequence* as *symmetric+progression+cause*, and *reason* as *asymmetric + elaboration + cause*.

In contrast, the distinction between *semantic* and *pragmatic* source of coherence (Sweetser, 1990; Sanders et al., 1992; Knott, 1996) makes useful distinctions to preserve coherence when texts are condensated. In a semantic relation, like that in (3.a), both related segments contribute to the full understanding of the described situation. In contrast, a segment related to another by a pragmatic relation, as the underlined segment in example (3.b), does not contribute to describe the situation, but elicits a relation between the illucotionary forces of the relad segments. Therefore, if the target is to produce a semantically coherent text, a pragmatic relation is useful to identify accessory segments, while it will signal relevant segments if the target is to compare sources of information that may be contradictory.

(3) (3.a) [..] the judge is believed to have added 25 per cent to each sentence <u>specifically because the police had carried out the attack while operating in their official capacity</u>.

(3.b) I would be so grateful if you could tell me the name of it, <u>because it looks so unusual</u>.

Despite the utility of this distinction, we have not included *source of coherence* as a basic feature of meaning because it is beyond our NLP capabilities to distinguish pragmatic and semantic relations. Indeed, these relations tend to be signalled by the same shallow cues; for example, most causal discourse markers can be used for marking semantic or pragmatic relations indistinctively.

Finally, structural features are not equivalent to SDRT *coordinating* and *subordinating* (Asher and Vieu, 2001) or to RST *nuclearity*. Indeed, *asymmetric* does not always imply *subordinating* or *satellite*; as in example (2.a), where the underlined segment holds an asymmetric relation with its matrix clause but is not discursively subordinated. In our annotation schema, discursive subordination is obtained compositionally; typically, a subordinating relation is an asymmetric relation of elaboration. For some representation purposes, interactions with some matter meaning may determine whether such a relation is considered subordinating or not; for example, it can be stipulated that causal relations are never subordinating, independently of their structural and argumentative features.

3.2 Descripive adequacy of the presented features

Our feature-based account is specially adequate to deal with ambiguity, both for human and automated analyses. In relation-based approaches, unclear cases are treated in an "all or nothing" basis, that is, either they are assigned no label or a label of dubious descriptive adequacy. In a feature-based approach, meaning is partitioned so that underspecification can affect only part of the meaning of the relation.

A feature-based account is specially adequate to integrate heterogeneous kinds of information in a systematic way. For example, discourse markers are very informative of matter and structural dimensions, but most of them do not convey strong argumentative information. Consequently, relations between segments marked by a discourse marker may obtain their argumentative information from the topical or intentional structures of text, or from the syntactical structures of the utterance, while the rest of features can be determined by the discourse marker.

If a single dimension of discourse is considered, the example in Figure 2 can only be adequately described by a DAG. Webber et al. (2003) avoid to treat this kind of cases as a DAG by exploiting the anaphoric properties of averbial discourse markers like *then*, as shown in the left of Figure 2. We explain it by resorting to different dimensions of discursive meaning, as seen in the example in the right, where solid lines represent the matter dimension, dashed lines represent the argumentative dimension and the structural dimension is represented by the relative level of terminal discourse units.

4 Annotation of Spanish text

We have applied the presented framework to the annotation of a small journalistic corpus in Spanish. The purpose of this application was twofold: first, to test the descriptive adequacy of the proposed components of meaning, then, to settle a schema for corpus annotation.

4.1 Corpus

The corpus to be annotated is of journalistic genre, with some variation in subgenres: television critic, everyday stories, and opinion. We chose a varied corpus in order to check the

a. John loves Barolo.
b. So he ordered three cases of the '97.
c. But he had to cancel the order
d. because *then* he discovered he was broke.

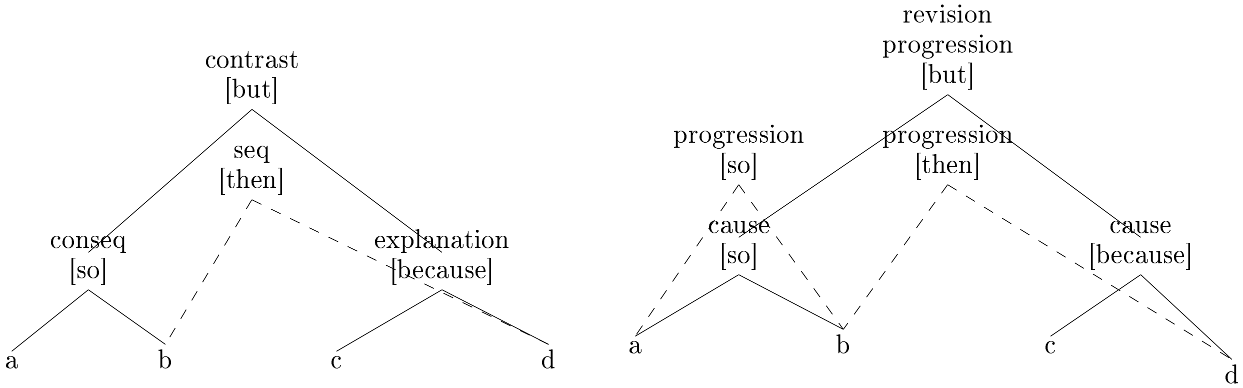

Figure 2: Two approaches to avoid treating the above example as a DAG: in the left, Webber et al. (2003, pp. 551) resort to the anaphoric properties of some discourse markers; in the right, we propose to resort to different co-exising dimensions of the meaning of discourse relations.

applicability of features in different kinds of text. The corpus is annotated in XML, and can be found at `http://lingua.fil.ub.es/~lalonso/discor/`).

The corpus consists of 6 articles totalling 3541 words (from 207 to 1042 words) in 154 sentences. Texts have been manually segmented in discourse minimal units, discourse segments and discourse markers, as defined in Section 2.1. There are 468 intrasentential segments, of which 84 are required by the argumental structure of the verb (as in example 2.a) and the rest are different kinds of adjuncts, of which 226 are dominated by a discourse marker. 261 discourse markers have been found, corresponding to 101 different forms, the most frequent are *y* (*and*, 50 occurrences), *para* (*for, to*, 20 occurrences), *como* (*like, as*, 15 occurrences) and *pero* (*but*, 9 occurrences).

4.2 Annotation procedure

Three judges have taken part in this preliminary annotation, two naive judges, with no background in linguistics, and a linguist. The purpose of having naive judges annotate the corpus was to check the intuitivity of the proposed features. After each text was annotated, the annotation was discussed between judges, but the annotated text was not modified.

Judges made decisions basically relying on their intuitions. However, as an aid to guide decision-taking, an annotation manual was created (Alonso et al., 2004), and a preparatory text was annotated collaboratively. The manual was enhanced and refined during annotation, with issues that were specially controversial, like the annotation of discontinuous discourse markers (*so... that...*) or an effective procedure to systematize the assignment of matter features to syntactical constructions like relative clauses or absolute participles.

The procedure for annotation is as follows: once a judge has read the whole text, discourse segments are characterized one by one in order of occurrence in the text, by the following features: **node(s) of attachment**, **features of meaning** and **glosses** for the selected features.

4.2.1 Node of attachment

The value for this attribute depends on the kind of discourse unit: segments are assigned the identification number(s) of the node(s) where they are attached in the tree-like structure of discourse. In contrast, discourse markers are assigned the identification number of the nodes they relate. These nodes can be minimal discourse segments or discourse units constituted by continuous spans of discourse segments.

However, segments with a relating function (as in example 1) are assigned the same kind of value as a discourse marker, that is, the identification numbers of the segments they relate. Contrastively, discourse markers that do not relate discourse segments, like those relating the text with the author (for example *luckily* or *of course*), can be attached to the discourse segment where they are found.

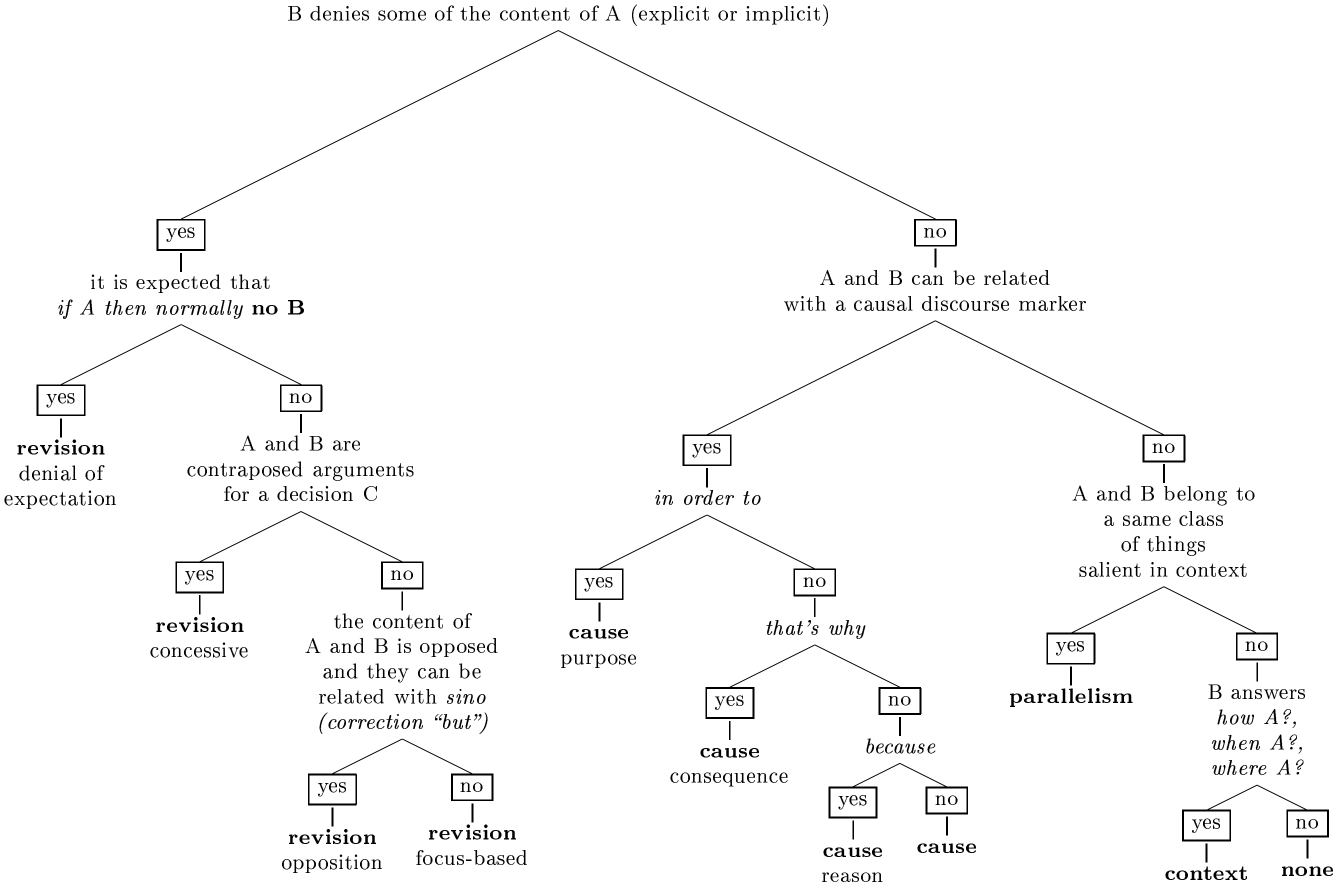

Figure 3: Decision tree for annotating features belonging to the matter dimension. It reflects the characterization of features presented in Table 1: decisions are taken in order of markedness, so that it is first discriminated whether the most marked feature (*revision*) is present or not. Finer-grained distinctions within *revision* and *cause* are also directed by the decision tree.

Each segment is related to at least one other segment. When there is no clear attachment node or when the segment is the starting point or main claim of a text, it is considered the top of a local structure, and is attached to itself.

Each segment carries the information for its own attachment to other segments, but none for the attachment of other segments to it. The most marked segment in a relation is the one that carries the information, usually, segments under the scope of a discourse marker or characterized by redundant lexic or referential expressions. If there is no difference in markedness between segments, those coming later in discourse are the ones to carry the information. Note that markedness has no relation with RST's nuclearity, since marked segments can be the nucleus or the satellite of a relation.

4.2.2 Features of meaning

The presence or absence of each of the eight proposed features was determined via decision trees that reflect the theoretical characterization of features. Figure 3 displays the decision tree for the matter dimension.

When judges were unsure about a feature for a relation, they left it underspecified. This resulted in an average of 2.6 features per relation, which increases slightly (2.7) for segments containing a discourse marker and decreases to 2.2 for relations at the beginning of a paragraph.

4.2.3 Glosses

If a *matter* or *argumentative* feature was applicable for a relation, a gloss was provided for it. Glosses were aimed to guarantee consistency and keep track of decision procedures, but they also served to encode finer-grained distinctions within each of the proposed features, which will be exploited to study the realization of discursive meanings that are widely accepted in the literature in a bigger corpus. The following glosses were provided:

	units	features	revision	cause	parallel	context	prog.	elab.	sym.	asym.	**average**
text 1	181	531	0.55	0.69	0.75	0.32	0.64	0.59	0.52	0.49	0.57
text 2	74	194	0.59	0.17	0.56	0.47	0.42	0.46	0.54	0.46	0.46
text 3	261	716	0.74	0.64	0.51	0.36	0.54	0.51	0.60	0.58	0.56
text 4	49	131	0.89	0.51	0.50	0.61	0.52	0.48	0.71	0.72	0.62
text 5	153	405	0.57	0.53	0.49	0.38	0.39	0.35	0.42	0.48	0.45
text 6	163	437	0.85	0.74	0.59	0.42	0.56	0.52	0.51	0.52	0.59
total	881	2414	**0.70**	**0.55**	**0.57**	**0.43**	**0.51**	**0.48**	**0.55**	**0.54**	**0.54**
discourse markers			**0.72**	**0.53**	**0.53**	**0.39**	**0.44**	**0.43**	**0.39**	**0.42**	**0.48**
beginning paragraph			**0.03**	**-0.0**	**0.16**	**0.01**	**0.29**	**0.18**	**0.03**	**0.11**	**0.10**

Table 2: Average kappa agreement between judges for the six annotated texts.

revision elicit the information that is denied, implicitly or explicitly, in four different forms, in order of markedness, the first three based in Lagerwerf (1998), the least marked based in Umbach (2004):

- **denial of expectation** elicit expectation
- **concessive** elicit *tertium comparationis*
- **opposition** the related segments can be rephrased with a correction *but* (*sino*)
- **focus-based** if none of the others apply

cause one of four discourse markers with which the relation can be paraphrased: *in order to* (purpose), *because* (cause), *that's why* (reason), and *therefore* (consequence)

parallelism the common class of things to which the two segments belong, which has to be salient in the context, either because it is lexicalized or by regular abstraction procedures, like hyperonymy

Glosses for **progression** and **elaboration** summarize the topic or intention that is introduced or elaborated, respectively. This procedure increased the consistency of the annotation for these two features

5 Preliminary results of annotation

5.1 Consistency of annotation

The consistency of the annotation was evaluated by kappa agreement between the values of the features assigned to each node. As can be seen in Table 2, the average agreement, $\kappa = .54$ is quite low, and does not guarantee a good reproducibility of the results (Carletta, 1996). However, it has to be taken into account that this agreement has been obtained in the preliminary phase of annotation, during the process when annotation criteria were being established, and also that the judges were not professional annotators. Even with trained professional annotators, Carlson et al. (2003) present $\kappa = .6$ in the initial stages of annotation, reaching $\kappa = .75$ at the end of the project.

5.2 Discussion of results

Interestingly, kappa agreement for discourse markers is significantly lower than the average, except in the case of the *revision* feature, which presents the highest agreement per feature. However, the number of features for characterizing discourse markers was higher than the average, which means that judges recognized them as highly informative of discourse organization but were not able to distinguish clearly the components of their meaning.

As can be seen by the low agreement for segments occurring at the beginning of paragraph, they were very difficult to characterize for judges, which seems to indicate that the coherence mechanisms that apply at paragraph level are qualitatively different from those at sentential and inter-sentential level. This seems a strong argument to treat the low level organization of discourse as an autonomous level of language.

Although judges had been allowed to relate segments to more than one discourse unit by more than one relation, this never happened in the corpus. This supports the claim that discourse can be represented as a hierarchical tree. In our framework, the structure of discourse is described as a superposition of the hierarchical trees that represent each dimension, where segments have only one relation in each dimension, although they might be attached to different segments in different dimensions (see Figure 2).

6 Conclusions and Future Work

We have annotated a corpus with a feature-based approach to the description of the organization of low level discourse. We have carried out an annotation of a small corpus in Spanish, and we have established an annotation framework in the form of systematic procedures for decision taking. These procedures are a direct mapping of the theoretical characterization of the basic components of meaning of discourse relations.

The consistency of this preliminary annotation is comparable to that of the preliminary stages of other annotation initiatives. The fact that two of the three annotators involved had no background in linguistics supports the validity of the proposed features as basic components of meaning of discourse relations.

Future work will be aimed to the application of this annotation framework to a Spanish corpus of a reasonable size, most probably a subset of the CLiC-TALP corpus, which is currently being annotated with syntactic functions.

7 Acknowledgements

This research has been partially funded by MCyT program - BFF2001-5440 and also thanks to grant PB98-1226 of the Spanish Research Department.

References

Laura Alonso, Jennafer Shih, Irene Castellón, and Lluís Padró. 2003. An analytic account of discourse markers for shallow NLP. In Manfred Stede and Henk Zeevat, editors, The Meaning and Implementation of Discourse Particles, whorkshop at ESSLLI'03.

Laura Alonso, Daniel Alonso, Ezequiel Andújar, and Robert Sola. 2004. Anotación discursiva de corpus en español. Technical Report GRIAL report series 2004-2, Departament de Lingüística General, Universitat de Barcelona.

Nicholas Asher and Laure Vieu. 2001. Subordinating and coordinating discourse relations. In Intl. Wkshp. on Semantic, Pragmatics and Rhetorics, San Sebastián.

Jean Carletta. 1996. Assessing agreement on classification tasks: the kappa statistic. Computational Linguistics, 22(2):249–254.

Lynn Carlson and Daniel Marcu. 2001. Discourse tagging manual. Technical Report ISI-TR-545.

Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2003. Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. In Jan van Kuppevelt and Ronnie Smith, editors, Current Directions in Discourse and Dialogue. Kluwer Academic Publishers.

Robin Cooper, Staffan Larsson, Colin Matheson, Massimo Poesio, and David Traum. 1999. Coding instructional dialogue for information states. Technical Report D1.1, The TRINDI Consortium.

Discourse Resource Initiative. 1997. Standards for dialogue coding in natural language processing. Technical Report 167, Dagstuhl-Seminar.

Barbara J. Grosz and Candace L. Sidner. 1986. Attention, Intention, and the Structure of Discourse. Computational Linguistics, 12(3):175–204.

Michael A. K. Halliday and Ruqaiya Hasan. 1976. Cohesion in English. Longman, London.

Andrew Kehler. 2002. Coherence, Reference, and the Theory of Grammar. CSLI Publications.

Alistair Knott. 1996. A Data-Driven Methodology for Motivating a Set of Coherence Relations. Ph.D. thesis, University of Edinburgh, Edinburgh.

Luuk Lagerwerf. 1998. Causal Connectives Have Presuppositions; Effects on Coherence and Discourse Structure. Den Haag, Holland Academic Graphics.

William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organisation. Text, 3(8):234–281.

Johanna D. Moore and Martha E. Pollack. 1992. A problem for RST: the need for multi-level discourse analysis. Computational Linguistics, 18(4):537–544, December.

Livia Polanyi, Chris Culy, Martin van den Berg, Gian Lorenzo Thione, and David Ahn. 2004. A rule based approach to discourse parsing. In SIGDIAL 2004.

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

Ted J. M. Sanders, Wilbert P. M. Spooren, and Leo G. M. Noordman. 1992. Toward a taxonomy of coherence relations. Discourse Processes, 15:1–35.

E. Sweetser. 1990. From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure. Cambridge University Press.

Carla Umbach. 2004. Contrast and information structure: A focus-based analysis of 'but'. Journal of Semantics, 21(2):155–175, May.

Arie Verhagen. 2001. Subordination and discourse segmentation revisited, or: Why matrix clauses may be more dependent than complements. In Ted J. M. Sanders, Joost Schilperoord, and Wilbert P. M. Spooren, editors, Text Representation. Linguistic and psychological aspects, pages 337–357. John Benjamins.

Bonnie Lynn Webber, Matthew Stone, Aravind K. Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. Computational Linguistics.

Bonnie Lynn Webber. 1978. A Formal Approach to Discourse Anaphora. Ph.D. thesis, Harvard University. Reprinted in Outstanding Dissertations in Linguistics Series, Garland Publishers, 1979.