

# Discourse Studies

<http://dis.sagepub.com/>

---

## Comparing rhetorical structures in different languages: The influence of translation strategies

Iria da Cunha and Mikel Irukieta  
*Discourse Studies* 2010 12: 563  
DOI: 10.1177/1461445610371054

The online version of this article can be found at:  
<http://dis.sagepub.com/content/12/5/563>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Discourse Studies* can be found at:**

**Email Alerts:** <http://dis.sagepub.com/cgi/alerts>

**Subscriptions:** <http://dis.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://dis.sagepub.com/content/12/5/563.refs.html>

>> [Version of Record](#) - Sep 28, 2010

[What is This?](#)



# Comparing rhetorical structures in different languages: The influence of translation strategies

Discourse Studies

12(5) 563–598

© The Author(s) 2010

Reprints and permission: [sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/1461445610371054

<http://dis.sagepub.com>

SAGE

**Iria da Cunha**

Université d'Avignon et des Pays de Vaucluse, France and Universitat Pompeu Fabra, Spain

**Mikel Iruskieta**

University of the Basque Country (UPV/EHU), Spain

## Abstract

The study we report in this article addresses the results of comparing the rhetorical trees from two different languages carried out by two annotators starting from the Rhetorical Structure Theory (RST). Furthermore, we investigate the methodology for a suitable evaluation, both quantitative and qualitative, of these trees. Our corpus contains abstracts of medical research articles written both in Spanish and Basque, and extracted from *Gaceta Médica de Bilbao* ('Medical Journal of Bilbao'). The results demonstrate that almost half of the annotator disagreement is due to the use of translation strategies that notably affect rhetorical structures.

## Keywords

annotation, discourse analysis, evaluation, medical research articles, rhetorical relations, Rhetorical Structure Theory, textual corpus, translation strategies

## 1. Introduction

Writing abstracts of research articles both in a lingua franca (English, French, etc.) and in local languages (Catalan, Spanish, Basque, etc.) is nowadays usual among the scientific community. In fact, it has become a requisite for the publication in some scientific journals. As a result, it is possible to obtain bilingual corpora to investigate how the

---

### Corresponding author:

Iria da Cunha, Université d'Avignon et des Pays de Vaucluse, Laboratoire Informatique d'Avignon, 339, chemin des Meinajaries, 84911 Avignon, France and Universitat Pompeu Fabra, Roc Boronat, 138, 08018 Barcelona, Spain.

Email: [iria.dacunha@upf.edu](mailto:iria.dacunha@upf.edu)

rhetorical structures of abstracts are shown in each language and how translation strategies affect discourse structure. Some authors have carried out studies about the evaluation of rhetorical structure annotation (Carlson et al., 2001; Marcu, 2000a; Marcu et al., 1999) and about the comparison of rhetorical structures in different languages: Chinese–English (Cui, 1986; Kong, 1998; Ramsay, 2000, 2001), English–Dutch (Abelen et al., 1993), English–French (Delin et al., 1996; Salkie and Oates, 1999), Portuguese–French–English (Scott et al., 1998) and English–Japanese (Marcu et al., 2000), among others. However, to our knowledge, no studies exist on the way that translation strategies affect the process of rhetorical annotation and on the evaluation of annotator agreement.

In this work, we use Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) since it is a language-independent theory. RST is a descriptive theory for textual organization that has been proven to be very useful in describing a document by characterizing its structure with relations maintained among its discursive or rhetorical elements (e.g. Circumstance, Elaboration, Motivation, Evidence, Justification, Cause, Purpose, Antithesis, Condition, List, Contrast, etc.). As Taboada and Mann (2006) state: ‘RST addresses text organization by means of relations that hold between parts of a text. It explains coherence by postulating a hierarchical, connected structure of texts, in which every part of a text has a role, a function to play, with respect to other parts in the text.’ RST determines a set of relations among the discursive units of texts. As a rule, one of the units is more essential to the speaker’s purpose (nucleus), while the other one (satellite) provides some rhetorical information about it. This is the more usual structural model between these two units (almost always adjacent units, although there are some exceptions). These relations are named ‘nuclear’ relations (e.g. Circumstance, Elaboration, Motivation, Evidence, etc.). In the case of relations with more than one central unit with regard to the author’s purposes, the relation is named ‘multinuclear’ and a coordinated relation is established (e.g. List, Joint, Contrast, etc.). For a more detailed explanation of RST, we recommend reading the article by Mann and Thompson (1988) or the RST web site by Mann (2005).

RST is used to inquire into several theoretical and applied subjects explained in Taboada and Mann (2005) as, for example, automatic generation of texts, automatic summarization, textual analysis, automatic translation, writing teaching, acquisition of discursive knowledge, spoken discourse analysis, information extraction, etc. Some relevant works on these subjects are, among others, Bouayad-Agha (2000), Burstein and Marcu (2003), da Cunha (2008), da Cunha et al. (2007), Ghorbel et al. (2001), Haouam and Marir (2003) and Marcu (2000a). In addition, some rhetorical parsers in different languages are also based on this theory: Sumita et al. (1992) in Japanese, Marcu (1998) in English, and Pardo and Nunes (2008) and Pardo et al. (2004) in Brazilian Portuguese. There is a current project to develop this parser for the Spanish language (da Cunha and Torres-Moreno, 2010). A rhetorical parser is a system that automatically analyzes a text, giving as output the rhetorical tree of this text in terms of RST. This kind of parser has three stages: rhetorical segmentation, determination of RST relations and development of rhetorical trees. They are usually based on lexical-syntactic rules and statistical techniques.

However, though widely used, some objections have been made to RST. Stede (2008), for example, criticizes its ambiguity, since many assumptions that annotators carry out cannot be made explicit in a single tree. The difficulty of obtaining the same rhetorical tree of a text from different annotators would prove this subjectivity:

An RST-style analysis of a text, on the other hand, cuts ‘vertically’: It tries to capture the essence of coherence within a single representation structure, making a series of quite different simplifications along the way. We do not doubt that this can be an insightful instrument for studying text – RST has been quite successful for a variety of purposes. But there are inherent limitations on the explanatory power when information from different realms is conflated in a single tree structure: On the one hand, one cannot do full justice to the separate realms; on the other hand, the single tree structure becomes ambiguous, because when crafting it, many underlying assumptions cannot be made explicit. (Stede, 2008: 329)

All the considerations taken into account until now lead us to formulate the following interesting questions:

- Is it possible to compare the rhetorical structures of a parallel corpus of medical texts in two very different languages such as a Romance language (Spanish) and a Non-Indo-European language (Basque) by means of the same theory? Do these texts share a similar superstructure?
- Taking into account the difficulty of two annotators carrying out the same rhetorical analysis with RST relations, how do translation strategies affect the agreement on the rhetorical structure of parallel texts? Which linguistic differences exist in both rhetorical structures?
- Which is the best evaluation method in order to determine the factors affecting the evaluation of rhetorical structure (translation strategies or linguistic differences; theoretical abstraction level or ambiguity of the rhetorical structure)?

In this article we aim to answer these questions. With this intention, an experiment has been designed. First, the corpus was annotated with rhetorical relations (one author annotated the Basque corpus and the other annotated the Spanish one). This corpus contains 20 abstracts in Spanish and Basque, included in medical research articles from the *Gaceta Médica de Bilbao*<sup>1</sup> (‘Medical Journal of Bilbao’). Afterwards, both annotations were compared and the differences among them were observed. The methodology used in this experiment is explained in section 2. In section 3, we give the details of the results of the quantitative and qualitative evaluations on spans, nuclearity and rhetorical relations. Conclusions are presented in section 4.

## 2. Methodology

The methodology of our research included several phases. First, a corpus of analysis was built. Second, departure criteria with regard to the segmentation of the text into units and to the specific relations used were defined. Third, the corpus texts were labeled by the annotators (one in Spanish and one in Basque). Fourth, quantitative analysis was carried out. Fifth, qualitative analysis was performed.

### 2.1. Corpus

Nowadays, no parallel Spanish–Basque corpora are available for research purposes. Research groups have to develop their own corpus in order to carry out contrastive

research in these two languages. For this reason, we had to create a specific corpus to perform our analysis. There are no previous studies comparing rhetorical structures in Spanish and Basque. As mentioned, our corpus contains 20 abstracts in Spanish and Basque included in medical research articles from the *Gaceta Médica de Bilbao* written by medical specialists between the years 2000 and 2008.

The first reason to choose this corpus was that this journal requests that authors submit the articles in Spanish and the corresponding abstracts in Spanish, Basque and English. As most of the authors of the texts of our corpus are Basque and a relevant portion of the Basque population is bilingual, we assume that they themselves wrote both the abstracts in Spanish and Basque. Nevertheless, in some cases, the author may have asked for some help to write the Basque abstract. We think this fact is not really relevant, because the journal gives the authors very detailed guidelines about the information that they have to include in their abstracts (in the three mentioned languages). Authors are asked to use in their abstracts the IMRD structure (Swales, 1990): Introduction, Methods, Results and Discussion:

The summary must contain approximately 150 words and it must include:

- a) the purpose of the study,
- b) the used procedures and the principal findings,
- c) the most relevant conclusions, with emphasis on what is new or relevant in the article.<sup>2</sup>

We think these two facts (bilingualism and journal guidelines) guarantee that both abstracts (Spanish and Basque) include the same information and a similar structure.

The second reason to choose this corpus is to analyze the relations among macrostructures and genres and, in this way, to highlight a rather open question of RST. As Taboada and Mann (2006) state: 'A more exhaustive study of different genres would throw light on the relationship between macrostructures or genres and RST structures.' We have selected a specialized corpus that contains medical texts with a very specific genre: the research article. In the future, we plan to analyze a general corpus to compare it with this specialized corpus.

Appendix Table 1 shows the information of the corpus texts (title, author[s] and year of publication).

## 2.2. Departure criteria

In order to avoid circularities as much as possible, we first define what is an EDU (Elementary Discourse Unit) in an abstract way and, second, we segment all the text only focusing on syntactic clues (see section 2.2.1.) before carrying out the rhetorical analysis.

**2.2.1. EDU segmentation.** Mann and Thompson (1988) proposed a definition of discourse unit based on a theory-neutral classification. Their motivation was to describe a theoretical frame for RST. To this end, they proposed an abstract definition and they escaped from a circular definition:

Unit size is arbitrary but the division of the text into units should be based on some theory-neutral classification. That is, for interesting results, the units should have independent functional integrity. In our analyses, units are essentially clauses, except that clausal subjects and complements and restrictive relative clauses are considered parts of their host clause units rather than separate units. (Mann and Thompson, 1988: 6)

Although Marcu (1999) uses RST as well, his definition of discourse unit has a different motivation: the conformation of a corpus of tagged documents for the research community. Thus, the annotation should offer all the possible information. As he states:

One (probably) uncontroversial choice would be to take sentences as the elementary units of discourse. Unfortunately, if we do so, we leave lots of rhetorical information outside the scope of our analysis. (Marcu, 1999: 9)

Marcu's definition of unit can be controversial in some aspects because of its circular nature, but for Marcu this is a secondary question given that it does not interfere with his main motivation.

Our goal is far from both Mann and Thompson's (1988) and Marcu's (1999) proposals because, first, we want to compare the rhetorical structure of translations at a propositional level and, second, we want to analyze some problems that appear during the annotation process. Therefore, in this work, we do not consider it necessary to carry out such a detailed analysis as Marcu.

With regard to EDU segmentation, we follow more or less the most common set of guidelines for segmenting text in RST. Carlson and Marcu (2001) departed from them in some aspects and we have revised some questions from their manual. Some specifications were made so that we would be able to clearly differentiate syntactic and discursive levels. In this work, we consider that EDUs must include a finite verb (that is, they have to constitute a sentence or a clause) and must show, strictly speaking, a rhetorical relation. These established specifications are the following ones:<sup>3</sup>

a) In Carlson and Marcu (2001), complements of attribution verbs (speech acts and other cognitive acts) are treated as EDUs, as example 1a shows:<sup>4</sup>

1a. [Bush indicated] [there might be 'room for flexibility' in a bill] [. . .]

In contrast, our approach does not consider these complements of attribution verbs as EDUs, and we would segment the same passage as example 1b shows:

1b. [Bush indicated there might be 'room for flexibility' in a bill] [. . .]

The clause 'there might be "room for flexibility" in a bill' constitutes a direct object (from a traditional grammar-oriented approach) or an actant II (from a dependency grammar-oriented approach) of the verb 'to indicate' and, because of that, we consider it only at this level (syntactic).

We do not consider the Attribution relation for three types of reasons: a) a definitional reason: it does not make explicit any kind of writer's intention, so Attribution does not

have the same status as other RST relations (Stede, 2008); b) a language level reason: it can be identified only by syntax rules (Skadhauge and Hardt, 2005); and c) a procedural reason: it implies circularity in EDU definition. As Stede (2008: 316) states:

*Attribution* thus does not have the same status as, say, relations of causality or contrast: The relationship between an event of saying and the specific contents of that saying is different from a coherence relation linking two complete propositions.

b) Carlson and Marcu (2001) specify that the clauses that depend to ‘so that their clients can’ are treated as various EDUs and these are considered as satellites in a Purpose relation. In turn, the satellite constitutes a multinuclear List of coordinated clauses, as we can see in example 2a:

- 2a. [Equipped with cellular phones, laptop computers, calculators and a pack of blank checks,] [they parcel out money] [so that their clients can find temporary living quarters,] [buy food,] [replace lost clothing,] [repair broken water heaters,] [and replaster walls.]

In contrast, we would treat all these clauses as a single EDU:

- 2b. [Equipped with cellular phones, laptop computers, calculators and a pack of blank checks,] [they parcel out money] [so that their clients can find temporary living quarters, buy food, replace lost clothing, repair broken water heaters, and replaster walls.]

c) In Carlson and Marcu (2001), relative clauses, nominal postmodifiers and clauses that break up other legitimate EDUs are treated as embedded discourse units, while we do not consider these units as such. Several examples follow:

Relative clauses:

- 3a. [A separate inquiry by Chemical cleared Mr. Edelson of allegations] [*that* he had been lavishly entertained by a New York money broker.]  
 3b. [A separate inquiry by Chemical cleared Mr. Edelson of allegations *that* he had been lavishly entertained by a New York money broker.]

Nominal postmodifiers with non-finite clause:

- 4a. [The results underscore Sears’s difficulties] [*in implementing* the ‘everyday low pricing’ strategy] [that it adopted in March, as part of a broad attempt] [*to revive* its retailing business.]  
 4b. [The results underscore Sears’s difficulties *in implementing* the ‘everyday low pricing’ strategy that it adopted in March, as part of a broad attempt *to revive* its retailing business.]

Appositives:

- 5a. [The fact] [*that* this happened two years ago] [and there was a recovery] [gives people some comfort] [*that* this won’t be a problem.]  
 5b. [The fact *that* this happened two years ago and there was a recovery gives people some comfort *that* this won’t be a problem.]

Parentheticals:

- 6a. [The Tass news agency said the 1990 budget anticipates income of 429.9 billion rubles] [(\$US693.4 billion)] [and expenditures of 489.9 billion rubles] [(\$US790.2 billion).]  
 6b. [The Tass news agency said the 1990 budget anticipates income of 429.9 billion rubles (\$US693.4 billion) and expenditures of 489.9 billion rubles (\$US790.2 billion).]

In this work, we only segment units appearing in parentheses when they clearly constitute an EDU, or an element maintaining some discourse relation with another element and containing a finite verb.

Coordinated clauses in embedded units:

- 7a. [She signed up,] [starting as an 'inside' adjuster,] [who settles minor claims] [*and* does a lot of work by phone.]  
 7b. [She signed up,] [starting as an 'inside' adjuster, who settles minor claims *and* does a lot of work by phone.]

d) In Carlson and Marcu (2001), phrases that begin with a strong discourse marker, such as *because*, *in spite of*, *as a result of*, *according to*, are treated as EDUs, as examples 8a and 9a show:

- 8a. [But some big brokerage firms said] [they don't expect major problems] [*as a result of* margin calls.]  
 9a. [Today, no one gets in or out of the restricted area] [*without* De Beers's stingy approval.]

In this work, we consider that sentences starting by these markers are EDUs only if a finite verb also exists. Therefore, we would segment the previous examples as follows:

- 8b. [But some big brokerage firms said they don't expect major problem *as a result of* margin calls.]  
 9b. [Today, no one gets in or out of the restricted area *without* De Beers's stingy approval.]

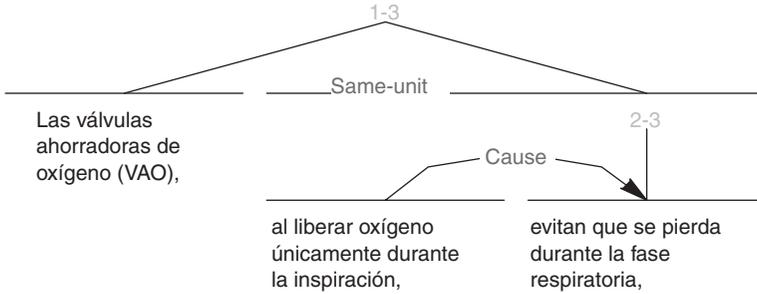
e) Carlson and Marcu (2001) establish several criteria to determine EDUs' boundaries. In this work, we only use these criteria if the marked EDU contains a finite verb. Some examples are offered below:

Parenthesis:

- 10a. [If the government can stick with them,] [it will be able to halve this year's 120 billion ruble] [(\$US193 billion)] [deficit.]<sup>5</sup>  
 10b. [If the government can stick with them,] [it will be able to halve this year's 120 billion ruble (\$US193 billion) deficit.]

Dashes:

- 11a. [This will require us to define] [– *and* *redefine* –] [what is 'necessary' or 'appropriate' care.]  
 11b. [This will require us to define – *and* *redefine* – what is 'necessary' or 'appropriate' care.]



**Figure 1.** Rhetorical tree showing a Same-unit relation

With regard to the utilization of other punctuation marks (comma, full-stop, semicolon, etc.) like boundary marks, we agree with Carlson and Marcu (2001: 30):

Commas and periods are not independent justification for an EDU boundary. If a unit is a legitimate EDU and it ends with a comma or period, the punctuation is included as part of that EDU.

Finally, it is important to highlight that an EDU can be truncated by another one (that is, it can include another EDU). If this occurs in our work, as in Carlson and Marcu (2001), the two fragments of the first EDU are segmented and they are linked later with a Same-unit relation, which is not a relation but a convention. For example, Figure 1 would be labeled as follows:

12. [*Las válvulas ahorradoras de oxígeno (VAO),*] [*al liberar oxígeno únicamente durante la inspiración,*] [*evitan que se pierda durante la fase respiratoria,*] [...]
- ENGLISH TRANSLATION: [Oxygen Conserving Valves (OCV),] [because of their release of oxygen only during inhalation,] [avoid losing oxygen during the breathing phase,] [...]

**2.2.2. Rhetorical relations.** Concerning the detection of rhetorical relations and nuclearity (that is, with regard to the decision of considering a segment as nucleus or satellite), the following tasks were carried out:

a) The list of rhetorical relations of the RST was determined. There are various classifications of rhetorical relations: the classic one by Mann and Thompson of 24 relations (Mann and Thompson, 1988), the extended one by Mann and Thompson of 30 relations (Mann, 2005) and Marcu's classification of 136 relations (Carlson et al., 2001), among others. The extended classification (Mann, 2005) was chosen for the annotation of the parallel corpus. As Marcu et al. (1999: 55) point out, reduction in the relations' taxonomy does not have a significant impact on annotators' agreement:

The results [...] show that a significant reduction in the size of the taxonomy of relations may not have a significant impact on agreement ( $k_{\gamma\gamma}$  is only about 4% higher than  $k_{\gamma}$ ). This suggests that choosing one relation from a set of rhetorically similar relations produces some, but not too much, confusion.

b) We looked for a real representative example of each relation and nuclei and satellites were marked. Examples are taken from the corpus used in da Cunha (2008), containing Spanish medical articles that were extracted from the journal *Medicina Clínica* ('Clinical Medicine').<sup>6</sup> Once the Spanish examples were selected, they were translated into Basque and their nuclei and satellites were marked.

Appendix Table 2 includes the list of relations used in this work, specifying if they are multinuclear relations (N-N) or nuclear relations (N-S). For each relation, an example in Spanish and Basque is provided, where its nuclei (N) and satellites (S) are marked.

### 2.3. Rhetorical annotation

Once departure criteria were established, both annotators labeled the 20 texts of the corpus with RST relations (one in Spanish [A1] and another one in Basque [A2]). The annotation was divided into two main stages: EDU segmentation and rhetorical analysis.

**2.3.1. EDU segmentation.** In this stage, each annotator segmented the 20 abstracts of the corpus into EDUs by using the RSTTool (O'Donnell, 2000).<sup>7</sup> This task was done separately and without any contact among annotators.

Once the data on the agreement of the performed segmentations by both annotators was collected, we carried out a small discussion in order to homogenize the segmentation of Spanish and Basque abstracts. This homogenization was carried out in order to minimize the noise that could arise from a different segmentation. By these means, we aimed at obtaining, first, a more detailed quantification of the nuclearity and of the relations of rhetorical trees and, secondly, an evaluation of the factors affecting the structure. This comparison was performed manually (measuring precision and recall), due to the current lack of automatic tools comparing rhetorical trees in different languages. Mazeiro and Pardo (2009) have developed the RSTeval tool, which does compare rhetorical trees but in the same language, so it could not be used in this study.

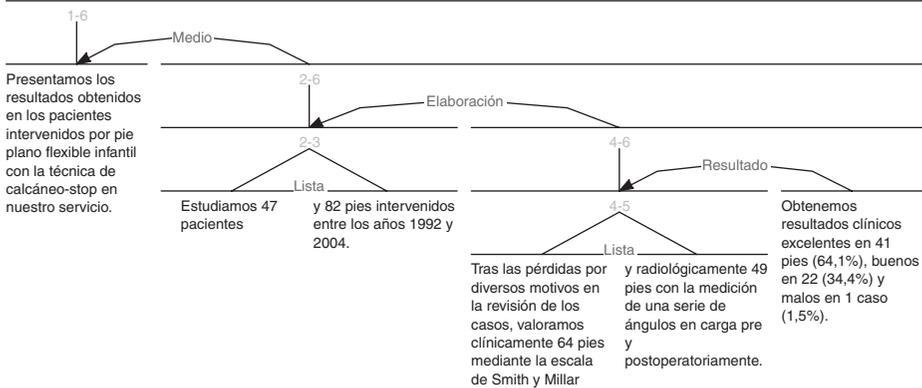
Since our comparison had to be manually done, we considered it appropriate to carry out this task of EDU homogenization so that annotators could label the same segments, establish relations among them, build the rhetorical trees and, finally, carry out the comparison among them in a more accurate way.

**2.3.2. Rhetorical analysis.** In this stage, each annotator labeled the homogenized segmentation of the studied abstracts, marking rhetorical relations among EDUs and determining which of these EDUs were nuclei or satellites. To this end, the RSTTool and the extended classification of rhetorical relations were used.

### 2.4. Quantitative analysis

After the annotation, a quantitative analysis about the two aspects detailed in the previous section was performed.

**2.4.1. EDU segmentation.** The contrast between the EDU segmentation of both annotators was carried out by evaluating precision and recall. To measure precision, we observed the coincidence between the selected EDUs by A2 and the selected EDUs by A1. To



**Figure 2.** Rhetorical tree in Spanish by A1

measure recall, we compared the number of detected EDUs by A2 with the number of detected EDUs by A1. This analysis was carried out, on the one hand, for each individual text and, on the other hand, for the set of texts of our corpus.

**2.4.2. Rhetorical analysis.** To quantify the agreement between the rhetorical analyses by both annotators, we used Marcu's (2000b) method. Specifically, we obtained data concerning detected spans (i.e. sets of related EDUs), nuclearity and rhetorical relations.

To compare both rhetorical analyses, precision and recall were measured again. To measure precision, we counted the number of detected spans, nuclei and satellites, and rhetorical relations marked by A2 coinciding with the ones selected by A1. To measure recall, we counted the total number of the same elements detected by A2, with regard to the total number detected by A1. Once again, this analysis was performed for each text and for the texts of our corpus taken together. For instance, Figure 2 shows a rhetorical tree fragment in Spanish carried out by A1, whereas Figure 3 shows the rhetorical tree of the same passage in Basque, carried out by A2. The English abstract passage of the author that corresponds with this text is provided in here, in order to make the example more understandable to the reader:<sup>8</sup>

English translation:

Unit 1: [We report our experience and the results obtained with surgical treatment of infantile flexible flin foot using the calcaneus-stop technique.]

Unit 2: [From 1992 through 2004, 47 patients]

Unit 3: [and 82 feet were studied.]

Unit 4: [After our revision, 64 feet were evaluated clinically using the Smith and Millar scale]

Unit 5: [and 49 feet were evaluated radiologically by several preoperative and postoperative radiological variables.]

Unit 6: [The clinical results were excellent in 41 feet (64.1%), good in 22 feet (34.4%) and bad in only case (1.5%).]

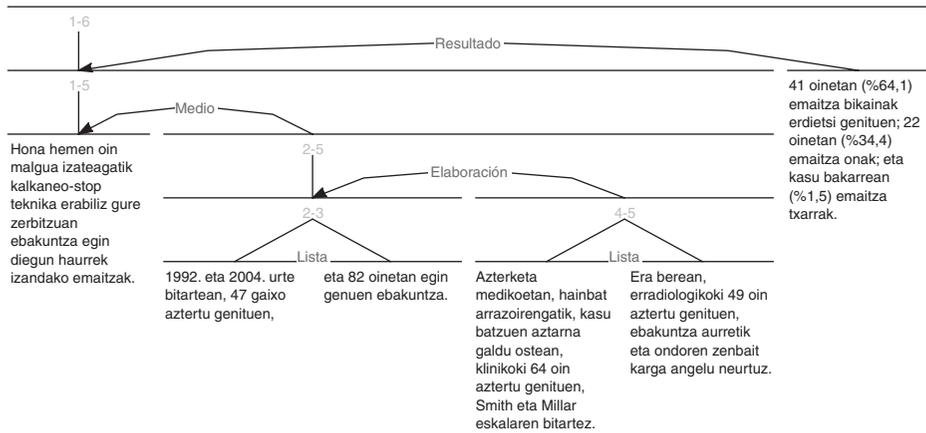


Figure 3. Rhetorical tree in Basque by A2

Table 1 below exemplifies Marcu’s (2000b) evaluation methodology. It includes a comparison of detected spans, nuclearity and relations annotated by A1 and A2. We have used the NUCLEUS<sup>9</sup> label to refer to the nuclei of nuclear relations, and the relation name (e.g. Result, Elaboration, Means, List, etc.) to refer either to the satellites of nuclear relations or to the nuclei of multinuclear relations. It is necessary to take into account that, since we homogenized the EDUs in the segmentation stage (see section 2.3.1.), the detected EDUs by A1 and A2 always coincided. In Table 1 we have indicated in grey the differences between both annotators, where nuclei are denoted by ‘N’ and satellites by ‘S’.

Table 1. Quantitative evaluation using Marcu’s (2000b) method

| Element | EDU |    | Span |    | Nuclearity |    | Relation    |             |
|---------|-----|----|------|----|------------|----|-------------|-------------|
|         | A1  | A2 | A1   | A2 | A1         | A2 | A1          | A2          |
| 1-1     | X   | X  | X    | X  | N          | N  | NUCLEUS     | NUCLEUS     |
| 2-2     | X   | X  | X    | X  | N          | N  | LIST        | LIST        |
| 3-3     | X   | X  | X    | X  | N          | N  | LIST        | LIST        |
| 4-4     | X   | X  | X    | X  | N          | N  | LIST        | LIST        |
| 5-5     | X   | X  | X    | X  | N          | N  | LIST        | LIST        |
| 6-6     | X   | X  | X    | X  | S          | S  | RESULT      | RESULT      |
| 4-5     | -   | -  | X    | X  | N          | S  | NUCLEUS     | ELABORATION |
| 4-6     | -   | -  | X    | -  | S          | -  | ELABORATION | -           |
| 2-3     | -   | -  | X    | X  | N          | N  | NUCLEUS     | NUCLEUS     |
| 2-6     | -   | -  | X    | -  | S          | -  | MEANS       | -           |
| 2-5     | -   | -  | -    | X  | -          | S  | -           | MEANS       |
| 1-5     | -   | -  | -    | X  | -          | N  | -           | NUCLEUS     |

**Table 2.** Quantitative evaluation results of rhetorical trees showed in Figures 2 and 3

|            | Recall | Precision |
|------------|--------|-----------|
| Spans      | 100%   | 80%       |
| Nuclearity | 100%   | 70%       |
| Relations  | 100%   | 70%       |

After the data were formalized with this method, we measured precision and recall, in the way explained above. Table 2 shows the results of this evaluation. The three factors obtain 100 percent of recall, whereas precision oscillates between 80 percent (spans) and 70 percent (nuclearity and rhetorical relations).

## 2.5. Qualitative analysis

As for qualitative analysis, we also focused on questions concerning EDU segmentation and rhetorical analysis.

**2.5.1. EDU segmentation.** After we quantified the differences of EDU segmentation by both annotators, we observed the specific cases on which they differed and we investigated the possible reasons for disagreement.

We observed that, when homogenizing EDUs, some aspects contradicted the established guidelines of segmentation. This is due to the fact that translation strategies also affect segmentation. For instance, some passages are considered as a single EDU in Spanish, but they have been segmented into two units in order to carry out the homogenization:

13a. [*Se realiza el estudio de la proteína 14-3-3, que resulta ser positivo.*]

ENGLISH TRANSLATION: [The study of 14-3-3 protein is carried out, which obtains positive results.]

13b. [*14-3-3 proteinaren azterketa egin zaio,*] [*eta emaitza positiboak lortu dira.*]

ENGLISH TRANSLATION: [The study of 14-3-3 protein is carried out,] [and its results are positive.]

Example 13a above shows that A1 annotated the Spanish passage as a single EDU, since relative clauses are not considered as EDUs. However, in example 13b, we observe that in Basque this relative clause was translated like a main sentence, related to the previous one by means of a discourse marker, the coordinative conjunction *eta* ('and'). In order to homogenize the segments, we decided to divide the Spanish EDU into two EDUs, as follows:

13c. [*Se realiza el estudio de la proteína 14-3-3,*] [*que resulta ser positivo.*]

ENGLISH TRANSLATION: [The study of 14-3-3 protein is carried out,] [which obtains positive results.]

**Table 3.** Qualitative partial evaluation of spans and nuclearity<sup>a</sup>

| Element |     | Span |    | Nuclearity |    |
|---------|-----|------|----|------------|----|
| A1      | A2  | A1   | A2 | A1         | A2 |
| 4-5     | 4-5 | X    | X  | S          | S  |
| 2-3     | 2-3 | X    | X  | N          | N  |
| 2-6     | 2-5 | X    | X  | S          | S  |
| 1-6     | 1-5 | X    | X  | N          | N  |
| 4-6     | 1-6 | -    | X  | S          | S  |

<sup>a</sup>The nuclei and the satellites are denoted by N and S, respectively.

**Table 4.** Qualitative partial evaluation of relations

| Annotated relations |             |
|---------------------|-------------|
| A1                  | A2          |
| Elaboration         | Elaboration |
| List                | List        |
| Means               | Means       |
| List                | List        |
| Result              | Result      |

Both annotators marked the same relation for this passage: the Result relation. This is due to the fact that there is the verb ‘result’ into the second EDU, and it produces more effect than the syntactic structure or the discourse marker. Probably, if there was another verb, the Elaboration relation would be considered in Spanish because of the relative clause, and the List relation would be considered in Basque because of the conjunction.

**2.5.2. Rhetorical analysis.** Though the evaluation method of Marcu (2000b) exemplified in section 2.4.2 is considered to be valid, the method only considers the absolute agreement in all factors. Thus, a disagreement on the segmentation or a disagreement on the lower spans will affect significantly the agreement on the upper rhetorical relations of a tree. For example, if we follow Marcu’s (2000b) method, disagreement with regard to spans, nuclearity and relations is observed. However, the five relations that were marked by both annotators coincide. In fact, there are differences concerning the detected nodes, but not with regard to the detected relations. We consider it necessary to also carry out this type of approach, more optimistic in a certain way and that we call ‘qualitative partial evaluation’, because we believe this approach to be necessary in order to detect and analyze the linguistic differences in rhetorical structure that are originated by translation strategies. Tables 3 and 4 include the data of this evaluation, concerning, in the first place, spans and nuclearity and, in the second place, relations.<sup>10</sup>

**Table 5.** Qualitative partial evaluation results of rhetorical trees showed in Figures 2 and 3

|            | Recall | Precision |
|------------|--------|-----------|
| Spans      | 100%   | 80%       |
| Nuclearity | 100%   | 100%      |
| Relations  | 100%   | 100%      |

Table 5 shows the qualitative partial evaluation results of the example. We notice that precision and recall are 100 percent in all cases, except for precision in spans, which is 80 percent.

Since we could obtain quantitative results concerning spans and nuclearity with Marcu's (2000b) method, we only focused on the qualitative partial evaluation of rhetorical relations. We think this qualitative evaluation is an effective way to detect the linguistic differences affecting rhetorical structure.

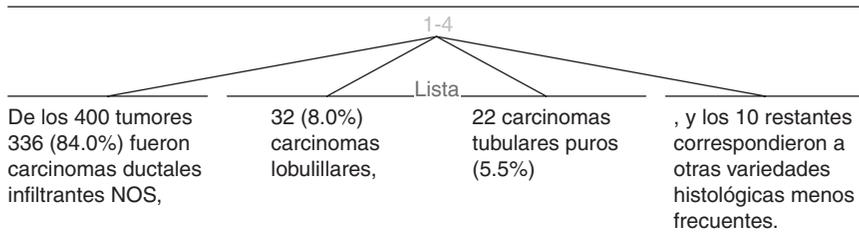
In the qualitative partial evaluation we systematically analyzed the causes of the disagreement between annotators. On the one hand, we observed the phenomena that could cause differences concerning the annotation agreement, mentioned by Mann and Thompson (1988): ambiguity of text structure, simultaneous analyses and analytic mistakes, among others. On the other hand, we analyzed the phenomenon reflected in Marcu et al. (2000: 10), consisting of changing the type of rhetorical relation when translating:

Hence, the mappings in (4) provide an explicit representation of the way information is re-ordered and re-packaged when translated from Japanese into English. However, when translating text, it is also the case that the rhetorical rendering changes. What is realized in Japanese using a CONTRAST relation can be realized in English using, for example, a COMPARISON or a CONCESSION relation.

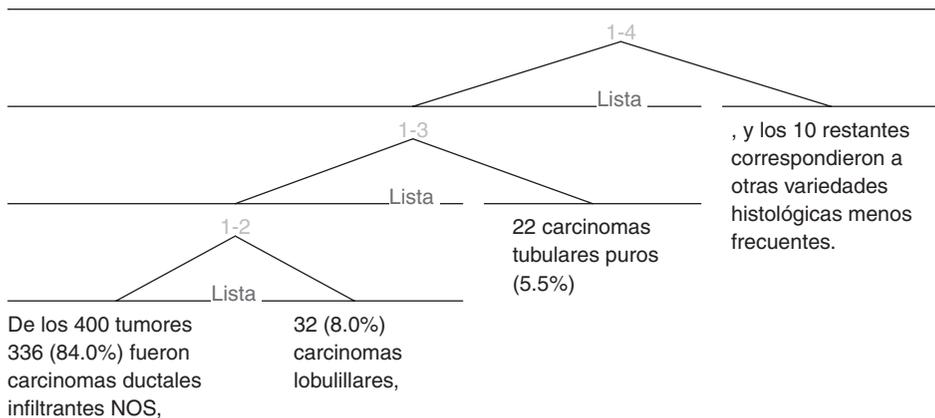
In this way, we detected the possible causes of discrepancies among annotators and the influence that translation strategies have on rhetorical structure (as explained in section 3.2.).

In order to count all the relations, we decided to consider each nuclear relation as one relation, while we considered multinuclear relations as binary ones. For example, a List relation with four nuclei is represented by joining its nuclei in a binary way, obtaining three multinuclear relations, each one with two nuclei. Figures 4 and 5 show respectively the Same-level annotation and the binary annotation of this List relation.

By these means, apart from correctly counting multinuclear relations, we could compare, for example, a) three units or spans of a List relation with three nuclei (by A1) with b) a List relation with two nuclei and one Elaboration relation (by A2). If we had not done it in that way, we would not have been able to compare a List relation by A1 with a List relation and an Elaboration relation by A2, and the evaluation could have lost precision. Moreover, it would not be correct to count as relations all the nuclear elements of a List relation, since multinuclear relations would then be more relevant than the others in the qualitative partial evaluation.



**Figure 4.** Same-level annotation of List relation



**Figure 5.** Binary interpretation of List relation

### 3. Results

In the previous sections the methodology of our experiment was presented. In this section we present segmentation and nucleus-satellite issues, with their corresponding results of agreement, and a discussion of the used translation strategies.

#### 3.1. Segmentation issues

The number of segmented EDUs by A1 in Spanish texts is 206, while the number of segmented EDUs by A2 in Basque texts is 238. We think there are more EDUs in Basque than in Spanish because Basque nominalization and subordination work with different syntactic procedures (Arakama et al., 2005). Arakama et al. (2005) state that some comprehension problems arise with literal translations of Spanish relatives. To avoid this problem, there is more than one translation strategy, one of them being the splitting of sentences. Language typology has an influence when nominalization is done, because Basque typology uses more verbs than nominalization, given that the ellipsis of verbal arguments is common in Basque (due to verb concordance). Thus, literal translation has no sense or comprehension problems arise.

Both annotators agreed on 152 EDUs. Following the explained methodology in section 2.4.1., we obtained precision (63.9%) and recall (86.6%) of the performed segmentation. The sources of disagreement are linguistic differences, being mainly motivated by translation strategies (85 cases) from Spanish to Basque, which we explore in detail in this section.

We noticed that, sometimes, linguistic differences between texts in Basque and Spanish cause a different segmentation of the same passage by annotators (see example 14).

14a. [*Hemos estudiado retrospectivamente 23 infecciones protésicas de rodilla tratadas en nuestro hospital entre el año 1996 y el 2004 de las cuales hemos excluido 6 por diferentes motivos.*]

ENGLISH TRANSLATION: [We retrospectively have studied 23 prosthetic knee infections that were treated in our hospital between 1996 and 2004 of which we have excluded 6 for different reasons.]

14b. [*1996. eta 2004. urteen bitartean gure ospitalean izandako 23 infekzio protesiko aztertu ditugu.*] [*Horien artean, 6 kasu baztertu ditugu hainbat arrazoiengatik.*]

ENGLISH TRANSLATION: [We have studied 23 prosthetic knee infections that were treated in our hospital between 1996 and 2004.] [Of these, we have excluded 6 for different reasons.]

In example 14a, we observe that A1 has established a single EDU in Spanish while, in example 14b, we notice that A2 has segmented the same passage in two EDUs. This disagreement on the segmentation phase is due to two facts: a) the relative clause is not considered as an EDU and b) the syntactic structure of the relative clause has been translated into Basque as a different sentence by using punctuation.

When the evaluation of the segmentation was carried out, the same difficulty mentioned by Carlson and Marcu (2001: 2) was found: they declare that the boundary between discourse and syntax can be very blurry. We think this fact is more prominent when structures of two languages are compared:

The first step in characterizing the discourse structure of a text in our protocol is to determine the elementary discourse units (EDUs), which are the minimal building blocks of a discourse tree. Mann and Thompson (1988, p. 244) state that ‘RST provides a general way to describe the relations among clauses in a text, whether or not they are grammatically or lexically signalled.’ Yet, applying this intuitive notion to the task of producing a large, consistently annotated corpus is extremely difficult, because the boundary between discourse and syntax can be very blurry.

Indeed, translation strategies are one of the causes influencing segmentation decisions. Consider example 15 below:

15a. [*Se han estudiado un total de 442 cánceres de mama unifocales de 2 cm o menos en la pieza histológica (pT1) operados entre enero de 1993 y diciembre de 2005.*]

ENGLISH TRANSLATION: [We have studied a total of 442 unifocal breast cancers of 2 cm or less in the histological part (pT1) operated between January 1993 and December 2005.]

15b. [*Guztira, foku bakarreko 442 bularreko minbizi aztertu dira, pieza histologikoan (pT1) 2 cm edo gutxiago dituztenak.*] [*Guztiak 1993ko urtarrilaren eta 2005eko abenduaren artean operatu ziren.*]

ENGLISH TRANSLATION: [We have studied a total of 442 unifocal breast cancers of 2 cm or less in histological part (pT1).] [All of them *underwent surgery* between January 1993 and December 2005.]

In this example, the non-finite verb (the participle form *operado* ['operated']) was translated into Basque like a finite verb (*operatu ziren* ['underwent surgery']). Besides, the sentence was separated by a full stop. These two facts strongly affect the segmentation in both languages.

We observe various translation strategies affecting the performed segmentation by both annotators, which we explore in detail in section 3.3. It is noteworthy that there is almost a total segmentation agreement concerning EDUs that were not influenced by translation strategies. Segmentation errors of annotators were minimal in these cases.

### 3.2. Nucleus-satellite issues

Disagreement with regard to the choice of nucleus and satellite is an interesting point of RST. On the one hand, the choice depends on the way the information is presented or the linguistic forms are employed (Marcu, 1999). On the other hand, the choice also depends on the context or the point of view of the whole text (Bateman and Rondhuis, 1997). Stede (2008: 317) criticizes RST because trees do not make the source of the choice explicit:

The final RST tree does not indicate whether some relation at the level of minimal units is there because its definition is optimally fulfilled or because text global factors make it seem advantageous to select one particular nucleus, which is incidentally performed by that particular relation.

As described in section 2.4.2. above, we measured precision and recall to assess the agreement between the two annotators on spans, nuclearity and rhetorical relations. Table 6 shows an overall result for the 20 texts of the corpus. We noted that results in terms of recall are similar, which is due to EDU homogenization, explained in section 2.3.1. However, results regarding precision vary. Despite this fact, the precision achieved is substantially high in all cases: the agreement between the annotated spans is 92.5 percent, the agreement on nuclearity is 82.1 percent and the agreement regarding the relations is 68.3 percent.

**Table 6.** Results of the quantitative evaluation

|            | Recall | Precision |
|------------|--------|-----------|
| Spans      | 98.6%  | 92.5%     |
| Nuclearity | 98.6%  | 82.1%     |
| Relations  | 98.6%  | 68.3%     |

Concerning rhetorical analysis, we mainly observed two types of situations:

1) Ambiguity or different interpretations when choosing relations: Annotators labeled differently some relations that could be ambiguous. For instance, in example 16, while A1 annotated a relation of Background, A2 annotated a relation of Elaboration for the same passage.

16a. *[Han participado 92 pacientes ingresados en un Área Médica del Hospital de Basurto (Bilbao).]N [Todos los pacientes fueron entrevistados para elaborar la historia patopsicobiográfica necesaria para aplicar la Clasificación Psicósomática de Pierre Marty.]S\_ Elaboración*

ENGLISH TRANSLATION: [92 patients admitted in a Medical Area Hospital de Basurto (Bilbao) have been involved.]N [All these patients were interviewed to develop the patopsicobiographic history that is needed to apply the Psychosomatic Classification of Pierre Marty.]S\_ Elaboration

16b. *[Basurtoko (Bilbo) Ospitaleko Medikuntza Arlo batean ospitaleratuta dauden 92 gaixok parte hartu dute.]S\_Fondo [Pierre Martyren Saïlkapen Psikosomatikoa aplikatzeko beharrezkoa den historia patopsikobiografikoa egiteko asmoz, elkarrizketa egin zitzaïen gaixo guztiei.]N*

ENGLISH TRANSLATION: [92 patients admitted in a Medical Area Hospital of Basurto (Bilbao) have been involved.]S\_Background [All these patients were interviewed to develop the patopsicobiographic history that is needed to apply the Psychosomatic Classification of Pierre Marty.]N

In this case, a disagreement regarding the nuclearity of the relation entails a different interpretation about the existing relation between two EDUs. In the example above the nucleus of the Spanish text is the first EDU (the participants of study) (16a), whereas the nucleus of the Basque text is the second EDU (the research methodology) (16b).

Consider other examples:

17a. *[Se estima que el 80% de los usuarios acuden por iniciativa propia a los servicios de urgencia]N\_Lista [y que el 70% de las consultas son consideradas leves por el personal sanitario.]N\_Lista*

ENGLISH TRANSLATION: [It is calculated that 80% of visitors come to emergency services by their own initiative]N\_List [and that 70% of consultations are considered like mild by the health staff.]N\_List

17b. *[Erabiltzaileen %80ak bere kabuz erabakitzen dute larrialdi zerbitzu batetara jotzea]N [eta kontsulta hauen %70a larritasun gutxikotzat jotzen dituzte zerbitzu hauetako medikuek.]S\_ Elaboración*

ENGLISH TRANSLATION: [80% of visitors come to emergency services by their own initiative]N [and 70% of consultations are considered like mild by the health staff.]S\_ Elaboration

In example 17 there was also a disagreement concerning nuclearity. However, in this case, the disagreement affects the nature of the relation: A1 annotated a paratactic relation of List (17a), while A2 annotated a hypotactic relation of Elaboration (17b).

18a. *[Por lo demás existen buenos indicadores de proceso]S\_Antítesis [pero se aprecia un escaso registro de la capacidad funcional del paciente al alta, que dificulta la comparación de los resultados de la atención sanitaria.]N*

ENGLISH TRANSLATION: [In addition, there are good indicators of the process]S\_Antithesis [but we see a poor record of the patient's functional ability to discharge, which makes the comparison of health care results difficult.]N

18b. [*Gainerakoan, prozesu adierazle egokiak daude,*]N [*baina altan dagoen gaixoaren lanen funtzionalaren erregistro urria antzematen da, eta horrek osasun arretaren emaitzen alderaketa zailtzen du.*]S\_Concesión

ENGLISH TRANSLATION: [In addition, there are good indicators of the process]N [but we see a poor record of the patient's functional ability to discharge, and this makes the comparison of health care results difficult.]N\_Concession

In example 18 the disagreement is due to the different meanings of the relation. Both annotators selected a hypotactic relation of presentation but, while A1 annotated an Antithesis relation (18a), A2 annotated a Concession relation (18b).

In this example, the disagreement is not due to the translation, since linguistic forms involved in the relation are identical, including the translation of the discourse marker 'but' (*pero* in Spanish and *baina* in Basque). Thus, we wonder which the source of the disagreement is: is it really a problem of relations definition or maybe a more general problem? This situation was considered by Stede (2008: 318):

Consider as one example the definitions of Antithesis and Concession. The constraints on the nucleus and the intentions of the writer (i.e., the 'effect') are identical. Antithesis has no constraint on the satellite, whereas Concession offers the constraint that 'writer is not claiming that satellite does not hold'. (Since Antithesis has no constraint here, does it properly subsume Concession?) Finally, the constraints on the nucleus/satellite combinations are largely paraphrastic with the one exception that Antithesis adds that 'one cannot have positive regard for both situations' (in nucleus and satellite). In total, the differences are not very restrictive, so that in many contexts both definitions are equally applicable. But, in the presentational/subject-division of the relations suggested by Mann and Thompson, Antithesis appears in the former, and Concession in the latter, despite their effects being identical. So it is not clear on what grounds the grouping is made in this case.

2) Differences regarding Spanish–Basque translation strategies: the linguistic differences between these two languages sometimes imply that annotators interpret the same passage differently (see examples 19 and 20).

19a. [*Escogiendo la especialidad más barata existente en el mercado*]S\_Circunstancia [*podríamos alcanzar un ahorro de 6.463.400,35€.*]N

ENGLISH TRANSLATION: [Choosing the cheapest specialty in the market]S\_Circumstance [we could achieve a saving of 6,463,400.35€.]N

19b. [*Merkatuak eskaintzen digun espezialitate merkeena aukeratuko bagenu*]S\_Condición [*6.463.400,35€-ko aurrezpena lortuko genuke.*]N

ENGLISH TRANSLATION: [If we chose the cheapest specialty in the market]S\_Condition [we would achieve a saving of 6,463,400.35€.]N

The gerund form (*escogiendo* ['choosing']) may indicate the relation of Circumstance in Spanish. But in Basque no gerund is included in the sentence; the conditional mark (*ba-* ['if']) in the verb (*bagenu* ['(we) chose']) justifies the annotation of the relation of Condition.

20a. [En los 7 ítems se han encontrado diferencias estadísticamente significativas entre el grupo de pacientes oncológicos con los pacientes afectados de otro tipo de patologías ( $p < 0.05$ ).]N [Estos ítems diferencian a los pacientes con neoplasias de otro tipo de pacientes, y permiten una valoración global de los mismos, ofreciendo una idea de las expectativas del proceso.]S\_Elaboración

ENGLISH TRANSLATION: [In the 7 items we have found statistically significant differences between the group of cancer patients and patients suffering from other pathologies ( $p < 0.05$ ).]N [These items differentiate patients with tumors from other patients, and they allow an overall assessment of the patients, providing an idea of the process prospects.]S\_Elaboration

20b. [7 ítemak aztertuta, estatistikoki desberdintasun aipagarriak aurkitu ziren gaixo onkologikoen eta bestelako patologiak dituzten gaixoen artean ( $p < 0.05$ ).]N\_Unión [Horrez gain, ítem horiek neoplasiak dituzten gaixoak eta bestelako gaixoak bereizten dituzte, horiei buruzko balorazio orokorra egiteko aukera ematen dute, eta prozesuaren igurkapenen gaineko argibideak ematen dizkigute.]N\_Unión

ENGLISH TRANSLATION: [Having studied the 7 items, we have found statistically significant differences between the group of cancer patients and patients suffering from other pathologies ( $p < 0.05$ ).]N\_Joint [In addition, these items differentiate patients with tumors and other patients, they allow an overall assessment of the patients, and they provide an idea of the process prospects.]N\_Joint

In Spanish, the relation of Elaboration was annotated due to the presence of the anaphora. The semantic relation between both EDUs shows an elaboration of the same topic. Nevertheless, in Basque, the additive connector *horrez gain* ('in addition') does not allow inclusion of both EDUs in the same argumentative scale (Cuartero, 1995), since it introduces a new topic in the speech. This fact causes A2 to select a multinuclear relation. Therefore, it is evident that a different translation strategy affects the rhetorical analysis of the text.

We studied this phenomenon systematically, which we explain in detail in section 3.3.

### 3.3. Discussion of translation strategies

As we have said in section 3.1, translation strategies are one of the causes influencing segmentation decisions. We observe various translation strategies affecting the performed segmentation by both annotators. Specifically, the authors of the texts used two main strategies to translate from Spanish into Basque. These two strategies constitute the 74.28 percent of all the translation strategies.

- Relative subordinate clauses in Spanish have been translated as separate sentences in Basque.
- Missing elements from ellipsis and anaphors in Spanish are retaken in Basque, forming new sentences.

The consequences of these translation strategies are:

- There are more EDUs in Basque than in Spanish. Specifically, in our corpus, there are 13.45 percent more EDUs in Basque than in Spanish.

- This difference between EDUs in the two languages significantly affects the agreement on the segmentation, and therefore it affects in a gradual way the other annotation levels and evaluated factors (spans, nuclearity and relations) as well. This fact makes quantitative and qualitative evaluation more difficult to perform.

As we have said in section 3.2, translation strategies may be the cause of a different rhetorical analysis. We include in Table 7 the used strategies to translate from Spanish into Basque, with their frequencies.

Three of these translation strategies are mentioned in Arakama et al. (2005): completing ellipsis and/or dividing sentences, using a finite verb and deleting relative clauses. Another of these strategies is used when the translator wants to provide more coherence to the translation: using discourse markers (Zabala, 1996).

We provide some examples herein:

a) Completing ellipsis and/or dividing sentences:

21a. *[Todos los pacientes presentaban una insuficiencia ventilatoria, en 10 casos de tipo obstructivo y en los restantes de tipo no obstructivo o mixto.]*

ENGLISH TRANSLATION: [All patients had ventilatory failure, 10 cases of obstructive type and the remaining of non-obstructive or mixed type.]

21b. *[Gaixo guztiek zeukaten aireztapen gutxiegitasuna:] [hamar kasutan butxaketa-motakoa zen] [eta gainerakoetan ezbutxaketakoa edo mistoa zen.]*

ENGLISH TRANSLATION: [All patients had ventilatory failure;] [10 cases were of obstructive type] [and the remaining were of non-obstructive or mixed type.]

In this example, the translation strategy was in Basque to complete the ellipsis of verbs describing the cases of ‘ventilatory failure’.

b) Using a finite verb:

22a. *[Estudiamos 47 pacientes y 82 pies intervenidos entre los años 1992 y 2004.]*

ENGLISH TRANSLATION: [We studied 47 patients and 82 feet *undergoing surgery* between 1992 and 2004.]

22b. *[1992. eta 2004. urte bitartean, 47 gaixo aztertu genituen,] [eta 82 oinetan egin genuen ebakuntza.]*

ENGLISH TRANSLATION: [Between 1992 and 2004, we studied 47 patients] [and we *operated* 82 feet.]

**Table 7.** Translation strategies determining different rhetorical relations

| Translation strategies                           | Spanish | Basque | Total |
|--|---------|--------|-------|
| a) Completing ellipsis and/or dividing sentences | 1       | 5      | 6     |
| b) Using a finite verb                           | 0       | 5      | 5     |
| c) Using discourse markers                       | 2       | 7      | 9     |
| d) Deleting relative clauses                     | 0       | 6      | 6     |
| e) Other strategies                              | 0       | 5      | 5     |
| Total  | 3       | 28     | 31    |

The Spanish participle (*intervenidos* [‘undergoing surgery’]) was translated into Basque by a structure with a finite verb and its direct object (*ebakuntza egin genuen* [‘(we) operated’]).

23a. [*Nuestros resultados sugieren la presencia de alteraciones respiratorias crónicas con el resultado de un déficit ventilatorio, varias décadas después del tratamiento con colapsoterapia; comprobando una buena respuesta al tratamiento con ventilación domiciliaria.*]

ENGLISH TRANSLATION: [Our results suggest the presence of chronic respiratory disorders with the result of a ventilatory deficit, several decades after treatment with Collapse Therapy; *proving* a good response to treatment with home ventilation.]

23b. [*Gure emaitzek iradokitzen dute kolapsoterapiarekin egindako tratemendutik hamarkada batzuk gerago arnas alterazio kronikoak daudela aireztapen déficit baten emaitzarekin;] [eta egiaztatu da etxeko aireztapenarekin egindako tratamenduak erantzun ona izan duela.]*

ENGLISH TRANSLATION: [Our results suggest the presence of chronic respiratory disorders with the result of a ventilatory deficit, several decades after treatment with Collapse Therapy;] [and a good response to treatment with home ventilation *has been proved.*]

In this example, the Spanish gerund (*comprobando* [‘proving’]) was translated into Basque by the finite verb (*egiaztatu da* [‘(it) has been proved’]).

c) Using discourse markers:

24a. [*Como cirugía primaria presenta una mortalidad del 0,5%*] [*y un 8,8% de complicaciones perioperatorias, destacando la hemorragia (4,8%) y la dehiscencia anastomótica (1,7%).*]

ENGLISH TRANSLATION: [As primary surgery, it presents a mortality of a 0.5%] [and a 8.8% of perioperative complications, standing out hemorrhages (4.8%) and dehiscence of anastomosis (1.7%).]

24b. [*Kirurgia mota honetan, heriotza tasa % 0,5ekoa da.*] [*eta ebakuntza osteko arazoak, berriz, % 8,8koak dira: odoljariora (% 4,8) eta dehiszentzia anastomotikoa (% 1,7).*]

ENGLISH TRANSLATION: [In this type of surgery, the mortality rate is 0.5%] [while the perioperative complications are 8.8%: haemorrhages (4.8%) and dehiscence of anastomosis (1.7%).]

The use of the Basque counterargument connector *berriz* (‘while’) shows a contrast, not a contradiction. This connector means that A2 labels this passage with a Contrast relation, while A1 labels the same passage with List relation, because he did not have any discourse marker.

d) Deleting relative clauses:

25a. [*Creemos que es importante dar a nuestros pacientes una información previa a la exploración lo más precisa posible, que sea capaz de resolver todas las posibles dudas que les plantee y que les permita afrontarla con tranquilidad.*]

ENGLISH TRANSLATION: [We think that it is important to give our patients a pre-scan information as accurate as possible, being able to resolve all the possible doubts raised by it and allowing them to deal with it peacefully.]

25b. [*Garrantzitsua iruditzen zaigu azterketa egin baino lehen, gaixoei informazio zehatza aurreratzea.*] [*Horrela, bere zalantzak argituz, hobeto egingo diote aurre azterketari.*]

ENGLISH TRANSLATION: [We think that it is important to give our patients a pre-scan information as accurate as possible.] [In this way, resolving their doubts, they will deal better with the medical examination.]

**Table 8.** Data of the partial qualitative evaluation

|                            | Absolute data | %     |
|----------------------------|---------------|-------|
| Total relations            | 224           | 100%  |
| Agreement on relations     | 157           | 71%   |
| Disagreements on relations | 65            | 29%   |
| Translation source         | 31            | 13.8% |
| Interpretation source      | 34            | 15.2% |

In this example, the literal translation of the relative clause used in Spanish was avoided in Basque and it was translated by an independent sentence with a finite verb (*aurre egingo diote* [(they) will deal with]).

Once all the cases have been described, we conclude that the use of the detected translation strategies is due to the fact that Basque sentences have the semantic load at the end of the sentence, since it is an SOV language. In order to facilitate the understanding, the translator has to locate the semantic load earlier in the sentence or has to reduce the size of it. In this corpus more sentences in Basque than in Spanish were used to facilitate the understanding of the semantic content. Precisely for this reason (to shorten sentences), some translation strategies were used in Basque. The use of these strategies definitely increases the linguistic differences that affect the rhetorical structure, changing the relations among EDUs and, thus, changing sometimes the meaning of the text or, at least, the presentation of the information. If the meaning of the text is different, it is normal that the disagreement between the annotators increases and, thanks to the partial qualitative evaluation, this great increase in the disagreement becomes an indicator of translation techniques.

Table 8 shows the data of the partial qualitative evaluation that we performed in this work.

Finally, Table 9 provides recall and precision of the quantitative evaluations, and recall of the qualitative evaluation. It is noticed that the precision of both evaluations is very similar (68.3% in the quantitative evaluation and 71% in the qualitative evaluation).

As it is shown in Table 9, the precision of the qualitative evaluation from the comparison of the 20 rhetorical trees of the corpus is more optimistic than the quantitative one, but not too much (only 2.7% more). However, this situation is not constant, since in some trees the difference between evaluations ranges approximately from -10% to +10%.

Although the use of translation strategies definitely affects rhetorical structures, it does not seem to affect the texts' superstructure, since both annotators have constructed a very similar superstructure for both languages. The macrostructure of a text is, according to van Dijk (1980, 1989), an abstract representation which tends to the overall understanding of the meaning of the text, while the superstructure is the organizational structure of the text, which can vary depending on the type of the text. Van Dijk (1989) described the superstructure of various types of texts, for example scientific texts, and he stated that:

*En los discursos científicos se presenta una variante especial de las superestructuras argumentativas [ . . . ]. La estructura básica del discurso científico no (sólo) consiste en una CONCLUSIÓN y su JUSTIFICACIÓN, sino también en un PLANTEO DEL PROBLEMA y una SOLUCIÓN. (van Dijk, 1989: 164)*

ENGLISH TRANSLATION: Scientific discourse provides a special variant of argumentative superstructures [ . . . ]. The basic structure of scientific discourse is not (only) a CONCLUSION and its JUSTIFICATION, but also a PROBLEM STATEMENT and a SOLUTION. (van Dijk, 1989: 164)

**Table 9.** Final results of quantitative evaluation and partial qualitative evaluation

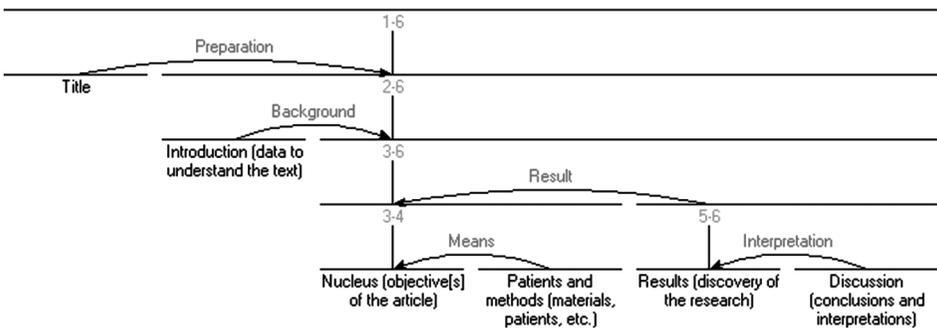
|           | Quantitative |           | Qualitative |
|-----------|--------------|-----------|-------------|
|           | Recall       | Precision | Precision   |
| Relations | 98.6%        | 68.3 %    | 71%         |

For example, van Dijk (1989) analyzed the superstructure of the *Experimental Report*, finding in it some *observations*, an *explanation*, a *hypothesis*, an *experiment*, etc. In this work we also analyze a scientific discourse but, as we have already discussed, our corpus of analysis includes abstracts of original articles, specifically from the medical field. These abstracts maintain the same superstructure of the articles that are related to them and, therefore, they have four main sections: *Introduction*, *Patients and methods*, *Results* and *Discussion*. This structure was labeled exactly by both annotators, by means of RST relations as Background, Means, Result and Interpretation. Figure 6 shows a diagram of this structure.

### 4. Conclusions

To conclude, we think that this work represents a new contribution concerning RST, since it extends our understanding about the comparison of rhetorical trees in various languages, specifically the comparison between Spanish and Basque, that had not been made before. We have mentioned some problems of quantitative evaluation, and an original qualitative evaluation has also been presented. Our work shows that, though there are differences regarding rhetorical analysis performed over the same corpus (with parallel texts in two languages) by two annotators, these are mainly due to the translation strategies being used. However, these strategies do not affect the superstructure of medical abstracts in a decisive way.

Another conclusion of this work is that translation strategies influence the interpretation of RST rhetorical relations. The translator did sometimes not use the same linguistic structures when translating from one language into another. Since the rhetorical structures were not maintained, the two annotators of our study interpreted differently a same passage written in two languages.



**Figure 6.** Main superstructure labeled by both annotators

Likewise, the comparison of rhetorical trees of parallel texts has allowed us to observe two situations: a) when translating an abstract, its rhetorical structure is not taken into account as much as its syntactic structure, and b) in the cases where it is not convenient to translate syntactic structures literally, the used translation strategies provide some clues about how languages usually structure their discourse (which is an issue to take into account for automatic translation of rhetorical structures).

As future work, we would like to compare the top spans of rhetorical structures in order to determine the level of agreement concerning the superstructure, and to analyze the linguistic factors determining the disagreement on rhetorical structure. Although the abstracts are quite short, we think their length is enough to evaluate the agreement of the annotators. Furthermore, we would like to study the reasons for the oscillations between the quantitative and qualitative evaluations, and to also add to this study a third language, English, since, as we have already mentioned, *Gaceta Médica de Bilbao* also includes the abstracts of the authors in that language. We consider that it is important to observe which types of translation strategies have been used and the existing differences among them. As English and Spanish are linguistically more similar, the applied translation strategies should be reduced and, therefore, this variable would decrease when comparing closer languages. In addition, we would like to confirm if medical abstracts in English have the same superstructure. Moreover, we plan to carry out a compilation of discourse markers in Spanish, Basque and English, starting from an empirical analysis of medical abstracts written in these three languages. The main goal of this last study would be to analyze the correlations among rhetorical relations and discourse markers, in the same way that Irukieta et al. (in press) have done.

## Notes

1. <http://www.gacetamedicabilbao.org/web/es/>.
2. The English translation is ours (see <http://www.gacetamedicabilbao.org/web/es/autores.php>).
3. The following examples are proposed by Carlson and Marcu (2001).
4. Throughout this article, examples marked with ‘a’ show the segmentation included in Carlson and Marcu (2001), and examples marked with ‘b’ show the segmentation that we would establish in our work.
5. ‘Deficit’ is part of the unit ‘it will be able to halve this year’s 120 billion ruble’.
6. [http://dialnet.unirioja.es/servlet/revista?tipo\\_busqueda=CODIGO&clave\\_revista=2426](http://dialnet.unirioja.es/servlet/revista?tipo_busqueda=CODIGO&clave_revista=2426).
7. <http://www.wagsoft.com/RSTTool/>.
8. For the purpose of this article, we have tried to do, for the English translation, the EDU segmentation as similar as possible with regard to the one proposed in Spanish and Basque.
9. Marcu (2000b) names them ‘spans’.
10. Note that numerical elements are included in one column in Table 1, while in Table 3 these elements are included in the first two.

## References

- Abelen, E., Redeker, G. and Thompson, S.A. (1993) ‘The Rhetorical Structure of US-American and Dutch Fund-Raising Letters’, *Text* 13(3): 323–350.
- Arakama, J.M., Arrieta, A., Lozano, J., Robles, J. and Urrutia, R.M. (2005) *IVAPeko Estilo Liburua*. Zarautz: IVAP.

- Bateman, J.A. and Rondhuis, K.J. (1997) 'Coherence Relations: Towards a General Specification', *Discourse Processes* 24: 3–50.
- Bouayad-Agha, N. (2000) 'Using an Abstract Rhetorical Representation to Generate a Variety of Pragmatically Congruent Texts', in *Proceedings of the 38<sup>th</sup> Meeting of the Association for Computational Linguistics. Student Workshop*, 16–22.
- Burstein, J. and Marcu, D. (2003) 'A Machine Learning Approach for Identification of Thesis and Conclusion Statements in Student Essays', *Computers and the Humanities* 37(4): 455–467.
- Carlson, L. and Marcu, D. (2001) *Discourse Tagging Reference Manual*. ISI Technical Report ISITR-545. Los Angeles, CA: University of Southern California.
- Carlson, L., Marcu, D. and Okurowski, M.E. (2001) 'Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory', in *Proceedings of the 2<sup>nd</sup> SIGDIAL Workshop on Discourse and Dialogue*. 1–10.
- Cuartero, J.M. (1995) 'El estatuto categorial de *además* y sus propiedades distribucionales', *Dicenda* 13: 103–118.
- Cui, S. (1986) 'A Comparison of English and Chinese Expository Rhetorical Structures', Unpublished Master's thesis, UCLA.
- da Cunha, I. (2008) *Hacia un modelo lingüístico de resumen automático de artículos médicos en español*. Barcelona: IULA. [CD-ROM] (Sèrie Tesis; 23).
- da Cunha, I. and Torres-Moreno, J.-M. (2010) 'Automatic Discourse Segmentation: Review and Perspectives', in *Proceedings of the International Workshop on African Human Languages Technologies*. Djibouti: Institute of Sciences and Information Technologies.
- da Cunha, I., Wanner, L. and Cabré, M.T. (2007) 'Summarization of Specialized Discourse: The Case of Medical Articles in Spanish', *Terminology* 13(2): 249–286.
- Delin, J., Hartley, A. and Scott, D. (1996) 'Towards a Contrastive Pragmatics: Syntactic Choice in English and French Instructions', *Language Sciences* 18(3–4): 897–931.
- Ghorbel, H., Ballim, A. and Coray, G. (2001) 'ROSETTA: Rhetorical and Semantic Environment for Text Alignment', in P. Rayson, A. Wilson, A.M. McEnery, A. Hardie and S. Khoja (eds) *Proceedings of Corpus Linguistics 2001*, pp. 224–233.
- Haouam, K. and Marir, F. (2003) 'SEMIR: Semantic Indexing and Retrieving Web Document using Rhetorical Structure Theory', *Lecture Notes in Computer Science*: 596–604.
- Iruskieta, M., Diaz de Ilarraza, A. and Lersundi, M. (in press) 'Correlaciones en euskera entre las relaciones retóricas y los marcadores del discurso', *Proceedings of 27th AESLA International Conference: Ways and Modes of Human Communication*. Ciudad Real: Universidad de Castilla-La Mancha.
- Kong, K.C.C. (1998) 'Are Simple Business Request Letters Really Simple? A Comparison of Chinese and English Business Request Letters', *Text* 18(1): 103–141.
- Mann, W.C. (2005) *RST Web Site*. Available at: [www.sfu.ca/rst](http://www.sfu.ca/rst) (accessed 15 August 2009).
- Mann, W.C. and Thompson, S.A. (1988) 'Rhetorical structure theory: Toward a functional theory of text organization', *Text* 8(3): 243–281.
- Marcu, D. (1998) 'The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts', PhD thesis, University of Toronto.
- Marcu, D. (1999) *Instructions for manually annotating the discourse structure of texts*. Available at: <http://www.isi.edu/~marcu>.
- Marcu, D. (2000a) *The Theory and Practice of Discourse Parsing Summarization*. Cambridge, MA: Massachusetts Institute of Technology.

- Marcu, D. (2000b) 'The Rhetorical Parsing of Unrestricted Texts: A Surface-Based Approach', *Computational Linguistics* 26(3): 395–448.
- Marcu, D., Amorrtortu, E. and Romera, M. (1999) 'Experiments in Constructing a Corpus of Discourse Trees', in *Proceedings of the ACL Workshop on Standards and Tools for Discourse Tagging*: 48–57.
- Marcu, D., Carlson, L. and Watanabe, M. (2000) 'The Automatic Translation of Discourse Structures', *Proceedings of the First Annual Meeting of the North American Chapter of the Association for Computational Linguistics*, 9–17.
- Mazeiro, E. and Pardo, T.A.S. (2009) 'Metodologia de avaliação automática de estruturas retóricas', in *Proceedings of the 7th Brazilian Symposium in Information and Human Language Technology (STIL 2009)*. São Carlos, São Paulo.
- O'Donnell, M. (2000) 'RSTTOOL 2.4 – A markup tool for rhetorical structure theory', in *Proceedings of the International Natural Language Generation Conference*: 253–256.
- Pardo, T.A.S. and Nunes, M.G.V. (2008) 'On the Development and Evaluation of a Brazilian Portuguese Discourse Parser', *Journal of Theoretical and Applied Computing* 15(2): 43–64.
- Pardo, T.A.S., Nunes, M.G.V. and Rino, L.H.M. (2004) 'DiZer: An Automatic Discourse Analyzer for Brazilian Portuguese', *Lecture Notes in Artificial Intelligence*: 224–234.
- Ramsay, G. (2000) 'Linearity in Rhetorical Organisation: A Comparative Cross-Cultural Analysis of Newstext from the People's Republic of China and Australia', *International Journal of Applied Linguistics* 10(2): 241–258.
- Ramsay, G. (2001) 'What are they Getting At? Placement of Important Ideas in Chinese Newstext: A Contrastive Analysis with Australian Newstext', *Australian Review of Applied Linguistics* 24(2): 17–34.
- Salkie, R. and Oates, S.L. (1999) 'Contrast and Concession in French and English', *Languages in Contrast* 2(1): 27–56.
- Scott, D., Delin, J. and Hartley, A. (1998) 'Identifying Congruent Pragmatic Relations in Procedural Texts', *Languages in Contrast* 1(1): 45–82.
- Skadhauge, P. and Hardt, D. (2005) 'Syntactic Identification of Attribution in the RST Treebank', in *Proceedings of the Sixth International Workshop on Linguistically Interpreted Corpora*. Jeju Island, 57–62.
- Stede, M. (2008) 'Disambiguating Rhetorical Structure', *Journal of Research in Language and Computation* 6: 311–332.
- Sumita, K., Ono, K., Chino, T., Ukita, T. and Amano, S. (1992) 'A Discourse Structure Analyzer for Japanese Text', in *Proceedings of the International Conference on Fifth Generation Computer Systems*, 1133–1140.
- Swales, J. (1990) *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Taboada, M. and Mann, W.C. (2005) 'Applications of Rhetorical Structure Theory', *Discourse Studies* 8(4): 567–588.
- Taboada, M. and Mann, W.C. (2006) 'Rhetorical Structure Theory: Looking Back and Moving Ahead', *Discourse Studies* 8(3): 423–459.
- van Dijk, T.A. (1980) *Macro-Structures. An Interdisciplinary Study of Global Structures in Discourse, Cognitions and Interaction*. Hillsdale, NJ: Lawrence Erlbaum.
- van Dijk, T.A. (1989) *La ciencia del texto*. Barcelona: Paidós.
- Zabala, I. (1996) 'Testu-lotura: lotura tematikoa eta erreferentzia-sareak testu teknikoetan', in *Testu-loturarako baliabideak: euskara teknikoa*, pp. 15–44. Bilbao: EHU.

**Appendix Table 1.** Information about the analyzed corpus<sup>a</sup>

| Reference | Title   | Author(s)   | Year |
|-----------|---|---|------|
| Text 1    | Pharmacoepidemiologic and pharmaco-economic study of arterial hypertension  | L.C. Abecia   | 2008 |
| Text 2    | Serious psychomatic criteria in oncology  | R. Ruiz, A. Aljelani, U. Shelick, U. Usobiaga, J. Muro, J. Bilbao, F. Franco  | 2007 |
| Text 3    | The 'basal-like' (c-erb-B2 -, ER - and PR - negative) tumour phenotype defines a biologically highly aggressive subgroup of surgical pT1 stage breast cancers | J. Schneider, A. Tejerina, C. Perea, A. Tejerina R. Lucas, J. Sánchez   | 2007 |
| Text 4    | Real incidence of axillar nodal invasion in T1 breast cancer among our population   | J. Schneider, A. Tejerina, J. Sánchez, J. Lucas   | 2007 |
| Text 5    | Prosthetic infection of knee  | O. Sáez-de-Ugarte-Sobroñ, I. Gutiérrez-Sánchez, A. Cruchaga-Celada, F. Labayru-Etxebarria, I. Garcia Sánchez, A. Álvarez-González | 2008 |
| Text 6    | Recurrent aphthous stomatitis (I): Epidemiologic, ethiologic and clinical features  | A. Eguía, R. Saldón, J. M. Aguirre  | 2003 |
| Text 7    | The surgery of the carotid bifurcation in cerebral ischemia of extracranial origin: A 10 year experience  | L. Estallo, A. Barba, L. Rodríguez, S. Gimena, A. G. Alfageme   | 2000 |
| Text 8    | Uncommon clinical features in Whipple's disease: An assay of four cases   | E. Ojeda, A. Cosme, J. Lapaza, J. Torrado, I. Arruabarrena, L. Alzate.  | 2005 |
| Text 9    | Evolution of the anthropometric measures in children's feet: Correlation indices with other variables   | R. De los Mozos, A. Alfageme, E. Ayerdi   | 2002 |
| Text 10   | Evolution of the anthropometric measures in children's feet: A stratified descriptive study   | R. De los Mozos, A. Alfageme, E. Ayerdi   | 2002 |
| Text 11   | Evolution of the anthropometric measures in children's feet: An overall descriptive study   | R. De los Mozos Bozalongo, A. Alfageme Cruz, E. Ayerdi Salazar  | 2003 |
| Text 12   | Stroke acute care and improvement possibilities   | J. Pérez-de-Arriba, G. Achutegui, L. Epelde, G. Viñegra, J.L. Elexpuru.   | 2005 |
| Text 13   | Morbidity and tolerance of the ultrasound-guided prostatic biopsy punction in 392 patients  | J.A. López-Lendoiro, P. Aísa, X. Aguirre, E. Añorbe, M. Paraiso   | 2002 |

**Appendix Table I.** (Continued)

| Reference | Title  | Author(s)  | Year |
|-----------|--|--|------|
| Text 14   | Surgical treatment of infantile flexible flatfoot using the calcaneus-stop technique   | I. Etxebarria-Foronda, I. Garmilla-Iglesias, A. Gay-Vitoria, J. Molano-Muñoz, D. Izal-Miranda, E. Esnal-Baza, A. Ruiz-Sánchez. | 2006 |
| Text 15   | The profile of the users from the emergency department from Galdakao's Hospital  | I. Bengoetxea Martínez   | 2004 |
| Text 16   | Fast progression dementia and myoclonus  | I. Villamil-Cajoto, A, M. J. González-Quintela, V. Villacian-Vicedo  | 2005 |
| Text 17   | Surgical and ultrasound correlation in full thickness tears of the shoulder rotator cuff   | J. de la Fuente-Ortiz-de-Zárate, J. Kutz-Peyroncelli, J. L. Imizcoz-Barriola   | 2004 |
| Text 18   | Surgical treatment for morbid obesity  | I. Díez-del-Val, C. Martínez-Blázquez, V. Sierra-Esteban, J. M. Vitores-López, J. Valencia-Cortejoso                           | 2005 |
| Text 19   | Progress of patients undergoing collapse therapy due to pulmonary tuberculosis   | K. Abu-Shams, J. Ardanaz, M. Murie, A. Sebastián, G. Tiberio, A. Arteché.  | 2000 |
| Text 20   | <i>Pseudomonas aeruginosa</i> infection-colonization in patients with bronchiectasias or COPD. Clinical features, microbiology and outcome | J. Garrós Garay, E. Ruiz de Gordejuela, G. Martín Saco, L. Gallego, J. Pérez Escjadillo, F. García Cebrián                     | 2002 |

<sup>a</sup>The titles in English have been extracted from the original articles, except for the titles of texts 7 and 19; we have translated these from Spanish into English.

**Appendix Table 2.** List of relations used in this study following the extended version and with representative examples in Spanish and Basque<sup>a</sup>

| Relation             | Example  |
|----------------------|--|
| CONTRAST<br>(N-N)    | S [Los antecedentes de primer grado se relacionan con un mayor riesgo de aparición del tumor,] <sub>N</sub> [mientras que los antecedentes familiares de segundo grado no influyen de manera importante.] <sub>N</sub>                           |
|                      | B [Lehen graduko aurrekariak tumorearen agertze arrisku handiagoekin lotzen dira;] <sub>N</sub> [bigarren graduko aurrekari familiarrak, ordea, ez dute modu garrantzitsuan eragiten] <sub>N</sub>   |
|                      | E [First-degree medical history is associated with an increased risk of developing the tumour,] <sub>N</sub> [while second-degree family medical history did not influence significantly.] <sub>N</sub>  |
| JOINT<br>(N-N)       | S [En todos los pacientes se realizó un seguimiento radiológico] <sub>N</sub> [y fueron dados de alta tras una radiografía del abdomen sin evidencia de cuerpos extraños.] <sub>N</sub>  |
|                      | B [Paziente guztiei erradiologiako jarraipena egin zaie] <sub>N</sub> [eta gorputz arrotzen ebidentzia gabekoABELaldearen erradiografiaren ostean guztiei alta eman zitzaie] <sub>N</sub>  |
|                      | E [All the patients underwent radiological monitoring,] <sub>N</sub> [and were discharged after a scan of the abdomen without evidence of strange bodies.] <sub>N</sub>  |
| LIST<br>(N-N)        | S [El 68% de los pacientes eran varones.] <sub>N</sub> [El 92% procedían de Colombia.] <sub>N</sub> [El 65% ingirieron fármacos antidiarreicos.] <sub>N</sub>  |
|                      | B [Pazienteen % 68a gizonezkoak ziren.] <sub>N</sub> [% 92ak kolonbiar jatorria zuen.] <sub>N</sub> [% 65ak beherakoaren kontrako botika irentsi zuen.] <sub>N</sub>   |
|                      | E [68% of patients were male.] <sub>N</sub> [92% came from Colombia.] <sub>N</sub> [65% ingested anti-diarrhea medication.] <sub>N</sub>   |
| SEQUENCE<br>(N-N)    | S [A todos ellos se les realizaron una historia clínica y un examen físico.] <sub>N</sub> [Se les preguntó por el país de procedencia.] <sub>N</sub> [Se registraron la frecuencia cardíaca, la temperatura y la presión arterial.] <sub>N</sub> |
|                      | B [Horiei guztiei egin zitzaie historia klinikoa eta azterketa fisikoa.] <sub>N</sub> [Jatorriko herrialdeaz galdetu zitzaie.] <sub>N</sub> [Bihotz-maiztasuna, temperatura eta presio arteriala erregistratu ziren.] <sub>N</sub>               |
|                      | E [We carried out a medical history and a physical examination to all of them.] <sub>N</sub> [We asked them their country of origin.] <sub>N</sub> [We registered their heart rate, temperature and blood pressure.] <sub>N</sub>                |
| DISJUNCTION<br>(N-N) | S [La mayoría de los pacientes que han perdido peso de forma apreciable roncan menos.] <sub>N</sub> [o han dejado de hacerlo por completo.] <sub>N</sub>   |
|                      | B [Pisua nabarmen galdu duten pazienteen gehiengoak zurrunga gutxiago egiten dute.] <sub>N</sub> [edo zurrunga egiteari utzi diote.] <sub>N</sub>  |
|                      | E [Most of the patients who have lost weight appreciably snore less.] <sub>N</sub> [or they have stopped completely.] <sub>N</sub>   |
| CONJUNCTION<br>(N-N) | S [Mendel no sabía que los genes se localizan en cromosomas] <sub>N</sub> [ni que los genes localizados uno cerca del otro en el mismo cromosoma se transmiten juntos.] <sub>N</sub>   |
|                      | B [Mendelek ez zekien geneak kromosometan kokatzen zirela.] <sub>N</sub> [ezta elkarrekin transmititzen zirela ere kromosoma batean bata bestetik hurbil kokaturiko geneak.] <sub>N</sub>  |
|                      | E [Mendel did not know that genes are located in chromosomes] <sub>N</sub> [nor that genes that are located near each other in the same chromosome are transmitted together.] <sub>N</sub>   |

(Continued)

Appendix Table 2. (Continued)

| Relation              | Example  |
|-----------------------|--|
| BACKGROUND<br>(N-S)   | S <i>[A los portadores de cuerpos extraños intraabdominales que contienen cocaína, con fines de contrabando, se les conoce con el síndrome del body packer.]<sub>S</sub> [Hemos estudiado la aparición de complicaciones en el seguimiento de individuos que ingieren estos paquetes de droga, con el fin de poder dar unas normas de actuación en estos casos.]<sub>N</sub></i> |
|                       | B <i>[Kokainadun sabelalde barneko gorputz arrotzen eramaleak, kontrabando helburudunak, “body packer” sindromea izenez ezagutzen dira.]<sub>S</sub> [Droga pakete hauek irensten dituzten norbanakoen jarraipenean konplikazioen agerpenak ikertu ditugu.]<sub>N</sub></i>  |
|                       | E <i>[Persons who transport strange bodies containing cocaine by internal concealment for smuggling purposes are referred to body packer syndrome.]<sub>S</sub> [We have analyzed the monitoring complications of persons that consume these packets of drug, with the objective of giving rules of conduct in these cases.]<sub>N</sub></i>                                     |
| CIRCUMSTANCE<br>(N-S) | S <i>[Parece necesario propiciar algún tipo de campaña informativa para sensibilizar a la población femenina ante el cáncer de mama.]<sub>N</sub> [mientras no se diluciden las incógnitas que plantean las costosas campañas de detección temprana.]<sub>S</sub></i>  |
|                       | B <i>[Bularreko minbiziaren aurrean beharrezkoa dirudi emakumezko biztanleriari zuzendutako nolabaiteko informazio-kanpainari bide ematea.]<sub>N</sub> [goiz antzemate kanpaina garestien auzia argitzen ez den bitartean behintzat.]<sub>S</sub></i>   |
|                       | E <i>[It seems necessary to carry out some sort of information campaign to sensitize the population to the female breast cancer.]<sub>N</sub> [until the factors of costly campaigns of early detection are not adequately considered.]<sub>S</sub></i>  |
| CONCESSION<br>(N-S)   | S <i>[El porcentaje de curación fue algo menor en los obesos que en los no obesos.]<sub>N</sub> [aunque esta diferencia no ha sido estadísticamente significativa.]<sub>S</sub></i>  |
|                       | B <i>[Sendatze-portzentajea zerbait hobeagoa izan da pertsona gizenetan ez-gizenetan baino.]<sub>N</sub> [nahiz eta diferentzia hori ez den estatistikoki esanguratsua izan.]<sub>S</sub></i>  |
|                       | E <i>[The cure rate was slightly lower in obese people than in non-obese people.]<sub>N</sub> [although this difference was not statistically significant.]<sub>S</sub></i>  |
| CONDITION<br>(N-S)    | S <i>[A efectos del presente estudio consideramos que ha habido acceso a la mamografía]<sub>N</sub> [si la mujer se ha realizado al menos una prueba en los 2 años previos a la realización del estudio.]<sub>S</sub></i>  |
|                       | B <i>[Ikerketa honen xedeetarako mamografía egin izan dela kontsideratu dugu]<sub>N</sub> [baldin eta emakumeak gutxienez froga bat egin izan badu ikerketa egin baino 2 urte lehenago.]<sub>S</sub></i>   |
|                       | E <i>[In this study, we consider that there has been access to mammography]<sub>N</sub> [if the woman has had at least one test in the 2 years preceding the survey.]<sub>S</sub></i>  |

(Continued)

Appendix Table 2. (Continued)

| Relation               | Example   |
|------------------------|---|
| ELABORATION<br>(N-S)   | S [Los pacientes suicidas que padecían una enfermedad orgánica eran 45.] <sub>N</sub> [La edad media de estos pacientes fue de 58,3 años (varones 57,6 años y mujeres 59,2 años) con unos límites de 16 a 90 años.] <sub>S</sub>                          |
|                        | B [Gaixotasun organikoa zuten pazienteak 45 izan dira] [16 eta 90 urte bitarteko paziente hauen bataz besteko adina 58,3 urtekoa izan zen (gizonezkoak 57,6 urte eta emakumezkoak 59,2 urte)] <sub>S</sub>  |
|                        | E [Suicidal patients suffering from organic disease were 45.] <sub>N</sub> [The average age of these patients was 58.3 years (men 57.6 years and women 59.2 years) with a range of 16 to 90 years.] <sub>S</sub>  |
| JUSTIFICATION<br>(N-S) | S [Se realizó cirugía en 7 pacientes (3.3%).] [en cinco de ellos porque presentaban obstrucción, en uno por rotura de uno de los paquetes y en otro por ausencia de progresión de dos de los paquetes que eran de tamaño superior al resto.] <sub>S</sub> |
|                        | B [7 pazientengan (% 3,3a) kirurgia burutu zen.] [hietetako bostek buxadura zutelako, beste bati paketeetako bat apurtu zitzaielako eta beste bati handiagoak ziren 2 paketeren kanporaketan garapenik agertzen ez zelako.] <sub>S</sub>                  |
|                        | E [Surgery was performed in 7 patients (3.3%).] [in five of them because they had obstruction, in one due to the breakage of one package and in another one because of lack of progression of two packages that were larger than the rest.] <sub>S</sub>  |
| PURPOSE<br>(N-S)       | S [Para que puedan cumplir su función con eficacia.] [los SUH precisan que exista un equilibrio apropiado entre la demanda asistencial y su capacidad de respuesta.] <sub>N</sub>   |
|                        | B [Eraginkortasunez haren funtzioa bete dezan.] [SUHak laguntza-eskaeraren eta haren erantzun-gaitasunaren arteko oreka egokia eduki behar du.] <sub>N</sub>  |
|                        | E [In order to fulfil their role effectively.] [ED needs a proper balance between care demand and its responsiveness.] <sub>N</sub>   |
| REFORMULATION<br>(N-S) | S [Se incluyeron sólo pacientes que se consideraba que estaban estables.] [es decir, que no habían precisado cambiar su medicación habitual en los últimos 15 días y clínicamente no referían un empeoramiento importante.] <sub>S</sub>                  |
|                        | B [Egonkor zeudela kontsideratzen ziren pazienteak bakarrik sartu genituen.] [hau da, azkeneko 15 egunetan ohiko medikazioa aldatu behar izan ez zutenak eta klinikoki okerrera egin ez zutenak.] <sub>S</sub>  |
|                        | E [We have included only patients who were considered as stable.] [that is, patients who did not need to change their regular medication in the last 15 days and who reported no significant worsening clinically.] <sub>S</sub>                          |

(Continued)

**Appendix Table 2.** (Continued)

| Relation                | Example  |
|-------------------------|--|
| RESULT<br>(N-S)         | S <i>[Se practicó una radiografía simple del abdomen en todos los enfermos.]<sub>N</sub> [Se observaron cuerpos extraños intra-abdominales en el 98,6% de los enfermos.]<sub>S</sub></i>   |
|                         | B <i>[Gaixo guztietan sabelaldearen erradiografia sinplea praktikatu da.]<sub>N</sub> [Sabelalde barneko gorputz arrotzak gaixoen % 98,6gan hauteman ziren.]<sub>S</sub></i>   |
|                         | E <i>[All patients underwent normal radiographs of the abdomen.]<sub>N</sub> [Intra-abdominal strange bodies were detected in 98.6% of the patients.]<sub>S</sub></i>  |
| SUMMARY<br>(N-S)        | S <i>[Se realizó una radiografía simple.]<sub>N</sub> [También se llevó a cabo una radiografía combinada mediante varias técnicas.]<sub>N</sub> [En resumen, se han aplicado diferentes pruebas radiológicas.]<sub>S</sub></i>   |
|                         | B <i>[Erradiografia sinplea egin zen.]<sub>N</sub> [Zenbait teknika bidezko erradiografia konbinatua ere egin zen.]<sub>N</sub> [Laburtuz, froga erradiologiako desberdinak aplikatu izan dira.]<sub>S</sub></i>   |
|                         | E <i>[A normal X-ray was performed.]<sub>N</sub> [We also carried out a combined X-ray by several techniques.]<sub>N</sub> [In short, we have applied various radiological tests.]<sub>S</sub></i>   |
| EVIDENCE<br>(N-S)       | S <i>[Presentaron datos clínicos de obstrucción intestinal 11 pacientes.]<sub>N</sub> [En todos ellos se observaron signos radiológicos de obstrucción.]<sub>S</sub></i>   |
|                         | B <i>[11 pazienteren hesteetako buxaduraren datu klinikoak aurkeztu ziren.]<sub>N</sub> [Horietan guztietan buxaduraren zeinu erradiologiakoak hauteman ziren.]<sub>S</sub></i>  |
|                         | E <i>[11 patients presented clinical data of intestinal obstruction.]<sub>N</sub> [Radiological signs of obstruction were detected in all of them.]<sub>S</sub></i>  |
| INTERPRETATION<br>(N-S) | S <i>[La utilización de técnicas como el lavado gástrico, la endoscopia, la extracción manual transanal o el uso de laxantes por vía rectal para intentar extraer los paquetes aumenta el riesgo de rotura de los mismos.]<sub>N</sub> [por lo que se desaconseja su uso.]<sub>S</sub></i> |
|                         | B <i>[Urdail-garbiketak, endoskopiaok, ondeste-bideko eskuzko erauzketak edo ondeste-bideko laxanteen erabilerak paketeak apurtzeko arriskua handitzen dute.]<sub>N</sub> [zeinarengatik ez dira horien erabilera gomendatzen.]<sub>S</sub></i>  |
|                         | E <i>[The use of techniques such as gastric lavage, endoscopy, manual transanal removal, or the use of rectal laxatives to try to extract the packages are factors that increase the risk of breaking them.]<sub>N</sub> [so we advise against their use.]<sub>S</sub></i>                 |
| OTHERWISE<br>(N-S)      | S <i>[Consideramos que el programa tenía cobertura total si incluía a todos los municipios.]<sub>N</sub> [si no, la cobertura del programa era considerada parcial.]<sub>S</sub></i>   |
|                         | B <i>[Programak kobertura osoa zuela kontsideratu dugu herri guztiak barnean biltzen bazituen.]<sub>N</sub> [bestela, programaren estaldura partzialtzat hartu izan da.]<sub>S</sub></i>   |
|                         | E <i>[We consider that the program had full coverage if it included all municipalities.]<sub>N</sub> [if not, the program's coverage was considered as partial.]<sub>S</sub></i>   |

(Continued)

Appendix Table 2. (Continued)

| Relation            | Example   |
|---------------------|---|
| ANTITHESIS<br>(N-S) | S <i>[Uno de los factores que se asocian al suicidio es, precisamente, la enfermedad física.]<sub>N</sub> [Sin embargo, la existencia de una enfermedad física no constituye una evidencia incontrovertible de que éste sea el factor único, ni siquiera el más importante, en determinar el acto suicida.]<sub>S</sub></i>   |
|                     | B <i>[Buru-hiltzeari lotutako eragile bat, hain zuzen ere, gaixotasun fisikoa izaten da.]<sub>N</sub> [Hala ere, gaixotasun fisikoa ez da ez halabeharrezko arrazoia ez faktore bakarra, ezta garrantzitsuena ere buru-hiltzearen ekingtza determinatzeko.]<sub>S</sub></i>   |
|                     | E <i>[One of the factors that is associated with suicide is precisely the physical illness.]<sub>N</sub> [However, the existence of a physical illness is not an incontrovertible proof that this is the only factor, nor even the most important, for determining the suicidal act.]<sub>S</sub></i>   |
| ENABLEMENT<br>(N-S) | S <i>[Al paciente no solo se le ha de diagnosticar y tratar la infección.]<sub>N</sub> [Es necesario ofrecerle pautas para que dicha infección no vuelva a aparecer.]<sub>S</sub></i>   |
|                     | B <i>[Pazienteari diagnostikatzea eta infekzioa tratatzea ez da nahikoa.]<sub>N</sub> [Beharrezkoa da jarraibideak eskaintzea infekzioa berriz ager ez dadin.]<sub>S</sub></i>  |
|                     | E <i>[It is not enough to diagnose and treat the infection of patients.]<sub>N</sub> [It is necessary to offer them guidelines in order to avoid the reappearance of this infection.]<sub>S</sub></i>   |
| CAUSE<br>(N-S)      | S <i>[La psiconeuroinmunología es un nuevo campo de la ciencia que está emergiendo]<sub>N</sub> [debido a un número cada vez mayor de datos que demuestran interrelaciones entre funciones inmunes y psiconeurales.]<sub>S</sub></i>  |
|                     | B <i>[Psiconeuroinmunologia garatzen ari den zientziaren eremu berria da.]<sub>N</sub> [Izan ere, gero eta datu gehiagok frogatzen dute funtzio immuneen eta psikoneuralen arteko erlazioak.]<sub>S</sub></i>   |
|                     | E <i>[Psychoneuroimmunology is a new field of science that is emerging]<sub>N</sub> [due to an increasing number of data that show interrelationships between immune functions and psychoneural functions.]<sub>S</sub></i>   |
| EVALUATION<br>(N-S) | S <i>[Hay trabajos que demuestran una mejoría en la distancia recorrida en la prueba de marcha debido al aprendizaje, sobre todo cuando las pruebas se repiten en un corto espacio de tiempo.]<sub>N</sub> [Teniendo esto en cuenta, puede considerarse que las pruebas de marcha son adecuadas para este tipo de estudios y reflejan el esfuerzo que el paciente hará en la vida cotidiana.]<sub>S</sub></i>     |
|                     | B <i>[Ikasketaren ondorioz ibilketa-proban ibilitako distantzian hobekuntza frogatzen duten lanak daude, batez ere denbora laburrean errepikatzen diren frogetan.]<sub>N</sub> [Hau kontuan izanik, pentsa daiteke ibilketa-probak ikasketaren tipoa haurtzat egokiak direla eta pazienteak eguneroko bizitzan egingo duen ahalegina erakusten dutela.]<sub>S</sub></i>   |
|                     | E <i>[There are works that show that there is an improvement regarding the distance that is covered in walking tests due to a learning process, especially when the tests are repeated in a short space of time.]<sub>N</sub> [Bearing this in mind, we consider that walking tests are adequate for this type of study and they show the effort that patients would make in their daily living.]<sub>S</sub></i> |

(Continued)

Appendix Table 2. (Continued)

| Relation             | Example   |
|----------------------|---|
| MOTIVATION<br>(N-S)  | S <i>[En contraste con las numerosas propuestas terapéuticas, sorprende que la pérdida de peso, mediante una dieta alimentaria hipocalórica, aparezca en un segundo o tercer plano y sean muy escasas las publicaciones dedicadas, exclusivamente, a los resultados de la misma, máxime cuando la gran mayoría de los pacientes son obesos.]<sub>S</sub> [Por este motivo, nos hemos decidido a comunicar nuestra experiencia con la dieta hipocalórica como tratamiento único en pacientes afectados de OSAS.]<sub>N</sub></i> |
|                      | B <i>Makina bat proposamen terapeutikorekin kontrastean, harrigarria da dieta hipokalarikoa bigarren edo hirugarren maila batean agertzea eta hain publikazio gutxi egotea proposamen horien datuei buruz; batez ere pazienteen gehiengoa pertsona gizenak direnean.]<sub>S</sub> [Zio horregatik, dieta hipokalarikoa tratamendu bakar gisa OSAS duten pazienteentzat izan dugun esperientzia komunikatzea erabaki dugu.]<sub>N</sub></i>  |
|                      | E <i>[In contrast to the many therapeutic proposals, it is surprising that weight loss, by a hypocaloric diet, appears in second or third place and that there are very few publications dealing exclusively with its results, especially since most of the patients are obese.]<sub>S</sub> [For this reason, we have decided to report our experience with hypocaloric diet as monotherapy in patients with OSAS.]<sub>N</sub></i>  |
| PREPARATION<br>(N-S) | S <i>[Pacientes y métodos.]<sub>S</sub> [Los 257 pacientes estudiados constituyen el 5% seleccionado de un total de 4.850 que se visitaron en la unidad de interconsulta psiquiátrica del Hospital Clínic i Provincial (HCP) de Barcelona desde junio de 1984 a junio de 1990.]<sub>N</sub></i>   |
|                      | B <i>[Pazienteak eta metodoak.]<sub>S</sub> [1984ko ekainetik 1990eko ekainerarte Bartzelonako Hospital Clínic i Provincial (HCP) psiquiatria sail arteko unitatean bisitatu ziren 4.850 pazientetik % 5 osatzen dute azterturiko 257 pazienteak.]<sub>N</sub></i>  |
|                      | E <i>[Patients and methods.]<sub>S</sub> [The 257 studied patients constitute the 5% of 4850 that visited the consultation-liaison psychiatry unit of the Hospital Clínic i Provincial (HCP) in Barcelona from June 1984 to June 1990.]<sub>N</sub></i>   |
| SOLUTION<br>(N-S)    | S <i>[Además de los problemas de infraestructura y de su mayor coste otro inconveniente de las fuentes portátiles es su corta autonomía.]<sub>N</sub> [En este sentido, se han diseñado diversos dispositivos destinados a economizar oxígeno manteniendo un aporte de gas suficiente.]<sub>S</sub></i>   |
|                      | B <i>[Azpiegitura arazoez eta hauen kosteez gain iturri eramangarrien beste eragozpen bat autonomia eskasia da.]<sub>N</sub> [Hori dela eta, gas hornikuntza nahikoa mantentzen duten oxigenoa aurrezteko zenbait gailu diseinatu dira.]<sub>S</sub></i>  |
|                      | E <i>[In addition to infrastructure problems and their greater cost, another disadvantage of portable sources is their short autonomy.]<sub>N</sub> [In that sense, various devices have been designed to save oxygen and maintain an adequate gas supply.]<sub>S</sub></i>   |

(Continued)

Appendix Table 2. (Continued)

| Relation               | Example   |
|------------------------|---|
| MEANS<br>(N-S)         | S <i>[Las tasas de mortalidad por muerte cardíaca súbita pueden reducirse,]<sub>N</sub><br/>[entre otros factores, por la correcta identificación de los pacientes con<br/>riesgo de sufrirla, por la rapidez con que se realicen las maniobras de<br/>reanimación y por la calidad del traslado a centros especializados.]<sub>S</sub></i> |
|                        | B <i>[Bat-bateko heriotza kardiakoaren heriotza-tasak murriz<br/>daitezke,]<sub>N</sub> [beste faktore batzuen artean, sufritzeko arriskua duten<br/>pazienteen identifikazio zehatzari esker, suspertze eragiketak buruturiko<br/>bizkortasunari esker eta gune espezializatuetarako lekualdaketa<br/>kalitateari esker.]<sub>S</sub></i>  |
|                        | E <i>[Mortality rates due to sudden cardiac death can be reduced,]<sub>N</sub><br/>[among other factors, by the correct identification of patients at<br/>risk of suffering it, by the speed of the resuscitation and by the<br/>quality of the move to specialized centers.]<sub>S</sub></i>   |
| UNCONDITIONAL<br>(N-S) | S <i>[Parece que la administración de este medicamento tiene efectos<br/>adversos,]<sub>N</sub> [aun incluso si se administra la dosis mínima.]<sub>S</sub></i>   |
|                        | B <i>[Botika hau hartzeak aurkako eraginak dituela dirudi,]<sub>N</sub> [nahiz eta dosi<br/>txikiena emanda ere.]<sub>S</sub></i>   |
|                        | E <i>[It seems that the administration of this drug has adverse effects]<sub>N</sub><br/>[even if the minimum dose is given.]<sub>S</sub></i>   |
| UNLESS<br>(N-S)        | S <i>[Los terapeutas deben admitir a cualquier paciente en el grupo,]<sub>N</sub> [a<br/>no ser que éste presente signos claros de actitud violenta que puedan<br/>perjudicar el correcto desarrollo de la terapia.]<sub>S</sub></i>  |
|                        | B <i>[Terapeutek edozein paziente onartu behar dute taldean,]<sub>N</sub> [non eta<br/>honen jarrera bortitzak ez duen terapiaren garapen zuzena kaltezen.]<sub>S</sub></i>   |
|                        | E <i>[Therapists must accept any patient in the group]<sub>N</sub> [unless he<br/>presents clear signs of violent behaviour that could harm the<br/>therapy success.]<sub>S</sub></i>   |

<sup>a</sup>In the second column, 'S' means Spanish, 'B' means Basque and 'E' means English.

Iria da Cunha Fanego holds a Hispanic Philology degree at the University of Santiago de Compostela, Spain and a PhD on Applied Linguistics at the Pompeu Fabra University (UPF), Spain. She was Assistant Professor at the UPF and researcher of the Institute for Applied Linguistics until 2008. At present, she holds a postdoctoral grant awarded by the Spanish Ministry of Science and Innovation to work at the Laboratoire Informatique d'Avignon, France. Her research fields are automatic summarization, discourse parsing and analysis of specialized discourse.

Mikel Iruskietia holds a Basque Philology degree at the University of the Basque Country (UPV/EHU). Since 2004, he has been a member of the IXA Research Group (Natural Language Processing Group) at the Faculty of Informatics (UPV/EHU), where he is doing a PhD on Applied Linguistics. He has also been professor of Basque at the same university since 2008. His research fields are semantic, syntactic and discourse parsing, and development of linguistic resources for Basque.