

---

# Thematic Patterning in English and Spanish: Contrastive Annotation of a Bilingual Newspaper Corpus for Linguistic and Computational Applications

---



## DOCTORAL THESIS

Lara Moratón Gutiérrez

Departamento de Filología Inglesa

Facultad de Filología

Universidad Complutense de Madrid

Madrid, 2015



La estructuración temática en inglés  
y español: anotación contrastiva de  
un corpus bilingüe para aplicaciones  
lingüísticas y computacionales

*Memoria que se presenta para optar al título de Doctor*

**Lara Moratón Gutiérrez**

*Dirigida por*

**Prof. Dra. Dña. Julia Lavid López**

**Departamento de Filología Inglesa  
Facultad de Filología  
Universidad Complutense de Madrid**

**Madrid, 2015**





*To my grandparents:  
Chele, Lines, Félix and Julia,  
the beginning of everything.*



## Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor Prof. Dra. Julia Lavid, who has been a strong and supportive advisor throughout this long process also believing and giving me the freedom to dive into new research areas, encouraging at the same time these new steps. For the continuous efforts she made to fit into my busy working timetables outside the faculty.

To all members of FUNCAP group (Functional Linguistics English-Spanish and its Applications) to which I belong as a doctoral student, especially to Dr. Jorge Arús-Hita, Dr. Juan Rafael Zamorano and Dra. Marta Carretero who have been always there to give a prompt hand in all annotation tasks and with whom I collaborated in the the projects CONTRANOT<sup>1</sup> and MULTINOT<sup>2</sup>.

The final part of my thesis could have not been possible without Sara Lana Serrano who translated all linguistic knowledge into rules that made the psalm to automatically annotate theme. With her I learnt to think differently - to focus on linguistics under a different light and prove that multidisciplinary approaches are deeply fruitful.

To all my family, friends and colleagues who have supported and encouraged me all these years, sobre todo a mis padres, Ana y Enrique porque es un orgullo y un placer compartir mi vida con ellos. A Óscar, mi marido, mi ‘equipo’ que hace brillar mis días y a nuestro pequeño Gorka que nos llena de sonrisas.

---

<sup>1</sup>The CONTRANOT project was financed by the Spanish Ministry of Science and Innovation (MICINN) for the period 2009-2012.

<sup>2</sup>The MULTINOT project is currently under way and it is financed by the Spanish Ministry of Economy and Competitiveness for the period 2013-2015.



## Abstract

Thematization is recognized as a fundamental phenomenon in the construction of messages and texts by different linguistic schools. This location within a text privileges the elements that guide the reader in the orientation and interpretation of discourse at different levels. Thematizing a linguistic unit by locating it in the first-initial position of a clause, paragraph, or text, confers upon it a special status: a signal of the organizational strategy which characterizes different text types playing a role as a variable in the distinction of registers, text types and genres.

However, in spite of the importance of the study of thematization for message and textual structuring, to date there are no linguistic studies that have undertaken the task of validating its aspects in a comparative manner, either for linguistic or computational purposes. This study, therefore, fills a research gap by implementing a methodology based on contrastive corpus annotation, which allows to empirically validate aspects of the phenomenon of Thematization in English and Spanish, it also seeks to develop a bilingual English-Spanish comparable corpus of newspaper texts automatically annotated with thematic features at clausal and discourse levels.

The empirically validated categories (Thematic Field and its elements: Textual Theme, Interpersonal Theme, PreHead and Head) are used to annotate a larger corpus of three newspaper genres —news reports, editorials and letters to the editor—in terms of thematic choices. This characterization, reveals interesting results, such as the use of genre-specific strategies in thematic position. In addition, the thesis investigates the possibility to automate the annotation of thematic features in the bilingual corpus through the development of a set of JAVA rules implemented in GATE. It also shows the efficacy of this method in comparison with the manual annotation results.



# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>List of Figures</b>	<b>xiii</b>
<b>List of Tables</b>	<b>xvii</b>
 <b>A. INTRODUCTION</b>	
Background . . . . .	3
Aims and Research Questions . . . . .	5
Methodology . . . . .	6
Thesis Organization . . . . .	6
 <b>B. THEORETICAL FRAMEWORK</b>	
<b>1 Introduction</b>	<b>11</b>
<b>2 Theoretical Account of Theme in Systemic-Functional Linguistics</b>	<b>13</b>
2.1 A Brief Historical Overview . . . . .	14
2.2 The Notion of Theme in Systemic Functional Linguistics . . . . .	17
<b>3 Theoretical Accounts of Theme in Spanish</b>	<b>21</b>
3.1 Lavid el al.'s Model of Theme in SFGS . . . . .	22
3.1.1 Multilayered Analysis of Theme . . . . .	23
3.1.2 Thematic Heads and Their Realization . . . . .	24
3.1.3 A Textual View of Theme . . . . .	26
<b>4 Genre, Register and Text Types from a Functional Perspective</b>	<b>29</b>
4.1 Text Types . . . . .	29
4.2 Genre: English for Specific Purposes and Systemic Functional Approaches	31
4.2.1 Genre in SFL Approaches . . . . .	31
4.2.2 Genre in ESP Approaches . . . . .	32

4.3	Register . . . . .	33
<b>5</b>	<b>Newspapers Genres used for this Study</b>	<b>35</b>
5.1	News Reports . . . . .	36
5.2	Editorials . . . . .	37
5.3	Letters to the Editor . . . . .	37
<b>6</b>	<b>The Relationship between Theme and Genre</b>	<b>39</b>
<b>7</b>	<b>State-of-the-art Issues in Corpus Annotation</b>	<b>43</b>
7.1	Definitions . . . . .	43
7.2	Challenges in Manual Corpus Annotation . . . . .	45
7.3	Manual Annotation Procedures . . . . .	47
7.3.1	Roles . . . . .	47
7.3.2	Process . . . . .	48
7.3.3	Annotation Tools . . . . .	49
7.4	Evaluation Measures . . . . .	52
7.4.1	Inter-annotation Agreement . . . . .	55
7.4.2	Relation Between Agreement and Reliability . . . . .	56
7.4.3	Guidelines and Data Reliability . . . . .	57
<b>8</b>	<b>Automatic Annotation and its Requirements</b>	<b>59</b>
 <b>C. EMPIRICAL VALIDATION OF THE MODEL OF ‘THEME’ IN SFGS</b>		
<b>9</b>	<b>Introduction</b>	<b>65</b>
<b>10</b>	<b>Selecting the Training Corpus</b>	<b>67</b>
<b>11</b>	<b>Delimiting the Categories to be Annotated and their Features</b>	<b>69</b>
11.1	Outer Thematic Field . . . . .	70
11.1.1	Textual Theme . . . . .	70
11.1.2	Interpersonal Theme . . . . .	72
11.2	Inner Thematic Field . . . . .	74
11.2.1	PreHead . . . . .	75
11.2.2	Thematic Head . . . . .	78
<b>12</b>	<b>Designing the Annotation Scheme and Guidelines</b>	<b>93</b>
<b>13</b>	<b>Performing Annotation Experiments</b>	<b>101</b>
13.1	Agreement Study 1: Identification of Thematic Spans . . . . .	102
13.2	Agreement Study 2: Labeling of Thematic Spans . . . . .	102
13.3	Measuring the Annotation Results . . . . .	103



13.3.1	Annotation Results for Agreement Study 1 . . . . .	104
13.3.2	Annotation Results for Agreement Study 2 . . . . .	108
<b>14</b>	<b>Summary and Discussion</b>	<b>125</b>
 <b>D. ANNOTATING A LARGER CORPUS</b>		
<b>15</b>	<b>Introduction</b>	<b>131</b>
<b>16</b>	<b>The Larger Corpus</b>	<b>133</b>
<b>17</b>	<b>Annotating Thematic Features at Clause Level</b>	<b>135</b>
<b>18</b>	<b>Results of the Thematic Annotation</b>	<b>137</b>
18.1	Introduction . . . . .	137
18.2	News Reports . . . . .	137
18.2.1	English News Reports . . . . .	138
18.2.2	Spanish News Reports . . . . .	148
18.2.3	Contrastive Analysis in English vs. Spanish . . . . .	159
18.3	Editorials . . . . .	169
18.3.1	English Editorials . . . . .	170
18.3.2	Spanish Editorials . . . . .	181
18.3.3	Contrastive Analysis in English vs. Spanish . . . . .	189
18.4	Letters to the Editor . . . . .	207
18.4.1	English Letters to the Editor . . . . .	207
18.4.2	Spanish Letters to the Editor . . . . .	214
18.4.3	Contrastive Analysis in English vs. Spanish . . . . .	222
18.5	From a Genre Perspective: NR, Eds and LEs . . . . .	235
18.5.1	Textual Themes Across Genres . . . . .	235
18.5.2	Interpersonal Themes Across Genres . . . . .	240
18.5.3	PreHead Across Genres . . . . .	243
18.5.4	Head Across Genres . . . . .	247
<b>19</b>	<b>Annotating Thematic Progression Patterns</b>	<b>253</b>
19.1	Corpus . . . . .	253
19.2	Models for the Study of Moves and Thematic Progression . . . . .	253
19.2.1	Moves . . . . .	254
19.2.2	Thematic Progression . . . . .	258
19.3	Analysis & Results of the Study of Moves and Thematic Progression . .	261
19.3.1	News Reports . . . . .	262
19.3.2	Editorials . . . . .	264
19.3.3	Letters to the Editor . . . . .	268
<b>20</b>	<b>Summary and Discussion</b>	<b>273</b>

## **E. TOWARDS AUTOMATIC THEME ANNOTATION**

<b>21 Introduction</b>	<b>283</b>
<b>22 The GATE Tool</b>	<b>285</b>
<b>23 Corpus and Metadata</b>	<b>289</b>
<b>24 Automatic Annotation of Theme in English</b>	<b>293</b>
24.1 Corpus Preprocessing . . . . .	293
24.2 Rule Based Automatic Annotation . . . . .	296
24.2.1 Rule for Thematic Field . . . . .	297
24.2.2 Rule for Textual Theme . . . . .	299
24.2.3 Rule for Interpersonal Theme . . . . .	300
24.2.4 Rule for PreHead . . . . .	302
24.2.5 Rule for Thematic Head . . . . .	302
<b>25 Semi-Automatic Annotation of Theme in Spanish</b>	<b>307</b>
25.1 Corpus Preprocessing . . . . .	307
25.2 Semi-automatic Annotation . . . . .	310
<b>26 Comparing the Results of Hand-coded versus Automatic Annotation</b>	<b>313</b>
26.1 English Manual and Automatic Annotation Results . . . . .	313
26.1.1 Sentence Splitting . . . . .	314
26.1.2 Thematic Field . . . . .	314
26.1.3 Textual Themes . . . . .	315
26.1.4 Interpersonal Themes . . . . .	317
26.1.5 PreHead . . . . .	319
26.1.6 Head . . . . .	323
26.2 Spanish Manual and Semi-automatic Annotation Results . . . . .	326
26.2.1 Sentence Splitting . . . . .	326
26.2.2 Thematic Elements . . . . .	326
<b>27 Summary and Discussion</b>	<b>329</b>

## **F. FINAL REMARKS**

<b>28 Conclusions</b>	<b>335</b>
-----------------------	------------

**REFERENCES****APPENDIXES**

<b>A Curriculum Vitae</b>	<b>365</b>
<b>B English and Spanish Summary</b>	<b>377</b>
B.1 English Summary . . . . .	377
B.2 Resumen en castellano . . . . .	386
<b>C Appendix</b>	<b>395</b>
C.1 Annotation Schemes and Guidelines . . . . .	396
C.2 Annotation Experiments . . . . .	408
<b>D Appendix</b>	<b>421</b>
D.1 Index of Texts . . . . .	421
D.2 Annotated News Reports in English . . . . .	434
D.3 Annotated News Reports in Spanish . . . . .	453
D.4 Annotated Editorials in English . . . . .	474
D.5 Annotated Editorials in Spanish . . . . .	505
D.6 Annotated Letters to the Editor in English . . . . .	543
D.7 Annotated Letters to the Editor in Spanish . . . . .	575
<b>E Appendix</b>	<b>597</b>
E.1 English Automatically Annotated Corpus (in CD attached) . . . . .	597
E.2 Spanish Semi-automatically Annotated Corpus (in CD attached) . . . . .	597



# List of Figures

2.1	Metafunctions and their reflexes in the grammar . . . . .	18
7.1	The generic annotation pipeline (after Hovy & Lavid, 2010) . . . . .	47
7.2	Basic annotation process (in Eckart de Castilho, 2014) . . . . .	48
11.1	The grammar of experience . . . . .	85
17.1	Screenshot of the manual annotation sheet . . . . .	136
18.1	Thematic elements in news reports (English vs. Spanish) . . . . .	159
18.2	Textual Themes in news reports (English vs. Spanish) . . . . .	160
18.3	PreHeads in news reports (English vs. Spanish) . . . . .	163
18.4	Types of PreHeads in news reports (English vs. Spanish) . . . . .	164
18.5	Syntactic realization of Head in news reports (English vs. Spanish) . .	166
18.6	Syntactic realization of NP Head in news reports (English vs. Spanish)	167
18.7	Types of semantic realizations in news reports (English vs. Spanish) .	168
18.8	Textual Themes in editorials (English vs. Spanish) . . . . .	190
18.9	Types of Textual Themes in editorials (English vs. Spanish) . . . . .	191
18.10	Types of binders in editorials (English vs. Spanish) . . . . .	191
18.11	Types of linkers in editorials (English vs. Spanish) . . . . .	193
18.12	Interpersonal Themes in editorials (English vs. Spanish) . . . . .	194
18.13	Types of Interpersonal Themes in editorials (English vs. Spanish) . . .	194
18.14	PreHeads in editorials (English vs. Spanish) . . . . .	196
18.15	Types of PreHeads in editorials (English vs. Spanish) . . . . .	197
18.16	Functions of PreHeads in editorials (English vs. Spanish) . . . . .	199
18.17	Syntactic realization of Heads in editorials (English vs. Spanish) . . .	201
18.18	Syntactic realization of NP Heads in editorials (English vs. Spanish) .	202
18.19	Syntactic realization of NP+App Heads in editorials (English vs. Span- ish) . . . . .	205
18.20	Semantic realization of Heads in editorials (English vs. Spanish) . . .	206
18.21	Textual Themes in letters to the editor(English vs. Spanish) . . . . .	223
18.22	Types of Textual Themes in letters to the editor (English vs. Spanish)	223
18.23	Binders in letters to the editor (English vs. Spanish) . . . . .	224
18.24	Linkers in letters to the editor (English vs. Spanish) . . . . .	226
18.25	Interpersonal Themes in letters to the editor (English vs. Spanish) . .	227

## LIST OF FIGURES

---

18.26	Types of Interpersonal Themes in letters to the editor (English vs. Spanish) . . . . .	227
18.27	PreHead in letters to the editor (English vs. Spanish) . . . . .	229
18.28	Types of PreHeads in letters to the editor (English vs. Spanish) . . . . .	229
18.29	Syntactic realisation of Heads in letters to the editor (English vs. Spanish) . . . . .	231
18.30	Semantic realization of Heads in letters to the editor (English vs. Spanish)	233
18.31	Textual Themes across genres in English . . . . .	236
18.32	Types of Textual Themes across genres in English . . . . .	237
18.33	Textual Themes across genres in Spanish . . . . .	238
18.34	Types of Textual Themes across genres in Spanish . . . . .	238
18.35	Mean length of English and Spanish texts and use of Textual Themes	239
18.36	Interpersonal Themes across genres in English . . . . .	240
18.37	Types of Interpersonal Themes across genres in English . . . . .	241
18.38	Interpersonal Themes across genres in Spanish . . . . .	242
18.39	Types of Interpersonal Themes across genres in Spanish . . . . .	242
18.40	PreHead across genres in English . . . . .	243
18.41	Types of PreHead across genres in English . . . . .	244
18.42	PreHead across genres in Spanish . . . . .	245
18.43	Types of PreHead across genres in Spanish . . . . .	246
18.44	NPs realization across genres in English . . . . .	248
18.45	NPs realization across genres in Spanish . . . . .	249
18.46	Experiential roles across genres in English . . . . .	250
18.47	Experiential roles across genres in Spanish . . . . .	251
19.1	Types of Thematic Progression patterns . . . . .	259
19.2	Split Theme . . . . .	261
23.1	Metadata file format . . . . .	290
23.2	Example of a file name . . . . .	290
23.3	Metadata example . . . . .	291
23.4	Folder with preprocessed files . . . . .	291
24.1	English Sentence splitting . . . . .	294
24.2	English Tokenizer . . . . .	294
24.3	English PoS Tagging . . . . .	295
24.4	English bootstrapping pipeline . . . . .	296
24.5	English Stanford parsing example . . . . .	298
25.1	Spanish Sentence splitting . . . . .	308
25.2	Spanish Tokenizer . . . . .	308
25.3	Spanish PoS Tagging . . . . .	309
25.4	Spanish Stemmer . . . . .	309
25.5	Spanish bootstrapping pipeline . . . . .	310

25.6	Schema for Thematic Field annotation and its elements . . . . .	311
25.7	Schema annotation process . . . . .	312
27.1	ANNIC English corpus screenshot . . . . .	331
C.1	Annotation scheme of Thematic Field English . . . . .	396
C.2	Annotation scheme of Thematic Field Spanish . . . . .	397
C.3	English tag-set definition for Theme categories . . . . .	398
C.4	Spanish tag-set definition for Theme categories . . . . .	399
C.5	English tag-set definition for Textual Themes . . . . .	400
C.6	Spanish tag-set definition for Textual Themes . . . . .	400
C.7	English tag-set definition for Interpersonal Themes . . . . .	401
C.8	Spanish tag-set definition for Interpersonal Themes . . . . .	402
C.9	English tag-set definition for PreHeads . . . . .	403
C.10	Spanish tag-set definition for PreHeads . . . . .	404
C.11	English tag-set definition for Head experiential roles . . . . .	405
C.12	Spanish tag-set definition for Head experiential roles . . . . .	406
C.13	Instructions of Agreement Study 1 Task 1 . . . . .	408
C.14	Annotation sheet of Agreement Study 1 Task 1 . . . . .	409
C.15	Annotation sheet of Agreement Study 1 Task 2 . . . . .	410
C.16	Instruction sheet of Agreement Study 2 Task 1 . . . . .	411
C.17	Annotation sheet of Agreement Study 2 Task 1 . . . . .	412
C.18	Instruction sheet of Agreement Study 2 Task 2 . . . . .	413
C.19	Annotation sheet of Agreement Study 2 Task 2 . . . . .	414
C.20	Instruction sheet of Agreement Study 2 Task 3 . . . . .	415
C.21	Annotation sheet of Agreement Study 2 Task 3 . . . . .	416
C.22	Instruction sheet of Agreement Study 2 Task 4 . . . . .	417
C.23	Annotation sheet of Agreement Study 2 Task 4 . . . . .	418
C.24	Instruction sheet of Agreement Study 2 Task 5 . . . . .	419
C.25	Annotation sheet of Agreement Study 2 Task 5 . . . . .	420





# List of Tables

7.1	Interpretation of Kappa . . . . .	54
10.1	Training Corpus A . . . . .	67
10.2	Training Corpus B . . . . .	68
12.1	Core tag-set for Theme types in English . . . . .	94
12.2	Core tag-set for Theme types in Spanish . . . . .	95
12.3	Experiential role types in English and Spanish . . . . .	96
12.4	Types of PreHead in English and Spanish . . . . .	98
12.5	Types of Textual Theme in English and Spanish . . . . .	98
12.6	Types of Textual Theme in English and Spanish . . . . .	98
13.1	Interpretation of Kappa . . . . .	104
13.2	Inter-annotator agreement: Thematic Field span in English . . . . .	105
13.3	Inter-annotator agreement: Thematic Field span in Spanish . . . . .	106
13.4	Inter-annotator agreement: Spans expressing TH, PH, IT, TT, PT, There in English . . . . .	106
13.5	Inter-annotator agreement: Spans expressing TH, PH, IT, TT, PT and Hay in Spanish . . . . .	107
13.6	Inter-annotator agreement: Labelling thematic spans . . . . .	109
13.7	Kappa agreement measures . . . . .	110
13.8	Inter-annotator agreement: Labelling thematic spans . . . . .	111
13.9	Kappa agreement measures . . . . .	111
13.10	Inter-annotator agreement: Labelling of experiential elements in English	112
13.11	Kappa agreement measures in Experiential roles in English . . . . .	113
13.12	Inter-annotator agreement: Labelling of experiential elements in Spanish	114
13.13	Kappa agreement measures in Experiential roles in Spanish . . . . .	115
13.14	Interannotator agreement in types of Textual Themes in English . . .	116
13.15	Kappa agreement measures in types of Textual Themes in English . .	116
13.16	Interannotator agreement in types of Textual Themes in Spanish . . .	117
13.17	Kappa Agreement Measures in Types of Textual Themes in Spanish .	117
13.18	Interannotator agreement in types of Interpersonal Themes in English	118
13.19	Kappa agreement measures in Interpersonal Themes in English . . . .	119
13.20	Interannotator agreement in types of Interpersonal Themes in Spanish	120

## LIST OF TABLES

---

13.21	Kappa agreement measures in Interpersonal Themes in Spanish . . . .	121
13.22	Interannotator agreement in types of PreHead Themes in English . . .	121
13.23	Kappa agreement measures in PreHead Themes in English . . . . .	122
13.24	Interannotator agreement in types of PreHead Themes in Spanish . . .	122
13.25	Kappa agreement measures in PreHead Themes in Spanish . . . . .	123
14.1	Results of Agreement Study 1 - English and Spanish . . . . .	126
14.2	Results of Agreement Study 2 - English and Spanish . . . . .	127
16.1	Bilingual corpus . . . . .	133
18.1	Corpus of news reports in English and Spanish. . . . .	138
18.2	Thematic elements in English news reports . . . . .	138
18.3	Types of Textual Themes in English news reports . . . . .	139
18.4	Types of PreHeads . . . . .	142
18.5	Circumstantials . . . . .	142
18.6	Syntactic realization of Head . . . . .	145
18.7	Thematic elements in Spanish news reports . . . . .	149
18.8	Textual Themes . . . . .	149
18.9	Types of PreHeads . . . . .	151
18.10	Syntactic realization of Head . . . . .	155
18.11	Corpus of editorials in English and Spanish. . . . .	170
18.12	Thematic elements in English editorials . . . . .	170
18.13	Types of Textual Themes in English editorials . . . . .	171
18.14	Types of PreHeads . . . . .	175
18.15	Syntactic realization of Head . . . . .	178
18.16	Thematic Elements in Spanish editorials . . . . .	181
18.17	Types of Textual Themes in Spanish Editorials . . . . .	182
18.18	Types of PreHeads . . . . .	185
18.19	Syntactic realization of Head . . . . .	186
18.20	Thematic elements in English vs Spanish editorials . . . . .	189
18.21	Corpus of letters to the editor in English and Spanish . . . . .	207
18.22	Thematic elements in English letters to the editor . . . . .	208
18.23	Types of Textual Themes in English letters to the editor . . . . .	208
18.24	Types of PreHeads . . . . .	210
18.25	Thematic elements in Spanish letters to the editor . . . . .	214
18.26	Types of Textual Themes in Spanish letters to the editor . . . . .	215
18.27	Syntactic realization of Head . . . . .	219
18.28	Thematic elements in English vs Spanish letters to the editor . . . . .	222
19.1	GSP news reports moves according to White (1997) . . . . .	256
19.2	GSP editorial moves according to Ansary & Babaii (2005) . . . . .	257
19.3	GSP letters to the editor moves according to Ghadessy (1995) . . . . .	258
19.4	GSP example of editorial Move 2 . . . . .	267

20.1	Contrastive table of moves and thematic progression patterns . . . . .	279
23.1	Corpus for automatic annotation . . . . .	289
26.1	Sentence splitting automatic/manual annotation . . . . .	314
26.2	Thematic Field . . . . .	315
26.3	Textual Themes in automatic vs manual annotation . . . . .	316
26.4	Clause complex automatic annotation of Interpersonal Themes . . . . .	317
26.5	Clause complex automatic annotation of PreHeads . . . . .	319
26.6	Clause complex automatic annotation of Heads . . . . .	324
26.7	Clause complex bootstrapping Spanish . . . . .	326
26.8	Manual vs semi-automatic annotation in Spanish . . . . .	327

## *LIST OF TABLES*

---

## **A. INTRODUCTION**



## Background

The motivation for the work reported in this dissertation is rooted in two main research areas: (Systemic)-Functional Linguistics and its application to language comparison and corpus analysis, on the one hand, and Natural Language Processing (NLP), on the other. In (Systemic)-Functional Linguistics, the phenomenon of thematization has been recognized as a fundamental mechanism in the construction of messages and texts, and several scholars have pointed out its discourse role as a signal through which writers guide their readers in the interpretation of messages and in the characterization of texts (see Halliday & Matthiessen, 2004; Downing, 1991; Lotfipour-Saedi & Rezai-Tajani, 1996). A productive line of research has focused on the relationship between thematic selection, registers and genres Ghadessy (1995), focusing almost exclusively on English. In addition, several scholars with an interest in comparing English to other languages, have also pointed out the potential of thematization for both clausal and discourse structuring (see Grzegorek, 1984; Hasselgård, 2004; Hidalgo Downing, 2003; McCabe & Alonso Belmonte, 2001; Hidalgo Downing, 2003; Lavid et al., 2010d, 2013a).

However, in spite of the importance of the study of thematization for message and textual structuring, to date, there are no linguistic studies that have undertaken the task of validating aspects of this phenomenon in a contrastive manner, both for linguistic and computational purposes. The current study, therefore, seeks to fill a research gap in this area, by implementing a methodology based on contrastive corpus annotation which will allow to empirically validate thematization in English and Spanish<sup>1</sup>.

In the field of Natural Language Processing (NLP), the phenomenon of thematization has also attracted the attention of researchers working in Discourse Parsing (Schwarz et al., 2008) and in Natural language generation (NLG). In NLG, for example, thematization plays a fundamental role in micro-planning tasks (Hovy & Lavid, 2007). However, the potential of thematization and other discourse features of high linguistic abstraction has been virtually unexplored in other computational areas such as Information Extraction (IE) and Document Classification (DC). In these two areas the predominant methods are string or word-based and even supervised data mining rarely employs discourse knowledge or information at higher levels of linguistic knowledge. One reason for this situation is the lack of quality corpora annotated in terms of discourse features (see Schwarz et al. 2008). Fortunately, this situation is progressively changing and enhancing corpora with linguistic features of discourse organization is

---

<sup>1</sup>This methodology has been reported in several publications by the author of this thesis and members of the FUNCAP research group (see Lavid et al., 2012, 2010b, 2009a; Arús et al., 2013, 2012; Moratón, 2009; Moratón et al., 2012; Moratón & Lavid, 2012)

becoming an increasingly active research area (see Taboada, 2013; Buráňová et al., 2000; Kruijff-Korbayová & Kruijff, 2004; Honnibal & Curran, 2007; Bhatia et al., 2008; Perfetti & Goldman, 1975, 1974).

The work reported in this dissertation is an attempt to advance knowledge on the topic of thematization in the two research areas mentioned before:

From the linguistic perspective, it provides empirical validation of the category of thematization -more specifically of the thematization model proposed by Lavid et al. (2010d) through corpus annotation. It also establishes correlations between thematic selection and progression patterns and the generic structures of three newspaper genres -reports, editorials and letters to the editor- in English and Spanish.

From the NLP perspective, the current work offers the NLP community a high-quality bilingual (English-Spanish) comparable corpus of newspaper texts annotated with thematic features at clausal and discoursal levels which may be a very useful resource not only for contrastive linguistic explorations, but mainly for a number of computational applications in the fields of Information Extraction and Retrieval (IR), Document Classification (DC) and Natural Language Generation (NLG), among others. It also explores the possibility to automate the annotation process of a highly complex linguistic feature such as Theme in a contrastive manner, thus opening new avenues of research in this area.

This dissertation is framed in the context of a larger research effort carried out at Universidad Complutense de Madrid (UCM) by the research group FUNCAP (Functional Linguistics English-Spanish and its Applications) to which I belong as a doctoral student. The FUNCAP research group at UCM has been active in the past years in the functional and corpus-based study of a number of discourse and lexicogrammatical phenomena in English and Spanish focusing on linguistic, methodological and computational issues, as shown by the results of the projects CONTRANOT<sup>1</sup> and MULTINOT<sup>2</sup>, led by Professor Julia Lavid, my thesis supervisor. It is in the context of these projects, where I have collaborated as an associated member of the UCM team, where the work presented here on thematic patterning in English and Spanish has been carried out.

---

<sup>1</sup>The CONTRANOT project was financed by the Spanish Ministry of Science and Innovation (MICINN) for the period 2009-2012.

<sup>2</sup>The MULTINOT project is currently under way and it is financed by the Spanish Ministry of Economy and Competitiveness for the period 2013-2015.



## **Aims and Research Questions**

The present study has three main general aims, each of them oriented to the research fields where the current dissertation is framed: Contrastive Functional Linguistics, Corpus Linguistics, and Natural Language Processing (NLP).

From the Contrastive Linguistics perspective, the current work seeks to advance knowledge in the study of the correlations between the linguistic category of genre and thematization in two Indoeuropean languages (English and Spanish) and across three different newspaper genres (news reports, editorials and letters to the editor). In pursuing this aim, the current dissertation wishes to extend previous research on this issue (see Eiler, 1986; Francis, 1989, 1990; Ghadessy, 1995; Fries & Francis, 1992; Bloor, 1991) and to contribute to a fruitful research line on the study of newspaper texts across genres and languages.

From the Corpus Linguistics perspective, the current dissertation aims at providing an empirical validation of the theoretical model of thematization put forward in Lavid et al. (2010d) through corpus annotation. This aim is clearly an innovation within the field of Corpus Linguistics.

From the NLP perspective, the final aim is to create a high-quality bilingual (English-Spanish) corpus, automatically annotated with thematic features, which can be used in a number of computational applications, such as Information Retrieval (IR) and automatic Document Classification (DC). The automatic annotation of the process (in English) is achieved through the development of JAVA rules based on the validated categories.

With these main aims and objectives in mind, the present thesis focuses on a number of specific research questions, which are investigated throughout the different sections of the manuscript. These are the following:

1. Can the theoretical phenomenon of thematization in English and Spanish be empirically tested through corpus annotation?
2. Are there any language-specific thematic selection and progression patterns in the three newspaper genres studied, i. e., news reports, editorials and letters to the editor? Are there any correlations between thematic patterns and generic moves in these three newspaper genres?
3. Is it possible to automate the manual annotation of thematic features in English and Spanish and how can this be achieved? If possible, is the result obtained reliable?

## Methodology

In order to investigate the research questions described above, the following methodological steps were carried out:

1. For the empirical validation of thematization through corpus annotation, a number of agreement studies were designed to investigate the reliability of certain thematic features extracted from Lavid's model of thematization in *Systemic-Functional Grammar of Spanish: A Contrastive Study with English* (henceforth SFGS), as presented in Lavid et al. (2010d:chapter 5).
2. For the exploration of the language-specific thematic selection and progression patterns in the three newspaper genres studied, a corpus of 220 newspaper texts consisting of editorials, news reports and letters to the editor in English and Spanish was manually annotated by the author of this dissertation with the thematic features validated in step 1.
3. For the exploration of the possibility to automate the manual annotation process -a labor-intensive and time-consuming task-, a number of rules was created on the basis of the validated categories and implemented in JAVA for their use within the GATE platform (Cunningham et al., 2011).

## Thesis Organization

The current Part (Part A) provides an introduction to the work reported, explaining the background and the context where the dissertation is framed, the aims and the research questions which have guided this investigation, the methodology used and the thesis organization.

The remainder of this thesis is organized in five Parts: Theoretical Framework, Empirical Validation of the Model of 'Theme' in SFGS, Annotating a Larger Corpus, Towards Automatic Theme Annotation and Final Remarks. In the following lines I briefly introduce the contents treated in these Parts.

**Part B. Theoretical Framework**, where I give a brief overview of the concepts used throughout the thesis. First, the notion of Theme is reviewed from the point of view of Systemic Functional Linguistics, focusing afterwards on the accounts of Theme in Spanish. Chapter 4 is dedicated to Genre, Register and Text types and its controversial definitions from a functional perspective. The genres under study

are approached in chapter 5. The strong relationship treated in the literature of these two aspects of Theme and Genre is reflected in chapter 6. Finally, an approach to state-of-the-art in corpus analysis and annotation (chapter 7) precedes the requirements for automatic annotation (chapter 8).

**Part C. Empirical Validation of the Thematization Model.** Now that the relevant notions of the study are set the work starts in this Part. First, an explanation of the selection of the training corpus is given followed by a revision in chapter 11 of the annotated categories and their more relevant features. The next steps on the validation of the annotated categories are performed through the design of annotation schemes (chapter 12) and the performance of several annotation experiments (chapter 13) are analyzed and discussed.

**Part D. Annotating a Larger Corpus** accounts for the annotation of a larger corpus from a language and genre perspective (chapter 18.1). The study of moves and thematic progression is also included in this part, its methodology is presented in chapter 19.2 and its results in chapter 19.3.

**Part E. Towards Automatic Theme Annotation** deals with the computational side of the thesis. Here I provide an introduction to the tools and procedures followed for the automatic annotation process, comparing the results obtained through manual and automatic annotation.

**Part F. Final Remarks**, summarizes and discusses the results obtained in view of the research questions presented at the beginning of the thesis, and provides pointers for future work.



## **B. THEORETICAL FRAMEWORK**



# 1

## Introduction

The work involved in the present dissertation has been written in the context of a number of theoretical and applied issues in the different fields of Systemic-Functional Linguistics, Corpus Linguistics and Natural Language Processing (NLP), respectively. More specifically, the literature on the following domains has been consulted and drawn upon:

- Theoretical accounts of Theme in the (Systemic)-functional literature
- Theoretical accounts of Theme in Spanish
- Theories of Genre, Register and text type from a functional perspective
- Genres used for this study
- The relationship between Theme and Genre
- State-of-the-art issues in corpus analysis and annotation
- Automatic annotation and its requirements

Given this wide range of issues, this part is divided into a number of chapters, each of which describes the main issues which serve as contextual framework for the current work.





# 2

## Theoretical Account of Theme in Systemic-Functional Linguistics

In this section I will first begin with a brief historical overview of the origins of the notion of Theme in the literature and I will then proceed to explain how this notion has been treated and discussed within the theory of Systemic-Functional Linguistics. This review will by no means be comprehensive, but purely illustrative of the main issues and challenges that the concept of Theme brings about when studying the structure of the clause and the organization of discourse. Secondly, I will also provide an account of a number of interesting and challenging issues which have emerged when comparing the thematic structure of different languages, such as English and Spanish, and will focus on the description of the theoretical model for English and Spanish provided by Lavid in SFGS (see Lavid et al., 2010d:chapter 10), as this is the theory which has been subject to empirical testing in this dissertation and which has also served as the basis for the corpus annotation tasks described in the remainder of this work.

## 2.1 A Brief Historical Overview

The first account of the notion of Theme appears in the linguistic literature when the linguist Henri Weil mentioned it in his doctoral dissertation entitled: *De l'ordre des mots dans les langues anciennes comparées aux langues modernes* in 1845 (published in English is 1978). He proposes that the order of words of an utterance is a reflection of the flow of ideas expressed by the speaker which may not always coincide with the syntax of the language, this means the ideas expressed about a fact is one entity and its syntactic representation is another.

There are in the proposition two different movements: an objective movement, which is expressed by syntactic relation; and a subjective movement, which is expressed by the order of words. It may be said that the syntax is the principal thing because it inheres the objects themselves, and because it does not vary with the points of view from moment to moment. But this is precisely the reason why we should attach the greatest importance to the succession of words.

(Weil, 1978:30)

Although the order of words is of high importance for Weil, he also points out that in modern languages this order is greatly determined by syntax, and it is normally present in the clause:

There is then a point of departure, an initial notion (...) which forms, as it were, the ground upon which the two intelligences meet; and another part of discourse which forms the statement (*l'énonciation*), properly so called. This division is found in almost all it is said.

(Weil, 1978:29)

Weil defines the first element in the sentence as the *point of departure* or the *ground upon which the two intelligences meet*; that is, a common knowledge both for the hearer and the speaker, and what will later be considered as *Given information* or *Theme*. The second part of the sentence is the *statement* or *l'énonciation* that is, what it is said about the first element, and what will later be *New information* or *Rheme*.

What is more, Weil also identifies the point of departure of the sentence with the subject as “(...) of a sensible act which serves as a model for the construction of the sentence. Our languages tend to make of the subject the point of departure for the thought itself” (1978:37). Here he identifies the thetic order as the more common order that different languages use to express an idea.

Both ideas will be further developed by the Prague School in later studies, as will be explained below.

In the Hispanic literature, Andrés Bello introduced the notion of *sustantivo sujeto* [noun subject my translation] in his *Gramática de la lengua castellana destinada al uso de los americanos* (1860) where he stated:

La palabra dominante en la oración es el sustantivo sujeto, a que se refiere el verbo atribuyéndole alguna cualidad, acción, ser o estado. Y en torno al sustantivo sujeto o al verbo se colocan todas las otras palabras, las cuales, explicándose o especificándose unas a otras, miran, como a sus peculiares últimos puntos de relación, las unas al sustantivo sujeto, las otras al verbo.

[The dominant word in the sentence is the noun subject, to which the verb which refers, attributing to it some quality of action, being or state. And around them are placed all the other words which explain or specify one another, referring to their ultimate points of relation, some to the noun subject, others to the verb.]

(1860: chap. XXII, my translation)

Although Bello and Weil point to the subject as the main element of a sentence, they also propose two different informational foci, one that refers to the subject, and the other to the verb. The verb attributes “some quality of action, being or state” to the subject (Bello, 1860).

After these early notions of Theme, scholars such as Vilém Mathesius -within what is known as the Prague School of Linguistics- focused on the conception of language as a system with two different but complementary aims, that of communication and that of expression. In his posthumous book entitled *A Functional Analysis of Present Day English on a General Linguistic Basis* (1961), he explained his functional conception of language:

In our conception language is a system of the means of expression, a system of signs, manifested in actual communication as the sum total of the possibilities available to the members of the same language community at a given time in a given place for the purpose of communication through speech, and identifiable from their realisations in particular utterances.

(Mathesius, 1975:13)

For him, the study of the language systems implies the study of three separated structural levels: semantic structure, grammatical structure and a system of organizational means of expression. The latter one is what he called Functional Sentence Perspective and consists of two sentence elements: the Theme and the Rheme<sup>1</sup>.

---

<sup>1</sup>For a more detailed description see Jiménez Juliá 1986: pp 23-27

Languages may be characterized by the degree of dependency that their thematic elements have on their grammatical elements.

According to Firbas (1987), Mathesius used three terms when referring to theme: *point of departure*, *Theme*, and *basis* (or *foundation*). Mathesius questions the dual aspect of Theme: that of *aboutness* (“the intrasentential relationship between what is being spoken about and what is being said about it” (in Firbas 1987); and the contextual aspect with regard to what is known. Firbas notices that Mathesius did not always distinguish *point of departure* from *Theme*, and never explained the difference. Later on, Mathesius dropped the use of *point of departure* and kept the use of the terms *theme* and *basis* (or *foundation*) as synonyms. Mathesius defined *Theme* in 1939 as: “that which is known or at least less obvious in the given situation and from which the speaker proceeds in his discourse.” *Core of the utterance* (rheme) was defined as “that which the speaker is communicating about or what relates to the starting point of the utterances” (in Firbas 1987).

Mathesius distinguishes a “usual” order (Theme-rheme) and a “marked” order to distinguish between the different criteria to refer to the organization of the clause.

The studies on Theme that followed were divided between two positions: the so called *combining approach* (Fries, 1983:133), that considers Theme and known or given information as two aspects of the same phenomenon; and the *separating approach* that disengages information structure from thematic structures, this line being followed by the Systemic-Functional Linguistics Schools.

**Firbas** writes: “it is also evident that although a thematic element is always linked with the ‘aboutness’ feature, it is not invariably linked with the context dependence. (...) being only the ‘aboutness’ feature constantly present” (1987:143).

Regarding this contextual feature, Firbas coined the term *Communicative Dynamism* (CD) by which the dichotomy “known versus unknown”, is a dynamic process in the sentence where the information advances gradually until its culmination in the complete expression (in Jiménez Juliá 1986):

By the degree of CD carried by a sentence element we understand the extent to which the sentence element contributes to the development of the communication, to which it ‘pushes the communication forward’ as it were.

(Firbas, 1972:270)

As Weil also stated, the beginning of the sentence establishes a common ground for the speaker and the listener to evolve until full information is rendered by the speaker to the hearer.

It was Firbas (1987) who looked back to the preceeding information and forward to the immediate relevant communicative purpose. That is, the sentence is embedded in the flow of communication and works under this communicative process.

Also within the Prague School, **Daneš** who provided interesting insights to the original theory: first, he brought in the discourse perspective, moving beyond isolated sentences with his account of *Thematic Progression* (TP). In this theory, the information in the discourse advances according to the information contained in the preceding context, and it is in this development where Theme is necessarily considered in the context of preceding information.

(...) it is evidently necessary to distinguish between the mass of information accumulated up to a certain point of text, and the portion of this mass contained (occurring) in the particular utterance following this point. This distinction involves a selection from the mass of known information for every utterance. (...) this selection is determined by the utterance Theme. Thus we must not be content with a statement that certain sentences elements convey the known information (in contrast to others, conveying the new one), but we ought to find out the principles exactly according to which this and not another portion of the mass of information has been selected. In other words, we have to inquire into the principles underlying thematic choice and thematic progression.

(Daneš, 1974:112)

Secondly he proposes to divide the concepts discussed into two pairs: *topic-comment* to distinguish between given and known information and *Theme-rheme* to signal the sentence point of departure and its further development. These two independent halves that do not have to always coincide.

Since this moment on, many attempts have been made to achieve a consensus on the definition of Theme.

## 2.2 The Notion of Theme in Systemic Functional Linguistics

The most influential account of Theme in the functional literature to date has been Halliday's modeling of this concept which appeared in a number of publications and in his *Functional Grammar* (Halliday, 1967, 1994). Halliday (1967) and later, Quirk et al. (1972) established a distinction between the notions of Theme and Rheme, on the one hand, and the notions of Given and New, on the other, which had been treated together in previous accounts by linguists of the Prague School (see above).

According to Halliday, Theme is the information from which the speaker proceeds, “which serves as the point of departure of the message” (Halliday 1967:12; 1985:39) and together with Rheme, structures the clause as a message.

... is the element which serves as the point of departure of the message; it is that which locates and orients the clause within its context. The remainder of the message, the part in which the Theme is developed, is called in Prague school terminology the Rheme.

(Halliday & Matthiessen, 2004:64)

Theme in English is realized through initial position in the sentence, but this is just a language-particular realization. Independently from this type of structuring, Halliday studies the distribution of information into material which the speaker presents as Given (in the sense of predictable or recoverable), and New (not-recoverable). Therefore, the functions of Given and New are not the same as those of Theme and Rheme.

In this holistic description of language Halliday distinguishes three basic functions of language: textual, interpersonal and experiential (ideational).

The following table extracted from *An Introduction to Functional Grammar* shows a schema of these metafunctions of language and its impact on grammar:

Metafunction (technical name)	Definition (kind of meaning)	Corresponding status in clause	Favoured type of structure
experiential	construing a model of experience	clause as representation	segmental (based on constituency)
interpersonal	enacting social relationships	clause as exchange	prosodic
textual	creating relevance to context	clause as message	culminative
logical	constructing logical relations	–	iterative

**Figure 2.1** Metafunctions and their reflexes in the grammar. (Halliday & Matthiessen, 2004:61)

The experiential or ideational metafunction considers the clause to construe a model of experience with a logical subject as actor, or doer of the action.

The interpersonal metafunction considers the sentence as an exchange focusing on social relationships, having the grammatical subject (subject) as that of which something is predicated.

At the grammatical level, that is, within the textual system of languages, clauses are understood as message dealing with a psychological subject (theme) in that this acts as structural link with respect to their Rhemes, which at the same time develop it, and to which the message is concerned.

Considering the clause as a message its structure responds to a Theme accompanied by a Rheme, and whatever is chosen as Theme is put first extending up to and including the first topical, or experiential element: a Participant, an Attribute, a Circumstance or a Process.

Halliday and Matthiessen use the phrase *unmarked Theme* to refer to a typical declarative clause, where the subject and Theme overlap. Halliday & Matthiessen (2004:73) write “the Subject is the element that is chosen as Theme unless there is good reason for choosing something else.” “A Theme,” they add, “that is something other than the Subject, in a declarative clause, we shall refer to as a *marked Theme*. The most usual form of marked Theme is an adverbial group, (...) functioning as Adjunct in the clause.” (Halliday & Matthiessen, 2004:73). They also add that when an adjunct appears is to express a setting for the clause or a feature of contrast. And, what this marked Theme comprises is the marked element, being what would have been the unmarked Theme part of the Rheme (Halliday & Matthiessen, 2004:78).

They give an exhaustive accounting of unmarked Themes in other type of clauses: in questions that include the *wh-* element or in a yes/no interrogative, the finite verbal operator will play this role; in imperative clauses, the Predicator; and in exclamative clauses the nominal group or adverbial group.

Many detractors redefined this extensive body of work. Some of these detractors co-opted Halliday’s terms. One study conducted by Chafe & Li (1976:30) defines *given* information as the knowledge a speaker assumes as understood on the part of his or her addressee at the time of the utterance; while *new* information is the knowledge the speaker assumes he is introducing to the addressee’s.

Another well-known adaptation of Halliday’s work is that of Downing (1991), found in her article *An alternative approach to Theme: a systemic-functional perspective*. She agrees with Halliday in that Theme is the point of departure to the message, understood in the initial element. But she disagrees with Halliday’s concept of the Theme as “what the clause is about.” In her words, consider the initial element as point of departure is counter-intuitive, since the clause frequently has an adjunct in this first position.

Other authors such as Huddleston (1988; 1991) question the notion of Theme to the point of rejection. Downing and Huddleston, as Gómez-González (1994) points out, share the idea that the Hallidayan “what a clause as a message is about” does not necessarily constitute its point of departure. Gómez-González considers that Downing and Huddleston’s conception of “aboutness” is taken from a “referential” perspective, that is, an intuitive context dependent point of view.

In this sense, I align myself with Fawcett who rightly criticizes that ‘this general

concept has led some scholars—including Halliday—to assume that ANY element that occurs early in the clause must be some type of ‘Theme’ (Fawcett, 2007). As Fawcett says, many elements that appear in initial position such as linkers or binders, or mood marking elements are not true ‘Themes’. This phenomenon is addressed by a theory proposed by Lavid et al. (2010d) that will be used in this study and thoroughly explained in chapter 11.



# 3

## Theoretical Accounts of Theme in Spanish

As shown in the previous chapter, the notion of Theme has been object of controverted studies when applied to the English language. When applied to other languages such as Spanish the various understandings of Spanish Theme have their difficulties and problematic areas. These can be summarized as follows:

- the distinction between the notion of Theme and its language-specific realization
- the boundaries of Theme within the clausal structure
- the assignment of thematic status to the process element or to the unexpressed Subject

The first has been treated by scholars such as Munday (1998, 2000); Taboada (1995); McCabe (1999); McCabe & Alonso Belmonte (2001). In these works, English Theme is studied in a contrastive manner with Spanish highlighting the different realizations of Theme in both languages. Other authors such as Fernández Lorences (2010); Downing (1991) or Hidalgo Downing (2003) have focused on Theme boundaries within the clausal structure. In addition, existing accounts of Theme in Spanish have also highlighted the problem of assigning thematic status to the process element or to the unexpressed

Subject, as illustrated by the the work of Arús, Lavid, & Moratón (2012); Lavid (2010); Lavid, Arús, & Moratón (2010a).

However, as discussed by the author of this dissertation and colleagues: “The question remains open, however, as to what to do about the verbal inflection and how to integrate its presence in a theoretically sound model of Theme” (Arús, Lavid, & Moratón, 2012:p.176).

As acknowledged elsewhere, the account of Theme provided in SFGS (Lavid et al., 2010d: chapter 5) is “the currently most fully worked out version of Theme” (see Maxwell-Reid, 2011) and for this reason it was adopted in that study and will also be adopted in the current dissertation. In addition, and as will be shown below, the account of Theme provided in SFGS integrates two different levels of analysis: the local clause-level rank and the more global discourse context where thematic choices are motivated by different variables. As explained in Lavid et al.: “only by integrating these two levels we will be in a position to gain a better understanding of this textual category” (2010d:390). Given its relevance for the rest of this dissertation, Lavid et al.’s model of Theme is described in detail in section 3.1 below.

## 3.1 Lavid et al.’s Model of Theme in SFGS

As discussed elsewhere (Arús et al., 2013), Lavid et al.’s model of Theme proposal, manages to integrate and solve unresolved issues in previous accounts of Theme in Spanish and offers an integrated view of the phenomenon of Theme at clausal and discourse levels which, in spite of its complexity, make it particularly interesting for high-quality annotation. Some of the perceived advantages of the model are the following:

- The model proposes to use several layers of analysis to account for a morphologically-rich language such as Spanish.
- The model proposes to break up the Theme to account for a diversity of functions within the thematic structure of the clause. Particularly interesting is the operationalization of the Thematic Head, i.e., which is defined as the first element with a function in the experiential configuration of the clause which is more central to the unfolding of the text by allowing the tracking of the discourse participants (Lavid et al., 2010d:299).
- The model offers an understanding of Theme as a textual category, with choices at clause level motivated by their role in discourse organization. This is particularly interesting for our study of the correlations between Genre and thematic choices.

These three features are discussed in detail in the different subsections below.

### 3.1.1 Multilayered Analysis of Theme

Lavid et al.'s model of Theme is not a simple one. On the contrary, from the clausal perspective, it takes full consideration of the lexico-grammatical and morphological features of a Romance language such as Spanish and it also provides a characterization of the textual phenomenon of Theme which can account for this complexity. From the discourse perspective, it gives motivation for the behavior of the thematic features proposed at clausal-level.

From the clausal perspective, Theme is not restricted to one or two elements of the clause, but it works the notion of the “[c]omplex functional zone in clause-initial position serving a variety of clausal and discourse functions” (Lavid et al., 2010d:299). Within the Thematic Field it is possible to consider textual and interpersonal meanings, as well as Circumstances, preceding the first experiential element of nuclear transitivity, and to defer the identification of the thematic climax, and therefore its end, to a lower stage of analysis.

The Thematic Field consists of two main components: the Inner Thematic field and the Outer Thematic field. The Inner Thematic field consists of elements selected from the experiential structure of the clause and includes the Thematic Head and the PreHead. The Thematic Head is the nuclear part of the Inner Thematic field and is defined as “the first element with a function in the experiential configuration of the clause which is more central to the unfolding of the text by allowing the tracking of the discourse participants” (Lavid et al., 2010d:299). As illustration, Lavid et al. explain that in (1) the Thematic Head is ‘María’, while in (2) it is ‘Pablo’, because “in (1) the scene is described from the Recipient’s perspective, whereas in (2) it is presented from that of the Agent. In the former, the speaker aligns himself with the Recipient, so to speak, while in the latter he sides with the Agent” (Lavid et al., 2010d:393):

- (1) María recibió una carta de Pablo  
 Mary received a letter from Pablo

<i>María</i>	recibió una carta de Pablo
Thematic Head	Rhematic field

### 3. Theoretical Accounts of Theme in Spanish

---

- (2) Pablo envió una carta a María  
Pablo sent a letter to María

<i>Pablo</i>	envió una carta a María
Thematic Head	Rhematic field

In these examples, the local, clausal function of the Thematic Heads is to present some experiential element “as most closely aligned with the speaker’s advantage point of view” (Lavid et al., 2010d:299) and the result is that the thematised material is foregrounded or given prominence, both with respect to the remaining clause content, and to the rest of the textual material.

The Outer Thematic field is configured by elements which surround and complete the Inner Thematic field, such as linkers, binders and other textual markers; interpersonal elements which express the attitude and the evaluation of the speaker and constructions which are not integrated in the main predication, such as Participial clauses and Absolute Themes. Example (3) below illustrates the different layers of analysis reflecting the internal structures of the OTF and the ITF:

- (3) Pero, tristemente, después de la paliza, el chico falleció  
But, sadly, after the beating, the boy died

<i>Pero,</i>	<i>tristemente,</i>	<i>después de la paliza</i>	<i>el chico</i>	falleció
Textual	Interpersonal	PreHead	Head	Rheme
OTF		ITF		
Thematic Field				Rhematic field

This multilayered analysis has important advantages over previous accounts since it is capable of accounting for the complex thematic structure of a morphologically-rich language such as Spanish.

#### 3.1.2 Thematic Heads and Their Realization

An interesting feature of this model is the notion of Thematic Head as the nuclear element within the Inner Thematic field which can be realized overtly through inde-

pendent lexical and grammatical forms, such as Nominal Groups or nominal clauses, but which can also be encoded implicitly “through verbal prefixes or suffixes, depending on the function of the Head in the interpersonal structure of the clause” (Lavid et al., 2010d:299). When functioning as Subject, it is realized “as a verbal suffix indicating the person and number of the participant, as in (4) below, where the verbal suffix ‘-ó’ of the verb form ‘comprendió’ [understood] encodes the Thematic Head of the clause (Lavid et al., 2010d:394):

- (4) Acto seguido comprendió que esa voluntad era inútil  
He then realized that will was useless

<i>Acto seguido comprendi-</i>	<i>-ó</i>	<i>que esa voluntad era inútil</i>
PreHead	Head	Rheme
ITF		
Thematic Field		Rhematic field

But Thematic Heads can also function as Complements, as in (5) below, where the clitic ‘la’ refers to the participant Emma, who was mentioned in the previous co-text:

- (5) La engañaron el sello y el sobre  
Her deceived the stamp and the envelope  
She was deceived by the stamp and the envelope

<i>La</i>	<i>engañaron el sello y el sobrel</i>
Head	Rheme
ITF	
Thematic Field	Rhematic field

It is also possible to find a Thematic Head functioning as a Complement realized by a full lexical element and its clitic doubling, as in (6) below:

- (6) A la niña la llamaron Inés  
To the girl, her called Inés  
They called the girl Inés

<i>A la niña</i>	la llamaron Inés
Head	Rheme
ITF	
Thematic Field	Rhematic field

#### 3.1.3 A Textual View of Theme

As advanced above, Lavid et al.’s model offers an understanding of Theme as a textual category, with choices at clause level motivated by their role in discourse organization. This is particularly interesting for the study of the correlations between Genre and thematic choices. As Lavid et al. explain:

...the selection of certain clausal elements as part of the *Thematic* or the *Rhematic field* is not a random choice on the part of the writer, but is the result of the different functions performed by the *Thematic* and the *Rhematic* segments within the discourse. Thus, while elements within the *Rhematic* field elaborate the field/subject matter of the text, those elements within the *Thematic field* typically have a guiding or orienting function, signalling different possible points or angles of expansion of the text. Elements selected for *Thematic Heads* typically signal participants in discourse; elements expressing time within the *Pre-Head* can signal a new setting in time or a temporal scope along a temporal line of development; linkers, binders and other textual markers in the *Outer Thematic field* usually signal rhetorical transitions in the text.

(Lavid et al., 2010d:407)

In addition, and as shown by Lavid (2000a): “the selection of certain types of Themes different text types is not random, but correlates with the chaining strategy or line of development selected by the writer to organize his/her text”. Lavid explains that “texts organized around a spatial chaining strategy tend to initiate their clauses with Themes expressing location while those whose primary chaining strategy is temporal tend to select circumstantial Themes expressing time to help the reader construct episodic sequences.” This idea was further explored in connection with the notion of Genre in subsequent publications both for English (Lavid, Arús, & Moratón, 2012) and contrasting English and Spanish (Lavid, Arús, & Moratón, 2013c)

In the first study, the analysis of thirty-three English newspaper texts -equally divided into two groups of seventeen news reports and sixteen commentaries,- revealed that each newspaper Genre prefers certain thematic features and that the differences

between the two Genres are statistically significant. The study concluded that these thematic preferences can be attributed to Genre-related variables such as the communicative purpose or the subject matter of the text.

In the second study, the contrastive (English-Spanish) analysis showed that most of the clausal and discourse thematic differences found in the newspaper Genres analyzed can be related to the different communicative purposes which characterize these Genres. Thus, for example, it was found that news reports writers choose Sayers as Thematic Heads in clause complexes, while writers of commentaries choose Carriers as Thematic Heads. These two choices can be interpreted as textual devices: in the case of news reports to reflect the news reporter's tendency to remain neutral and attribute opinions to external sources, and in the case of commentaries to give their opinion in what seems to be an objective and impartial way. The contrastive analysis also revealed the use of concrete nouns realizing Thematic Heads in news reports in contrast to the preference for abstract nouns in commentaries, and the preference for longer and more complex Noun Groups in commentaries than in news reports (see Lavid, Arús, & Moratón, 2013c:280).

Given the fact that one of the aims of this dissertation is to examine the relationship between Theme and Genre in English and Spanish newspaper texts, the following sections will be dedicated to these notions. I will begin with the notion of Genre and its relationship with related notions such as Register and Text Types, mainly as studied within the Systemic-Functional literature. Later I will focus on the relationship between Theme and Genre, focusing not only on Theme selection but also in Thematic Progression (TP). In both sections the overview will not be comprehensive but purely illustrative and selecting those studies which are more relevant for the current dissertation.

### *3. Theoretical Accounts of Theme in Spanish*

---



# 4

## Genre, Register and Text Types from a Functional Perspective

In the past recent years the study of Genre has gained immense popularity, from very different perspectives, and has become a very active research area. This does not mean that this notion has reached a clear cut definition or common agreement, nor that it has been made completely distinguishable from the concepts of Text type and Register. However, there are interesting proposals which capture distinctions between these related notions and deserve attention. In the following sections, I will, therefore, provide a brief overview of how these related concepts have been defined and used by different authors, with a special focus on how Genre has been conceptualized in the Systemic-Functional and in the ESP approaches. Afterwards, it is provided a brief description of the rationale behind the selection of the three newspaper Genres used for this study.

### 4.1 Text Types

Text types traditionally refer to rhetorical categories, such as narration, description, exposition and argumentation. The number and the labels used to identify different

types can vary widely in the literature (Faigley & Meyer, 1983): for instance, Werlich (1976) proposes five text types (narration, description, exposition, argumentation, and instruction); de Beugrande & Dessler (1981) defines seven text types (descriptive, narrative, argumentative, scientific, didactic, literary, and poetic); and Adam (1992) refers to five text types (récit, description, argumentation, explication, and dialogue). Some scholars use the label “text types” to indicate instrumental or practical Genres, as opposed to literary Genres (see Görlach, 2004). Others use “Genres” as synonymous with “text types” (see Stubbs, 1996; Karlgren, 2000). Finally, some scholars use the term “text types” without any further explanation on the interpretation of the term (see Kilgariff & Grefenstette, 2003). Douglas Biber defined the term as follows:

I use the term ‘text type’ (...) to refer to groupings of texts that are similar with respect to their linguistic form, irrespective of Genre categories. (...) For example, a science fiction text represents a Genre of fiction (relating to author’s purpose), but it might represent an abstract and technical text type (in terms of its linguistic form), similar to some types of academic exposition and different from most other fictional texts. In a fully developed typology of texts, ‘Genres’ and ‘text types’ must be distinguished, and the relations among them identified and explained.

(1988:70)

Under this scope Biber distinguishes eight text types: “Intimate personal interaction,” “Informational interaction,” “Scientific exposition,” “Learned exposition,” “Imaginative narrative,” “General narrative exposition,” “Situated reportage”, and “Involved persuasion.” Being these types defined according to its internal linguistic characteristics, Biber (1995:320) points to its empirical classification using a statistical procedure called *cluster analysis*. According to it, linguistic similarities are identified and represented in a chart that groups texts on the basis of their similarities with respect to their dimension characteristics. These internal differences allow for its maximal distinction in terms of their linguistic characteristics. However, this classification is made at the level of individual texts, which may be part of the same “Genre” (in terms of external criteria). This theory has raised many criticisms on the validity of its empirical validation but it is so far the most suggestive work on the area.

## 4.2 Genre: English for Specific Purposes and Systemic Functional Approaches

The notion of Genre has been developed within different traditions, both literary and linguistic, with a lot a common ground but with different emphasis. Within the linguistic traditions I will focus on the two approaches that have informed the work reported in this dissertation, namely, Systemic-Functional Linguistics (SFL), on the one hand, and English for Specific Purposes (ESP), on the other.

### 4.2.1 Genre in SFL Approaches

In this section I will focus on the treatment given to Genre in Systemic-Functional Linguistics (SFL)<sup>1</sup>. This is in line with the overall functional approach undertaken in this dissertation and also because the SFL approach has contributed richly to how the notion is understood and applied in textual analysis and language teaching over the last twenty-five years.

Systemic Functional Linguistics (SFL) operates from the premise that language structure is integrally related to social function and context. Accordingly, the concept of Genre is used to describe the cultural context in a language, and it is viewed as a goal-oriented social activity. In fact, SFL approaches to Genre have been influenced most widely by the work of J.R. Martin, who defined Genres as ‘staged, goal-oriented social processes through which social subjects in a given culture live their lives’ (Martin, 1997:43)

As further explained in Martin, Christie, and Rothery, Genres function as *social processes* “because members of a culture interact with each other to achieve them; as goal-oriented because they have evolved to get things done; and as staged because it usually takes more than one step for participants to achieve their goals” (1987:59)

According to Martin, whose approach builds on Halliday’s work, the concepts of Genre and Register are closely related: while Register functions on the level of context of situation, Genre functions on the level of context of culture. However, as explained by S. Eggins, the context of culture is an abstract and general concept and cannot be generally identified in language. Its realization in the language is mediated in two ways: first, by specifying a common staging organization, or schematic structure; second, by the second level of context (i.e. Register) through specifying the Genre potential of a

---

<sup>1</sup>For an overview of Genre history, theory, research and pedagogy, see Anis S. Bawarshi and Mary Jo Reiff 2010. *Genre: An Introduction to History, Theory, Research, and Pedagogy*. 2010 by Parlor Press and The WAC Clearinghouse, ISBN 978-1-60235-171-4

particular in terms of the possible configurations of Register variables allowed within a given culture at a given time (Eggins, 1994:35). In Eggins' model there are two fundamental concepts in the description of the schematic structure: constituency and labeling. Constituency refers to the part/whole relationship between elements of some structure, and labeling is the act of assigning functional labels to these constituents. These constituents can be obligatory, optional or recursive, and a Genre is defined according to the elements of a schematic structure. Texts belonging to the same Genre perform the same stages of organization or schematic structure (Eggins, 1994:37).

A highly relevant contribution to generic analysis in the SFL community is the model of *Generic Structure Potential* (GSP) proposed by Halliday & Hasan (1989), which encapsulates commonality and variation of schematic structures between texts of the same Genre (more on this model in chapter 19.2).

#### 4.2.2 Genre in ESP Approaches

The field of English for Specific Purposes conceptualizes Genre as “conventionalized communicative event” and focuses on academic and research English and their use for applied ends (see Swales, 1990; Bhatia, 1993).

Most English for Specific Purposes (ESP) studies of Genre adopt a text-based perspective in which a Genre is seen as a class of communicative events that share common purposes that are recognized by the members of the discourse community that owns it. These purposes determine the structure and linguistic choices in the Genre. The focus of research in ESP studies is the interaction of communicative purpose, rhetorical structure, and linguistic choice: texts belonging to a Genre are seen as sharing some communicative purpose, textual features, and content.

Similarly, Genre researchers of in SFL (mainly the so called Sydney school) are interested in the relation between context, communicative purpose, form, and function, and they concentrate on the analysis of textual features. However, there are differences between these two approaches: while both share the fundamental view that linguistic features are connected to social context and function, SFL approaches ‘generally focus on pre-Genres and tend to define context at a fairly macro-level’ (Bawarshi & Reiff, 2010:44). As explained above, SFL approaches locate Genre at the level of *context of culture*, while ESP approaches locate Genres “within more specifically defined contexts (what Swales first termed *discourse communities*), where the Genres’ communicative purposes are more specified and attributable” (Bawarshi & Reiff, 2010:44).

In ESP Genre analysis, the point of departure is the notion of ‘communicative

purpose’ and, therefore, a typical ESP approach to Genre analysis will begin by identifying a Genre within a discourse community and defining the communicative purpose the Genre is designed to achieve. In a subsequent step, as explained by Bawarshii & Reiff (2010:46), the Genre’s organization —its schematic structure— is examined and often characterized by the rhetorical “moves” it undertakes. This is followed by an examination of the textual and linguistic features (style, tone, voice, grammar, syntax) that realize the rhetorical moves. According to these authors, the trajectory of the analysis –from a Genre’s schematic structure to its lexico-grammatical features- is a not a linear process, but one “that tends to move from context to text, with context providing knowledge of communicative purpose and discourse community members’ Genre identifications” (Bawarshi & Reiff, 2010:46).

One of the best known ESP researchers –Vijay Bhatia– outlines seven steps to analyzing Genres, which reflect the trajectory described above. These steps include placing a given Genre-text in its situational context; surveying the existing research on the Genre; refining the researcher’s understanding of the Genre’s discourse community; collecting a corpus of the Genre; conducting an *ethnography* of the institutional context in which the Genre takes place and deciding the level of linguistic analysis to explore (lexico-grammatical features, text-patterning and/or structural interpretation. “In the final step, Bhatia advises researchers to seek a specialist informant from the research site to verify findings” (Bawarshi & Reiff, 2010:36).

In the present study news reports, editorials and letters to the editor will be taken as Genres, and an effort has been made to ensure both Genre and Register comparability (see below) in the corpus compilation phase.

### 4.3 Register

While the notion of Genre is associated with the context of culture, the notion of Register is associated with the context of situation and defined, according to Halliday, as “The total environment in which a text unfolds” (1989) and includes three main variables: the field, the tenor and the mode of discourse. The field of discourse refers to what is happening, the nature of the social action that is taking place. The tenor of discourse refers to who is taking part, the participants, their statuses and roles. The mode of discourse refers to the symbolic organization of the text type (according to Biber), the channel of communication (spoken/written, mono- logic/dialogic, +/- visual contact, computer-mediated...)

In this study the selection of news reports, editorials and letters to the editor have

#### *4. Genre, Register and Text Types from a Functional Perspective*

---

undergone a modest Register analysis to ensure comparability of the English and Spanish texts compiled. In the following section I will provide a brief account of the reasons motivating the selection of this corpus.

# 5

## Newspapers Genres used for this Study

In this dissertation I have focused on three main newspaper genres as the basis for the study of thematic patterning in English and Spanish. In spite of the changes introduced by the online availability of most journalistic texts, the three newspaper genres of news reports, editorials and letters to the editor have not varied in mission or goal. These three genres are still what they have traditionally been: objective information, the voice of the newspaper and the voice of the readers, respectively. Their selection as the basis for this study is not random but clearly motivated by the following reasons:

- Accessibility and availability: The three genres are widely available online and in different languages on the Internet. Availability of similar newspaper texts in different languages ensures a certain degree of comparability for contrastive corpus studies, which need to use comparable texts for the study of cross-linguistic and cross-cultural differences between two or more languages.
- Systematicity and coherence in text categorization: Unlike other online resources, the three genres contain text which has been professionally written and edited, usually with a certain level of quality and standardization in terms of spelling, Genre conventions and other features (Andersen, 2012). Its large and regular text production ensures for a systematical compilation which results in the increase

of the size of the corpus whenever this is needed, not only in English or Spanish, but in any other language in which newspapers are published.

- Variability in communicative purpose and textual conventions: The three genres are characterized by different communicative purposes and each has its own features and generic conventions. They are, therefore, good candidate texts for the corpus analysis and annotation tasks carried out in this dissertation.

In the following sections each of these particular genres is briefly characterized.

## 5.1 News Reports

Bell (1991:14) distinguishes two classes of news: *hard news* and *soft news* that include *feature articles* and *special-topic news*. *Hard news* is “time-bound to immediacy, in contrast to *soft news* which do not have such a temporal dependency.”

White (1997) also points out that hard news reports are extensively studied as a Genre, taking into account that mass-media news reporting is “one of the most influential textual domains in contemporary society,” reflecting its “social and ideological objectives” (1997:101). This author also distinguishes between two main types of news reports: the *event story* and the *issues report*. The latter, as the object of the present dissertation, is defined as “grounded in a communicative event – such as speeches, interviews and press releases.” They are communication-based rather than event-based, striving to remain neutral and objective while presenting a diversity of opinions, voices, and perspectives of the event, incident, or issue under discussion. These must be impartial and objective and this is reflected in its semantics and lexico-grammatical choices. As explained elsewhere, both English and Spanish news reports share these features and in both languages news reporters minimize interpersonal involvement in the construction of the text, attributing citations to external sources (Lavid, Arús, & Moratón, 2013c). It is these different strategies what, according to Thompson, attract readers read news reports:

The main expectation that a reader brings to a newspaper report on current affairs is that it will give a reasonably objective “true” picture of the event and of the evidence, reactions, and opinions of the people involved.

(Thompson, 1994:151)



## 5.2 Editorials

An editorial is defined by Sinclair (1995) as “an article in a newspaper that gives the opinion of the editor or publisher on a topic or item of news” (in Ansary & Babaii, 2005). In other words as Roy defines it, it is “the voice of the newspaper. An editorial seeks to encourage critical thinking, to mould opinion, and to promote action.” (2011:109). The newspaper states its opinion and its policy on any number of affairs in the editorial. It is normally written by the editor or by any staff member under the editor’s instructions. Although these are not usually signed they enact some type of ideology with the aim of furnishing the reader with background information about the news of the day, interpreting or explaining current events through the newspapers point of view. Even though editorials are a clear vehicle of opinion, Roy (2011:109) advises to writers of editorials to avoid the first person just as in news stories. Despite the distinction between objective news reports and editorial opinion they do not represent a simple dichotomy, but rather a spectrum on which texts can be located on the basis of how much and what type of interpersonal features they exhibit (Lavid, Arús, & Moratón, 2013c).

## 5.3 Letters to the Editor

Letters to the editor (henceforth LE) have not received the same attention as other journalistic genres by discourse analysts and genre researchers. This is due to the fact that LEs form a difficult Genre due to its variation in styles, linguistic features and purposes.

It is difficult to find a canonical definition for letters to the editor. Mémet (2005) defines them as “reactions to publications” and Pounds (2010) extends this definition:

In opinion letters to the paper the writers take position by expressing their views on current events and states of affairs and by appealing for action. In so doing they also reveal their feelings towards such events and provide their evaluation of the people and circumstances involved (...) undermining or aligning with the positive aspects of the critized situation.

(Pounds, 2010:50)

Studies on the topic have mainly focused on the argumentative structure of letters in English (see Ghadessy, 1983), and on the argumentative and the specific attitudinal meanings associated with the various components of such structure in two languages (see Pounds, 2010; Mémet, 2005; Bronia, 2005; Vázquez y del Árbol, 2005). Recent

work carried out by the author of this dissertation and Professor Julia Lavid tries to fill a gap in this area, by analyzing the relationship between intersubjective positioning and thematization in English and Spanish in this Genre (see Lavid et al., 2013c; Lavid & Moratón, 2014).

In the following section I will provide a brief overview of previous work related to the relationship between Theme and Genre.

# 6

## The Relationship between Theme and Genre

As mentioned in the Background chapter , the phenomenon of Theme and its fundamental role in the construction of messages and texts has attracted the attention of many scholars within the Systemic-Functional approach, but also within other related functional approaches. Therefore, it is not surprising that, given its fundamental role for textual organization, a considerable number of studies have touched upon or developed in depth the connections and correlations between thematic patterning (including thematic selection and progression) and textual variation in terms of Registers, Genres and Text types.

As Vande Kopple (1991:331) suggests, to focus on the study of textual, interpersonal, and ideational Themes, their position and progression “might help in the distinction of Genres.” In SFL among the first scholars who tried to establish an association between a text’s method of development –the theme types selected- and its thematic progression, on the one hand, and the Genre to which the text belongs, on the other, was Fries (1983, 1994, 1995a,b), who set four different hypotheses (1995a:7):

1. Different patterns of thematic progression correlate with different genres. i.e., patterns of thematic progression do not occur randomly but are sensitive to genre.
2. The experiential content of Themes correlates with what is perceived to be the

method of development of a text or text segment.

3. The experiential content of Themes correlates with different genres.
4. The experiential content of Themes of a text correlates with different generic elements of structure within a text.

While Fries did not empirically test these hypotheses himself, different scholars have investigated one or several of these in different contexts. Thus, for example, Gill Francis -drawing upon essential groundwork by scholars such as Fries (1983), Jordan (1985) and Martin et al. (1984) - explored how three newspaper genres can be distinguished from each other in terms of the information typically contained in the Themes of their component clauses and in their thematic progression patterns (Francis, 1989). Her analysis showed interesting coincidences between the theme types selected and the genres she studied, but her hypothesis was not tested since she did not perform any statistical test to prove that her results were not random. Nor could she identify characteristic thematic progression patterns associated with the newspaper genres studied.

Other scholars have also tried to establish a link between both a texts's method of development, its thematic progression and its genre: the two volumes by Ghadessy (1983, 1995), and his own article on thematic organization as a criterion for establishing English text types (1997), are good exemplars of these first attempts, but the descriptive comparisons provided focus on similarities and differences among different Registers/Genres in a general sense, without performing statistical tests that would prove that the differences are significant.

The empirical studies by Julia Lavid advanced significantly in this area by: a) examining the relationship between text types, characterized by specific discourse strategies, and the phenomenon of thematization in English (Lavid, 2000a); b) extending the analysis to other languages (Lavid, 2000b), and c) proving the existence of statistically-significant correlations between those strategies and the semantic types of Theme selected to signal them (Lavid, 2000a,b, 1998). More recent work by this author and colleagues focusing on the signaling role of thematic choices in newspaper genres has extended the scope and the empirical validity of those previous studies (see Lavid, Arús, & Moratón, 2012) and contributed significantly to a fruitful research line on the role of interpersonal and textual features in the thematic organization of texts, represented by studies such as the one by McCabe & Heilman (2007), McCabe (1999) or Eiler (1986), among others.

Other studies have focused on how Thematic Progression (TP) can help in the identification of different genres. For example, Taylor carried out a computer-assisted

---

study of eighteen high school textbooks, finding an interesting correlation between the selection of propositions as Themes in history books and conditions in science texts. He also noted the value of the data obtained in the analysis of thematic organization. Lorés (2004) analyzed the thematic progression and method of thematic development of research articles, identifying two characteristic patterns. Khany and Nejad also focused on research articles and identified characteristic thematic patterns within the rhetorical moves, pointing out that “the rhetorical moves deal with the larger segments of texts and the rhetorical knowledge renders a uniform orientation to the segments” (2010:50). For these authors, Theme contributes significantly to the organization of discourse but to a smaller extent than the generic moves. Therefore, they claim, “it seems important to see how these structural categories interact in the research articles and, more specifically, how thematic elements function within the generic moves.” (Khany & Nejad, 2010:50).

Recent work by the author of this dissertation and colleagues has contributed significantly to this research line by revealing interesting correlations between the generic moves of news reports, editorials and letters to the editor, on the one hand, and the TP patterns found in those genres, both in English and in Spanish (Moratón et al., 2012; Moratón & Lavid, 2013). The studies suggested that the occurrence of certain TP patterns in each generic move is not random, but is a useful textual strategy to signal the transitions between those moves. This issue will be taken up further in chapter 19.



# 7

## State-of-the-art Issues in Corpus Annotation

In this section I will outline a number of state-of-the art issues in the task of corpus annotation both from an empirical and a computational perspective. I will first begin with general definitions and will later proceed to describe the main challenges underlying manual and automatic corpus annotation as background to the empirical work presented in this dissertation.

### 7.1 Definitions

Within the field of Corpus Linguistics, corpus annotation was first defined by Geoffrey Leech as “the practice of adding interpretative, linguistic information to an electronic corpus of spoken and/or written language data” (1997).

Currently the corpus object of annotations is understood by McEnery, Xiao, and Tono as:

(...) the prototypical corpus is a machine-readable collection of language used in authentic settings/contexts: one that is intended to be representative for a particular language, variety, or register (in the sense of reflecting all the possible parts of the intended language/variety/register), and that is intended to be balanced such that the sizes of the parts of the corpus correspond to the proportion these parts

make up in the language/variety/register.

(2013:264)

This corpus is where linguistic annotations are carried out, that is “(...) any sort of annotation which one might like to add to linguistically derived data. This might mean adding free-format notes to specific points or spans in the data, or it might mean applying systematic codes, which can then be analyzed statistically” (Carletta et al., 2002).

Leech advocates for ‘best practice’ in the process of corpus annotation by proposing seven annotation maxims:

1. The annotations should be removable, so the raw corpus can be shown.
2. The annotated corpus should allow the maximum flexibility to be manipulated by the user, that is, if it has more than one layer of annotation these should be able to extract these annotations one by one.
3. Guidelines should be available to the end user providing full details of the annotation scheme used by the annotators.
4. It should be made clear how and by whom the annotation was carried out, either if it has been automatically or manually annotated.
5. The end user should be made aware that the corpus annotation is not infallible, the corpus is a tool not exempt of possible mistakes and always prone to be improved.
6. Annotation schemes should be based on agreed theory principles.
7. Annotation schemes are not a standard as such, these emerge through consensus.

Similarly, other well-known corpus linguists advocated for corpus annotation as an activity that gives ‘added value’ to a corpus in terms of reusability, stability and reproducibility (McEnery et al., 2006). These terms will be separately treated in more detail due to its importance in the following subsection.

In the NLP camp, researchers defined corpus annotation as “the process of enriching a corpus by adding linguistic and other information, inserted by humans or machines (or a combination of them) in service of a theoretical or practical goal” (Hovy & Lavid, 2010:13). Accordingly, corpus annotation can be done manually (by humans), automatically (by computer programmes) or by a combination of both.



Manual annotation is a fundamental process for the creation of technical devices with human language processing capabilities, which “would not have been possible today without the legions of researchers that manually analyzed human language for centuries” and is “still the principal tool to improve our understanding of language” (Eckart de Castilho, 2014:12). Traditionally, manual analysis was done with pen and paper, but today it is increasingly being supported by software (e.g. annotation editors). In the field of NLP, moreover, manual annotation has become a highly-controlled process, subject to scientific and rigorous methodologies that will ensure the consistency and reliability of the annotation process since the main concern is obtaining ‘high-quality’ and processable data on which computer algorithms can be trained for optimal performance.

Automatic annotation is carried out “when huge bodies of text need to be analyzed or when language processing capabilities are to be integrated into devices” (Eckart de Castilho, 2014:15). Automatic annotation requires a sufficiently large body of manually analyzed data, the design of machine learning algorithms, and the evaluation of their results against the manually created gold standard. Therefore, in order for automatic analysis to be successful and achieve optimal performance, a manually-created ‘gold-standard’ is necessary to ensure the quality of the annotations.

In the following sections I will focus first on the main challenges which have emerged and are still subject to debate in manual corpus annotation and will proceed with some state-of-the art issues in automatic annotation and its integration with manual annotation.

## 7.2 Challenges in Manual Corpus Annotation

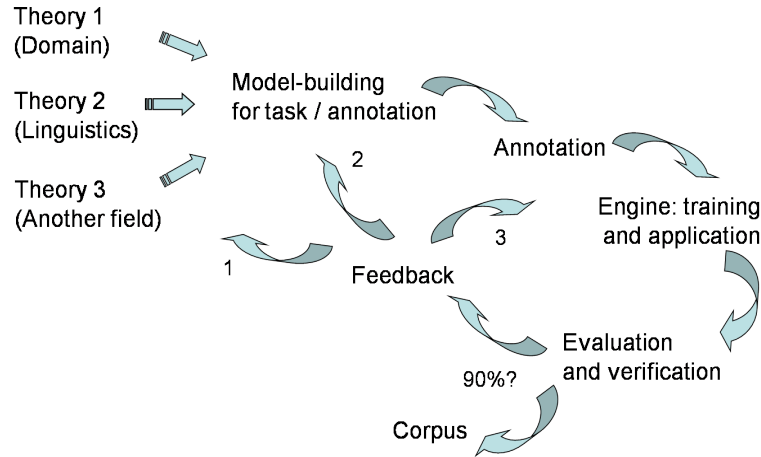
Given the necessity of ‘high-quality’ data for further processing in the NLP community, researchers in this field have developed rigorous scientific methodologies to ensure the reliability and the consistency of the manual annotation process.

The standard methodology in NLP when annotating a corpus consists of two principal stages: first, to have humans manually annotate texts (the ‘training corpus’) with the desired tags; second, to train computer algorithms or to apply rules on the corpus to perform the same job. As explained by (Hovy & Lavid, 2010:16), the manual annotation process for NLP consists of the following steps:

1. Identifying and preparing a selection of the representative texts as starting material for the ‘training corpus’ (sometimes called ‘training suite’).

2. Instantiating a given linguistic theory or linguistic concept, to specify the set of tags to use, their conditions of applicability, etc. This step includes beginning to write the annotator instructions (often called the Codebook or Manual).
3. Annotating some fragment of the training corpus, in order to determine the feasibility both of the instantiation and the annotator Manual.
4. Measuring the results (comparing the annotators' decisions) and deciding which measures are appropriate, and how they should be applied.
5. Determining what level of agreement is to be considered satisfactory (too little agreement means too little consistency in the annotation to enable machine learning algorithms to be trained successfully). If the agreement is not (yet) satisfactory, the process repeats from step 2, with appropriate changes to the theory, its instantiation, the Manual, and the annotator instructions. Otherwise, the process continues to step 6.
6. Annotating a large portion of the corpus, possibly over several months or years, with many intermediate checks, improvements, etc.
7. When sufficient material has been annotated, training the automated NLP machine learning technology on a portion of the training corpus and measuring its performance on the remainder (i.e., comparing its results when applied to the remaining text, often called the 'held-out data', to the decisions of the annotators).
8. If agreement is satisfactory, the technology can be applied to additional, unannotated, material of the same type, thereby assisting future analysis. If agreement is not satisfactory, the process repeats, possibly from step 2, or possibly from step 6 if more training data is required (see Hovy & Lavid, 2010:16).

These steps are graphically represented in figure 7.1, the generic annotation pipeline:



**Figure 7.1** The generic annotation pipeline (after Hovy & Lavid, 2010)

Although all these steps are relevant for the work presented in this dissertation, in the following subsections I will focus on two main issues in manual corpus annotation –annotation procedure and reliability and evaluation measures- since these pose interesting challenges for one of the main aims of this dissertation, namely, the empirical validation of aspects of Theme through corpus annotation.

## 7.3 Manual Annotation Procedures

In this section I will refer to a number of fundamental issues in manual annotation, including: the annotation roles, the annotation process and the annotation tools.

### 7.3.1 Roles

The main roles in the manual annotation include:

**Annotators** perform manual annotation based on annotation guidelines. This task is slow and error-prone, depending on the complexity of the guidelines. The annotators should have domain skills to be able to interpret the guidelines correctly, and this makes overall corpus production very expensive. The less-specialized annotators allow for a reduction in expenses. Therefore, the better the interface, and the better the guidelines, the better development of the project.

**Editors or curators.** To curate is “to collect the best examples and or eliminate errors” (Petrillo & Baycroft, 2010). In some cases the role of curator is indistinguishable from that of the editor. The curator is responsible for measuring the Inter-Annotator

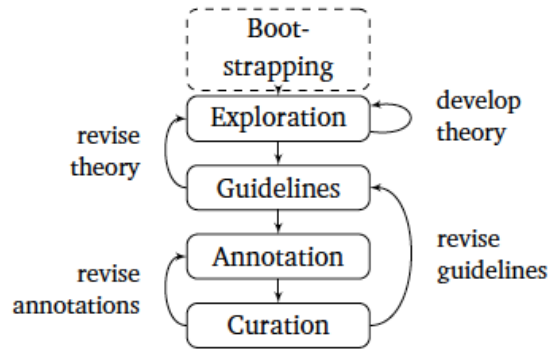
Agreement (IAA), annotation adjudication, gold-standard production, annotator training, and make expert choices in disagreement between annotators (Bontcheva et al., 2010). In projects where an annotation can be “right” or “wrong” the curator produces a gold standard annotation with the correct choices. They have close relationship to the guidelines: either having written them, or as a result of the work they do to improve their accuracy.

**Explorer.** According to Eckart de Castilho (2014) the explorer makes a preliminary exploration of the corpus data in order to generate a new hypothesis or corroborate an existing one. This exploration takes place before the annotation process begins and the guidelines are created. In some cases, after the exploration of the corpus, he/she undertakes the role of guideline author.

Other authors as Dipper et al. (2004) also contemplate the role of a **language engineer** who may explore the data from a statistical point of view with the aim of exploiting the results of the annotation process.

### 7.3.2 Process

The process of manual annotation and the roles that support it is depicted by Eckart de Castilho (2014) in the following figure:



**Figure 7.2** Basic annotation process (in Eckart de Castilho, 2014)

As the figure above shows, the process involves five different stages: bootstrapping, exploration, writing of annotation guidelines, annotation, and curation. This process, although defined in a linear way is normally iterative at some stages. As the arrows in the diagram show the means by which the hypothesis will be formulated, the process may have to be reviewed and adapted more than once. The exploration process may be necessary to revisit once the guidelines are written, in order to ensure the clarity and legibility of the instructions.

Once the annotation as such begins, the guidelines may not be clear enough for the annotator to use them, and a further composition will be needed. The process will undergo this stage again once more when the curator enters the cycle, in case the results of the Inter-Annotator Agreement reveal that the quality of the process is not reliable enough.

There is a brief explanation of the actions that take place in each stage of the annotation development.

**Bootstrapping** is not purely a stage of manual analysis by itself, but rather it is thought to be the preparation of the corpus for manual annotation. It is usually an automatic process; that is, the plain corpus is pre-annotated automatically, and then a parser, tagger, or gazetteer is applied to prepare the corpus for a higher level of manual annotation. Or as it is the case of the present study to be annotated according to a rule-based system.

**Exploration** phase consists on the inspection of the corpus by the explorer who will pose a hypothesis and gather relevant examples to include in the process of writing the guidelines.

**Annotation** takes place once the guidelines have been written. As defined in section 7.1, linguistic annotation is “any sort of annotation which one might like to add to linguistically derived data.” The annotation process may start by a single researcher who sets the ground for the annotation tasks by a larger group at the same time that the explorer examines the corpus and prepares for the formulation of the hypothesis, both together help in the process of writing the guidelines.

**Curation** phase unifies the annotations under the expert choices made by the curator. In this stage the curator, either manually or automatically, decides which of two or more not co-occurrent annotations is correct. In this stage the Inter-Annotation Agreement is made the guidelines may be revisited and reformulated to achieve an improved annotation quality.

### 7.3.3 Annotation Tools

The selection of a software to support linguistic annotations is another key feature in this study.

The need for a tool that includes the ability to design and program, as well as a flexible display and interface, is not an easy task. The recent interest in more sophisticated corpus processing mean that both linguists and computer scientists are demanding more complex processing pipelines for their analysis tasks (Eckart de Castilho et al., 2009).

Another relevant feature to take into account when selecting the annotating tool is, as Carletta et al. (2002) rightly suggests, that it must use a well-understood, documented data format, or if it uses proprietary formats, that these can be converted to general available formats in order to share these data with others and to inspect the coded data. Of course, there is the possibility to create ad-hoc tools but these are time consuming impeding the scientific progress. This raises the issue of integrating/harmonizing different types of annotation that have possibly been produced by different tools.

In the following paragraph I review some of the most complete annotations tools that serve the needs of this study. Here, a deeper look will be taken at GATE Tool.

### **GATE General Architecture for Text Engineering**

is an infrastructure for developing and deploying software components that process human language. It is nearly 15 years old and is in active use for all types of computational task involving human language. GATE excels at text analysis of all shapes and sizes.

(Cunningham et al., 2011)

GATE has a twofold advantage. For expert programmers it provides a software architecture for processing data, and for non-expert programmers it provides an easy to use Integrated Development Environment (IDE), focused on information extraction and allowing for different pipeline combinations of the available resources as well as introducing your own implementations. Tag spans are identified by clicking and sweeping and may contain both simple values and attribute-value pairs for fine-grained information. They are stored using character offsets from the start of the file (Carletta et al., 2002).

The outstanding characteristics of GATE can be summarized in the following points:

1. JAPE language (Java Annotation Patterns Engine), provides a domain-specific language for rule-based analysis tasks. JAVA can be used in the right-hand side of JAPE rules, making the language much more powerful (Eckart de Castilho, 2014). Ultimately this allows to create *ad-hoc* APIs to be adapted to your particular needs, as it is the case in the present study.
2. GATE presents integrated resources known as CREOLE (a Collection of REusable Objects for Language Engineering). These resources, available within GATE or downloadable from public repositories can be conveniently

assembled and exported as Java Archive (.JAR), and a XML configuration file. This allows for the sharing and reuse of the created resources in other applications. The available resources to be combined within GATE can be classified in:

*Language Resources (LRs)*: refers to lexicons, corpus, and ontologies  
*Processing Resources (PRs)*: algorithm-based like parsers or generators  
*Visual Resources (VRs)*: elements that affect the graphic interface

3. A built-in annotation editor allows manual text correction and annotation. It can correct automated annotated texts and inspect the results of other annotations. Eventually it makes changes to the original text, something that in most processing frameworks is not possible.
4. ANNIC (ANnotation In Context) sub-system allows searching a corpus using JAPE patterns. This querying engine allows for maximum exploitation of annotated data with statistical analysis.

### **UIMA Unstructured Information Management Architecture**

It is an open framework designed to integrate language analysis developed at IBM Research location encouraging the reuse of components.

The motivation driving the development of UIMA was the integration of language analysis developed at different IBM Research locations into a common framework and to encourage the reuse of components.

UIMA is a platform-independent XML-based descriptors for analysis workflows and analysis components that are predominantly used in the framework to facilitate the deployment and interoperability of UIMA workflows. Plug-ins are provided for the Eclipse development environment. It was designed to support the interaction of both expert and non-expert programmers so that both linguistic and computing work can be supported by the platform (Ferrucci & Lally, 2004).

### **U-Compare**

is an integrated text mining/natural language processing system based on the UIMA Framework. It provides an easy-to-use interface with no installation needed. Available components are included in the U-Compare single-click-to-launch package, ready to use by just drag-and-drop in the online interface. External tools can also be added to the workflow. It provides corpus readers and

writers with syntactic and semantic and annotation tools as well as statistical tools for data exploitation (Kano et al., 2009).

## 7.4 Evaluation Measures

According to Hovy and Lavid: “The underlying premise of (manual) corpus annotation is that if people cannot agree enough, then either the theory is wrong (or badly stated or instantiated), or the annotation process itself is flawed. In any case, training of computer algorithms is impossible on inconsistent input” (Hovy & Lavid, 2010:23). As explained by (Lavid, 2011:208): “... this methodology is also used in the field of content analysis, where observers/analysts are supposed to interpret textual, visual or audible materials”. Thus, scholars such as Hayes and Krippendorff state the following:

When relying on human observers, researchers must worry about the quality of the data—specifically, their reliability. Are the data being made and subsequently used in analyses and decision making the result of irreproducible human idiosyncrasies or do they reflect properties of the phenomena (units of analysis) of interest on which others could agree as well? The answer to this question can affirm or deny the usability of one’s data, and respectable scholarly journals typically require quantitative evidence for the reliability of the data underlying published research.

(Hayes & Krippendorff 2007:77, quoted in Lavid 2011)

According to Krippendorff, reproducibility is the strongest and most feasible kind of reliability to test among other possible kinds, such as stability and accuracy (Krippendorff, 2004). Reproducible data are those obtained within a tolerable margin of error when using a coding instrument by different observers on the same set of phenomena (Hayes & Krippendorff, 2007). Reliability is measured by the agreement observed among independent observers. “The more observers agree on the data they generate, and the larger the sample of units they describe, the more comfortable we can be that their data are exchangeable with data provided by other sets of observers (c.f. Hayes et al., 2005), reproducible, and trustworthy” (Hayes & Krippendorff, 2007:77). Transferring this method to the field of Corpus Linguistics, states Lavid, “the coding instrument are the annotation scheme and guidelines, and researchers use agreement measures to test stability (intra-annotator agreement) and reproducibility (inter-annotator agreement)” (2011:209).

The most frequent measures used in the NLP community for inter-annotator agreement are *simple agreement*, *Scott’s pi*, various forms of *Kappa* (Cohen et al., 1960), *Krippendorff’s alpha* (Krippendorff, 2004), and *Metric agr* (Wiebe et al., 2005).



**Percent Agreement or Observed Agreement** measures the percentage of cases that two (or more) annotators agree over all the decisions they have both made. It is the percentage of annotations on which two coders agree when annotating the same sample independently. It is easy to understand and quick to determine (Artstein & Poesio, 2008). This measurement has two main drawbacks:

- The problem of agreeing by chance is not corrected and it becomes more severe when there are few classification categories in the coding scheme. Agreement becomes more difficult to achieve as the number of available categories increases (Sullivan, 2014).
- It does not correct for the distribution of items among categories: “we expect a higher percentage agreement when one category is much more common than the other” (Artstein & Poesio, 2008).

**Scott’s  $\pi$**  solves the issue of chance agreement present in Percentage Agreement but it can only be applied to nominal or categorized data annotated by two coders. This measurement was further improved by Fleiss who adopted a calculation to be applied to any number of raters in a nominal dataset. But it still assumes that coders have evenly distributed their type of annotations, if not, the results may be biased (Sullivan, 2014).

**Cohen’s Kappa** was based on Scott’s approach to nominal data, but includes a different approach to chance agreement. It is one of the most used measurements to date but it has some drawbacks as it cannot handle missing data, unequal sample sized or different informational types (Sullivan, 2014). Carletta explains that when the value of  $K$  is 1, it means that there is a perfect agreement; but when this is not the case, the interpretation of the result is critical (1996:252). This is explained in detail by Viera & Garrett (2005) who define this measurement as:

The calculation is based on the difference between how much agreement is actually present (“observed” agreement) compared to how much agreement would be expected to be present by chance alone (“expected” agreement).

(2005:361)

These authors also illustrate their explanation with table 13.1:

**Table 7.1** Interpretation of Kappa (from Viera & Garrett, 2005:361)

<i>k</i> -value	Agreement
<0	<Chance
0.01-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-0.99	Almost perfect

Carletta’s proposals were enormously influential and  $K$  has become the standard for measuring agreement in NLP for a wide variety of annotation tasks. Although some time has passed and this measurement has received some criticisms, it is still one of the most reliable ways of analyzing inter-annotator agreement.

**Krippendorff’s Alpha** improves the other measures by allowing a larger number of raters and additional data types but it is also more complex in its usage (Sullivan, 2014). Krippendorff (2004) defines it as “Krippendorff’s alpha as the standard reliability measure. It is general in that it can be used regardless of the number of observers, levels of measurement, sample sizes, and presence or absence of missing data”.

**Metric *agr*** This is defined in Wiebe et al. (2005) as follows: “*agr* is a directional measure of agreement that measures what proportion of  $A$  was also marked by  $b$ .” It is useful when the aim is to delimit boundaries or segments of texts where coders could be selecting different expressions in the identification process.

In addition to this multiplicity of measures, the annotation manager has to determine the tolerances, i.e., when is the agreement good enough? Some projects use “the 90% agreement rule” basing it on the observation that if humans can agree on something at  $N\%$ , systems will achieve  $(N-10)\%$ ; others consider 80% realistic for word senses. But, as Hovy and Lavid explain, “for many complex problems, 90% agreement is simply unreachable”. And they continue:

the intended use of the annotated corpus is probably the single most important guideline(...). If the goal is to train machine learning systems, then there should be enough annotated data, at high enough agreement, to enable the systems to be

trained to do their job effectively. If the goal is to identify all relevant phenomena, test the theory instantiation, and hence validate the underlying theory, then perhaps it doesn't matter what the agreement level is, as long as poor agreements are seriously investigated.

(Hovy & Lavid, 2010)

### 7.4.1 Inter-annotation Agreement

Since the last decade of the last century, big efforts have been made to put semantics and discourse analysis at the same level of other areas of Computational Linguistics in terms of empirical research, to achieve this goal reliable hand-coded data is needed. Inter-annotator or coder agreement has become almost the standard procedure to test the accuracy of manual annotations. These data could then be used either to support an empirical claim or to develop and test a computational model. (Artstein & Poesio, 2008)

On a common basis the most reliable annotated corpora is obtained through blind double annotation followed by adjudication of disagreements. This means that the reliability is inferred for a large-scale annotation by measuring the agreement on a small subset of the data which is the one annotated by more than one coder.

However, there are other authors as Dligach et al. (2010) that claim that the most efficient and reliable corpora is obtained through more single handed coded annotations, as the title of their article suggests *Annotate more accurately or to annotate more* they defend the idea that reliability coding relies on the extension of the annotated data not so much on the number of coders that made the annotations. This study aligns with the classical view of double hand coded annotation as Beigman Klebanov & Beigman (2009) define, that raises the problem of how to test the reliability of this coded annotations. To begin with, it has to be considered what kind of data will be annotated, there are two possibilities:

- data in which items are labeled with categories that is, categories are assigned to units (Artstein & Poesio, 2008)
- data in which the coding is as subjective as delimiting boundaries or segment values (Carletta, 1996)

If these coders produce similar results it can be inferred that they have internalized a similar understanding of the annotation guidelines, and it can be expected that they would perform consistently in future annotations:

Reliability is thus a prerequisite for demonstrating the validity of the coding scheme—that is, to show that the coding scheme captures the “truth” of the phenomenon being studied, in case this matters: If the annotators are not consistent then either some of them are wrong or else the annotation scheme is inappropriate for the data.

(Artstein & Poesio, 2008:557)

Nonetheless, as Artstein & Poesio (2008) also highlight good agreement cannot ensure validity, it could be the case that both coders are wrong. In these cases the figure of the expert annotator appears, Carletta (1996) discusses the different options to consider in relation to naive versus expert annotators. As she explains, in total agreement with Klippendorff, the far reaching distinction between the annotations made by experience or naive coders does not make the real point in the discussion, what is more dealing with subjective coding such as the annotation of the second type of data mentioned above, the key question relies in how well the guidelines instruct the coders to perform their annotations.

### 7.4.2 Relation Between Agreement and Reliability

The concepts of agreement and reliability have been used interchangeably in different applications, but they must be distinguished from one another.

Tinsley & Brown (2000) define inter-rater agreement as “the extent to which the different judges tend to assign exactly the same rating to each object (...) when judgments are made on a numerical scale, inter rater agreement means that the judges assigned exactly the same values when rating the same person” (2000:98).

On the other hand, inter-rater reliability “provides an indication of the extent to which the variance in the ratings is attributable to differences among the objects rated” (2000:98). It quantifies the relationship between two coders over the evaluation of multiple subjects.

“To be clear, agreement is what we measure; reliability is what we wish to infer from it. In content analysis, reproducibility is arguably the most important interpretation of reliability” (Krippendorff, 2004:215).

The quantitative relationship between agreement and reliability is analogous to the relationship between accuracy and precision in a numerical dataset. Each metric provides a different viewpoint of the same story (Sullivan, 2014).

But this agreement coefficient can only become an index of reliability when in Krippendorff (2004:215) words it fulfills:

1. Preconditions for measuring reliability:

- work with reliable data (defined under Krippendorff’s four recommendations below) achieved through double-blinded annotation of the same sample
  - coders should be available, use the same instructions, work independently, and annotate identical units of analysis
2. Identifiable units of analysis through descriptions generated by a coding procedure, nomatter who performs the coding. That is the importance of the guidelines to reproduce the experiment. Reproducibility is about data making, not about coders. A coefficient for assessing the reliability of data must treat coders as interchangeable and count observable coder idiosyncrasies as disagreement.
  3. An index of the degree of reliability must have at least two designated values: one to know when reliability is perfect, and the other to know when the conclusions drawn from imperfect data are valid by mere chance.

### 7.4.3 Guidelines and Data Reliability

Bayerl & Paul (2011) highlight the relevance of the work by Artstein & Poesio (2008) in which they point out that the “main reason for the analysis of annotation quality is to obtain a measure of the trustworthiness of annotations” (Bayerl & Paul, 2011:700). Therefore, it is only possible to rely on the conclusions drawn from the annotated data knowing whether the annotations are consistent and reproducible.

This degree of certainty depends on degree to which agreement is attained by the annotators in the coding process. The overall development of the study is influenced by several factors that may affect the coding process, from slips of the coders’ attention, misinterpretations of coding instructions, personal biases, or genuinely disparate interpretations of material.

Krippendorff (2004) recommends four steps for establishing data reliability, in order to reproduce the coding process with different coders and under different circumstances (given that the alteration of the conditions do not affect the results):

1. The coding of the sample data has to be generated by coders “widely available, follow explicit and communicable instruction and work independently of each other” (2004). The sample data has to be representative with enough duplication or overlapping examples for coders to be interchangeable.
2. The way in which dat agreement is measured has to account for variables in the experiments: number of coders, type of data, size of the sample, missing data, ... etc.

3. The criteria for the quality of data may depend on the type of study carried out. Making sure that the interpretation of the boundaries chosen can be similarly interpreted.
4. Relevant differences should be tested.

# 8

## Automatic Annotation and its Requirements

This study combines automatic, semi-automatic, rule-based and, manual annotation. As mentioned in chapter 7.1, **manual annotation** is increasingly supported by computational tools but has been traditionally done on paper. In order to ensure that its results are consistent and reliable agreement studies are carried out in the NLP community, as was explained in chapter 7.4 above. If agreement among the annotators is high, it is an indication that the annotation scheme and guidelines are clear and can be used to produce quality annotations. If, on the contrary, there is a low level of agreement between annotators, it is necessary to review the annotation scheme and the guidelines and to investigate the cases of poor agreement (Eckart de Castilho, 2014; Hovy & Lavid, 2010).

**Semi-automatic annotation** relies on software to facilitate the task of manually adding the labels as described in the guidelines. This helps to avoid mistakes which can be frequent in handwritten annotations. The annotated documents are thus processable not only by humans but also by automated agents. Additional layers of annotation can be added either manually or automatically.

**Automatic annotation** can be partial or fully automatic. When partial, some labels are left for human coders to annotate due to unsolved ambiguities or because the system does not know how to solve them. In the second situation, all tags are

automatically coded without any manual intervention.

The kinds of automatic annotations have been traditionally called “annotation levels” but more recently they have been coined as “annotation layers”.

Traditionally EAGLES (1996) identified three layers of annotation:

1. Morphosyntactic annotations:

- Lemma tagging (specially valuable for morphologically rich languages) (Leech, 1997).
- Part-of-speech tagging, one of the most extended types of linguistic annotation and one of the most accurate. However, if a mistake is made in this layer of annotation the rest of the subsequent annotations are compromised. It is also referred to as POS tagging or grammatical tagging. It consists of the annotation of the grammatical class (e.g., noun, verb, etc.) of each word or token in a text, along with their morphological attributes (Pareja-Lora, 2012).

2. Syntactic annotations:

Once morphosyntactic units (or categories) in a text have been coded, syntactic annotations add information about these relationships between them (Pareja-Lora, 2012). These syntactic relationships can take different forms, as phrase structure grammar, dependency grammar, systemic functional grammar, etc... (McEnery & Wilson, 2001). Contrary to morphosyntactic annotations, the resources to perform syntactic annotation or parsing do not reach the levels of correctness and accuracy of POS tagging, as the syntactic linguistic features of each language play a relevant role in its performance (see Lavid & Moratón, 2013b).

3. Semantic annotations:

This type refers to components whose character is principally task-oriented or algorithmic, such as lemmatizers, generators, translators, parsers and speech recognizers.

Pareja-Lora (2012) mentions two main types:

- Semantic role labelling: it consists of identifying and annotating the predicate-argument structure of clauses and sentences revealing the syntax-semantic look of texts (2012:42).



- 
- Annotation of word senses:
    - Word Sense Disambiguation (WSD): refers to “the association of content words in a text with their corresponding senses in an electronic dictionary or in other computational linguistic resources” , here words are tagged with a sense (meaning) label (2012:42).
    - Word Domain Disambiguation (WDD): associates the above mentioned tag, not to meaning but at domain level, reducing ambiguities identifying polysemy.
    - Named Entities: are defined as words or sequences of words that are ‘unique identifiers’ of entities, time or quantities, and are annotated by means of Named Entity Recognition and Classification (NERC) processes.

Regarding the format of the annotation it may vary from binary annotations to annotations based on open standards. There are initiatives such as the The Text Encoding Initiative (TEI) and the Encoding Standard for XML (XCES), whose goal is the definition of standards for the structural and linguistic representation of textual documents. The approach most widely accepted is the use of XML based formats for the representation of all associated information in a document: metadata, structure and linguistic annotations. These annotations may be embedded, that is, next to the element coded; or standoff, placed in a different location than the coded elements.

In this study I have opted to use a single platform that embeds all annotation layers, supporting semi-automatic annotation in XML based format, to ensure reproducibility (more in chapter 21)



## C. EMPIRICAL VALIDATION OF THE MODEL OF ‘THEME’ IN SFGS



# 9

## Introduction

This chapter focuses on one of the aims of the current dissertation, namely, to provide an empirical validation of the model of Theme as developed by Lavid et al. in SFGS through corpus annotation. As advanced in the Introduction, this aim is clearly an important methodological innovation with respect to current linguistic practices where many linguistic hypotheses remain untested. The idea, as proposed by Lavid, is to use the results of manual corpus annotation “as a mechanism to test hypotheses about linguistic phenomena empirically” (Lavid et al., 2013b). But how can this be achieved? Hovy and Lavid propose that, at least, two aspects of theories can be tested through manual corpus annotation: theory formation and theory redefinition. In addition, these authors claim, “corpus annotation makes it possible to enrich theories with quantitative information” (Hovy & Lavid, 2010). As Artstein & Poesio (2008) point out agreement studies are a prerequisite for the validity of the annotation guidelines, if different coders agree which items are labeled with certain categories this means that they have the same understanding of the guidelines and that these are therefore reliable.

In the present study, before performing the manual annotation on the whole corpus (chapter 18.1), it was considered useful to carry out a set of agreement studies in order to refine the definitions of the categories given, and decide which of them to include in the core and the extended tag-sets of the annotation schemata. The results of the

agreement annotation make more reliable both the larger annotation of a corpus, and the definition of rules to automatically annotate these categories, as well as, for next annotators to come, to count on more accurate definitions on the categories needing less refine work than with a mere manual annotation of a smaller number of texts (Dligach et al., 2010).

In the following sections I will describe the methodological steps which have been taken to use corpus annotation for the validation of a number of aspects of the model of Theme proposed in SFGS. The steps follow standard practice in corpus annotation in the field of NLP and have been presented in previous work by the author of this dissertation and colleagues (see Arús et al., 2012). They are explained in detail in the following sections and summarised below:

1. *Selecting the corpus.* This step involves the compilation of those texts to create what is known as the ‘training corpus’ on which the annotations will be hand-coded both by human experts and by trained annotators.
2. *Delimiting the category(es) to be annotated and its/their features.*
3. *Designing the annotation scheme and guidelines.* This involves instantiating all or part of the features of the selected theoretical model and developing a core and an extended tag-set to be used in the process of annotating the training corpus.
4. *Performing annotation experiments* on the training corpus, in order to determine the feasibility and consistency both of the instantiation and the annotator manual.
5. *Measuring the results of the annotations* by comparing the degree of agreement between the annotators’ decisions. This step also involves deciding which measures are appropriate, how they should be applied, and determining the level of agreement which will be considered satisfactory. Each of these steps is described in detail below.

# 10

## Selecting the Training Corpus

In this study two training corpora were compiled, which were named A and B for convenience:

Training corpus A: This corpus was used for Agreement Studies 1 and 2 and its tasks 1 and 2. It consisted of a total of 6 texts evenly divided into English and Spanish (number of clause complexes 143 in the English dataset, and 79 in the Spanish one). The type of experiment carried out in this training corpus involved the identification of spans of text in the unit of analysis. Therefore the training corpus is composed of a random selection of complete texts. They belonged to two groups of texts collected from published sources between 2008 and 2013.

**Table 10.1** Training Corpus A

Training corpus A	Texts	Clauses
English	3	143
Spanish	3	79
Total	6	222

Training corpus B: This corpus was used for Agreement Study 2 (tasks 3, 4 and 5) and consisted on a compilation of clauses especially selected for the annotation of these tasks. These tasks implied the labeling of subtypes of main categories. For this reason only occurrences of the main categories were selected and given to the coders. The reason for this is to avoid the coders go through examples that did not meet the criteria to be annotated. Selecting clauses of the main category ensure that the coders were exposed to the evaluated problem: the distinction of different subtypes of a main category. The clauses were selected from the same corpus as Training corpus A. They are distributed as follows:

**Table 10.2** Training Corpus B

Training corpus B	Language	Clauses
PreHead	English	50
	Spanish	50
Textual Theme	English	50
	Spanish	50
Interpersonal Theme	English	40
	Spanish	50



# 11

## Delimiting the Categories to be Annotated and their Features

In this chapter, I summarize the most relevant theoretical notions used for the annotation of the thematic categories to be annotated in the larger corpus and validated in this Part (see Lavid, Arús, & Moratón, 2012; Arús, Lavid, & Moratón, 2012; Lavid, Arús, & Moratón, 2013a,c; Arús, Lavid, & Moratón, 2013). As explained in Hovy & Lavid (2010:19):

...instantiating the theory encounters the problem that no theory is ever complete, and few if any are developed to an equal degree for all variants of the phenomena they address. Since theories tend to focus on some phenomena over others, uncertainty arises about exactly which categories to define as tags for annotation, how to define them exactly, and what to do with the residue not covered by the theory.

As said above, the analysis draws upon the model of Theme proposed in SFGS (Lavid et al., 2010d). The clausal features selected for analysis capture the three main types of meaning represented by the category of Theme in the linguistic literature: experiential, interpersonal and textual. In this model the phenomenon of Theme is divided between the Inner Thematic Field and Outer Thematic Field. The former includes the Thematic Head and any possible, but not necessary, PreHeads. The Outer

Thematic Field, which is not an obligatory segment, includes Textual and Interpersonal thematic elements preceding the Inner Thematic Field.

In the following sections I will provide an outline of the model of Theme presented in SFGS and the features selected for annotation.

## 11.1 Outer Thematic Field

The Outer Thematic Field (OTF) is defined in Lavid et al. (2010d:302) as the elements that surround and complete the Inner Thematic Field (ITF) consisting of the following types:

**Textual Themes:** Textual markers which help creating the logical connections in the text

**Interpersonal Themes:** Express the attitude and evaluation of the speaker with respect to his/her message

**Absolute Thematic constructions:** Constructions not integrated in the main predication usually separated by a comma or a pause

### 11.1.1 Textual Theme

Textual Themes are those elements in clause-initial position which are instrumental in the creation of the logical connections in the texts. In Halliday & Matthiessen (2004:373) they refer to *parataxis* as the “relation between two like elements of equal status, one initiating and the other continuing” and *hypotaxis* as the “the relation between a dependent element and its dominant, the element on which it is dependent”.

These *paratatic* and *hypotactic* relationships are reflected in the following labels:

**Linkers** as paratactic nexus

- (1) [TT-Link: **However**,] Hilco faces competition from other suitors for the UK business and the structure of any deal is likely to be complex, partly because HMV has many loss-making stores that a new owner is likely to jettison.

(Report: *Hilco UK emerges among favourites to give HMV a spin.*  
Sentence 7)

- (2) [TT-Link: **No obstante**,] Lagarde señaló que “queda mucho por hacer”, sobre todo en la reducción de las “barreras a la competitividad” y en la puesta en práctica de los planes de privatización.

[However, Lagarde signalled "all there is left to be done", especially in the reduction of the barriers to competitiveness and start practising privatization plans].

(Report: *La 'Troika' concederá a Grecia seis meses en los que no le reclamará más medidas.* Sentence 9)

### Binders as hypotactic nexus

- (3) [TT-Bind: **But**] we in Britain should tread carefully.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 22)

- (4) [TT-Bind: **Pero**] no está claro que PP y PSOE tengan las ideas muy claras sobre la estructura y funciones de los organismos independientes.

[But it is not clear that PP and PSOE have very clear ideas on the structure and functions on the independent agencies] (Editorial: *Europa vigila.* Sentence 11)

### Correlatives (not only ... but; either ... or)

- (5) [TT-COR: **Not only**] didn't he call [TT-Cor: **but**] also forgot completely about us

These labels will be used in the classification of textual elements in manual annotation. In this phase these categories will be tested in the intra-annotator agreement experiments as described in Lavid, Arús, & Moratón (2013b).

Other authors as Caffarel et al. (2004) define Textual themes as "(sic)..it does connect each message to the previous text and in doing so makes the text cohesive". Thus these authors include in textual themes Conjunctions and Continuatives. Continuatives are oral expressions like *Ummm, yeah* that we have not considered in this study, as the texts we are using do not reflect oral language.

According to Bosque & Demonte (2000) in their *Gramática Descriptiva del Castellano* [Descriptive Grammar of Spanish]:

La hipotaxis predomina en los registros más elaborados, la parataxis, en los más coloquiales

[Hypotaxis is found in more elaborated registers whereas parataxis is more common of colloquial registers]

The results obtained in the annotation carried out in the present study confirm this statement, news reports present a very low number of hypotactic relationships whereas

editorials and letters to the editor record a higher presence of these type of textual themes, as can be seen and explained in chapter 18.5.1.

### 11.1.2 Interpersonal Theme

Interpersonal Themes (IT) are defined by Lavid et al. (2010d:302) as the elements preceding the Thematic Head and expressing the attitude and the evaluation of the speaker with respect to his/her message.

For example:

- (6) [IT: **Doubtless,**] movement is difficult in rural areas after heavy snowfall; but is it really necessary for transport and schools in towns and cities to be so badly disrupted by what is a natural seasonal occurrence, albeit one experienced less regularly today than in the past?  
(Editorial: *This winter's tale*. Sentence 8)

- (7) [IT: **Realmente,**] este hecho supone un sacrificio que no le gusta a nadie, pero que han de aceptar buena parte de los trabajadores del país para poder mantener el empleo, en una Catalunya en la que hay más de 700.000 parados.

[*Actually, this fact is a sacrifice that does not like anyone, but should have to be accepted by most part of the workers to maintain their jobs, in Catalunya where there are more than 700.000 unemployed.*]

(Editorial: *Grave error en Nissan*. Sentence 3)

As Matthiessen (1995:41) points out “*Interpersonal Themes* indicate to the reader that the ideational configurations that follow are to be taken from the writer’s angle”. Considering this point of view here is the definition of the three first IT labels used in the study:

**Vocatives** This is any item, typically (but not necessarily) a personal name, being used to address. (Halliday & Matthiessen, 2004:81)

- (8) [IT-Voc: **Tom!** ] This is a nice surprise

- (9) [IT-Voc: **Sir** ] could you follow me, please?

**Comment adjuncts** Express the speaker or writer’s judgement or his/her attitude to the content of the message. (Halliday & Matthiessen, 2004:81)

- (10) [IT-Com: **Doubtless,**] movement is difficult in rural areas after heavy snowfall; but is it really necessary for transport and schools in towns and cities

to be so badly disrupted by what is a natural seasonal occurrence, albeit one experienced less regularly today than in the past?

(Editorial: *This winter's tale*. Sentence 8)

- (11) [IT-Com: **Realmente**,] este hecho supone un sacrificio que no le gusta a nadie, pero que han de aceptar buena parte de los trabajadores del país para poder mantener el empleo, en una Catalunya en la que hay más de 700.000 parados.

[*Actually, this fact is a sacrifice that does not like anyone, but should have to be accepted by most part of the workers to maintain their jobs, in Catalunya where there are more than 700.000 unemployed.*]

(Editorial: *Grave error en Nissan*. Sentence 3)

**Modal adjuncts** In Halliday's words (2004:145): "What the modality system does is to construe the region of uncertainty that lies between 'yes' and 'no'".

- (12) [IT-Mod: **Perhaps**] this was because the statement in which the agency justified its action will have made even less comfortable reading for the shadow chancellor than his opposite number.

(Editorial: *Osborne needs to budget for economic growth*. Sentence 2)

- (13) Pero [IT-Mod: **quizá**] beneficie al otro independentismo, el de Esquerra, que goza de una imagen más limpia.

[*But maybe this benefits the other independentism, the one of Esquerra, which has a clearer look.*]

(Editorial: *Escuchas y fallos*. Sentence 13)

This idea led to consider that other constructions that announce that what is coming next is seen under the point of view of the author, should be included as part of these Interpersonal Themes. To decide which expressions have to be included I have followed what Hyland proposes in *Stance and Engagement: a model of interaction in academic discourse* (2005). In this article, Hyland studies the interactions between readers and writers in research articles. He defines a categorization of Interpersonal expressions depending on whether they respond to the writer's intention to present him/herself and connect to others (Engagement) or to the way writers present themselves and convey their judgements and opinions (Stance).

**Engagement** In Hyland's words "writers acknowledge and connect to others, recognizing the presence of their readers, pulling them along with their argument,

focusing their attention, acknowledging their uncertainties, including them as discourse participants, and guiding them to interpretations.” (2005:176)

- (14) [IT-Eng: **It is clear that**] many problems would arise if the legislation as now tabled were to be implemented.

(Letter to the Editor: *Settle this Plebgate stand-off now*. Sentence 7)

- (15) [IT-Eng: **Creo desde mi modesta opinión que**] es compatible aplicar medidas dirigidas a reducir el déficit público como ya se está haciendo, eso sí con un calendario más flexible para cumplirlo, con políticas de crecimiento.

[*I think, from my humble opinion, that it is possible to apply measures to reduce public deficit, as it is in fact been done, but of course with more flexible schedules and accompanied by growing policies.*]

(Letter to the Editor: *Otras políticas sí son posibles*. Sentence 9)

**Stance** Hyland’s definition “This can be seen as an attitudinal dimension and includes features which refer to the ways writers present themselves and convey their judgements, opinions, and commitments. It is the ways that writers intrude to stamp their personal authority onto their arguments or step back and disguise their involvement.” (2005:176)

- (16) [IT-Sta: **We thought that**] consumers were entitled to assume that meat – and indeed all food – was subject to random spot checks to confirm that it was what it said it was.

(Editorial: *The market needs watching*. Sentence 6)

- (17) [IT-Sta: **Esperemos**] que el presidente del Gobierno ofrezca hoy alguna pista en este sentido durante el debate sobre el estado de la nación.

[*Let’s hope that the Prime Minister gives some hint today in the debate on the state of the nation.*]

(Editorial: *Exportamos bien*. Sentence 12)

## 11.2 Inner Thematic Field

The Inner Thematic field consists of elements selected from the experiential structure of the clause, and can include two components as explained in Lavid, Arús, & Zamorano-Mansilla (2010d:299).

PreHead

Thematic Head

### 11.2.1 PreHead

The PreHead is the non-nuclear part of the Inner Thematic Field, and although not really optional, is not always present. This element is any circumstantial and/or finite element preceding the Thematic Head which do not exhaust the thematic potential of the clause. In Spanish it can also be realized as the lexical part of the process or middle markers.

#### 11.2.1.1 Circumstantial elements

The Circumstantial elements can be encoded as simple Circumstance (PH-Circumstance), realized by Adverbial or Prepositional Groups, but also as (PH-CCL), when realized by a dependent clause; or as a Finite (PH-Fin) in the form of auxiliaries preceding the verbal form in questions or imperatives, since these are the main choices as specified in Lavid, Arús, Carretero, Moratón, & Zamorano (2014).

- (18) [PH-Circ: **Last year,**] Barclays raised funds from the China Development Bank (CDB) and Singapore's Temasek.

(Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi*. Sentence 12)

- (19) [PH-Circ: **En el último mes,**] el yen se ha apreciado el 13% frente al dólar y el 27% frente al euro, lo que ha propiciado fuertes descensos en la bolsa de Tokio y caídas en los beneficios de las empresas. El índice Nikkei de la Bolsa de Tokio bajaba, a 45 minutos para el cierre, un 6,23%, situándose a 7.172,89 puntos.

[*Last month the yen has increased its value in a 13% in respect to the dollar and 27% in relation to the euro, this provoked Tokyo stock exchange to decrease and a downfall in the benefits of several companies.*]

(Report: *El G7 reafirma su interés en que el sistema financiero sea "fuerte y estable"*. Sentence 6)

- (20) [PH-CCL: **In a conference call with investors,**] Mr Liddell said Microsoft will "tailor our business to whatever the economy brings."

(Report: *Microsoft results offer hope to tech sector*. Sentence 14)

- (21) [PH-CCL: **Ayer, tras la derrota política sufrida en una Asamblea madrileña donde el PP goza de cómoda mayoría absoluta -tan cómoda y relajante**

**que algunos de sus diputados usan tabletas y teléfonos móviles para distraerse con juegos en línea durante las sesiones-**] los médicos decidieron desconvocar la huelga y seguir su lucha en los tribunales, persiguiendo irregularidades.

[Yesterday after the political defeat suffered in a Madrid Assembly where PP rules in majority - in such a relaxing and comfortable situation many of their members use tablets or mobile phones with games during parliamentary sessions - doctors decided to call off the strike and continue with their fight in court, pursuing irregularities.]

(Editorial: *La sanidad pública*. Sentence 11)

- (22) [PH-Finite: **Will**] those same civil servants be entrusted with the task a second time?

(Letter to the Editor: *Why Lord Heseltine was right in calling the Civil Service amateurish* Sentence 11)

- (23) [PH-Finite: **Funcionarán**] bajo la misma entidad legal pero serán gestionados diferentemente.

[They will work under the same legal entity but will be managed differently.]

(Report: *Catalunya Banc y NCG sufrirán pérdidas de hasta 20.000 millones en 2012* Sentence 15)

Circumstantial elements in thematic position may perform a variety of functions: “Temporal or Location Circumstantials often function as scene-setting frameworks” (Downing, 1991:127). According to this idea and the definitions given in Downing & Locke (2006:155) further detailed analysis (in PreHead analysis section 18.1) has been carried out identifying whether these PreHeads expressed *Time*, *Place*, *Evidence*, *Contingency*, *Manner*...

### 11.2.1.2 Lexical part of the process

Being Spanish a pro-drop language, it is common to find a verb as the first element of the clause, when the verb is in thematic position, as in these two examples (24) and (25), PreHead has to be present, as the non-inflectional part of the verb, therefore although not an obligatory element PreHead is not always really an optional element: when the verb is in thematic position, as in these two examples, the PreHead has to be present. When this happens the grammatical information of the verb functions as Head. This is explained in Lavid et al. (2010d:300) in the section devoted to Thematic Head where:



An implicit Thematic Head is encoded through verbal prefixes or suffixes, depending on the function of the Head in the interpersonal structure of the clause. If the implicit Thematic Head functions as Subject, the realization is a verbal suffix indicating the person and number of the participant.

Particularizing this in the PreHead the definition given in the guidelines for the Inter-annotator agreement tasks is that PreHeads may be realized as the “Lexical part of the process” (see chapter 12)

- (24) [PH: **Agrav**]-a la conducta de los acusados haciéndola merecedora de penas de prisión: tres años para el principal acusado, el general de Sanidad Vicente Navarro, y año y medio para el comandante y el capitán médicos José Ramón Ramírez y Miguel Ángel Sanz.

*[This aggravates the conduct of the defendants making it worthy of sentences prison, three years for the main accused, General Health Vicente Navarro, year and a half to the commander and captain physician Jose Ramon Ramirez and Miguel Angel Sanz.]*

(Editorial: *Yak-42: justicia a medias*. Sentence 4)

- (25) [PH: **E**]-s el primer recorte trimestral en el consumo desde el último trimestre de 1991 y el mayor desde la primavera de 1980.

*[It is the first quarterly cut in consumption from the last quarter of 1991 and the highest since the spring of 1980.]*

(Report: *La crisis financiera lleva a Estados Unidos al borde de la recesión*. Sentence 10)

In example (24) the identification of the lexical part of the verb is straight forward as we are dealing with a regular verb of the first conjugation. But in example (25) this identification does not come so easily as this is an irregular verb. The identification of the lexical part of the verb as **E-** and the grammatical part as **-s** is a convention adopted to facilitate the annotation and identification of these elements in irregular verbs.

### 11.2.1.3 Middle markers

In Spanish PreHead may also include other elements such as the *se* marker, when functioning as pseudo-reflexive and pronominal or passive.

Pseudo-reflexives, or ‘romance’ reflexives do not qualify for Thematic Head because it is not a participant (Lavid et al., 2010d:114-118). The part that identifies the PreHead is the pseudo-reflexive plus the lexical part of the verb:

- (26) [PH: **Se llam**]-a Agustín.

*[His name is Agustín]*

Pronominal or passive *se* can be considered as part of the verb but as it is not a participant it functions as a PreHead:

- (27) [PH **Se trat**]-a del segundo informe independiente del FMI sobre el proceso de reformas de la banca española después de que se aprobara el rescate de la misma en julio del año pasado.

*[This is the second independent IMF report on the process of reforms on Spanish banks after the bailout was approved in July last year.]*

(Report: *Lagarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída.* Sentence 22)

### 11.2.2 Thematic Head

The Thematic Head is defined as the ‘first element with a function in the experiential configuration of the clause which is more central to the unfolding of the text by allowing the tracking of the discourse participants’ (Lavid et al., 2010d:299). In the following subsections I will present the contexts where Thematic Heads can occur and the different subtypes and realizations which have been considered for the annotation tasks.

#### 11.2.2.1 Thematic Head in Declarative Clauses

According to Downing (2006:181) a declarative clause is the basic clause type, with Subject-Finite ordering, being Finite realized by any modal or tensed verb.

In **English** declarative clauses the Thematic Head usually conflates with the Subject or with the Complement of the main clause in preverbal position (Lavid, Arús, & Moratón, 2012:9), as in the following examples below:

- (28) [H: **The pound**] took a fresh beating yesterday as the prediction of a German economic resurgence triggered a sterling sell-off and experts said the currency was at risk of a “large-scale devaluation”

(Report: *Pound takes another hit on German resurgence.* Sentence 1)

- (29) [H: **Warnings that the American economy is already in the grip of a new recession, and fears that this will prove long and deep,**] were reinforced after official figures showed that it contracted at an annual rate of 0.3 per cent in the past quarter.

(Report: *US economy officially on brink of recession.* Sentence 2)

In the **Spanish** clause, the Thematic Head may be explicit (30) encoded through independent lexical and grammatical forms, such as Nominal Groups or nominal clauses. Or implicit, depending on factors such as the informational status of its referent its definiteness, etc.... An implicit Thematic Head is encoded through verbal prefixes (see middle markers 11.2.2.1) or suffixes (31), depending on the function of the Head in the interpersonal structure of the clause. If the implicit Thematic Head functions as Subject, the realisation is a verbal suffix (the part that carries the grammatical information of the verb) indicating the person and number of the participant.

- (30) [H: **Los abogados de Morgan Lewis**] afirman que no hubo conductas reprobables más allá del "error personal" que el propio Strauss-Kahn, de 59 años, ha reconocido: no presionó a la mujer para obtener sexo ni intervino en su decisión de cambiar de trabajo ni favoreció una compensación mayor de la usual cuando abandonó el FMI en agosto.

*[Morgan Lewis attorneys assembled there was reprehensible behavior beyond the "personal error" that Strauss-Kahn himself, 59 years, has recognized: he did not pressed the woman for sex or intervened in their decision to change work or favoured a greater compensation than usual when leaving the IMF in August.]*

(Report: *Una noche sin consecuencias para Strauss-Kahn*. Sentence 7)

- (31) Viaj-[H: **ó**] también a Estambul y a Jerusalén.

*[He also travelled to Istambul and Jerusalem.]*

(Editorial: *El papado de Benedicto XVI*. Sentence 33)

Regarding **middle markers** as Thematic Head, the following realizations can be found in Spanish:

- Participants functioning as direct (*me, te, nos, os, lo, la, los, las*), as in example (32), or indirect objects (*le, les, se*), in the mood structure of the clause.

- (32) [H: **La**] alimentaban la difusión de numerosos casos de abusos sexuales cometidos por eclesiásticos durante decenios, y un creciente clamor popular que exigía el fin de estas conductas y de su encubrimiento.

*[It was fed by the spread of numerous cases of sexual abuse by priests for decades, and a growing public outcry demanding the end of these behaviors and their concealment.]*

(Editorial: *El papado de Benedicto XVI*. Sentence 11)

- Reflexives (*me, te, se, nos, os, se*), which can be direct or indirect objects.

- (33) [H: **Me**] dirijo a usted para mostrarle mi indignación por una entrevista que se publicó el domingo 8 de febrero en la sección de Cultura al cantaor flamenco Manuel Agujetas.

*[I am writing to show my outrage at an interview, published Sunday February 8 at the Culture section, to the flamenco singer Manuel Agujetas.]*

(Letters to the Editor: *el arte no entiende de razas ni de sexos*.  
Sentence 1)

- Impersonal *se*, when the *se* marker functions as a Participant in the clause. It is used in cases where there is no agent, or when the agent is not an essential part of the message. As in:

- (34) [H: **Se**] augura un 2013 cargado de protestas y movilizaciones sociales por la defensa de los derechos fundamentales y la dignidad de las personas.

*[It is predicted to be a 2013 full of protests and social movements in defense of the fundamental rights and dignity of people.]*

(Editorials: *Indignación social*. Sentence 2)

### 11.2.2.2 Thematic Equatives

Besides these typical Thematic Heads, there are also other more marked realizations such as thematic equatives defined by Halliday & Matthiessen (2004:69) as:

A thematic equative (which is usually called a ‘pseudo-cleft sentence’ in formal grammar) is an identifying clause which has a thematic nominalization in it. Its function is to express the Theme-Rheme structure in such a way as to allow for the Theme to consist of any subset of the elements of the clause.

As explained in (Lavid et al., 2010c:345) this form identifies the Theme with the Rheme or in other words as Gómez-González (2001:107) puts it these constructions are said to highlight “cognitive prominence exclusively identifying a participant, circumstance, attribute or process”

This pseudo-cleft construction in English is accomplished in Spanish through the use of the subordinate form *Lo que*, as in the following examples:

- (35) [H: What Britain needs] is a robust and unambiguous tax system that requires corporations to pay their share, without the need for negotiations with HMRC, or

moral urgings from politicians, or gestures of largesse by pressurised managements.

(Editorial: *Blame the tax loopholes*. Sentence 9)

- (36) [H: Lo que requiere una investigación y la correspondiente depuración de responsabilidades] es la colaboración que la empresa del expistolero presta a la lucha contra el crimen con acceso a datos confidenciales.

[What needs an investigation and the corresponding purification of responsibilities is the collaboration that the company of the ex-gunman's pays to the fight against crime granting them access to confidential data.]

(Editorial: *El caso "Hellín"*. Sentence 4)

### 11.2.2.3 Predicated Themes

In contraposition to Thematic Equatives, Predicated Themes do not try to identify Theme with Rheme but to give a special relevance to the Theme being this one exclusive at this point of the sentence (Gómez-González, 2001:107).

In Spanish, this structure is only possible in declarative sentences. It is impersonal, and therefore with a subjectless *set* plus the identifier (the identified, 'it' in English, is implicit) but it may also begin with the relative pronouns: *que*, *quien*, *cual*, *como*, *cuando*, *donde*, *el/la/los/las cual(es)/que*, etc. (Taboada, 1995:64)

The following two sentences exemplify these uses in English and Spanish:

- (37) [H: It is Mr Paterson, above all, who] has questions to answer.

(Editorial: *The market needs watching*. Sentence 27)

- (38) [H: Porque fue también la prensa la que] hace 20 años abrió los ojos a la policía dando con el paradero de Hellín en Paraguay, donde se había refugiado tras huir de la cárcel.

[Because it was the press who 20 years ago opened the eyes of the police giving the whereabouts of Hellin in Paraguay, where he had taken refuge after fleeing from jail.]

(Editorial: *El caso "Hellín"*. Sentence 8)

### 11.2.2.4 Existential Clauses

In existential clauses, the Thematic Head is *there*, the only case in which the Thematic Head is not an experiential participant, as in the following example:

- (39) [H: There] was a hint of arrogance in Mr Cameron's offer that could well have smacked, to those on the receiving end, of colonial-style condescension.

(Editorial: *David Cameron's mixed signals on terrorism*. Sentence 16)

As McCabe does in her analysis (1999) here the identification of the existential *there* in English will be done in Spanish through the use of *hay*. As some authors point out *hay* “seems to occupy a subject-like position in the grammatical structure of the sentence, and the indefinite noun functions as an object”

- (40) [H: Hay] señales de un positivo cambio de tendencia.

[There are signals of a positive change of tendency.]

(Editorial: *Italia y el cambio de tendencia*. Sentence 27)

### 11.2.2.5 Interrogative Clauses

In interrogative clauses, according to Downing (2006:182), the Finite verb precedes the Subject remaining the rest of the sentence the same.

Regarding the selection of the Thematic Head in this type of clauses it is important to take into account the type of question: *Wh- questions* or *Yes/No questions*.

**Wh-interrogatives** In this type of questions the speaker is looking to fulfill this *wh-* part, the rest is already given information (Downing & Locke, 2006:185), thus the Thematic Head will be the *wh-* particle in English and the interrogative pronoun (*cu-words* (Bosque & Demonte, 2000)) in Spanish, as in the examples:

- (41) But [H: how] are insurers to know who all the safe drivers are?

(Letters to the Editor: *Women's premiums*. Sentence 2)

- (42) [H: ¿Quién] lo supera?

[Who could overpass it?]

(Letters to the Editor: *¿Cómo están ustedees?* Sentence 9)

**Yes/no questions:** in this type of questions it is the polarity that is at stake, that is, the speaker only wants a confirmation or denial of the content (Downing & Locke, 2006:185). The Thematic Head in this type of clauses will be fulfilled with the operator. Let's see the following examples:

- (43) [H: May] I add to Christopher Pratt's list of dangerous cycling habits (Letters, November 13) those cyclists who ride along the pavement and then veer across zebra crossings as if they had right of way?

(Letters to the Editor: *Correct etiquette for cyclists crossing the road*. Sentence 1)

In Spanish polarity questions are signaled through intonation, orthographically marked by the open question mark ‘?’ (McCabe, 1999). In these cases the Thematic Head will be selected as if it were a declarative clause, in the example below the grammatical part of the verb.

- (44) ¿Podr- [H: -án] seguir durmiendo los responsables del robo mayúsculo de tantos miles de millones, cuando sabiendo que con tan solo 2.500 millones podrían haber sido felices tantas familias?  
[Could the responsables for the egregious theft of so many billions be able to sleep, knowing that with only 2.500 million could have made happy so many families?]  
(Letters to the Editor: *La lotería y la salida de la crisis*. Sentence 3)

#### 11.2.2.6 Directive Clauses

Directives are used to express a command, in this study it will be limited to the ones used with an imperative verb, that in English takes the form of infinitive without *to* and in Spanish the conjugation proper of the imperative tense. In English the particle *Let* will also be considered.

- (45) [H: Let] justice be done.  
(Letters to the Editor: *If doctors and nurses can be struck off, why are managers immune?* Sentence 5)
- (46) [H: ¡Mantengan] el Instituto de Cardiología!  
[Keep open the cardiology institute!]  
(Letters to the Editor: *Un cierre incomprensible*. Sentence 9)

#### 11.2.2.7 Exclamative Clauses

Exclamative clauses typically have the Subject-Finite ordering of declarative clauses, in some cases with a fronted *wh*- element (Downing & Locke, 2006:191), in some others only the appropriate intonation will make it. In Spanish the fronted element are the *cu*- words (*qué*, *cuál*, *cómo*, *cuánto*) that weight or intensify something that affects the speaker (Bosque & Demonte, 2000). Although unlikely to be found in our corpus, except in letters to the editor, the Thematic Head will be fulfilled with the grammatical part of the verb or with the *wh*- particle.

### 11.2.2.8 Reporting Clauses

Reporting clauses are composed of the reporting signal and a quote containing the communicated message (García Riaza, 2012). This type of clauses are specially relevant in news reports where the reader expects a reliable picture of the account added to the evidence provided (Thompson, 1994). This evidence is given through the direct or indirect quotation of relevant or authorised voices on the topic. For the present study it is important to highlight the treatment of direct quotations in thematic position. The structure of a reporting clause in thematic position exhausts the thematic role by itself if followed by a verbal phrase, in this case it acts as PreHead. When after the quotation there is the subject of the sentence, then it occupies the position of Head as in the following examples in English and Spanish:

- (47) [PH: “We, like most companies, saw a slowdown at the tail end of the quarter in particular,”] [H: Chief Financial Officer Chris Liddell] said.

(Report: *Microsoft results offer hope to tech sector*. Sentence 12)

- (48) [H: “Estamos preocupados por la reciente excesiva volatilidad en el tipo de cambio del yen y sus posibles implicaciones adversas para la estabilidad económica y financiera”], indican los ministros de Finanzas y gobernadores de los bancos centrales del G7.

[“We are concerned about the recent excessive volatility in the exchange rate of the yen and its possible adverse implications for economic and financial stability”, say the finance ministers and central bank governors of the G7.]

(Report: *El G7 reafirma su interés en que el sistema financiero sea “fuerte y estable”*. Sentence 5)

### 11.2.2.9 Semantic Realization of Thematic Head

In the same fashion as it was done in Lavid, Arús, & Moratón (2010b, 2013b) an analysis on the experiential role of Thematic Heads within the transitivity structure of the clause was performed according to the definition of these roles given by Halliday & Matthiessen (2004:Ch.5). The analysis recorded whether the Thematic Heads functioned as *Actor*, *Goal*, *Beneficiary*, *Range*, *Senser*, *Phenomenon*, *Sayer*, *Verbiage*, *Receiver*, *Token*, *Value*, *Carrier*, *Attribute*, *There*, *Process* and *None*.

The following figure of the cover of *An Introduction to Functional Grammar* (1994) shows the *Grammar of experience* the classification of the experiential roles proposed by M.A.K. Halliday.



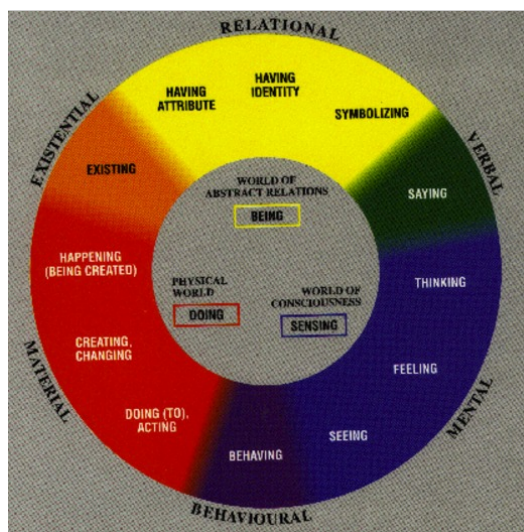


Figure 11.1 The grammar of experience

As Halliday points out this third line of construing meaning after the textual Theme and Rheme, also contemplated in this study, and Mode and Residue, takes care of the experiential line of organization. This experience construal in texts depicts another feature in the genre of a text, thus *verbal* clauses play an important role in newspapers as Halliday points out Halliday & Matthiessen (2004:174), and shown in this study on the analysis section devoted to experiential roles in news reports (18.2.3.4).

In what follows, I will provide some examples of Thematic Heads functioning as the different experiential roles mentioned above (e.g. Actor, Goal, etc..) depending on the transitivity patterns encountered in the corpus. The process types and its participants analyzed in this thesis are defined according to Halliday & Matthiessen (2004:179-197), these roles are illustrated by means of examples found in the corpus.

**Material process** construes an outer experience, a change in the flow of events. (Halliday & Matthiessen, 2004:179) The main participants are Actor and Goal.

- *Actor* is “the one that does the deed” in Halliday’s words. The following two examples in English and Spanish exemplify its use:

- (49) Last March, **Tuareg tribesmen and members of al-Qaeda in the Islamic Maghreb** seized control of the northern part of the country.  
(Editorial: *France moves in*. Sentence 6)

- (50) Además de las eléctricas, **Morales** ha nacionalizado una quincena de

empresas de hidrocarburos, cementos y minas, entre otras, desde que llegó al poder en el 2006.

[*Besides electric companies, Morales has nationalized around fifteen oil, cement and mining companies, among others, since he came to power in 2006.*] (Editorial: *Las expropiaciones de Evo Morales*. Sentence 5)

– *Goal* is the one that ‘suffers’ or ‘undergoes’ the process. As in:

- (51) At the end of the year, **controls on migration from Bulgaria and Romania** will be relaxed.

(Editorial: *Heads in the sand over Eastern immigration*.  
Sentence 6)

- (52) **La trinchera doméstica** ya la había zanjado la esposa del ex ministro francés, Anne Sinclair: “Fue cosa de una noche, nos queremos como el primer día”.

[*The domestic trench had already being settled by the wife of the former French minister, Anne Sinclair: “It was a one night thing, we love each other as the first day”.*]

(Report: *Una noche sin consecuencias para Strauss-Kahn*.  
Sentence 4)

**Mental process** construes the inner experience. The change takes place in our own consciousness. The main participants are *Senser* and *Phenomenon* (Halliday & Matthiessen, 2004:197-210).

– *Senser* is the human participant, the one that ‘feels’, ‘thinks’, ‘wants’ or ‘perceives’

- (53) **I** believe the blame lies with the Home Office, and more specifically with stubborn ministers and arrogant officials.

(Letters to the Editor: *How PCC electors were kept in the dark*. Sentence 6)

- (54) **Rajoy** tampoco sabía lo que hacían muchos dirigentes a los que apoyaba, no conocía el caso Gürtel ni las innumerables acciones judiciales en curso sobre la financiación ilegal del PP.

[*Rajoy did not know what many of his colleagues in the party were doing, he did not know about the Gürtel case nor about the innumerable on going trials about the illegal financing of the party.*]

(Letters to the Editor: *Rajoy dice la verdad*. Sentence 3)

- *Phenomenon* is what is ‘felt’, ‘thought’, ‘wanted’ or perceived’.

- (55) **In the business world, the construction of an opulent new headquarters** is often seen as a warning sign that an organization has lost its soul. (Editorial: *We all pay the cost of the BBC’s failures.* Sentence 20)

- (56) **El expediente español** sigue preocupando en Europa, pese a las notas positivas en la previsión macroeconómica, y la situación política interna se ha visto súbitamente agravada por una sucesión de revelaciones periodísticas sobre una supuesta contabilidad paralela del partido gobernante en España.

*[The Spanish case is still of great concern in Europe, despite the positive notes on macroeconomic forecasting, the internal political situation has suddenly worsened by a series of media revelations about an alleged parallel accounting of the Spanish ruling party.]*

(Editorial: *Cita en Berlín.* Sentence 5)

**Relational process** serves to characterize and identify. The characterization is done through the attributive relational process (“a is a kind of b”) where the main participants are Carrier and Attribute and the indentifying relational process (“a equals b”) participants are Token and Value (Halliday & Matthiessen, 2004:210-248).

- *Carrier* is the entity to which an Attribute is ascribed to. As in the following examples:

- (57) **Microsoft’s guidance for the current quarter** was marginally lower than Wall Street was expecting.

(Report: *Microsoft results offer hope to tech sector.* Sentence 10)

- (58) En cierta medida **el indicador** es viejo, porque no refleja el fuerte deterioro registrado en octubre, y podría quedarse corto, ya que los analistas hablan ya de una caída de tres puntos en el cuarto trimestre.

*[Up to a certain extent this marker is old, it does not reflect the strong decline registered in October, and could fall short, as analysts are already talking about a drop of three points in the fourth quarter.]*

(Report: *La crisis financiera lleva a Estados Unidos al borde de la recesión.* Sentence 27)

- *Attribute* is what is being ascribed to the Carrier. The following examples clarify this idea:

- (59) **With today’s investment, over 35 per cent of Barclays** is now owned by foreign investors.

(Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi*. Sentence 13)

- (60) **Inevitables** son, por tanto, las sospechas, que retrotraen a una de las etapas más oscuras de la historia política reciente; máxime cuando llueve sobre mojado.

[*Inevitable are therefore the suspicions that traced back to one of the darkest stages of the recent political history; especially when it rains it pours.*] (Editorial: *El caso “Hellín”*. Sentence 7)

- *Token* is the ‘operative form’ its main difference is that this identifying verb has no passive form (Halliday & Matthiessen, 2004:231).

- (61) **Algeria** was only the latest manifestation of an arc of instability from Yemen in the east to Mauritania in the west.

(Report: *As the terror threat shifts to Africa, we must meet it with great care*. Sentence 14)

- (62) Y **el cerrado apoyo de Rajoy** sólo significa, por su parte, que el Partido Popular ha renunciado a exigírsela.

[*And the close support from Rajoy just means that the Popular Party has waived demand it.*] (Editorial: *Camps y Torquemada*. Sentence 11)

- *Value* is the *identified* element if *Token* is the *identifier* (Halliday & Matthiessen, 2004:230). Of this type there are no occurrences in the English Corpus.

- (63) **Especialmente elevada** fue la nueva contratación en los servicios públicos, la educación y la sanidad con 105.000 puestos de trabajo (+1,1%), así como en las empresas privadas de servicios con 55.000 empleos (+1,0%) y en el área de la información y comunicación con 44.000 personas (+3,5%).

[*Epecially high was the new recruitment in public services, education and health with 105,000 jobs (+1.1 %) as well as private utilities with 55,000 jobs (+1.0 %) and in the area of information and communication*

with 44,000 people (+3.5 %).]

(Report: *Alemania alcanza nuevo récord histórico de empleo en el último trimestre 2012*. Sentence 6)

**Behavioural processes** are the outer manifestations of inner workings (Halliday & Matthiessen, 2004:171) typically of physiological or psychological nature as *breath, cough, smile...* Its main participant is the human Behaver. As it could be expected the analysis of this corpus did not retrieve any of these Behavioural processes.

**Verbal process** relationships enacted in the form of language. The main participants are Sayer and Verbiage (Halliday & Matthiessen, 2004:197)

– Sayer is the entity that emits the message

- (64) **Shakour Shaalan**, the fund’s executive director, said IMF staff, and notably female staff, “are not at all happy” with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet *le grand séducteur*.

(Report: *Dominique Strauss-Kahn, the head of the IMF, escapes dismissal over affair*. Sentence 4)

- (65) **Los abogados de Morgan Lewis** afirman que no hubo conductas reprobables más allá del “error personal” que el propio Strauss-Kahn, de 59 años, ha reconocido: no presionó a la mujer para obtener sexo ni intervino en su decisión de cambiar de trabajo ni favoreció una compensación mayor de la usual cuando abandonó el FMI en agosto.

[*Morgan Lewis attorneys assert that there was no objectionable behavior beyond the “personal error” that Strauss-Kahn, 59, himself has acknowledged: he did not pressure the women to get sex or intervened in their decision to change jobs or favored greater compensation than usual when he left the IMF in August.*]

(Report: *Una noche sin consecuencias para Strauss-Kahn*. Sentence 7)

– Verbiage corresponds to what is being said.

- (66) “**We are concerned about the recent excessive volatility in the exchange rate of the yen and its possible adverse implications for economic and financial stability,**” said the G7 statement, adding that all members “shared interest in a strong and stable international

financial system” and would “co-operate as appropriate.”

(Report: *G7 ‘preparing to drive down the yen’*. Sentence 10)

(67) “**Debemos estar a su lado**”, agregó.

[“We have to be next to him”, he added.]

(Report: *Hollande reconoce que Francia crecerá menos de lo previsto en 2013*. Sentence 11)

**Existential process** concerned with existence, they represent that something exists or happens simply recognized to ‘be’. The main participant is Existent *there* which only serves to indicate the feature of existence and prepares the addressee for something that is to be introduced (Halliday & Matthiessen, 2004:256). These existential clauses have also been treated in the present study in section 11.2.2.4. Refer to examples (39) and (40).

**Process** is typical of Spanish declarative sentences and rarely used in English sentences but for imperatives or yes/no questions. In the extended tag-set, process is referred to as “a whole process, whether material, mental, verbal, relational or existential” (chapter 12 and appendix C.11) (Lavid, Arús, & Moratón, 2013a:214).

(68) **Fund** candidates to explain their policies, and hold the voting in May when many families have a walk out after tea to vote – instead of looking out of the window, deciding it’s too cold, and staying in to watch the football instead.

(Letters to the Editor: *A disappointing day for democracy*. Sentence 7)

(69) Tien-**e** más de 75 millones de clientes en Europa, EEUU, Canadá, Latinoamérica, Asia y Australia, de los que 1,7 millones habitan en España, con depósitos de 13.800 millones de euros, informa EFE.

[It has more than 75 million customers in Europe, USA, Canada, Latin America, Asia and Australia, of which 1.7 million live in Spain, with deposits of 13,800 million euros, reports EFE.]

(Report: “*Holanda inyectará 10.000 millones en el grupo ING*”. Sentence 12)

#### 11.2.2.10 Syntactic Realization

Regarding Syntactic realizations of Head and PreHead two positions have been taken on this particular feature:

The first one, in the process of manual annotation, a manual analysis of the internal complexity of the different realizations of the Thematic Head and the PreHead has been carried out. A detailed study on the results can be found in the section corresponding to each genre and language (chapter 18.1), from the more complex *clauses* or *noun phrases* to the *proper noun* or *pronoun*.

The second one deals with automatic annotation, in these cases the syntactic analysis has been performed automatically through trained automatic annotators as *Stanford* for English (see section 24.1).

In the intra-annotator agreement tasks the annotators were not asked to perform syntactic analysis as this was not main object of the study.





# 12

## Designing the Annotation Scheme and Guidelines

The next step in this study was to design the annotation scheme and guidelines which would allow us to test the consistency of the categories selected from the model of Theme proposed in SFGS. An annotation scheme is, according to Leech “the document describing and explaining the scheme of analysis employed for the annotations” (Leech, 1997:6). This involves instantiating all or part of the features of the selected theoretical model and developing a core and an extended tag-set to be used in the process of annotating the training corpus. The process is a complex one which requires step-wise refinements and modifications during the whole annotation process. The main challenges identified in the creation of the annotation schema were, (a) the identification and definition of the thematic categories to include as part of the core annotation scheme; and (b) the creation of an extended annotation scheme that would help annotators overcome the difficulties posed by the shortcomings of the core annotation scheme.

The theoretical basis for the design of a preliminary annotation scheme, including both coarse and more fine-grained annotations is based on the theory developed by Lavid et al. (2010d). The coarse-grained annotations are specified in a core tag-set and the more fine-grained ones in an extended tag-set, one for English and another for Spanish.

The core tag-set included two guidelines one for the annotation of the Thematic Field, and the other for the annotation of the types of Theme.

The extended tag-set contained the more fine-grained categories: types of Textual Themes, Interpersonal Themes, PreHeads and the semantic realizations of the Thematic Head<sup>1</sup>. The definition given for Thematic Field in English is the following and it is accompanied by several examples:

Initiating clause span of varying length up to and including the first nuclear constituent in main clause, or one of the following:

- Predicated Theme construction
- “There” in Existential clauses.

In Spanish the definition reads as follows:

Clause initial material which goes from the beginning of the clause complex up to and including the first nuclear experiential constituent in the main clause. The FNC can be realized by either lexical or morphological means.

The second guideline in the core tag-set includes the definition of six tags, (Lavid, Arús, & Moratón, 2010a) (see also figures C.3 and C.4 in the appendixes) reflecting the range of possible thematic elements which can occur as part of the Thematic Field in English and Spanish declarative clauses. Table 12.1 shows the English tag-set and table 12.2 its correspondent for Spanish (for a full reference see appendix C.1).

**Table 12.1** Core tag-set for Theme types in English

Annotation Layer	Thematic Field	Description
Unit		Main Clause
Core Annotation Scheme		
	(TH) Thematic Head	First nuclear constituent (Participant or Process, not Circumstantial) in main clause, or ‘There’ in Existential clauses
	(PH) PreHead	Any Circumstantial element and/or Finite element preceding the Thematic Head

---

<sup>1</sup>Both the Core and Extended tag-sets were published by the author of this thesis and colleagues in Lavid, Arús, & Moratón (2010a, 2013a); Lavid, Arús, Carretero, Moratón, & Zamorano (2014); Arús, Lavid, & Moratón (2013)

Tags	(TT) Textual Theme	Elements which are instrumental in the creation of the logical connections in the text, such as linkers, binders and other textual markers
	(IT) Interpersonal Theme	Elements which express the attitude and the evaluation of the speaker with respect to his/her message, such as Vocatives and Modal Adjuncts, including mood and comment adjuncts
	(There) There Theme	This is a type of construction which occurs in existential clauses. It is realized by the ‘There’ element
	(PT) Predicated Theme	This construction is also known as “it-cleft” and consists of an ‘it+be+relative clause’

The core tag-set for Theme types in Spanish:

**Table 12.2** Core tag-set for Theme types in Spanish

Annotation Layer	Thematic Field	Description
Unit		Main Clause
Core Annotation Scheme		
Tags	(TH) Thematic Head	First nuclear element (not circumstantial) in main clause, realised by either lexical or morphological means
	(PH) PreHead	Elements preceding the Thematic Head, including: Circumstantials, pronominal ‘se’, lexical part of Verbal Group.
	(TT) Textual Theme	Elements which are instrumental in the creation of the logical connections in the text, such as linkers, binders and other textual markers

(IT) Interpersonal Theme	Elements which express the attitude and the evaluation of the speaker with respect to his/her message
(There) There Theme	This is a type of construction which occurs in existential clauses. it is realized by ‘Hay’ element
(PT) Predicated Theme	This construction consists on two parts: an initial thematic segment consisting of the copular verb followed by the element in Focus; and a rhematic segment realized by a relative-like clause

---

On the basis of these core tag-sets, a number of annotation experiments were performed to test the feasibility of the proposed annotation schemes. Given the intrinsic complexity of tagging semantic categories – as opposed, for instance, to POS (Part-Of-Speech) tagging, the core annotation schemes for Theme were not expected to be final. The annotation experiments were conceived as a probe to find out what was lacking in the core schemes with a view to the design of the extended schemes.

The extended tag-set include more fine-grained subtypes of some of the tags contained in the core tag-set, namely the types of Thematic Head experiential roles (table 12.3), types of PreHead (table 12.4), Textual Themes (table 12.5) and Interpersonal Themes (table 12.6).

The following table gives a general idea of the elements contained in the Thematic Head (TH) Experiential roles (for the sake of space both tag-sets, English and Spanish are provided in the same table). More details on experiential roles are given in chapter 11.2.2.9. (Also see figures C.11 and C.12 in appendix):

**Table 12.3** Experiential role types in English and Spanish

Experiential Roles	Definition
Actor	The participant doing the deed in a material processes
Goal	The participant impacted by a doing in a material process
Beneficiary	The participant benefiting (positively or negatively) from the doing in a material process

---

---

Range	The participant that construes the domain over which the process takes places, or construes the process itself, either in general or in specific terms
Senser	The participant sensing in a mental process
Phenomenon	The participant being sensed in a mental process,
Sayer	The participant saying, telling, stating, informing, asking, threatening, suggesting and so on in a verbal process
Verbiage	The content of saying in a verbal clause, when expressed as a nominal group
Receiver	The addressee of a speech interaction in a verbal process
Token	The participant representing the expression, symbol, form, name, function, position or actor in an identifying relational process. Identifying relational processes are reversible and the Token tends to appear in the first position with respect to the Value. The Token is also the participant that tends to go first in possessive and circumstantial identifying relational processes
Value	The participant representing the content, symbolized thing, meaning, referent, filler, holder of position or role in an identifying relational process. Identifying relational processes are reversible and the Value appears in initial position when the process is reversed. The Value is also the participant that goes first in possessive and circumstantial identifying relational processes when these are reversed
Carrier	The participant to which an Attribute is assigned in an attributive relational process, whether intensive, possessive or circumstantial. These relational processes are not easily reversed
Attribute	What is assigned to the Carrier in an attributive relational process, whether intensive, possessive or circumstantial. As attributive processes are not easily reversed, Attributes are not found in thematic position except in exclamations
There / Hay	The starting element in an existential process. It is not a participant
Process	A whole process, whether material, mental, verbal, relational or existential
None	

---

The types of PreHead are shown in the following table 12.4. Both English and Spanish definition are provided in the same table, but the coders were given separated definitions with according examples. These tag-set can be consulted in appendix figure C.9 for English and C.9 for Spanish:

**Table 12.4** Types of PreHead in English and Spanish

Types of PreHead	Definition
PH-Circ	Realized by adverbial or prepositional groups
PH-CCL	Realized by adverbial or prepositional groups
PH-Finite	Realized by a finite subordinate clause

The types of Textual Themes are presented in the next table 12.5. Both English and Spanish definition are provided in the same table, but the coders were given separated definitions with according examples. These tag-set can be consulted in appendix (figures C.5 and C.6):

**Table 12.5** Types of Textual Theme in English and Spanish

Types of Textual Theme	Definition
TT-Link	Linkers that realize the function of paratatic nexus
TT-Bind	Binders that realize the function of hypotactic nexus
TT-Cor	Correlatives that distribute in the form <i>not only ...but either</i>

The types of Interpersonal Themes are presented in the next table 12.6. Both English and Spanish definition are provided in the same table, but the coders were given separated definitions with according examples. These tag-set can be consulted in appendix (figures C.7 and C.8):

**Table 12.6** Types of Textual Theme in English and Spanish

Types of Interpersonal Theme	Definition
IT-Voc	Vocatives, any item used to address
IT-Com	Comment adjuncts

---

IT-Mod	Modal adjuncts
IT-Sta	Stance, the writer acknowledge and connect to others, recognizing the presence of the reader and guiding them to interpretations
IT-Eng	Engagement, the way writers stamp their personal opinion onto their arguments

---





# 13

## Performing Annotation Experiments

Annotation experiments (also called agreement studies) were performed on the training corpora in order to determine the consistency of the annotation schemes. In this study, the process consisted of two rounds in which the annotators worked individually and agreement results were measured (see chapter 13.3.1).

In the first round coders were asked to identify thematic spans, and in the second, to label specific thematic types. Afterwards results were discussed and a consensus was reached with respect to the spans realizing the thematic markables.

The annotation procedure was the following: two annotators (or coders), #1 and #2, were asked to analyze the thematic features of six texts individually, both having studied the model of theme in SFGS (see chapter 11) in depth and internalized the definitions for the core and the extended tag-set contained in the annotation scheme and guidelines (defined above in 12). The annotators were given coding sheets with instructions, each corresponding to a different type of task. In each agreement study there were different tasks: in Agreement Study 1 the two tasks focused on the identification of thematic spans; in Agreement Study 2 the five tasks focused on the labeling of the agreed thematic spans. These are included in appendix C.2.

## 13.1 Agreement Study 1: Identification of Thematic Spans

For this agreement study, two tasks were designed which focused on the identification of thematic spans.

**The first task** was aimed at the identification of the whole set of potential thematic markables. This was carried out by asking the two coders to identify the Thematic Field of each clause complex in each text. The annotation sheets given to the coders can be found in appendix C.2. The coding sheet provided to the coders contained a brief explanation on the aim of the task, then an embedded document with the guidelines related to the exercise, and a list of the texts where the annotation is to be carried out, with a link for an easier access. The texts are prepared to be annotated in an excel file. The units of analysis are placed on the left column and the right column is left for the coder to select and copy the identified span.

**The second task** focused on the identification of spans realizing only specific thematic types from the Thematic Field. Here the two coders were asked to identify in each clause complex the spans realizing Thematic Head (TH), PreHead (PH), Interpersonal Theme (IT), Textual Theme (TT), Predicated Theme (PT) and There Theme (There) (see chapter 12 and figures C.3 and C.4 in the appendix). For the annotation of each of these categories a separated sheet is given to the coders, for, in one unit of analysis, more that one category may appear in one unit of analysis.

## 13.2 Agreement Study 2: Labeling of Thematic Spans

Agreement Study 2 consisted of five tasks focused on the labeling of markables.

**The first task** focused on the labeling of the thematic markables agreed in the previous study. For this task the agreed thematic spans were highlighted in the coding sheet so that coders could carry out the classification task on the marked span selecting from a drop down list the appropriate label. There are also some ‘red herrings’ in this task, that is, some of the highlighted items did not correspond to any of the thematic elements of the core tag-set. These spans should be classified as ‘None’ with the aim of checking the coders’ knowledge on the different items. As in Agreement Study 1 Task 1, this sheet contains the same structure: one introductory sheet, the embedded file

with the guidelines and the linked texts to annotate (see appendix C.2 for the full set of annotations). The difference lies in that in the left column the unit of analysis have a selected span of text underlined and in the right column a drop down list shows the possible options the selected text may take.

**The second task** is the labeling of the Thematic Heads as one of the possible subtypes specified in the extended tag-set, i.e, as Actor, Goal, Beneficiary, Senser, Phenomenon, Sayer, Verbiage, Token, Value, Carrier, Attribute, Process, ‘There’ element or None. This task is organized as the previous one, with a short definition of the aim of the task, embedded guidelines (chapter 12) and a link to each of the text to annotate. As in the previous task a span is underlined and a drop down list shows the possible options to annotate (see appendix C.18 and C.19, for the full set of annotations).

**The third task** is the labeling of the types of Textual Themes as specified in the guidelines provided, the options available are linkers, binder or correlative Textual Themes (see C.5 and C.6 for the tag-sets and C.20, C.21 for the annotation templates in the appendixes).

**The fourth task** consists on the labeling of the types of Interpersonal Themes. The guidelines embedded in the instruction sheet shown the five types of possible labels to code: vocatives, comment and modal adjuncts, stance and engagement (chapter 12 and figures C.8, C.7 and C.22, C.23 in the appendixes).

**The fifth task** consists on the labeling of types of PreHead Themes: Circumstance, complex clauses and finite clauses. The full set of the annotation tasks is in appendixes C.9 and C.9, C.24 and C.25.

### 13.3 Measuring the Annotation Results

This step of the manual annotation process is essential for evaluating the results of the annotation process, and in the NLP community it is fundamental for the success of an annotation project (Lavid, Arús, & Moratón, 2013a). It involves deciding which measures are appropriate for inter-annotator and intra-annotator agreement, how they should be applied, and determining the level of agreement which will be considered satisfactory. In the Linguistics community, these measures have not been applied widely, but recent research on corpus annotation of different linguistic categories indicates that it is being progressively adopted in corpus linguistic projects with a focus on empirical issues (see Lavid, 2014; Boye, 2014; Glynn & Krawczak, 2014; Lavid & Moratón, 2014; Pérez Blanco, 2014; Pérez Veneros & Elorza, 2014; Zamorano-Mansilla & Carretero, 2014).

As explained by Lavid (2014), in the field of NLP too little agreement means too little consistency in the annotation to enable machine learning algorithms to be trained successfully. But in the field of Corpus Linguistics and other Systemic-Functional Linguistics too little agreement on certain linguistic features will give the analyst the possibility to investigate those cases to redefine the theoretical categories that were initially hypothesized (theory redefinition), and to propose new categories that were not proposed in the original theory (theory formation).

The agreement measures used in this thesis were the following:

For those tasks consisting of the identification of thematic spans the Agreement Metric (AGR) was used. As defined in chapter 7.4 Agreement Metric (AGR) “*agr* is a directional measure of agreement that measures what proportion of *A* was also marked by *b*” (Wiebe et al., 2005). The presentation of the data and results are developed in detail in section 13.3.1.

For those tasks involving classification, Kappa was used. The K value ranges from 0 (the agreement is no other than the expected by chance) to 1 (there is total agreement), as shown in table 13.1 below, where different values are given the appropriate interpretation.

**Table 13.1** Interpretation of Kappa (from Viera & Garrett (2005:362))

<i>k</i> -value	Agreement
<0	<Chance
0.01-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-0.99	Almost perfect

The results of this second type of measurements are thoroughly explained in section 13.3.1 below.

### 13.3.1 Annotation Results for Agreement Study 1

Agreement study 1 consisted of two tasks, as explained above 12. The results for each of these two tasks are presented below, with a separation between the English and the

Spanish results.

#### 13.3.1.1 Task 1: Identification of the Thematic Field

In this task, annotators had to identify the whole Thematic Field span in each of the 143 clause complexes of the English dataset and 79 for the Spanish one. The number of segments annotated by the two coders in each text is the same: both annotated the same number of sentences (143 in English and 79 in Spanish) without leaving any element blank. The agreement in this case is the same for both coders as they, obviously, differ in the same codifications.

**English** For the English data, the results are graphically presented in table 13.2 below:

**Table 13.2** Inter-annotator agreement: Thematic Field span in English

		Recall	Precision	Avg.
Coder 1	Coder 2	agr(1  2)	agr(1  2)	Avg.
Annotations	Annotations	0.97	0.97	
<b>Average</b>				0.97

Agreement between annotator (1) and annotator (2) was very high on average (0.97).

**Spanish** With respect to the Spanish dataset, agreement was high on average in the first task (0.71), i.e. the identification of the whole Thematic Field in each clause span. Disagreement mainly occurred when confronted with a complex Verbal group which included an auxiliary and a lexical part. Here annotators hesitated about including the lexical part or not within the Thematic Field.

**Table 13.3** Inter-annotator agreement: Thematic Field span in Spanish

		Recall	Precision	Avg.
Coder 1	Coder 2	agr(1  2)	agr(1  2)	Avg.
Annotations	Annotations	0.74	0.74	
Average				0.74

### 13.3.1.2 Task 2: Identification of Thematic Elements

**English** In the second task the annotators were asked to identify the spans realizing thematic categories of the core tag-set in our annotation scheme. The results for each of the thematic categories are collectively presented in table 13.4:

**Table 13.4** Inter-annotator agreement: Spans expressing TH, PH, IT, TT, PT, There in English

	TH	PH	IT	TT	PT	There
Average	0.9384	0.787	0.375	0.965	0.98	1

As illustrated above, agreement was high in the identification of the span expressing the Thematic Head (0.9384) and Textual Theme (0.965), but lower – although still substantial – in the identification of PreHead (0.787). By contrast, agreement was only fair (0.375) in the identification of Interpersonal Theme. As shown in (1) the main reason for the lower agreement in the identification of the Interpersonal Theme was the labeling (rather than a spanning problem), by one of the annotators, of some Textual Themes as Interpersonal Themes.

	Clause Complex	Coder1	Coder2
(1)	<i>For example</i> , although the United Nations Mission in Sudan	IT	TT
	<i>For example</i> , if 20% of the world's electricity were to be generated	IT	TT

Regarding the identification of the extent of text identified as Predicated Theme, both selected the appropriate occurrence but not the correct extension. The first coder

included in the span until ...oil price, whereas the second coder left it in ..Aramco,, being the correct Predicated Theme span: *It is Saudi Aramco, that*.

	<b>Coder</b>	<b>Clause Complex</b>
(2)	Coder1	<i>It is Saudi Aramco that has been responsible for the weakening oil price, pushing more oil on the market since the energy summit in Jedda.</i>
	Coder2	<i>It is Saudi Aramco, that has been responsible for the weakening oil price, pushing more oil on the market since the energy summit in Jedda.</i>

**Spanish** With respect to the Spanish dataset, agreement was high on average in identification of Thematic Head (0.91) in each clause span. Disagreement mainly occurred when confronted with a complex Verbal group which included an auxiliary and a lexical part. Here annotators hesitated about including the lexical part or not within the Thematic Head or the PreHead, that is the most affected by this decision taking (0.89). Identifying Textual Theme (0.97) and Interpersonal Theme (0.94) also yielded almost perfect results. The same happened with Predicated Theme (0.98) and Hay (0.98).

**Table 13.5** Inter-annotator agreement:  
Spans expressing TH, PH, IT,  
TT, PT and Hay in Spanish

	<b>TH</b>	<b>PH</b>	<b>IT</b>	<b>TT</b>	<b>PT</b>	<b>Hay</b>
<b>Average</b>	0.91	0.89	0.94	0.97	0.98	0.98

The few disagreements that occurred in the identification of the Thematic Head were due to discrepancies between annotators as to what part of the verbal morphology realized the Thematic Head, specially when the verb was irregular. For example, when confronted with the form ‘es’ (3rd person singular of the Present tense ‘to be’ in Spanish), annotators found it difficult to decide which would be the Head.

Examples (3) and (4) below illustrate more problems of this kind. In all cases one coder selects a span that comprises both Head and PreHead (*no cabe, en ello sigue,*) whereas the other differentiates and separates the lexical and the grammatical part of the verb (*no cab-e, en ello sigu-e,*) .

**Coder    Clause    Complex**

- (3)    Coder1    *No cab -e*, pues, hablar de un delito continuado.  
          Coder2    *No cabe*, pues, hablar de un delito continuado.

**Coder    Clause    Complex**

- (4)    Coder1    *En ello sigu -e* estos días el FMI, al que el presidente del Gobierno.  
          Coder2    *En ello sigue* estos días el FMI, al que el presidente del Gobierno.

Also some disagreements occurred in the identification of the Interpersonal Theme where one annotator coded them as PreHead. For example, one coder considered *Quizás* [perhaps] as an interpersonal element whereas the other annotator has coded it as PreHead.

- |     |  |                         |
|-----|--|-------------------------|
| (5) | <b>Clause    Complex</b>   | <b>Coder1    Coder2</b> |
|     | <i>Quizás</i> , para curarse en salud, el tribunal se cuida de delimitar (...) | IT            PH        |

Predicated Themes where correctly identified but the span selected between both coders was different, the first one chose up to *tal* and the other one did not include this particle in the extent elected.

**Coder    Clause    Complex**

- (6)    Coder1    *No ha habido tal*, aunque tampoco ha solicitado medidas cautelares, lo que habría indicado reforzamiento de los indicios contra el reo.  
          Coder2    *No ha habido tal*, aunque tampoco ha solicitado medidas cautelares, lo que habría indicado reforzamiento de los indicios contra el reo.

#### 13.3.2 Annotation Results for Agreement Study 2

The Second Agreement Study focused on the labeling of thematic markables which had been agreed on in the previous tasks. The results of the annotation experiments of the five different tasks are presented below:

##### 13.3.2.1 Task 1: Thematic Elements

As explained above, in this task 1 coders had to select the type of Thematic Element highlighted in the dataset. The types were: PreHead (PH), Thematic Head (TH),



Textual Theme (TT), Interpersonal Theme (IT) or None (None).

**English** In English both coders labelled 143 clause complexes. As contingency table 13.6 shows the agreement distribution in the identification of PreHeads is almost perfect, as both Coders marked 18 cases as this type of element and only in 1 case coder 1 marked an occurrence as Thematic Head where coder 2 did it as PreHead. In the case of Thematic Head there is also a great deal of agreement: both coders marked as Head 85 occurrences and disagreed in 6 cases where coder 1 label 5 as None and in 1 as PreHead. In the case of Textual Themes both coders agreed on 19 cases but in one of them coder 1 identified it as Interpersonal Theme meanwhile coder 1 did not labelled any occurrence as such. Finally, there were 14 cases where both coders chose None. They disagreed when coder 2 chose 5 cases as Thematic Head when coder 1 did it as None.

**Table 13.6** Inter-annotator agreement: Labelling thematic spans

Contingency table <b>Coder 1 * Coder 2</b>							
		Coder 2					Total
		PH	TH	TT	IT	None	
Coder 1	PH	18	0	0	0	0	18
	TH	1	85	0	0	5	91
	TT	0	0	19	1	0	20
	None	0	0	0	0	14	14
Total		19	85	19	1	19	143

Table 13.6 shows that the main disagreements occurred in labeling Thematic Heads; the confusions came with red herrings marked as None as exemplified below:

	Clause Complex	Coder1	Coder2
	Then <i>came</i> the addition of draft animals and gradual supplementation of animal prime movers(...)	TH	None
(7)	<i>A fundamental break</i> with this millennia-long pattern came only with widespread (...)	TH	None
	<i>“The managing director has expressed his regrets. I don’t think that we can ask him to do more at this time,” said , Mr Shaalan. ”We will continue to work with him.”(...)</i>	TH	None

As illustrated in table 13.7 the overall Kappa value is quite high – at 0.915 – indicating almost perfect agreement<sup>1</sup>.

**Table 13.7** Kappa agreement measures

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.915	0.031	17.355	0.000
<b># of valid cases</b>	143			

**Spanish** When looking at the Spanish dataset -where annotators labelled 79 cases each- the types of Thematic Head and None are the most controversial labels. In the identification of Head coders agreed on 54 cases and did not in 4. Disagreement occurred when one of the coders marked 1 PreHead and 2 None where the other pointed to 1 None. Both coders agreed on 6 cases in the None category and they disagreed in 3 cases in the Thematic Head category. As to Textual Themes, both coders agreed on 8 cases and disagreement happened only once: one coder marked it as Textual Theme whereas the other one chose Interpersonal instead.

<sup>1</sup>Symmetric measures: (a) Not assuming the null hypothesis; (b) Using the asymptotic standard error assuming the null hypothesis; Approx. T.: approximate t-test value; Asymptotic Standard Error (ASE): calculated in the same way as the standard errors (standard deviation of each parameter); Approx. Sig: Approximate Significance: the p-value (the smallest critical value alpha for which we would reject the null hypothesis based on these data).

	Clause Complex	Coder1	Coder2
	No <i>ha</i> habido tal, aunque tampoco ha solicitado medidas cautelares, lo que habría indicado reforzamiento (...)	TH	None
(8)	Un " <i>serio error de juicio</i> ", algo " <i>lamentable</i> ", pero sin sombra de acoso, favoritismo o abuso de autoridad (...)	TH	None
	<i>Pese a ello</i> , los letrados recomiendan que se establezcan mayores exigencias para su director gerente (...)	TT	IT

**Table 13.8** Inter-annotator agreement: Labelling thematic spans

Contingency table Coder 1 * Coder 2							
		Coder 2					Total
		PH	TH	TT	IT	None	
Coder 1	PH	5	0	0	0	0	5
	TH	1	54	0	0	2	57
	TT	0	0	8	1	0	9
	None	0	1	0	0	6	7
Total		6	55	9	1	8	79

The kappa value observed is at the range of almost perfect agreement at 0.839 :

**Table 13.9** Kappa agreement measures

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
Kappa agr. Measure	0.839	0.061	11.579	0.000
# of valid cases	79			

### 13.3.2.2 Task 2: Classification of Experiential Roles

The second task focused on the labeling of the Thematic Head experiential roles in clause complexes. Annotators had to choose the tags from the extended tag-set for

### 13. Performing Annotation Experiments

Thematic Head types, corresponding to different experiential roles conflating with Thematic Heads, (see appendixes C.1 and C.2 for the annotation templates). Numbers 1 to 15 correspond to:

<b>1</b>	Actor	<b>4</b>	Range	<b>7</b>	Sayer	<b>10</b>	Token	<b>13</b>	Attribute
<b>2</b>	Goal	<b>5</b>	Senser	<b>8</b>	Verbiage	<b>11</b>	Value	<b>14</b>	There
<b>3</b>	Beneficiary	<b>6</b>	Phenomenon	<b>9</b>	Receiver	<b>12</b>	Carrier	<b>15</b>	None

**English** In the English dataset, both coders selected the same labels for annotation and did not Beneficiary, Range nor Verbiage. The coder selections are shown in table 13.10 below:

**Table 13.10** Inter-annotator agreement: Labelling of experiential elements in English

Contingency table <b>Coder 1 * Coder 2</b>														
		Coder 2												Total
		1	2	5	6	7	9	10	11	12	13	14	15	
<b>Coder 1</b>	<b>1</b>	39	1	1	0	1	0	4	0	1	0	0	0	47
	<b>2</b>	0	12	0	0	0	0	0	0	0	0	0	0	12
	<b>5</b>	0	0	6	0	0	0	1	0	0	0	0	0	7
	<b>6</b>	0	0	0	1	0	0	0	0	0	0	0	0	1
	<b>7</b>	2	0	0	0	9	0	0	0	0	0	0	0	11
	<b>9</b>	0	0	0	0	0	1	0	0	1	0	0	0	2
	<b>10</b>	1	0	0	0	0	0	8	0	1	0	0	0	10
	<b>11</b>	0	0	0	0	0	0	1	5	0	0	0	0	6
	<b>12</b>	0	0	0	0	0	0	0	0	24	0	0	0	24
	<b>13</b>	0	0	0	0	0	0	0	0	0	6	0	0	6
	<b>14</b>	0	0	0	0	0	0	0	0	0	0	6	0	6
	<b>15</b>	0	0	0	0	0	0	0	0	0	0	0	11	11
<b>Total</b>		42	13	7	1	10	1	14	5	27	6	6	11	143

Total agreement occurred when selecting Goal (label 2), Phenomenon (label 7), Value (label 11), Carrier (label 12), Attribute (label 13), There (label 14) and the red-herring None (label 15). There were only a few cases of disagreement worth mentioning

in this task: one in the selection of Actor (label 1) where coder 1 considers 2 as Sayer and 1 as Token and coder 2 has chosen 1 Goal, 1 Sensor, 1 Sayer, 4 Range and 1 Carrier. But they agreed on the types in 39 cases. Another conflicting label is that of Verbiage (label 8), where coders chose disagreed between Actor and Carrier and Actor, Sensor and Value. Carrier (label 12) also caused disagreements: both coders agreed on 24 cases and did not when coder 2 marked 1 as Token, 1 as Receiver, and another one as Actor. Examples illustrate these choices below:

	Clause Complex	Coder1	Coder2
	<i>President Barack Obama</i> has promised an energy revolution in the world's largest economy (...)	Actor	Sayer
(9)	<i>China, a close ally of the government in Khartoum</i> , is now carefully weighing its oil interests and its strategic concerns in the South.	Actor	Sensor
	If the government in Khartoum persists (...) for the South in January 2011, <i>a return to full-scale civil war, with calamitous consequences for the peoples of Sudan and the entire region</i> , is a real possibility.	Token	Carrier

The Kappa value is rather high – 0.875 –, as shown in table 13.11, and agreement is, therefore, considered to be substantial.

**Table 13.11** Kappa agreement measures in Experiential roles in English

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.875	0.031	26.282	0.000
<b># of valid cases</b>	143			

Disagreement occurred in 15 cases, probably due to the inherent difficulty in disambiguating experiential roles conflating with Thematic Heads. The highest number of disagreements often reflected different interpretation of process types for non-clear-cut cases, which automatically involved the assignation of different participant roles.

The dividing line between material and metaphorical relational processes proved to be particularly problematic. Within relational processes, the differentiation between attributive and identifying, as well as the directionality of identifying processes, were also important sources of disagreement, as illustrated by example (10) below:

### 13. Performing Annotation Experiments

	Clause Complex	Coder1	Coder2
	<i>President Barack Obama</i> has promised an energy revolution in the world's largest economy (...)	Actor	Sayer
(10)	<i>China, a close ally of the government in Khartoum,</i> is now carefully weighing its oil interests and its strategic concerns in the South.	Actor	Senser
	If the government in Khartoum persists (...) for the South in January 2011, <i>a return to full-scale civil war, with calamitous consequences for the peoples of Sudan and the entire region,</i> is a real possibility.	Token	Carrier

**Spanish** When looking at the Spanish dataset, where both coders annotated 79 cases, it can be observed that disagreement in this task was very high, as shown in table 13.12 below:

**Table 13.12** Inter-annotator agreement: Labelling of experiential elements in Spanish

Contingency table Coder 1 * Coder 2													
		Coder 2											Total
		1	2	3	4	5	6	7	8	10	12	15	
Coder 1	1	20	4	0	0	2	1	12	0	4	0	0	43
	2	0	0	0	1	0	1	0	0	0	0	0	2
	3	0	0	1	0	0	0	0	0	0	0	0	1
	4	0	0	0	1	0	0	0	0	0	0	0	1
	5	0	1	0	0	0	0	0	0	0	0	0	1
	6	1	0	0	0	0	0	0	0	0	0	0	1
	7	0	0	0	0	0	0	1	0	0	0	0	1
	8	0	0	0	0	0	0	0	1	0	0	0	1
	10	0	0	0	0	0	0	0	0	1	4	0	5
	12	0	0	0	0	0	0	0	0	2	6	0	8
	15	0	0	0	0	0	0	0	0	0	0	15	15
Total		21	5	1	2	2	2	13	1	7	10	15	79

In case of Actors (label 1) both coders agreed in 20 cases but the number of disagreement is quite high with coder 1, choosing 12 Sayer where coder 2 chose Actor. Also coder 1 chose 2 Senser and Range in 4 cases, where coder 2 chose 1 Phenomenon.

Token (label 10) and Value (label 12) were also mixed between both coders: in 6 cases both chose Value but of these, 4 were considered Token by coder 2 and 2 by coder 1. These are relational process where there is not a clear cut distinction among them.

	Clause Complex	Coder1	Coder2
	<i>Brown</i> ha rechazado adelantar unas elecciones, como le pide la oposición conservadora, que presumiblemente perdería por goleada.	Actor	Sayer
(11)	En cualquier caso, <i>los ingresos de Microsoft</i> alcanzaron en los seis primeros meses de su ejercicio fiscal 30.129 millones de dólares, un incremento del 29 por ciento respecto al mismo período de 2006, cuando facturó 23.353 millones de dólares.	Actor	Token
	En el caso del general Navarro, y a expensas de posibles recursos, <i>la pena de tres años</i> supone su ingreso en prisión.	Token	Carrier

This disagreement is reflected in the Kappa value at 0.475, which is considered only moderate (see table 13.13 below).

**Table 13.13** Kappa agreement measures in Experiential roles in Spanish

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.475	0.062	10.108	0.000
<b># of valid cases</b>	79			

### 13.3.2.3 Task 3: Classification of types of Textual Themes

In this task coders had to classify the marked span according to the labels given. In task 3 the labels corresponded to the types of Textual Themes: Linkers, Binder, Correlatives or None.

**English** When looking at the English dataset, it can be observed that the disagreements mainly occur with the subtype Binder: although coders agree in 30 cases, there is a great disagreement where coder 1 labels 5 Linkers coder 2 labels them as Binders. The same can be observed in Linkers where there are 11 cases in which they both agree but 5 are considered Binders by coder 2 and 1 None by coder 1.

### 13. Performing Annotation Experiments

	Clause Complex	Coder1	Coder2
	And if there is, why did no one look at it?	Linker	None
(12)	But Russian oil output is already falling because of weak investment, as is Mexican production.	Binder	None
	Yet on Saturday David Cameron offered François Hollande two C-17 transport planes to help French forces counter an Islamist rebellion.	Binder	Linker

**Table 13.14** Interannotator agreement in types of Textual Themes in English

Contingency table Coder 1 * Coder 2						
		Coder 2				
		Linker	Binder	Cor	None	Total
Coder 1	Linker	11	0	0	1	12
	Binder	5	30	0	1	36
	Cor	0	0	1	0	1
	None	0	1	0	0	1
Total		16	31	1	2	50

In spite of these discrepancies the agreement is still substantial as the following table 13.15 shows: a Kappa value of 0.664

**Table 13.15** Kappa agreement measures in types of Textual Themes in English

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
Kappa agr. Measure	0.664	0.103	5.612	0.000
# of valid cases	50			

**Spanish** When looking at the Spanish dataset, agreement is total in all cases except in Binders where both coders agree in 17 cases but both label 1 case each as None. This is illustrated in table 13.16 below:



**Table 13.16** Interannotator agreement in types of Textual Themes in Spanish

Contingency table <b>Coder 1 * Coder 2</b>						
		Coder 2				Total
		Linker	Binder	Cor	None	
Coder 1	Linker	27	0	0	0	27
	Binder	0	17	0	1	18
	Cor	0	0	3	0	3
	None	0	1	0	0	1
Total		27	18	3	1	49

The results of the Kappa test, show in table 13.17, reflect this almost total agreement at 0.927

**Table 13.17** Kappa Agreement Measures in Types of Textual Themes in Spanish

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
Kappa agr. Measure	0.927	0.049	7.994	0.000
# of valid cases	49			

#### 13.3.2.4 Task 4: Classification of types of Interpersonal Themes

The fourth task of the Agreement Study 2 focuses on the labeling of the types of Interpersonal Themes. The coders had to choose among: Stance, Engagement, Modal, Comment and None. A detailed definition of each category was provided in chapter 11.1.2 above. The dataset is included in appendixes C.1 and C.2 .

**English** In the English data, both coders annotated 39 cases, leaving each one blank. Table 13.18 graphically displays the results of the annotations, which clearly show a high degree of mismatch between the codings.

**Table 13.18** Interannotator agreement in types of Interpersonal Themes in English

Contingency table <b>Coder 1 * Coder 2</b>							
		Coder 2					Total
		Stance	Eng	Modal	Comment	None	
<b>Coder 1</b>	<b>Stance</b>	10	0	0	1	1	12
	<b>Engagement</b>	9	2	1	0	0	12
	<b>Modal</b>	0	0	3	0	0	3
	<b>Comment</b>	3	0	0	7	0	10
	<b>None</b>	2	0	0	0	0	2
<b>Total</b>		24	2	4	8	1	39

As shown in the table, both coders agreed on 10 cases. Where coder 1 chose Stance coder 2 also selected 9 Engagement, 3 Comment and 2 None. On the other hand, coder 2 choose 1 Comment and 1 None. Engagement was coded by both annotators only in 2 occasions. In 9 clause complexes coder 2 selected Stance and 1 Modal where coder 1 chose Engagement. Regarding Modals both agree in 3 cases and only in 1 occasion coder 2 marked as Engagement an occurrence considered by coder 1 as Modals. With respect to the category Comment both coders agree in 7 cases and disagree in 4 cases, distributed in 3 for coder 2 and 1 for coder 1, both marked as Stance. Finally None has not achieved any level of agreement, coder 1 has selected 2 where coder 2 marked Stance and coder 2 has 1 element as None where coder 1 has Stance.

The categories of Stance and Engagement are the ones that posed more problems to the coders.

	Clause Complex	Coder1	Coder2
(13)	<i>It is therefore understandable that</i> the Obama administration is worried about the possibility of a referendum on the UK's continued membership of the 27-nation group.	Sta	Eng
	<i>For goodness sake</i> let's have elections combined – if you go out to elect a councillor or MP you may as well pick a police commissioner too.	Sta	Eng
	<i>It is worth remembering too that</i> Americans as a whole did not want a Republican House.	Sta	Eng
	<i>True,</i> the GOP has retained control of the House of Representatives, but largely thanks to its diligent gerrymandering of Congressional districts; in the overall House vote, the Democrats won half a million more.	Sta	Com
	<i>Let us hope that,</i> for once, lessons really have been learnt by officialdom.	Sta	None

This is reflected in the Kappa value obtained (table 13.19) that, contrary to other tasks, is only at 0.405 considered a Fair level of agreement.

**Table 13.19** Kappa agreement measures in Interpersonal Themes in English

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.405	0.100	5.109	0.000
<b># of valid cases</b>	39			

**Spanish** Both coders did 49 annotations, leaving blank one each. In Spanish the situation found in the accomplishment of this task is similar to the English one where the disagreement is quite relevant.

### 13. Performing Annotation Experiments

**Table 13.20** Interannotator agreement in types of Interpersonal Themes in Spanish

Contingency table Coder 1 * Coder 2							
		Coder 2					Total
		Stance	Engagement	Modal	Comment	None	
Coder 1	Stance	14	1	2	1	0	18
	Engagement	16	6	0	1	0	23
	Modal	0	0	4	0	0	4
	Comment	3	0	0	7	0	10
	None	0	0	0	1	0	1
Total		30	7	7	4	1	49

Stance has been selected by both coders in 14 clause complexes, but coder 1 has chosen 16 times Engagement where coder 2 annotated Stance as well as 3 times Comment. On the other hand coder 2 has 1 Engagement, 2 Modal and 1 Comment where coder 1 marked Stance. In the case of Engagement the cases of disagreement are very salient. They both agreed on 6 cases and disagreed on 16 that coder 2 marked and Stance and 1 as Comment. Coder 1 only marked 1 case as Stance. Regarding Modal they agree on 4 clause complexes, in 2 cases coder 1 marked Stance where coder 2 did Modal. In Comments the agreement is of 7 cases, in 3 of them coder 1 selected Stance, Comment and None.

As in English the most controversial types are Engagement and Stance.

Clause Complex		Coder1	Coder2
(14)	<i>Se puede afirmar que</i> el dinero facilitado, se gasta en viajes, comidas y otros fastos.	Sta	Eng
	<i>Se hace obvio que</i> los docentes siguen siendo ignorados a pesar de los acuerdos alcanzados durante las I Jornadas de Secundaria.	Sta	Eng
	<i>Pero es difícil admitir que</i> indagar sobre la existencia y naturaleza de la misión que parece haber condicionado la conducta del principal acusado sea tarea ajena al tribunal, bien directamente, bien por deducción de testimonio.	Sta	Eng
	Y, <i>desde luego</i> , deben presentarse una vez bien estudiadas y con la viabilidad de su aplicación asegurada.	Com	Eng

This disagreement is reflected in the kappa value (table 13.21) at 0.291 which is considered a Fair agreement.

**Table 13.21** Kappa agreement measures in Interpersonal Themes in Spanish

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.291	0.094	3.766	0.000
<b># of valid cases</b>	49			

### 13.3.2.5 Task 5: Classification of types of PreHeads

This task deals with the classification of PreHeads into three main types: Circ, CCL, Finite and None. (See chapter 11.2.1 for a complete description of each type and appendixes C.1 and C.2 for the dataset)

**English** Both coders annotated 49 cases, leaving one case each blank. Circumstantials account for 23 agreements. Only in one case coder 1 marked one instance as None where coder 2 did it as Circ. The result of the annotations is displayed on table 13.24 below:

**Table 13.22** Interannotator agreement in types of PreHead Themes in English

Contingency table Coder 1 * Coder 2						
		Coder 2				
		Circ	CCL	Finite	None	Total
Coder 1	Circ	23	0	0	0	23
	CCL	0	22	0	1	23
	Finite	0	0	2	0	2
	None	1	0	0	0	1
Total		24	22	2	1	49

This almost perfect agreement is reflected in the kappa measure at 0.927 as illus-

trated on table 13.25 below:

**Table 13.23** Kappa agreement measures in PreHead Themes in English

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
<b>Kappa agr. Measure</b>	0.927	0.049	7.629	0.000
<b># of valid cases</b>	49			

**Spanish** Both coders annotated 49 cases, leaving one blank. In Spanish the rate of agreement here is also very high. As shown in table 13.24 coders agreed on labeling Circ in 17 cases, and there was only 1 disagreement that was coded as Finite by coder 2. Both coders agree in 8 cases when labeling CCL, and again only in 1 case coder 2 selected Finite. Finally, coders agreed when coding Finite (in 22 cases) and disagree in the commented above by coder 2. Neither coder 1 nor coder 2 selected any case under the category None.

**Table 13.24** Interannotator agreement in types of PreHead Themes in Spanish

Contingency table <b>Coder 1 * Coder 2</b>					
		Coder 2			
		Circ	CCL	Finite	Total
<b>Coder 1</b>	<b>Circ</b>	17	0	1	18
	<b>CCL</b>	0	8	1	9
	<b>Finite</b>	0	0	22	22
<b>Total</b>		17	8	24	49

These almost perfect agreement is reflected on the Kappa value at 0.934.

**Table 13.25** Kappa agreement measures in PreHead Themes in Spanish

Symmetric Measures	Value	Asymp. Std. Error(a)	Approx. T. (b)	Approx. Sig.
Kappa agr. Measure	0.934	0.045	8.849	0.000
# of valid cases	49			





# 14

## Summary and Discussion

The two agreement studies performed on the training corpora revealed interesting aspects of the annotation process of the selected thematic features of the Theme model of SFGS. The identification tasks in Agreement Study 1 proved to be more difficult than the labeling tasks in Agreement Study 2. On the main, the tasks were more challenging for the Spanish data than for the English data, probably due to the complex morphology of the Spanish verbal group.

In task 1 of Agreement Study 1 a reliable result was achieved in both languages (AGR 0.97 and 0.74), but agreement was lower in Spanish than in English. One possible explanation for this is the more complex morphology of the Spanish verbal group which makes it more difficult for coders to delimitate the correspondent span of the Thematic field.

In task 2 a high level of agreement was achieved in the selection of the spans corresponding to Thematic Head, PreHead, Interpersonal Theme, Textual Theme, Predicated Theme and ‘There’ Theme. Agreement on identifying the Thematic Head in English is at 0.93 and in Spanish at 0.91, although there is a slight difference in the values reflecting a better agreement in English. The category of PreHead shows a better agreement in Spanish 0.89 than in English 0.78. The identification of Interpersonal Themes shows higher agreement in Spanish (0.94) than in English (0.37), and Textual

Themes are similarly identified by both coders in both languages (English 0.96 and Spanish 0.97). Almost perfect agreement was achieved in the identification of Predicated Themes and ‘There’ Theme in both languages. These results are shown in table 14.1 below:

**Table 14.1** Results of Ag Study 1 - English and Spanish

<b>Agreement Study 1</b>						
<b>Task 1</b>	English	0.97				
	Spanish	0.74				
<b>Task 2</b>	<b>TH</b>	<b>PH</b>	<b>IT</b>	<b>TT</b>	<b>PT</b>	<b>There</b>
English	0.93	0.78	0.37	0.96	0.98	1
Spanish	0.91	0.89	0.94	0.97	0.98	0.98

In Agreement Study 2, coders had to perform five classification tasks consisting of labeling specific thematic spans that had been agreed upon in Agreement study 1.

In task 1, the coders had to select the type of Thematic element highlighted in the dataset. The Kappa value obtained is of 0.915 for English and 0.839 for Spanish, which indicates almost perfect agreement between coders.

In task 2 -labeling of experiential roles realizing Thematic Head- the kappa value obtained in English was of 0.875 and 0.475 for Spanish. The moderate level of agreement presented by the Spanish coding is due to discrepancies presented by the coders in the identification of relational processes where Token and Value types were mistaken.

In task 3 -annotation of subtypes of Textual Themes- the Kappa values obtained were of 0.664 in English and 0.927 Spanish, with a lower agreement in English than in Spanish. In this case the problem lied in the selection by one of the coders of one category where the other marked it as None or as the opposite one. This may point to the need for further refinement and clarification of the definitions in the annotation guidelines.

In task 4, coders had to label different subtypes of Interpersonal Theme. In this task both languages present a very low Kappa values (0.405 in English and 0.291 in Spanish), which can be considered as moderate and fair, respectively, according to Viera & Garrett (2005). The subjectivity and novelty of some of the categories (Stance and Engagement) points to the need for refinement, clarification and even redefinition before they can be reliably used for automatic annotation purposes.

---

In task 5, the annotation result reveal a high level of agreement in the selection of the types of PreHead, with Kappa values of 0.927 in for English and 0.934 for Spanish. These results are summarize in table 14.2:

**Table 14.2** Results of Ag Study 2 - English and Spanish

<b>Agreement Study 2</b>		
	<b>English</b>	<b>Spanish</b>
<b>Task 1</b>	0.915	0.839
<b>Task 2</b>	0.875	0.475
<b>Task 3</b>	0.664	0.927
<b>Task 4</b>	0.405	0.291
<b>Task 5</b>	0.927	0.934

These results indicate that the annotation schemes designed for the thematic categories of the model of Theme in SFGS are reliable and can, therefore, be consistently used for the annotation of a larger corpus of bilingual texts.

From the theoretical and methodological point of view, the results also indicate that the categories of the model of Theme of SFGS instantiated in this empirical study are reproducible in the majority of the cases. Those cases where poor disagreement occurred between annotators indicate that:

a) the definitions probably need more clarity and accuracy (theory redefinition). This is the case of Spanish experiential roles, namely Value and Token and some of the subtypes of Interpersonal Theme (Stance and Engagement). It could also be considered the redefinition, in a lesser extent, of the Textual Theme types in English (Binder and Linker).

b) the annotators cannot handle certain categories, probably because the theoretical coverage falls short (theory formation). This is clearly the case of Interpersonal Stance and Engagement, both categories were problematic in both languages. This may indicate that the category itself is ambiguous and cannot be properly identified by the coders. The difficulties lied in the distinction among these two elements, one solution may well be to merge these two categories into a single category, and review the option to split them in the future with better specifications in the guidelines if a possibility to disambiguate these elements is found.

## *14. Summary and Discussion*

---

## **D. ANNOTATING A LARGER CORPUS**



# 15

## Introduction

In this 4th Part, I will focus on the manual annotation of a larger corpus with the thematic tags which were validated through the agreement studies described in part C. This is one of the aims of the current dissertation and a preliminary step for the development of automatic annotation which will be explained in Part 5.

In addition, and in order to complete the thematic characterization of the three newspaper genres selected for this study, I will also report on the manual annotation of Thematic Progression (TP) patterns within the generic moves that organize these three genres.

Excel was selected as annotation tool for the Part dedicated to the Empirical Validation of the Model of Theme and for the present one Annotating a Larger Corpus. In Empirical validation the coders were provided the annotation sheets in this format and the Annotation of the Larger Corpus is carried out also in Excel. The selection of this tool allowed two purposes: the manual annotation of all levels of analysis and the accuracy of quantitative data account Rasinger (2008). In the last part of this study the use of a different support to automatically annotate the same data will give the opportunity to compare both methodologies, the more traditional one using Excel and the computational one using GATE.

For this task I have followed the empirical evidence provided by Dligach et al.,

according to which, the greatest value for an annotation project lies in single annotating more data when there unlimited supply of unlabeled data and limited funding. In fact, although these authors recognize that “there may, however, be other considerations that still argue in favor of double annotation” their main thesis is that, instead of double annotating and adjudicating –the common accepted wisdom in annotation projects-, “it is often better to single annotate more data because it is a more cost-effective way to achieve a higher performance” (Dligach et al., 2010).

In the following chapters I will focus on the features of the larger corpus to be annotated, the thematic features –both at clausal and discourse levels- which were hand-coded, and the quantitative results obtained.



# 16

## The Larger Corpus

The larger corpus used in this study is a bilingual comparable English-Spanish corpus consisting of a total of two hundred and twenty texts (sixty news reports, sixty editorials and one hundred letters to the editor) evenly divided into English and Spanish. The bilingual sample collected is outlined in table 16.1.

**Table 16.1** Bilingual corpus

Bilingual Corpus	Texts	Clauses	Words
Reports	60	708	21717
Editorials	60	1112	28192
Letters to the Editor	100	731	17392
Total	220	2551	67301

All texts were collected from three well-known British and Spanish newspapers between 2009 and 2013 using an opportunistic sampling technique, i.e. the texts chosen were “readily available” (see Hausser, 2001:291) and online. The English texts were extracted from three online newspapers: Times online (<http://www.timesonline.co>).

uk), The Independent (<http://www.independent.co.uk/>) and The Telegraph (<http://www.telegraph.co.uk/>). The Spanish texts were gathered from the online versions of the following national newspapers: El País (<http://elpais.com/>), El Mundo (<http://www.elmundo.es/>) and La Vanguardia (<http://www.lavanguardia.com/>). The collected sample is an extension of version of a previous one used by the author of this work and colleagues in the context of the CONTRANOT and MULTINOT projects Lavid, Arús, & Moratón (see 2009b, 2010a, 2011, 2012); Arús, Lavid, & Moratón (see 2012); Lavid, Arús, & Moratón (see 2013a,c); Arús, Lavid, & Moratón (see 2013); Moratón & Lavid (see 2012) and in my own Master Thesis (Moratón, 2009)<sup>1</sup>.

---

<sup>1</sup>The annotated corpus is included in appendix D and it is also available as text archives (txt format) which can be uploaded onto the GATE annotation platform (available in CD format)

# 17

## Annotating Thematic Features at Clause Level

The unit of analysis used for the annotation of thematic features in this study is the complex clause. The decision to use this unit of analysis responds to theoretical and practical reasons, as suggested by Taboada & Zabala (2008). My theoretical inspiration was Halliday and Webster's definition of the scope of a clause complex: “..in a written text one can go by the orthography: a clause complex is equated with a written sentence (which is the written surrogate of the clause complex to start with)” (Halliday & Webster, 2009:76).

The practical reasons are the consistency across coders and the applicability to different languages (Taboada & Zabala, 2008). Coders need to clearly recognize the units of analysis and these units have to be comparable in English and Spanish. Another third practical reason is the need to be computationally replicable. Automatic segmentation of clause complexes is called *segmentation* and is commonly performed by Sentence Splitters which “do not work perfectly but achieve a reasonable accuracy” (Halliday & Webster, 2009:115).

The last advantage for this selection is that the identification of these units of analysis allows to count them and analyze their distribution providing “handles on things that would otherwise be obscure” (Chafe, 1994:58). Therefore all clause complexes occurring in the texts of the bilingual sample were numbered and inserted in tables

## 17. Annotating Thematic Features at Clause Level

for the analysis of the types of thematic elements and semantic and syntactic features. The clause complexes are separated as units of analysis using Microsoft Word (2011, v14.3.2) for Mac (OSX 10.6.8) through search with regular expressions; and reviewed in order to avoid mistakes. Then, the texts segmented into clause complexes are copied into an Excel file (2011, v14.3.2) as the figure below shows. Each full sentence occupies a cell in the template, next to it there is a space reserved for the analysis of the distinct features of the study.

Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair

Cl#	Thematic Field				SYNTACTIC REALIZATION				SYNTACTIC REALIZATION				Text	
	TEXTUAL	STANCE	ENGAGEMENT	PRE-HEAD	STANCE	ENGAGEMENT	GRA.	HEAD	STANCE	ENGAGEMENT	GRA.	EXPERIENTIAL ROLE		
1								Dominique Strauss-Kahn, the French head of the International Monetary Fund,			NP+App	Proper Noun+Det Adj Noun+Prep Proper Noun	Actor	Dominique Strauss-Kahn, the French head of the International Monetary Fund, escaped dismissal for a one-night stand with a subordinate today, but was denounced by board members for a "serious error of judgment."
2								The fund's managing director			NP	Det Noun Adj Noun	Goal	The fund's managing director was allowed to remain in his post after an investigation cleared him of accusations of harassment, favoritism and abuse of power over his affair with Priscilla Nogué, a Hungarian economist.
3	but							his reputation			NP	Det Noun	Carrier	But his reputation took a hammering in the statement by the IMF board, which said the relationship "was regrettable and reflected a serious error of judgment on the part of the managing director."
4								Shakour Shaulan, the fund's executive director			NP+App	Proper Noun+Det Noun Adj Noun	Sayer	Shakour Shaulan, the fund's executive director, said IMF staff, and notably female staff, "are not at all happy" with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet le grand séducteur.
5								The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said			Clause		verbiage	"The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said Mr Shaulan. "We will continue to work with him."
6								Mr Strauss-Kahn, 59,			NP+App	Proper Noun Noun	Sayer	Mr Strauss-Kahn, 59, said in a statement: "I very much regret the incident and I accept responsibility for it."
7								The investigation			NP	Det Noun	Goal	The investigation was launched when Martin Blythe, a senior Argentine economist, said Mrs Nogué, his wife, had been seduced at the Davos international forum.
8								A US law firm			NP	Det Adj Noun Noun	Receiver	A US law firm was asked to probe accusations that Mrs Nogué, who left the IMF in August, had been forced out by Mr Strauss-Kahn, possibly with an excellent pay-off.
9								Morgan, Lewis & Bockius			NP+NP+NP	Proper Noun+Proper Noun+Proper Noun	Sayer	Morgan, Lewis & Bockius said there was no evidence to support the claims.
10								The decision			NP	Det Noun	Goal	The decision was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remedy and regulate capitalism in the wake of the financial crisis.
11								French commentators			NP	Adj Noun	Sayer	French commentators hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is now behind us. We love each other as much as we did on the first day."

Figure 17.1 Screenshot of the manual annotation sheet

# 18

## Results of the Thematic Annotation

### 18.1 Introduction

This section describes the result of the manual annotation of the Thematic field and its elements in each genre of study.

For the sake of clarity the results are organized in four sections. The first three sections are devoted to each genre, giving a detailed account of its most relevant thematic features, first in English, then in Spanish and finally from a language contrastive perspective. Then, the fourth section offers an account of the results from a contrastive point of view, this time the comparison is carried out between genres.

### 18.2 News Reports

In White (1998:243) words news reports are “grounded in communicative events such as speeches, interviews and press releases”, which “act primarily to represent, not activity sequences, but the points of view of various external sources”

Table 18.1 shows the corpus of news reports in English and Spanish. The first column lists the number of texts per language, the second column presents the total number of complex clauses in each corpus, the third column presents the average num-

ber of clauses per text, the column entitled *Words* provides information on the number of words in each corpus, and the finally the last column presents the average number of words per clause in each language.

**Table 18.1** Corpus of news reports in English and Spanish.

NR	Texts	Clauses	Cl/text	Words	W/cl
English	30	349	11.63	9767	27.98
Spanish	30	359	11.96	11950	23.28
Total	60	708	-	21717	-

### 18.2.1 English News Reports

The total number of clauses is 349 and 3.15% are rhematic clauses with no thematic construction. The rest of the complex clauses (96.85%) are declarative thematic complex clauses. Although our analysis will focus on the thematic clause complexes, attention will be paid to the appearance of these rhematic clauses in the texts where relevant. The first step consists in the identification of the thematic field in each clause. Then, in this thematic field I will analyze the thematic elements.

The percentage of each type of Theme is calculated considering the total amount of thematic clause complexes and not the total number of clauses in the corpus. For example, the percentage of PreHeads in English news reports is counted on the number of thematic clause complexes in the corpus, which is 338.

The results of the annotation of thematic elements in English news reports are shown in the following table:

**Table 18.2** Thematic elements in English news reports

English		Thematic Field		
Types	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
	6.21%	0.30%	16.57%	100%

The thematic elements are divided into those occurring in the Inner thematic field and those occurring in the Outer thematic field. The Outer thematic field includes two

types of thematic elements: Textual Themes (TT) and Interpersonal Themes (IT). The Inner thematic field comprises PreHeads (PH) and Heads (H).

Textual Themes appear only in a 6.21% of clause complexes. Interpersonal themes are almost inexistent (0.30%). In the Inner thematic field PreHead accounts for 16.57%. As dealing with thematic clauses all of them have Head (100%). This means that 76.92% of the clause complexes in the English news reports corpus are realized by Heads not accompanied by any of the other possible optional elements that can appear in the Thematic Field. This points to a likely simple thematic construction of Themes in English news reports. In the following sections these data are analyzed in detail.

### 18.2.1.1 Textual Themes in English News Reports

Textual Themes are those elements which were defined as instrumental in the creation of the logical connections in the text. The proportion found in this corpus is relatively low (6.21%), probably due to the fact that news reports are short in length textual connections rely on paragraphing. The average length of clause complexes per text is of 11.73 cl/text, with such a short text extension the need to use Textual elements as connectors is very limited.

The types of Textual elements are analyzed according to the types explained in the Methodology section 11.1.1. The following table shows the elements found in English news reports corpus:

**Table 18.3** Types of TT in English news reports

Types of Textual Themes	
Binder	90.48%
Linker	9.52%
Total	100%

*Binders* are described by Halliday as the “binding of elements of unequal status” and *Linkers* as “the linking of elements of equal status” (Halliday & Matthiessen, 2004:384). In spite of the factual nature of this genre, one should expect to find more connective devices than adversative or contrastive connectors but the data shows that there is a statistically significant difference ( $P < 0.0001$ )<sup>1</sup> between the occurrences of these two

<sup>1</sup>The calculation of the statistical significance has been carried out using Chi-square calculation.

types of Textual Themes, with the use of *Binders* being more frequent (90.84%) than the use of *Linkers* 9.52%. Even though this fact may be influenced by the mentioned short length of the texts, the only connectors needed are the ones that convey the adversative sense whereas the sense of addition is achieved by the juxtaposition of elements. Another fact that strengthens this idea is that the amount of these elements in the text is very scarce.

These *Binders* are primarily realized by the adversative connectors *However* 38.10% and *But* 23.81% and the contrastive connector *Although* 14.29%. Some examples are provided below:

- (1) **However**, the Saudis will be listening to the grinding cogs of the slowing Asian industrial juggernaut with mounting concern.  
(Report: *Stop ministers seeing official stats early – MPs*. Sentence 14)
- (2) **However**, Hilco faces competition from other suitors for the UK business and the structure of any deal is likely to be complex, partly because HMV has many loss-making stores that a new owner is likely to jettison.  
(Report: *Hilco UK emerges among favourites to give HMV a spin*. Sentence 7)
- (3) **But** his reputation took a battering in the statement by the 1M F board, which said the relationship ‘was regrettable and reflected a serious error of judgment on the part of the managing director.’  
(Report: *Opec hawks want to cut oil production to keep up price*. Sentence 8)
- (4) **Although** the Bank of Japan has remained tight-lipped on the question of currency intervention, Shoichi Nakagawa, Japan’s Finance Minister, said today that he saw “excessive volatility” in the yen’s exchange rate and that he was he “watching with great interest.”  
(Report: *G7 ‘preparing to drive down the yen’*. Sentence 8)

On the other hand, *Linkers* are realized by the copulative and addition connectors *And* 4.76% and *In addition* 4.76%:

- (5) **And** this month Mr Dilnot warned politicians to avoid mixing up debt with the deficit after the Prime Minister claimed, erroneously, in a party political broadcast that the Coalition was “paying down Britain’s debts”.  
(Report: *Stop ministers seeing official stats early – MPs*. Sentence 14)
- (6) **In addition**, Qatar Holding will receive £66 million for having helped arrange the capital raising.

---

See Butler (1985).



(Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi*. Sentence 8)

### 18.2.1.2 Interpersonal Themes in English News Reports

Interpersonal themes are those elements that express the attitude and the evaluation of the speaker with respect to his or her message. This type of Themes is very scarce in News reports, where evaluation and attitude are infrequent. As shown on table 18.2, there is only one element (0.30%) in the entire news reports English corpus with this characteristic.

- (7) **It is not appropriate that**, ministers should have lengthy prior access to certain statistics but other interested parties do not, it says.

(Report: *Stop ministers seeing official stats early – MPs*. Sentence 3)

### 18.2.1.3 PreHead in English News Reports

The element PreHead is a non-nuclear part of the Inner Thematic Field, an optional element realized by any circumstantial and/or finite element preceding the Thematic Head which do not exhaust the thematic potential of the clause.

In English news reports only 16.57% of occurrences are PreHeads. As shows table 18.2. The types of PreHeads are annotated according to the labels described in chapter 11:

**PH-Circ** adverbial groups or prepositional phrases.

**PH-CCL** circumstantial clauses.

**PH-Finite** finite verbal forms.

Table 18.24 accounts for the distribution of the types of PreHeads found in English news reports, calculated on the total number of PreHeads found.

**Table 18.4** Types of PreHeads

PreHead Types	%
PH-Circ	57.14%
PH-CCL	42.86%
PH-Finite	0.00%
Total	100%

There is no occurrence of *PH-Finite* forms in this type of texts, which indicates that there are only declarative sentences, in English finite forms in first position are found in imperative or questions but not in declarative sentences. The distribution of *PH-CCL* and *PH-Circ* occurrences is quite even. *PH-CCL* is 42.86% and *PH-Circ* 57.14%. What is worth noticing is that of the 42.86% *PH-CCL* elements an amount of 45.83% are direct quotations. This is a characteristic typical of these type of texts, as White (1997) says, it is what news reports do: “(…) presenting a diversity of opinions, voices, and perspectives of the event (…)”

Another interesting analysis that provides valuable information on PreHeads is the study of the functions realized by PreHeads in terms of the type of circumstance expressed.

**Table 18.5** Circumstantials

PreHead Circ	%
Time	33.93%
Quote	19.64%
Place	14.29%
Evidence	10.71%
Other	21.43%
Total	100%

According to the analysis shown in table above, the most common realization is that of *Time* 33.93% followed by *Quotes* (19.64%) that present different voices relevant to the issue and *Place* circumstantials 14.29%).

“Circumstantial elements in thematic position may perform a variety of functions,

depending on their semantic nature. Temporal or Location Circumstantials often function as scene-setting frameworks” (Downing, 1991:127)

For White (1998:143) news reports are “grounded in communicative events such as speeches, interviews and press releases”, which “act primarily to represent, not activity sequences, but the points of view of various external sources”. And that is why direct quotations are regularly found (see table 18.5). Both analysis reveal that in the English news reports the circumstantials of *Time* are realized mainly by *Prepositional Phrases* 47.37% and *Adverbial Phrases* 10.53%. Below are some examples of *Time* realized by *Prepositional Phrases* :

- (8) **In December**, Mr Dilnot rebuked the Coalition for claiming that health spending had increased in real terms since 2010.  
(Report: *Stop ministers seeing official stats early - MPs*. Sentence 12)
- (9) **On Friday**, Hilco was locked in talks with Deloitte which are expected to continue this week.  
(Report: *Hilco UK emerges among favourites to give HMV a spin*. Sentence 5)

Other examples of *Time* expressions are realized by *Adverbial Phrases* :

- (10) **Meanwhile**, wages have remained stagnant.  
(Report: *Lower income earners being squeezed hardest*. Sentence 9)
- (11) **Now**, the council proposes to sell off a cluster of houses for just £1 to residents - on the condition they will be brought up to a decent standard.  
(Report: *Pound land: Derelict houses in Liverpool to be sold for just £1*. Sentence 4)

The “diversity of opinions and voices” mentioned by White (1997) are represented by the use of direct quotations of people relevant to the issue on the piece of news. In this annotation the use of *Quotes* in PreHead position reveals that it is very likely to find the author of the quote functioning as Head, as shown in the examples (12) and (13):

- (12) **“We, like most companies, saw a slowdown at the tail end of the quarter in particular,”** Chief Financial Officer Chris Liddell said.  
(Report: *Microsoft results offer hope to tech sector*. Sentence 12)
- (13) **“This is a healthy company,”** he said.

(Report: *Dutch Government to inject €10bn of fresh capital into ING savings bank.* Sentence 10)

The third most frequent type of PreHead is that of *Place*, another resource of news reports to set the information the text is conveying. White (1998) says that News Stories present independent satellites in their body to contextualize the events in a spatial context. The most frequent syntactic realization of this circumstance are *Prepositional Phrases* 87.50%, as in the following example:

- (14) **In a conference call with investors**, Mr Liddell said Microsoft will “tailor our business to whatever the economy brings.”

(Report: *Microsoft results offer hope to tech sector.* Sentence 14)

#### 18.2.1.4 Thematic Head in English News Reports

Thematic Head is defined as ‘first element with a function in the experiential configuration of the clause which is more central to the unfolding of the text by allowing the tracking of the discourse participants’ (Lavid et al., 2010d:299). All thematic sentences have Head whereas rhematic sentences, do not have Theme, nor Thematic Head. The total number of Thematic Sentences in news reports English Corpus is of 338, which accounts for the total number of elements on which the analysis has been carried out. In this section a detailed analysis into the Thematic Head is carried out, focusing on its semantic and syntactic realizations. Only by looking at these specific features it is possible to obtain relevant information about this thematic element. Since this element is common to all Thematic Complex clauses in the corpus, a quantitative analysis based only on its frequency would not lead to a deep understanding of it.

#### Syntactic Realizations of Thematic Head in English News Reports

According to Bell (1991) “*hard news* distinctiveness lies in its exclusion of first person presentation, citations of sources and numbers with a precision and syntactic complexity”. As this study shows, the realization of Heads is done by *Noun Phrases*, more specifically by *Proper Nouns* and *Determiners plus Nouns*, as examples (15), (10) and (17), (18) show.

As shown in table 18.6 below, *Noun Phrases* appear more frequently as Syntactic realizations of Head (67.46%) than *Noun Phrases plus Apposition* (23.96%). In fact, there is a statistically-significant difference between the use of these two elements ( $P < 0.0001$ ). The fact that 77.21% of all clause complexes in the corpus are realized by Heads as a single Thematic Element, and the statistically-significant relevance in

the prevalence of *Noun Phrases* over more complex syntactic formations confirms what was mentioned in the introduction of this section: the simple thematic construction of the Thematic field in English news reports clause complexes.

*Quotations* as Thematic Heads are also present in the text (1.78%) but they are relegated in favour of *Noun Phrases*, more specifically of *Noun Phrases* realized mainly by *Proper Nouns* (27.19%), that is, *what* is said is given less relevance than *who* says it. This will be further developed in section 18.2.1.4 where I deal specifically with the semantic realization of Heads:

**Table 18.6** Syntactic realization of Head

Syntactic forms	%
Np	67.46%
Np+App	23.96%
Np+Np	2.07%
Quote	1.78%
Nonfinite	1.48%
Other	3.25%
Total	100%

If looking in detail at *Noun Phrases*, it is interesting to note the prevalence of *Proper Nouns* 27.19% and *Determiner plus Nouns* 22.81% as the preferred realization over the rest of the options found. The nature of the texts requires the identification of the entity about which the information is going to be given, and the use of these two strategies clearly identifies the discourse participant in a simple and clear way. Below are some examples of the use of *Noun Phrases* realized by *Proper Nouns*:

- (15) **Mr Bos** said the deal was structured to ensure a state exit as soon as “this financial hurricane” receded. (Report: *Dutch Government to inject €10bn of fresh capital into ING savings bank*. Sentence 9)
- (16) **Barclays** is seeking to raise up to a further £1.5 billion from existing and other investors The move boosted Barclays’ share price by 7.1 per cent to 220p in early trading. (Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi*. Sentence 4)

Examples of *Noun Phrases* realized by *Determiner plus Noun*:

- (17) **The investigation** was launched when Mario Blejer, a senior Argentine economist, said Mrs Nagy, his wife, had been seduced at the Davos international forum.

(Report: *Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair.* Sentence 7)

- (18) **The move** follows ING's revelation after trading on Friday that it expected a €500 million net loss in the third quarter.

(Report: *Dutch Government to inject €10bn of fresh capital into ING savings bank.* Sentence 13)

Another frequent realization of Thematic Heads is through *Noun Phrases* followed by *Appositions* (23.96%). These provide more elaborated material or extended information and its frequency can be explained by the clear need in News Articles to identify and provide additional information on the discourse participants. These formations are evenly distributed between *Noun Phrases followed by Prepositional Phrases* (*Np + PP*) 38.27% i.e. (19), (20) and *Proper Nouns followed by Noun Phrases* 25.93% i.e. (21), (22). These *Appositions* give further details on the issue at stake, as in the following examples:

- (19) **Revenue for the quarterly period** rose 9 per cent to \$15.06 billion.

(Report: *Microsoft results offer hope to tech sector.* Sentence 3)

- (20) **Anxieties over the global outlook** were stoked after the latest snapshot of economic sentiment in the euro-zone showed confidence among businesses and consumers at its lowest since 1993.

(Report: *US economy officially on brink of recession.* Sentence 12)

- (21) **Toan Tran, an analyst at Morningstar**, said: "If you are a Microsoft investor you should be happy that Microsoft thinks they will weather the storm relatively well."

(Report: *Microsoft results offer hope to tech sector.* Sentence 7)

- (22) **Taro Aso, the Japanese Prime Minister**, said: "A harsh storm seen only once in 100 years is raging. Under such circumstances, I am certain that what is most important is to remove uncertainties from the lives of people."

(Report: *US economy officially on brink of recession.* Sentence 18)

Not as relevant, but interestingly enough, the third element found in the realization of *Noun Phrases followed by Appositions* (*Np+App*) involve the use of *Proper Nouns*

accompanied by different formations 22.22%, reinforcing again the idea of giving more information on the *who*, in Swales words (1990)

- (23) **Mr Cameron, whose delegation includes representatives of more than 100 companies**, added that the issue of whether bribes were paid to win a contract was a question for India and Westland's owner, the Italian company Finmeccanica.

(Report: *Cameron backs Agusta Westland despite Indian bribery scandal*. Sentence 3)

- (24) **Council Liberal Democrat group leader Cllr Richard Kemp, who has previously called into question the council's decision to get involved with private firms in the delivery of social housing**, said he hoped the scheme would succeed, adding: "I think this is a good idea, provided there's a solid basis for it."

(Report: *Pound land: Derelict houses in Liverpool to be sold for just £1*. Sentence 13)

### Semantic Realizations of Thematic Heads in English News Reports

News reports should strive to remain objective and should use neutral language while presenting a diversity of opinions, voices, and perspectives of the event, incident, or issue under discussion (White, 1998). These opinions, voices and perspectives are accomplished through the use of Sayers as the preferred experiential role of English news reports 34.32% as shown in the examples below:

- (25) **The bank announced** this morning that it will raise up to £2 billion pounds from Qatar Qatar Investment Authority (QIA), an existing shareholder, and its associated holdings and £300 million from Challenger, an investment vehicle of a member of Qatar's royal family.

(Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi*. Sentence 2)

- (26) **Mr Tilmant said** the bank had the trust of its customers and had not seen a large outflow of funds.

(Report: *Dutch Government to inject €10bn of fresh capital into ING savings bank*. Sentence 11)

*Actor* is the second role most frequently used 28.99%. These choices express the account of actions that the discourse participants carries out.

- (27) **The Government will buy** €10 billion of subordinated bonds in ING using a €20 billion fund set aside for fighting the current crisis.

(Report: *Dutch Government to inject €10bn of fresh capital into ING savings bank.* Sentence 3)

- (28) **Dominique Strauss-Kham, the French head of the International Monetary Fund, escaped** dismissal for a onenight stand with a subordinate today, but was denounced by board members for a “serious error of judgment.”

(Report: *Dominique StraussKahn, head of the IMF, escapes dismissal over affair.* Sentence 1)

It is important to highlight the role of *Verbiages* in this type of texts. Although this genre presents opinions and voices on a certain issue, the number of *verbiages* is not very relevant 4.14% in Head position. *Verbiages* are mainly found in PreHead position 19.64%, as it was pointed out in section 18.2.1.3. This means that the Head position prefers, with a statistically significant difference ( $P < 0.0001$ ), to convey the person who speaks (Sayer) rather than the message spoken (*Verbiage*). In fact, syntactic analysis revealed that *Proper Nouns* were found in a 27.19% of simple realizations and were involved in 48.15% of complex realizations.

#### 18.2.1.5 Rheme in English News Reports

Due to the nature of this type of texts, there are not many Rhematic expressions (see 18.2.1); only 3.15% of the total number of clause complexes in the news reports English corpus. Although this amount of non-thematic clauses is not relevant it is interesting to highlight that all of them are *Quotations*, as in the following examples:

- (29) **“We are currently working on plans to grow our retail business, our SME business, we are building a team for large corporates in the UK which we didn’t have before.”**

(Report: *Santander eyes a double bank buy with £2bn tag.* Sentence 6)

- (30) **“Worsening consumer finances are likely to further rein in spending on the high street and to complete this circle latest survey data showed that retail sector workers were the most downbeat about their job security and workplace activity in February.”**

(Report: *Lower income earners being squeezed hardest.* Sentence 7)

#### 18.2.2 Spanish News Reports

The total number of complex clauses found in the Spanish news report corpus is of 359 clauses. Of this amount 1.39% are rhematic clauses with no thematic construction.



The rest are declarative complex clauses (98.61%). Although my analysis focuses on the thematic clause complexes, attention will be paid to the appearance of these rhematic clauses in the texts if relevant.

Table 18.7 shows the distribution of Thematic elements in the Spanish news reports. As can be seen, the number of Interpersonal themes is almost non-existent, with only one occurrence in the whole corpus. Cohesive devices in the form of Textual themes are 14.69% and the number of PreHeads accounts for 34.75%. According to this data, realizations of Heads with no accompaniment of Outer thematic elements or PreHeads, is of 50.28%.

**Table 18.7** Thematic elements in Spanish news reports

Spanish	Thematic Field			
	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
	14.69%	0.28%	34.75%	100%

### 18.2.2.1 Textual Themes in Spanish News Reports

Textual themes are the cohesive markers that give logical connection to the text. The amount of these elements is not very high in Spanish news reports (14.69%). The low frequency of these elements is due, on the one hand, to the nature of the texts, and on the other, to their extent. The mean use of clause complexes per text in the Spanish news report is of 11.96 clauses per text. Table 18.8 shows the preference of Spanish news reports for different types of Textual Themes:

**Table 18.8** Types of TT in Spanish News Reports

Types of Textual Themes	
Binder	34.62%
Linker	63.46%
Correlative	1.92%
Total	100%

As it could be expected for this genre, there is a statistically-significant preference ( $P=0.0033$ ) for *Linkers* type of Textual Theme 63.46%, these *Linkers* which display the information in a lineal fashion, showing an exposition of facts and opinions. The use of *Binders* is very scarce with 34.62% of Textual elements. The preferred realizations of *Linkers* are addition linkers: *Además* 27.27% and *Asimismo* 15.15%.

- (31) **Además**, el director gerente del organismo, Dominique Strauss-Kahn, ha prometido una “importante ayuda financiera” para Hungría con el objetivo de estabilizar la economía de este país centroeuropeo.

*[In addition, the managing director of the agency, Dominique Strauss-Kahn, has promised a “major financial aid” for Hungary in order to stabilize the economy of this Central European country.]* (Report: *El FMI anuncia ayudas financieras para Ucrania y Hungría*. Sentence 2)

- (32) **Asimismo**, las ventas al por menor registraron una subida del 15,2% en diciembre, en contraste con el 14,9% de noviembre, lo que supuso un repunte final del 14,3% en 2012, aun así 2,8 puntos porcentuales por debajo de 2011, relató la citada fuente.

*[Also, retail sales recorded an increase of 15.2% in December in contrast to 14.9% in November, which was a rise of 14.3% in 2012, even 2.8 percentage points below than in 2011, said the same source.]*

(Report: *La economía china crece un 7,8% en 2012, su menor subida en 13 años*. Sentence 9)

*Binders* (34.62%) are realized in the same amount by the adversative connectors *Pero* (22.22%) and *No obstante* (22.22%):

- (33) **Pero** este parón es importante si se tiene en cuenta que Estados Unidos crecía a un ritmo anual del 2,8% en el segundo trimestre.

*[But this break is important if you consider that U.S. is growing at a annual rate of 2.8% in the second quarter.]*

(Report: *La crisis financiera lleva a Estados Unidos al borde de la recesión*. Sentence 6)

- (34) **No obstante**, se encuentra por encima del 7,5% de crecimiento fijado por el régimen comunista durante la Asamblea Popular de marzo del pasado año, por lo que los analistas creen que allana el camino para que se reponga durante este año.

*[However, it is above 7.5% of growth fixed by the communist regimen during the communist People’s Congress in March last year, so the analysts believe that paves the way to be reset during this year.]*

(Report: *La economía china crece un 7,8% en 2012, su menor subida en 13 años.* Sentence 3)

### 18.2.2.2 Interpersonal Themes in Spanish News Reports

As table 18.7 shows only one element was found in the data, accounting for only 0.28% of the total number of clause complexes in the corpus. The reason for this small amount of Interpersonal elements is due to the objective nature of these texts. This is shown in example (35):

(35) **Al menos** no lo hacen de manera limpia.

[At least they do not do it in a clear way.]

(Report: *Inspectores sin fronteras.* Sentence 8)

### 18.2.2.3 PreHeads in Spanish News Reports

The number of PreHeads in the Spanish news reports is of 34.75%. It is important to point out that Spanish allows for the use of *verbs* in first position in declarative clauses which. As explained in section 11.2.1, in the thematic analysis followed in this study, the lexical part of the verb is considered PreHead and its grammatical part Head. The following table shows the types of PreHead selected in the Spanish news reports:

**Table 18.9** Types of PreHeads

PreHead Types	%
PHCirc	73.17%
PHCCL	9.76%
PH-Finite	17.07%
Total	100%

The number of PreHeads realized by adverbial groups or prepositional phrases (*PH-Circ*) is 73.17%, and shows a statistically-significant difference ( $P < 0.0001$ ) between the number of PreHeads realized by circumstantial clauses (*PH-CCL* 9.76%). This difference points to simple syntactic realizations of these elements, indicating a clear and direct language typical of news reports whose purpose is to inform. The second most preferred selection is of *Verbs* (17.07%). Due to the language specific characteristic, the

number of Finite elements in PreHead position is very relevant even though all clause complexes are declaratives.

Looking in detail at the specific functions realized by *PH-Circ*, it is observed that the nature of this genre determines the need of a temporal location of facts (31.71%) expressed almost uniquely through the use of simple syntactic realizations of *Prepositional Phrases* (73.53%):

- (36) **En el segundo trimestre fiscal** la empresa ha logrado además una facturación récord por las ventas del Windows Vista y el Office 2007 y ha elevado sus previsiones para el conjunto del ejercicio fiscal ignorando las previsiones de una caída del consumo en EE UU y Europa, sus principales mercados.

*[In the second fiscal quarter, the company has also achieved a billing record for sales of Windows Vista and Office 2007 and has raised its forecast for the entire fiscal year ignoring forecasts a fall of consumption in the U.S. and Europe, its key markets.]*

(Report: *Los resultados de Microsoft brillan gracias al Windows Vista*. Sentence 4)

- (37) **En los últimos doce meses**, Unnim ha realizado dotaciones por importe de 474,2 millones, cinco veces más que en 2011, para sanear su cartera inmobiliaria, cubrir el deterioro de activos financieros y por los costes de reestructuración laborales, entre otros apartados.

*[In the last twelve months, Unnim made provisions amounting to 474.2 million, five times more than in 2011, to clean up his property portfolio, cover impairment of financial assets and restructuring costs of labour, between other things.]*

(Report: *Las pérdidas de Unnim Banc casi se duplican en 2012*. Sentence 4)

The second type of PreHead is the Circumstantial expression of *Evidence* defined as the element that relates to the source of information in verbal processes and is expressed by *as x says, or according to x* (Downing & Locke, 2006:158). In Spanish there are corresponding structures “*Para...*”, “*Según...*”. This element acquires special relevance in this genre as a means to convey relevant information on the topic treated.

- (38) **Para la OPEP**, los precios del petróleo han sufrido un “dramático derrumbe sin precedente en velocidad y magnitud”, por lo que se encuentra “preocupada”, ya que los niveles actuales podrían “poner muchos proyectos en peligro de suspensión o retraso”, lo que generaría “cortes en el suministro”.

*[For OPEC, oil prices have suffered a “dramatic collapse without precedent in speed and magnitude”, so it is “concerned” because current levels could “put many*

*projects in jeopardy of suspension or delay ”, which may generate “power cuts.”]*

(Report: *El recorte de producción de la OPEP hunde el petróleo.*  
Sentence 4)

- (39) **Según informaciones del diario griego “Kathimerini”**, el objetivo de esta moratoria es dar a Grecia un respiro para poder llevar a la práctica las reformas fiscales y estructurales ya comprometidas.

*[According to the Greek newspaper “Kathimerini”, the aim of this moratorium is to give Greece a breathe to be able to implement the fiscal and structural reforms already committed.]*

(Report: *La ‘Troika’ concederá a Grecia seis meses en los que no le reclamará más medidas.* Sentence 2)

The second type of PreHeads most frequently used are *Verbs* 17.07%. This factor is relevant from a linguistic point of view because Spanish allows for the use of verbs in first position in declarative clauses without an independent realization of the subject. It is also important to note that it is only the lexical part of the verb that is present in this Thematic element, leaving the grammatical part of the verb where the subject resides in the Head (see Lavid et al., 2010d) as explained in section 11.2.2.

The number of *Verbs* in first position as in example (40) is of 17.07%. But there is also a number of combined elements realized by Circumstantial Elements followed by *Verbs* 13.85% as in example (42) where the verb follows a circumstantial element (in bold letters).

- (40) **Insistió** en que hay que acabar con la incertidumbre para crear confianza en los mercados y en que los gobiernos deben aprobar programas a medio plazo para que las cuentas públicas estén saneadas, pero no subrayó la urgencia o la inmediatez de los recortes como en el pasado.

*[She insisted on the need to end the uncertainty to build confidence in the markets and that governments should adopt medium-term programs to clean public accounts, but she did not stress the urgency or immediacy of cuts as in the past.]*

(Report: *Lagarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída.* Sentence 6)

- (41) Además, **incide** en que este ajuste se está llevando a cabo con un ‘fuerte aumento’ de los costes de financiación, derivados en parte de las tensiones financieras.

*[In addition, he insists that this adjustments that are being carried out involve a ‘sharp increase’ in financing costs, derived in part from financial stress.]*

(Report: *La Caixa anticipa una ‘gradual recuperación’ de la situación económica.* Sentence 7)

- (42) A este respecto, **informó** de que la nueva previsión de crecimiento se establecerá a finales del mes de marzo por parte del Alto Consejo de las Finanzas Públicas y determinará lo que tendrá que hacer Francia en materia presupuestaria en 2014, según las recetas de la Comisión Europea.

[In this regard, he reported that the new growth forecast will be established in late March by the High Council of the Public Finance and determine what France will have to make in terms of budget in 2014, according to the prescriptions of the European Commission.]

(Report: *Hollande reconoce que Francia crecerá menos de lo previsto en 2013*. Sentence 3)

What is interesting to note is not only the position of these elements as such but also the lexical information they convey: *verbs* in first position, not preceded by any circumstantial element express a communicative aim (*afirmar, incidir, señalar, insistir, advertir, justificar, alabar, añadir, mencionar, recalcar*) in in a 71.42% of the cases, in detriment of verbs expressing action (28.57%) like *aplicar, funcionar, destruir, encontrar, etc.*... Also in the compound formation where *verbs* are accompanied by Circumstantial elements, as in example (43), there is a majority of these communicative verbs (47.05%), such as *advertir, recalcar, indicar, recomendar, etc.*..

- (43) **En concreto, subraya** que el ámbito donde se han realizado más progresos ha sido el exterior, ya que los países de la periferia lograron reducir su déficit por cuenta corriente al 1,7%, gracias en parte a una mejora de la competitividad de precios y sin que los países del centro hayan impulsado políticas expansivas de su demanda interna.

[In particular, stresses that the more area where more progresses have been made is in the outside, as the countries of the periphery reduced their account current deficit on 1.7%, thanks in part to an improved price competitiveness and without center countries expansionary policies have boosted its internal demand.]

(Report: *La Caixa anticipa una 'gradual recuperación' de la situación económica*. Sentence 4)

#### 18.2.2.4 Head in Spanish News Reports

The number of thematic clauses in the Spanish news reports is of a 98.61% of the total amount of clause complexes in the corpus. All of them, as it could be expected for this type of text are declarative sentences.

### Syntactic Realizations of Thematic Head in Spanish News Reports

The following table shows the preferred syntactic realizations of Heads in Spanish news reports:

**Table 18.10** Syntactic realization of Head

Syntactic forms	%
Np	50.85%
Np+App	28.25%
Verb	13.56%
Quote	2.26%
Np+Np	1.41%
Other	3.13%
Total	100%

The selection of simple syntactic structures over complex formations is statistically-significant ( $P < 0.0001$ ). *Noun Phrases* are used 50.85% whereas *Noun Phrases plus Appositions* appear 28.25%. This choice of simple forms reinforces the idea of simplicity and clearness of news reports. *Noun Phrases* patently identify the discourse participant as they are realized by *Proper Nouns* (36.67%) and *Determiner plus Nouns* (35.00%) as shown in the following examples.

Examples of *Proper Nouns* in Spanish news reports:

- (44) **La OPEP** explica que el 17 de diciembre celebrará en Argelia una reunión extraordinaria en la que se revisará la decisión adoptada hoy.

[OPEC explains that on December 17 an extraordinary meeting in Algeria will be held in which the decision taken today will be reviewed.]

(Report: *El recorte de producción de la OPEP hunde el petróleo.*  
Sentence 7)

- (45) **Japón** anuncia mayor control en la venta de acciones.

[Japan announces more control over the share sales.]

(Report: *El G7 reafirma su interés en que el sistema financiero sea “fuerte y estable”.* Sentence 4)

Examples of *Determiner plus Nouns* in Spanish news reports:

- (46) **La entidad** emitirá 3.000 millones de libras (3.822 millones de euros) en opciones sobre acciones, destinadas a esos inversores árabes, y otros 4.300 millones de libras (5.479 millones) en títulos convertibles de forma obligatoria.

[The bank will issue 3,000 billion pounds (3.822 million euros) in options on actions aimed at those Arab investors, and 4,300 million pounds (5.479 million) in compulsory convertible securities.]

(Report: “*Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown*”. Sentence 2)

- (47) **La operación** está valorada en 135 millones de euros y reportará a la entidad plusvalías cercanas a los 133 millones de euros.

[The operation is valued at 135 million euros and will report to the entity gains of close to 133 million euros.]

(Report: *Banco Popular gana 133 millones con la venta de su negocio de recobro*. Sentence 2)

*Noun Phrases plus Appositions* (28.25%) help adding extra information to the issue in question as in examples (48) and (49). The main realization of these elements are of the form *Noun Phrase plus Prepositional Phrases* in a 59.00% of cases.

- (48) De hecho, **la delegación del PP en el Parlamento Europeo** ha registrado sendas preguntas en la Eurocámara dirigidas a la Comisión y al Consejo de la UE, con el fin de que estas dos instituciones valoren si “se pueden mantener” los programas de cooperación al desarrollo con Bolivia así como el acuerdo preferencial en materia de aduanas si La Paz prosigue con sus “ataques” a empresas europeas.

[In fact, the delegation of the European Parliament PP registered questions addressed to the Commission and the Council of EU, in order to assess whether these two institutions can maintain development cooperation programs with Bolivia as well as the preferential agreement in customs matters if La Paz continues with “its attacks” to European companies.]

(Report: *Bolivia promete a España llegar a una “solución satisfactoria” para indemnizar a las empresas expropiadas*. Sentence 8)

- (49) **La prima de riesgo de la tercera economía de la zona euro** se ha disparado más de 40 puntos en la apertura de este martes, hasta los 335 puntos básicos, mientras la española ha llegado a tocar los 412 puntos, un vertiginoso ascenso desde los 361 puntos del cierre del lunes.

[The risk premium of the third largest economy in the eurozone has soared more than 40 points at the opening on Tuesday, up 335 basic points, while the Spanish one has reached 412 points, a dizzying ascent from the 361 points of late Monday.]



(Report: *La prima de riesgo española se dispara a 400 puntos por los resultados en Italia.* Sentence 2)

Verbs are the third preferred realization of Thematic Heads with a 13.56% of occurrences. In Thematic Heads the relevant part of the verb to be taken into account is the grammatical form of the verb where tense and person are concealed (for more on this, see section 18.2.3.4).

### Semantic Realizations of Thematic Head in Spanish News Reports

The results on the annotating Experiential roles in Spanish news reports show that there is a clear tendency towards the selection of *Actors* (statistically significant difference  $P < 0.0001$ ) oppositely to *Sayer* as the second most frequently experiential role used. *Actors* appear in 36.16% of cases (for the definitions of these elements see chapter 11.2.2.9). The use of this element reveals the aim of a mere account of facts, as shown in the examples below:

- (50) **Los precios del crudo han caído** más del 50% desde el récord de 147 dólares registrado en julio.

[*Oil prices have fallen over 50% from the record of 147 dollars in July.*]

(Report: *El recorte de producción de la OPEP hunde el petróleo.*  
Sentence 11)

- (51) **Microsoft ha logrado esquivar** la desaceleración del consumo que ha lastrado las economías de EE UU y Europa desde el pasado verano gracias a las ventas del Vista.

[*Microsoft has managed to avoid the slowdown in consumption that has hindered the economies of the U.S. and Europe since last summer thanks to the sales of Vista.*]

(Report: *Los resultados de Microsoft brillan gracias al Windows Vista.* Sentence 1)

*Sayers* have also a relevant presence as realizations of Experiential role (20.90%). They serve as a mean to quote direct sources for the sake of maintaining the objectivity of this type of text:

- (52) **Los abogados de Morgan Lewis afirman** que no hubo conductas reprobables más allá del “error personal” que el propio Strauss-Kahn, de 59 años, ha reconocido: no presionó a la mujer para obtener sexo ni intervino en su decisión de cambiar de trabajo ni favoreció una compensación mayor de la usual cuando abandonó el FMI en agosto.

[Morgan Lewis attorneys assembled there was reprehensible behavior beyond the “personal error” that Strauss-Kahn himself, 59 years, has recognized: he did not pressed the woman for sex or intervened in their decision to change work or favoured a greater compensation than usual when leaving the IMF in August.]

(Report: *Una noche sin consecuencias para StraussKahn*. Sentence 7)

- (53) **La OPEP explica** que el 17 de diciembre celebrará en Argelia una reunión extraordinaria en la que se revisará la decisión adoptada hoy.

[OPEC explains that on December 17 an extraordinary meeting in Algeria will be held in which the decision taken today will be reviewed.]

(Report: *El recorte de producción de la OPEP hunde el petróleo*. Sentence 7)

#### 18.2.2.5 Rheme in Spanish News Reports

As it has been pointed out before, the presence of Rhemes in this type of texts is not very frequent; in fact, in this corpus there are only 5 occurrences of these rhematic clauses. The appearances found are mostly *Noun Phrases* that seem to behave as titles, but due to the scarce number of these elements, it is too risky to state something of the kind. Here are some examples:

- (54) Retroactividad  
[Retroactivity]

(Report: *El Eurogrupo aprobará hoy 1.865 millones de euros para la banca española*. Sentence 6)

- (55) Legitimidad democrática  
[Democratic retroactivity]

(Report: *La Caixa anticipa una ‘gradual recuperación’ de la situación económica*. Sentence 5)

- (56) Gradual recuperación  
[Gradual recovery]

(Report: *La Caixa anticipa una ‘gradual recuperación’ de la situación económica*. Sentence 11)

- (57) Visita a España  
[Visit to Spain]

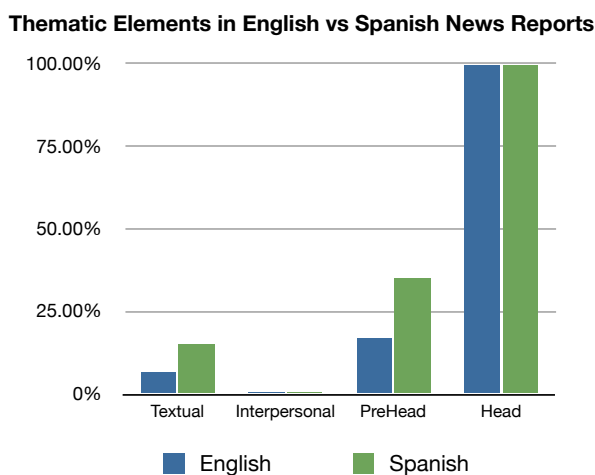
(Report: *Lagarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída*. Sentence 17)

### 18.2.3 Contrastive Analysis of News Reports in English vs. Spanish

This section presents the results of the analysis of news reports from a language contrastive perspective.

As it was shown in previous sections the total amount of complex clauses in English corpus is of 349 and in Spanish of 354, in both corpora rhematic sentences are almost nonexistent (3.15% in English and 1.39% in Spanish). As a common feature of these type of texts, all clause complexes are declarative sentences.

The following figure 18.2 shows the types of thematic elements found in the thematic field of English versus Spanish news reports. As it was pointed out in previous sections, the percentages of the different thematic elements are calculated over the total amount of the thematic clause complexes in each corpus. As such the number of Textual Themes found in thematic clauses in English news reports is of 6.21% and in Spanish of 14.96%. In the same manner the number of Interpersonal Themes is of 0.30% in English and 0.28% in Spanish. PreHeads are found in 16.57% of English clauses and in 34.75% in Spanish Corpus. Heads appear in all thematic clauses analyzed, as by definition all Thematic Clauses have Head if not they are not Thematic Clauses.



**Figure 18.1** Thematic elements in news reports (English vs. Spanish)

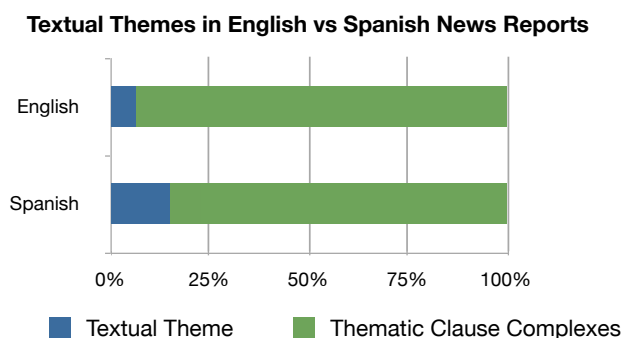
The following subsections will go in detail into each of these elements.

#### 18.2.3.1 Textual Themes: English vs. Spanish in News Reports

Textual themes in News reports are not very frequent since the textual organization relies in paragraphing rather than on the use of these elements. Even though the use

of Textual Themes in English and Spanish present interesting differences in the line of the findings in Moratón (2009).

There is a statistical significant difference ( $P=0.0002$ ) in the use of these elements in English (6.21%) versus Spanish (14.69%), where Spanish uses more Textual devices than English. This difference could be due to a disparity in the extension of the texts but, the average extension of these types of text is 11.63 clauses per text in English and 11.96 in Spanish, showing an irrelevant discrepancy in extension between them. The following figure shows in blue the amount of Textual Themes in English (first row in blue) and Spanish (second row in blue) the rest of the green column are the rest of thematic clauses with no textual elements in the news reports Corpus.



**Figure 18.2** Textual Themes in news reports (English vs. Spanish)

The answer may rely in the choices of this type of cohesive devices among these two languages:

As table 18.3 and 18.8 show in sections 18.2.1.1 and 18.8 for English and Spanish respectively, English choice for *Binders* (90.48%) over *Linkers* was significantly relevant. The same significative difference occurred in Spanish, but in this case the choice was for *Linkers* (63.46%) over *Binders*. This disparity in choice leads to a statistically significant difference ( $P<0.0001$ ) in the selection of *Binders* between these two languages, even though both select adversative connectors: the more formals *However* (38.10%) ie. (58) and *No obstante* (22.22%) ie.(59) and the less formals *But* (523.81%) ie.(60) and *Pero* (22.22%) ie. (61).

- (58) **However**, Hilco faces competition from other suitors for the UK business and the structure of any deal is likely to be complex, partly because HMV has many loss-making stores that a new owner is likely to jettison.

(Report: *Hilco UK emerges among favourites to give HMV a spin.*  
Sentence 7)

- (59) **No obstante**, Lagarde señaló que “queda mucho por hacer”, sobre todo en la reducción de las “barreras a la competitividad” y en la puesta en práctica de los planes de privatización.

*[However, Lagarde signalled “all there is left to be done”, especially in the reduction of the “barriers to competitiveness” and the start carrying out privatization plans]*

(Report: *La ‘Troika’ concederá a Grecia seis meses en los que no le reclamará más medidas.* Sentence 9)

- (60) **But** his reputation took a battering in the statement by the IMF board, which said the relationship “was regrettable and reflected a serious error of judgement on the part of the managing director.”

(Report: *Opec hawks want to cut oil production to keep up price.* Sentence 8)

- (61) **Pero** esto no es un problema para la Fed, porque en octubre empezaron a moderarse.

*[But this is not a problem for the Fed because in October they begun to moderate.]*

(Report: *La crisis financiera lleva a Estados Unidos al borde de la recession.* Sentence 32)

The choice of *Linkers* in both languages also presents a significant difference ( $P < 0.0001$ ). The formality in this type of connectors is mostly regarded in Spanish selection of *Además* (27.27%) and *Asimismo* (15.15%) i.e.: (64) and (65), while English uses the more formal *In addition* (4.76%) in the same amount as the less formal *And* (4.76%) i.e.: (62) and (63).

- (62) **In addition**, Qatar Holding will receive £66 million for having helped arrange the capital raising.

(Report: *Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi.* Sentence 8)

- (63) **And** this month Mr Dilnot warned politicians to avoid mixing up debt with the deficit after the Prime Minister claimed, erroneously, in a party political broadcast that the Coalition was “paying down Britain’s debts”.

(Report: *Stop ministers seeing official stats early - MPs.* Sentence 14)

- (64) **Además**, el Eurogrupo deberá elegir mañana a su nuevo presidente, en sustitución del luxemburgués Jean-Claude Juncker, un puesto para el que solo se ha presentado un candidato, el ministro de Finanzas holandés, Jeroen Dijsselboem, y que, según fuentes diplomáticas, tiene el camino bastante despejado.

*[In addition, the Eurogroup should choose tomorrow its new president, in substitution of Luxembourg Jean-Claude Juncker, a post for which only one candidate has*

presented, Dutch Finance Minister, Jeroen Dijsselboem and, according to Diplomatic sources he has a fairly clear path.]

(Report: *El Eurogrupo aprobará hoy 1.865 millones de euros para la banca española.* Sentence 8)

- (65) **Asimismo**, las ventas al por menor registraron una subida del 15,2% en diciembre, en contraste con el 14,9% de noviembre, lo que supuso un repunte final del 14,3 % en 2012, aun así 2,8 puntos porcentuales por debajo de 2011, relató la citada fuente.  
[Also, retail sales recorded an increase of 15.2% in December in contrast to 14.9% in November, marking a final recovery of 14.3% in 2012 and still 2.8 percentage points below 2011, said the aforementioned source.]

(Report: *La economía china crece un 7,8% en 2012, su menor subida en 13 años.* Sentence 9)

The differences in the use of Textual Themes in English and Spanish are very interesting. On the one hand, there is a larger number of cohesive elements used in Spanish than English not attributable to the extent of the texts. On the other hand, the preference for the use of one type of textual cohesion in English (*Binders*) and the opposite in Spanish (*Linkers*). Regarding these results, and to interpret these values in context it has to be taken into account that the number of textual elements is far more numerous in Spanish than in English. Considering this fact the use of *Binders* in both languages is very small so the key fact is the difference in the use of *Linkers* in Spanish far more than in English. In Spanish news reports there is a higher use of Linking cohesive devices than in English. To say that this is common to all genres and stating it as a language specific characteristic can only be confirmed after the analysis of the other two genres.

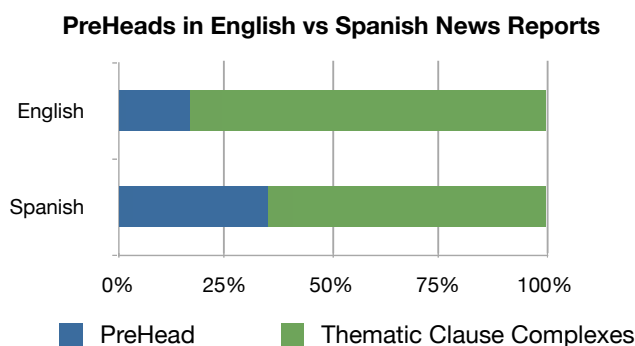
### 18.2.3.2 Interpersonal Themes: English vs. Spanish in News Reports

Due to the nature of news reports and the need of the author to remain objective the use of these elements in both languages is insignificant, having found only one occurrence in each language as shown in sections 18.2.1.2 and 18.2.2.2.

### 18.2.3.3 PreHead: English vs. Spanish in News Reports

There is a relevant difference in the selection of PreHead by each language, in fact it presents a statistically significant difference represented in a P value of less than 0.0001. English use of PreHeads accounts for 16.57% and in Spanish of 34.75%. The use of PreHeads is mainly devoted to the expression of circumstances, but it also includes the

lexical part of the verb when its grammatical part functions as thematic head (Lavid et al., 2010d). In English the presence of verbs in PreHead position occurs in directive clauses or yes/no questions. On the contrary, Spanish allows for the use of verbs in first position in declarative clauses. And this is a fact that will have to be taken into consideration in the analysis of the data.



**Figure 18.3** PreHeads in news reports (English vs. Spanish)

PH-Circ, the circumstantialials realized by adverbial groups and prepositional phrases, present a significant difference between English and Spanish ( $P=0.0499$ ). According to these data Spanish uses more syntactic simple structures (73.17%) to express circumstances than English in news reports (57.14%). This idea seems to be reinforced looking at the use of Circumstantial clauses (PH-CCL) by language, where English uses a 42.86% of these constructions, Spanish does so only in a 9.76% ( $P<0.001$ ). But in this case it is worth noticing that, as it was said in section 18.2.1.3, the use of circumstantial clauses in English in this type of texts it is not only devoted to expressions of circumstance but also to the use of quotation, and the amount of PH-CCL that work as quotations is of 45.83% almost half of them. In Spanish, on the contrary, there is only 0.81% quotes found in PreHead position. So let us look at the expression of circumstance as such, leaving aside the use of quotations in PreHead position, this topic will be treated at the end of this section. Thus the amount of “pure” Circumstances used in English and in Spanish is of 23.21% and 8.94% respectively.

PreHead Types in English vs Spanish News Reports

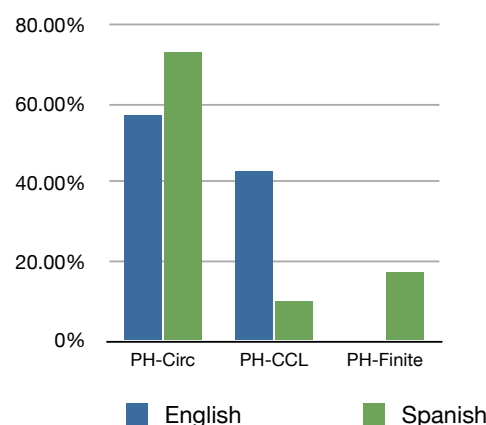


Figure 18.4 Types of PreHeads in news reports (English vs. Spanish)

Although the number of quotations has been separated from the expressions of circumstance in each language the difference between the use of dependent clauses functioning as circumstantial clauses is still significant ( $P=0.0182$ ) between English and Spanish. Now it can be said, having taken into account all particularities of the data analyzed, that English uses a more complex way of expressing circumstance in news reports than Spanish.

Bell (1991:155) states that the news story's overturning of temporal sequence is one of the most noteworthy differences with other newspaper genres. The importance of immediateness and temporal location of events is key for news reports; accordingly, both languages present time circumstantials as the most frequently used function of PreHead (English 33.93% and Spanish 31.71%). Both languages opt for simple Prepositional Phrase in its construction (English 47.37% and Spanish 73.53%) as the following examples show:

- (66) **On Friday**, Hilco was locked in talks with Deloitte which are expected to continue this week. (Report: *Hilco UK emerges among favourites to give HMV a spin*. Sentence 5)

- (67) **En los últimos doce meses**, Unnim ha realizado dotaciones por importe de 474,2 millones, cinco veces más que en 2011, para sanear su cartera inmobiliaria, cubrir el deterioro de activos financieros y por los costes de reestructuración laborales, entre otros apartados.

*[In the last twelve months, Unnim made provisions amounting to 474.2 million, five*



*times more than in 2011, to clean up his property portfolio, cover impairment of financial assets and restructuring costs of labor between other charts.]*

(Report: *Las pérdidas de Unnim Banc casi se duplican en 2012.*

Sentence 4)

It is in the use of verbs where one of the most relevant difference is found ( $P=0.0010$ ). This typological characteristic allows Spanish not to overtly mark the subject, encoding it in the grammatical part of the verb, but this subject is not part of the PreHead, but of the Thematic Head. The lexical part of the verb is the one that remains under this thematic element. And as it was explained in section 18.2.2.3, 75.41% verbs functioning as PreHeads in Spanish are of a communicative nature: *Afirmar, incidir, señalar, insistir, advertir, justificar, alabar, añadir, mencionar, recalcar...* The use of these verbs in first position may replace the function that Quotations do in English PreHeads. The selection of quotations as PreHead is also significantly different ( $P=0.0010$ ) between both languages (English 19.64% and Spanish 0.81%). English uses these direct quotations as a way to present “the points of view of various external sources” (White, 1998) typical of news reports. Spanish seems to use a different device to achieve the same goal: instead of quoting in PreHead position what it does is placing a communicative verb to “advance” the quotation that will come in the rheme as in:

- (68) En concreto, **subraya** que el ámbito donde se han realizado más progresos ha sido el exterior, ya que los países de la periferia lograron reducir su déficit por cuenta corriente al 1,7%, gracias en parte a una mejora de la competitividad de precios y sin que los países del centro hayan impulsado políticas expansivas de su demanda interna.

*[In particular, he stresses that the more general area where progresses have been made has been outside, as the countries of the periphery reduced their deficit on account current 1.7%, thanks in part to improved price competitiveness and without the core countries have driven expansionary demand policies internal.]*

(Report: *La Caixa anticipa una 'gradual recuperación' de la situación económica.* Sentence 4)

- (69) Asimismo, **añadió** que de aquí a 2016 pondrá en marcha un programa de ahorro de gastos estimado en 650 millones de euros que afecta al sector informático, inmobiliario y de inversiones, y sobre el que no precisó el impacto eventual que esos planes podrían tener sobre el empleo.

*[He also added that by 2016 he would launch a savings program estimated at 650 million euros affecting informatic rates, real estate and investment, and on the eventual need that those plans do not have impact on employment.]*

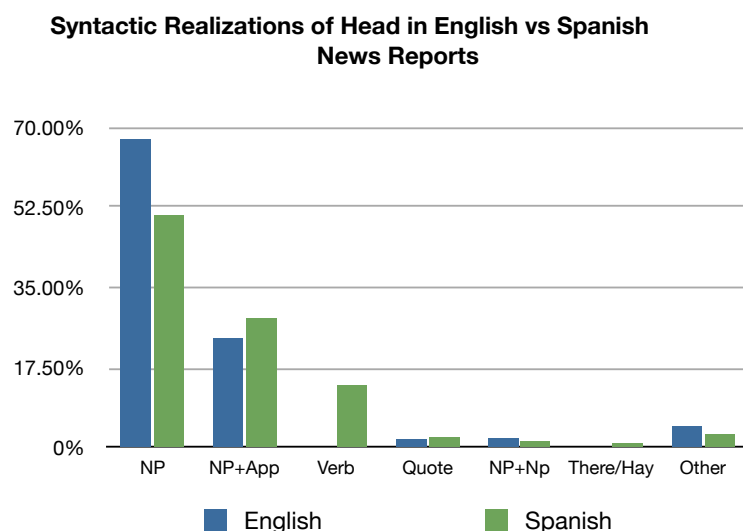
(Report: *El banco francés Crédit Agricole multiplica por cuatro sus pérdidas.* Sentence 7)

#### 18.2.3.4 Heads: English vs. Spanish in News Reports

The number of thematic clauses in the English Corpus is of 96.84% and in the Spanish news reports Corpus of a 98.61% of the total amount of clause complexes found in news reports. All of them, as it could be expected for this type of text are declarative sentences.

##### Syntactic Realizations

This section compares the syntactic preferences of the realizations of Head in English and Spanish news reports.



**Figure 18.5** Syntactic realization of Head in news reports (English vs. Spanish)

Although both languages select Noun Phrases as their preferred syntactic structure of Heads, there is a significant difference between its use in English and Spanish ( $P < 0.0001$ ). The main function of Noun Phrases in Heads in this genre is the identification of the discourse participant, English does it exclusively through the use of Noun Phrases (67.46%) (70) but this same purpose is done in Spanish by two different means through Noun Phrases (50.85%) and the grammatical part of the verb (13.56%) functioning as Head. Grammatical information of verbs appear in the 3rd person (13.56%) in most cases in singular (71) and in a few cases in its plural form (72).

- (70) **Wall Street** is scared about how much the US recession will erode American corporate profits over the next few years.

(Report: *Microsoft results offer hope to tech sector*. Sentence 4)

- (71) También alab-ó en particular a Portugal, que “ha hecho un trabajo muy bueno en la reducción del déficit” y ya ”ha conseguido dos tercios del camino”.

[*(She) Also praised in particular Portugal, who “has done a very good job in reducing the deficit” and it “has won two thirds of the way”.]*

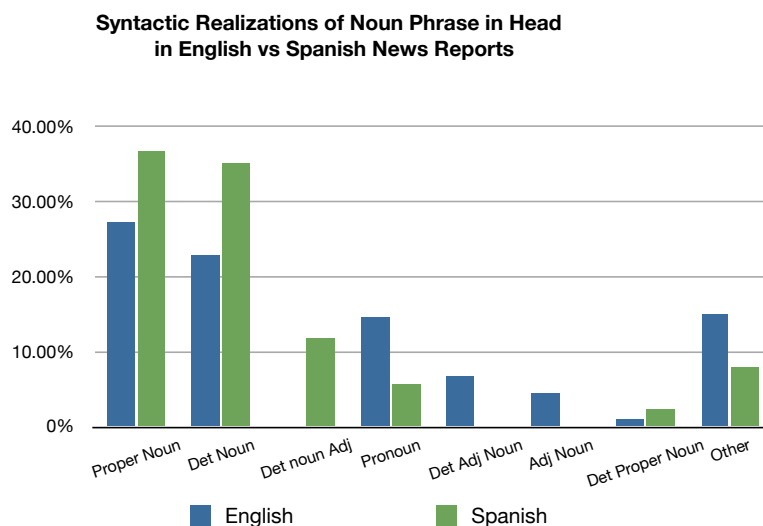
(Report: *Lagarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída*. Sentence 17)

- (72) “Estamos preocupados por la reciente excesiva volatilidad en el tipo de cambio del yen y sus posibles implicaciones adversas para la estabilidad económica y financiera”, **indican** los ministros de Finanzas y gobernadores de los bancos centrales del G7.

[*“We are concerned about the recent excessive volatility in the exchange rate of the yen and its possible adverse implications for economic and financial stability”, say the finance ministers and central bank governors of the G7.*]

(Report: *El G7 reafirma su interés en que el sistema financiero sea “fuerte y estable”*. Sentence 5)

Taking a look at the detailed realizations of these Noun Phrases, the aim of identifying the discourse participant is set clear in the use of Proper Nouns (English 27.19% and Spanish 36.67%) and Determiner plus Nouns (English 22.81% and Spanish 35%).



**Figure 18.6** Syntactic realization of NP Head in news reports (English vs. Spanish)

Due to the need of providing a high degree of truth and clarity, both languages use *Appositions* (English 23.96% and Spanish 28.95%) in the search of a better identification of the referent:

- (73) **The software giant, based in Seattle, Washington,** made net profits of 4.37billionforthethreemonthstoSeptember30,comparedwiththe4.29 billion it reported for the same period the year before.

(Report: *Microsoft results offer hope to tech sector*. Sentence 2)

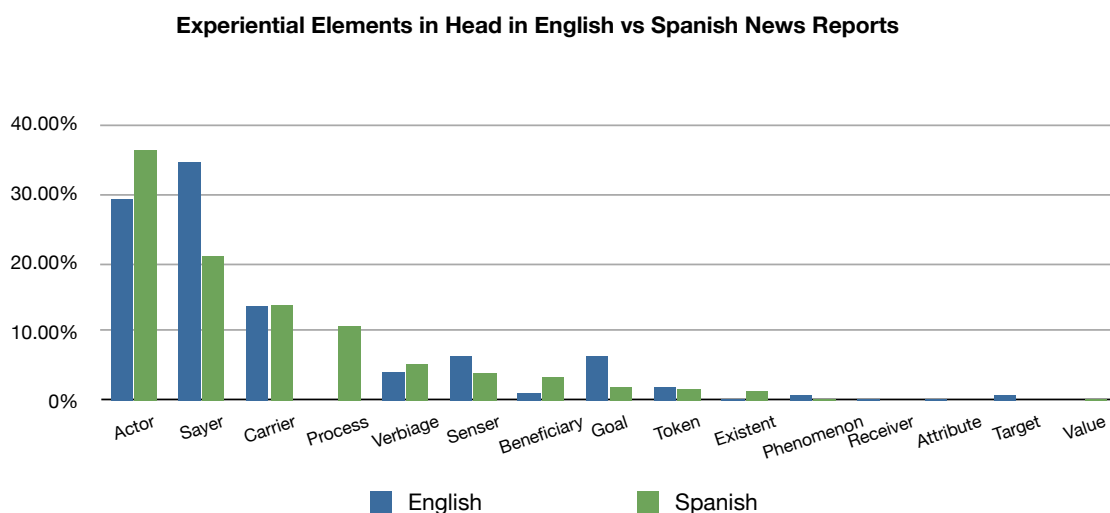
- (74) De este modo, **el Barclays, que tras la ampliación estará controlado en más de un 30% por esos inversores del Golfo,** rechaza de nuevo la ayuda ofrecida por el Gobierno británico para incrementar la liquidez de los bancos del país.

[Thus, the Barclays, which after its enlargement will be controlled by more than 30 % by Gulf investors, again rejected the help offered by the British Government to increase the liquidity of the country's banks.]

(Report: *Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brow*. Sentence 4)

### Semantic Realizations

This section studies from a contrastive point of view the experiential roles selected by English and Spanish in Head position.



**Figure 18.7** Types of semantic realizations in news reports (English vs. Spanish)

This figure 18.7 represents the preferred types of experiential roles in Head position

in both languages. As Francis (1990) found in his study, news reports present higher numbers of material and verbal processes and fewer relational ones. This is due to the informative nature of the text. Thus in the present study it has been found that *Actors* and *Sayer* are the preferred choices for both languages in news reports. The selection of *Actors* in these two languages does not make a significant difference (English 28.99% and Spanish 36.16%) although Spanish favours this role instead of the use of *Sayers* which is the one preferred in English. This selection of the experiential role of *Sayer* as Thematic Head in English (34.32%) is a language specific difference. Spanish conveys this role in a different form using processes or quotes in rhematic position, (see section 18.2.2.4 above)

- (75) **Morgan, Lewis & Bockius** said there was no evidence to support the claims.

(Report: *Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair*. Sentence 9)

- (76) **Los terroristas** han utilizado este dinero para afianzar sus posiciones en la región y cometer más atentados y secuestros.

[*Therrorists used that money to ensure their positions in the region and commit more attacks and kidnappings*]. (Report: *Terrorismo en el Sahel*. Sentence 4)

#### 18.2.3.5 Rheme: English vs Spanish in News Reports

The use of Rhemes in the news reports corpus is very low due to the nature of these texts, where the identification of the discourse participant is so relevant.

### 18.3 Thematic characterization of Editorials in English and Spanish

Van Dijk (1989) defines editorials as:

“(...) intended and structured to express and convey the opinions of the newspaper about recent news events. Since editorials are usually written about a single event or issue per day, they implicitly signal that the newspaper attributes particular social or political significance to such an event.”

Alonso (2007) defines their function in newspapers as to comment on recent events and accomplish the readers' agreement or to corroborate a pre-existing consensus through a series of crafted textual strategies. The strategies Alonso mentions will be addresses in the following sections.

The following table 18.11 shows the corpus of editorials in English and Spanish on which the study has taken place:

**Table 18.11** Corpus of editorials in English and Spanish.

Editorial Corpus	Texts	Clauses	Words
English	30	587	13448
Spanish	30	525	14744
Total	60	1112	28192

### 18.3.1 English Editorials

The total number of complex clauses found in the English editorials corpus is of 587 clauses. Of this amount 3.24% are rhematic clauses with no thematic construction. The rest of the complex clauses 97.76% are thematic complex clauses (3 directives, 5 questions and the rest declaratives). As in news reports I focus on Thematic clauses, therefore all percentages referred to the types of Thematic clauses are calculated over the total amount of these complex clauses. The following table shows the types of Thematic elements found in the editorials English corpus.

**Table 18.12** Thematic elements in English editorials

English	Thematic Field			
Types	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
	21.12%	11.27%	31.33%	100%

Thematic elements are divided into Inner thematic field and Outer thematic field. The Outer thematic field includes two elements: Textual Themes (TT) and Interpersonal Themes (IT). The Inner thematic field comprises PreHeads (PH) and Head (H). Textual Themes appear in a 21.12% clause complexes. Interpersonal themes are very frequent with a 11.27% which indicates a presence of the author in the texts. In the Inner thematic field PreHead accounts for 31.33%. As dealing with thematic clauses all of them have Head (100%). This means that an amount of 36.27% of the clause

complexes in English editorial corpus are realized by Heads as the only type of thematic element in the clause, what points to few simple thematic constructions of clause complexes in English editorials. The following sections analyze this data in detail.

### 18.3.1.1 Textual Themes in English Editorials

Textual themes are the cohesive devices used in the text. In editorial corpus there are 21.12% of these elements (table 18.12). As outlined in table 18.13 the number of *Binder* type of textual themes is of 66.67% whereas the number of *Linkers* is only of 33.33% these selection shows a statistical significant difference ( $P < 0.0001$ ) in the preference for *Binders*.

**Table 18.13** Types of TT in English Editorials

Types of Textual Themes	
Binder	66.67%
Linker	33.33%
Total	100%

As explained in section 11.1.1 binders are “the binding of elements of unequal status” (Halliday & Matthiessen, 2004:384). The number of these type of elements in editorials is quite high due to the large number of clauses per text (19.56). This is also due, as MacCabe points out (2007), than binder are “to provide a link between propositions based on how the writer views the expectedness of the upcoming proposition in light of the previous proposition”, that is another way in which the writer is present in this type or texts, providing an evaluative or explanatory point of view on the previous or upcoming proposition.

*Binders* are primarily realized by the adversative connector *But* 41.25% and *Yet* 16.25% followed by the conjunct *If* 10.00%.

The use of *But* introduces a counter argument in the unfolding of the events account, as in:

- (77) **But** what is staggering, and disappointing, is the Coalition’s failure to show moral support. (Editorial: *Our private schools deserve political praise*. Sentence 11.)

- (78) **But** we in Britain should tread carefully.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 22)

The same happens with the use of *yet*, as in the following examples:

- (79) **Yet** many of the economic difficulties he railed against are of Labour's creation – not least the country's indebtedness and the red tape strangling the small businesses he now holds in such high regard.

(Editorial: *Ed Miliband hasn't learnt any economic lessons.* Sentence 6.)

- (80) **Yet** according to Theresa May, the Home Secretary, this expression of Parliament's will is routinely being ignored by the courts.

(Editorial: *Courts and Parliament must work together.* Sentence 4.)

With the use of *if* the author introduces a hypothesis evaluating the facts presented in the text.

- (81) **If** Mr Miliband deems this worthy only of his sneers, he is making a serious mistake.

(Editorial: *Britain is ready to rock the boat in the EU.* Sentence 16.)

- (82) **If** Mr Cameron is serious about confronting the threat posed by Islamist terrorists, he should make sure our Armed Forces have the resources and manpower to deny them the safe havens they crave.

(Editorial: *Britain is ill-prepared for this deadly new world.* Sentence 19)

The use of *Linkers* relies in *And* 25% and *As* 17.50%.

*And* adds more information in a continuative way to the text and *As* has a more explicative sense, as examples (85) and (86) show:

- (83) **And** yesterday Mark Simmonds, Foreign Office minister for Africa, spoke of providing training and support for the Malian army.

(Editorial: *France moves in.* Sentence 3)

- (84) **And** it is interesting that supermarkets have had to take emergency action to provide their own assurances to their customers.

(Editorial: *The market needs watching.* Sentence 24)

- (85) **As** Con Coughlin pointed out on these pages yesterday, Britain has only eight of these aircraft, and France none; America has roughly 200.

(Editorial: *Britain is ill-prepared for this deadly new world.* Sentence 15)



- (86) **As** we report today, ministers have been forced by this furore to abandon their plan to opt out of new EU regulations requiring retailers to state exactly what parts of an animal are in their mincemeat.

(Editorial: *The market needs watching*. Sentence 16)

### 18.3.1.2 Interpersonal Themes in English Editorials

The use of Interpersonal themes in English editorial is of a 11.27% of clause complexes (table 18.12). This element reflect the presence of the author in the text, revealing a subjective point of view of the matter at stake.

The different devices used to express this subjectivity is performed through the following means (see chapter 11.1.2):

**Comment adjuncts** Expresses the speaker or writer's judgement or his/her attitude to the content of the message. (Halliday & Matthiessen, 2004:81)

**Modal adjuncts** In Halliday's words (2004:145): "What the modality system does is to construe the region of uncertainty that lies between 'yes' and 'no'".

**Engagement** In Hyland's words is how "writers acknowledge and connect to others, recognizing the presence of their readers..." (2005:176)

**Stance** Hyland's (2005:176) definition "This can be seen as an attitudinal dimension and includes features which refer to the ways writers present themselves and convey their judgements, opinions, and commitments."

The analysis revealed a statistical significant difference in the use of *Engagement* (50%) over the rest of interpersonal types ( $P=0.019$ ). As Hyland (2005) points out through Engagement strategies:

"The writer acknowledge and connect to others, recognizing the presence of the reader, focusing their attention, acknowledge their uncertainties and guiding them to interpretations"

These Engagement strategies reinforce the point of view of the editorial making opinions appear almost as facts rather than as mere points of view.

- (87) **Like it or not**, the borders will open – yet alarmingly, there is little sign that central or local government are taking the issue seriously.

(Editorial: *Heads in the sand over Eastern immigration*. Sentence 19)

- (88) **So it is a shame** we cannot say the same about our transport network or our schools.  
(Editorial: *This winters tale*. Sentence 3)

*Comment adjuncts* are the second preferred choice of Interpersonal themes in Editorials with a 26.56%. Following the same line than Engagement strategies, the author states clearly his/her position on the subject, as in the following examples:

- (89) **Sadly**, one suspects that if the banks were capable of being shamed out of this kind of antisocial behaviour, it would have happened long ago.  
(Editorial: *Blame the tax loopholes*. Sentence 5)
- (90) **Doubtless**, movement is difficult in rural areas after heavy snowfall; but is it really necessary for transport and schools in towns and cities to be so badly disrupted by what is a natural seasonal occurrence, albeit one experienced less regularly today than in the past?  
(Editorial: *This winter's tale*. Sentence 8)

*Stance elements* are “The way writers stamp their personal opinion onto their arguments” (Hyland, 2005). They appear in a 15.63%, this shows that the author does not very often present her/himself directly as someone who is stating his/her personal opinion onto their arguments. This may be due to the nature of the texts, the author present the opinion of the editorial line they represent but not their personal view.

- (91) In fact, **it seems that** the Republicans have already blinked in the most feared immediate confrontation, over a renewal of the US federal debt ceiling, without which the country might have defaulted.  
(Editorial: *Obama must seize the chance to make history*. Sentence 18)
- (92) **We thought that** consumers were entitled to assume that meat – and indeed all food – was subject to random spot checks to confirm that it was what it said it was.  
(Editorial: *The market needs watching*. Sentence 6)

*Modal Adjuncts*, along with *Stance elements* are the least frequent type of Interpersonal Themes used in the texts, with a presence of only 7.81% in English editorials.

- (93) **Perhaps** this was because the statement in which the agency justified its action will have made even less comfortable reading for the shadow chancellor than his opposite number.  
(Editorial: *Osborne needs to budget for economic growth*. Sentence 2)
- (94) **Instead**, it was a long and eminently sensible list of goals, enumerated by a man confident in his power and visibly liberated by the knowledge that he will never face

the voters again.

(Editorial: *A President at the very height of his power.* Sentence 3)

### 18.3.1.3 PreHead in English Editorials

PreHeads, as it was said in section 11.2.1, is a non-nuclear part of the Inner Thematic Field, an optional element realized by any circumstantial and/or finite element preceding the Thematic Head which do not exhaust the thematic potential of the clause.

In the English editorial Corpus there are a 31.33% of PreHeads. As shown in table 18.12.

Following the explanation given on section 11.2.1 the table below shows the the number of Circumstantial elements realized by Adverbial or Prepositional Groups (*PH-Circ* 55.06%), they present a statistically significant difference ( $P=0.0080$ ) in relation to the Circumstantial PreHeads realized by dependant clauses (*PH-CCL*) (40.45%) and Finite elements (*Ph-Finite* 4.49%). This shows the use of simple structures to express the circumstances that precede the Head.

**Table 18.14** Types of PreHeads

PreHead Types	%
PH-Circ	55.06%
PH-CCL	40.45%
PH-Finite	4.49%
Total	100%

The following examples illustrate the use of *Ph-Circ* in the texts:

- (95) **Today** the lobby group MigrationWatch plucked from the air another dubious figure, saying there will be around 50,000 new arrivals every year once restrictions are lifted. (Editorial: *Stop all the demonising of new EU migrants.* Sentence 3)

- (96) **In the business world**, the construction of an opulent new headquarters is often seen as a warning sign that an organisation has lost its soul.

(Editorial: *We all pay the cost of the BBC's failures.* Sentence 20)

Complex syntactic structures are provided in *PH-CCL* examples:

- (97) **As David Cameron spelt out in his Commons statement today**, the complex is far out in the desert, 18 hours by road from Algiers, and it had been overrun in what appears to have been a large-scale and carefully planned operation by Mokhtar Belmokhtar, an Algerian who may have ties to al-Qaeda.

(Editorial: *David Cameron's mixed signals on terrorism*. Sentence 2)

- (98) **When America's first black president took the oath of office on that frigid, unforgettable 20 January 2009 in Washington**, the mood was of exhilaration – but amid the deepest economic slump in three-quarters of a century and two unfinished wars, the prospects by any rational measure were sobering.

(Editorial: *Obama must seize the chance to make history*. Sentence 2)

The use of other types of clause complexes than Declaratives in editorials is reflected in the use of Finite expressions in PreHead position 4.49% (3 directives and 5 questions). Although the number of directives and questions are not relevant they do indicate the use of less formal language in this type of text, where the writer is freer to express his/her opinion as in the directive example below:

- (99) **Chalk up** another victory for the PR departments of the global investment banks.

(Editorial: *Blame the tax loopholes*. Sentence 1)

These PreHeads express different functions. It is interesting to highlight that there are two elements more frequently used, those expressing *Time* (33.71%) and *Contingency* (30.90%), with a statistical significant difference ( $P=0.0028$ ) over the rest of circumstantial expressions such as *Manner* (16.85%) or *Evidence* (5.06%), among others with less relevant presence in the texts.

Setting the action in a time frame:

- (100) **Over the next year or two**, the eurozone countries will almost certainly forge a closer relationship that will make a renegotiation of the UK's position inevitable.

(Editorial: *Our interests come first, not those of America*. Sentence 12)

- (101) **While many of the new arrivals contributed to the economy**, they also put intolerable stresses on unprepared public services and stoked such resentment that immigration became one of the most pressing (and most toxic) issues in politics.

(Editorial: *Heads in the sand over Eastern immigration*. Sentence 4)

*Contingency* is defined by Angela Downing (2006:156) as the element that “covers such meanings as cause, purpose, reason, concession and behalf”. These devices are expressed through syntactic complex structures (*PH-CCL*) as in:

- (102) **And although François Hollande is open to charges of precipitousness,** he was right that there was no time to waste.

(Editorial: *Mali's Islamists are too dangerous to be ignored.* Sentence 14)

- (103) **Even if Lance Armstrong had come across as the most contrite of penitents – which he did not** – an interview with Oprah Winfrey on American TV is not sufficient to allow the cyclist to draw a line under his disgrace and “move on”.

(Editorial: *A confession is not enough.* Sentence 1)

#### 18.3.1.4 Head in English Editorials

All thematic sentences have thematic Head whereas Rhematic sentences, which will be seen in detail in the next section, do not have Theme, therefore, Thematic Head. In the English editorial Corpus thematic clauses are a 97.76% of the total clause complexes in the corpus. As explained in section 18.3.1.

The novelty in this corpus is the appearance of Thematic Equatives (0.88%) and Predicated Themes (0.70%) (see chapter 11.2.2) more typical of informal languages.

Example of Thematic Equative:

- (104) **What Britain needs** is a robust and unambiguous tax system that requires corporations to pay their share, without the need for negotiations with HMRC, or moral urgings from politicians, or gestures of largesse by pressurised managements.

(Editorial: *Blame the tax loopholes.* Sentence 9)

Example of Predicated Theme:

- (105) **It is Mr Paterson, above all, who** has questions to answer.

(Editorial: *The market needs watching.* Sentence 27)

#### Syntactic realizations

Heads are most frequently realized by *Noun Phrases* (62.85%) followed, with a statistical significance difference, by *Noun Phrases plus Appositions* (22.36%). *Existential There* (5.11%) are the third most preferred syntactic form preferred by Heads, with a statistical significant difference from *Noun Phrases followed by Appositions*. The following table shows in detail the realizations of these *Noun Phrases*. The first thing that can be noticed when looking at the table are the wide range of different varieties of *Noun Phrases* realizations.

**Table 18.15** Syntactic realization of Head

Syntactic forms	%
Np	62.85%
Np+App	22.36%
There	5.11%
Clause	2.46%
Interrogative pronouns	2.11%
Other	5.11%
Total	100%

More in detail the use of **Noun Phrases** are mainly realized by *Pronouns* (35.85%) over the rest of used elements. This point to a frequent use of anaphora, as a references to an already named entity, as in:

- (106) Instead, **he** cited a British offer of technical and intelligence support, including hostage negotiation and rescue teams, which – he said with transparent regret – had been declined.

(Editorial: *David Cameron's mixed signals on terrorism*. Sentence 13)

- (107) **They** are generally young, and have more children, which is good news, given the ageing population.

(Editorial: *Stop all the demonising of new EU migrant*. Sentence 23)

The use of *Proper nouns* (17.09%) is the second most preferred choice in realization of Noun Phrases identifying the subject in the text.

- (108) **Sir David** points out that some 40 per cent of all patients in the average general hospital now have dementia.

(Editorial: *A dilemma that wont go away*. Sentence 4)

- (109) For the same reason, **David Cameron** is right to have extended logistical and other assistance to Paris.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care*. Sentence 4)

The third preferred structure is the one realized by *Determiner plus Nouns* (13.45%)

- (110) **The result** has been terrible suffering.

(Editorial: *Mali's Islamists are too dangerous to be ignored.*  
Sentence 7)

- (111) **The crisis** is also yet another reminder that Britain's influence with allies, and its own national security, require strong military capabilities of the sort that have been drastically shrunk over the past two years.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 21)

**Noun Phrases plus Appositions** are, as shown in table 18.15, the second preferred realization of Heads. More in detail these *Appositions* are *prepositional phrases* (50.39%) a syntactic form that adds more information on the *Proper nouns* or on the *Noun Phrase* that identifies the subject of the text.

- (112) **Military involvement in foreign conflicts** can never be undertaken lightly.

(Editorial: *Mali's Islamists are too dangerous to be ignored.*  
Sentence 1)

- (113) **A rebel victory at the town of Konna** opened the way to Mopti, Mali's second city, and thus to the whole of the south.

(Editorial: *Mali's Islamists are too dangerous to be ignored.*  
Sentence 12)

It is interesting to note the appearance of *Existential* “there” as a device to present facts as shared knowledge or commonly accepted facts (5.11%).

- (114) **There** are many reasons to suppose they are right.

(Editorial: *Stop all the demonising of new EU migrants.* Sentence 12)

- (115) **There** is also some evidence that those Libyan terrorists had been in contact with al-Qaeda in the Islamic Maghreb (AQIM), a group that forms part of the Islamist coalition in Mali and that was singled out in the Prime Minister's address to the House of Commons on Friday.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 17)

### Semantic realizations

The use of *Carriers* (32.39%) and *Actors* (31.69%) are the preferred choices of semantic realizations of Head in English editorials over the rest of experiential roles.

The difference of these two semantic roles (*Actor* and *Carrier*) with the rest of the roles selected present a statistically significant difference.

The use of *Carriers* in relational processes seems to be a strategy to give an impression of objective presentation of information. This feeling of objectivity is reinforced with the use of other devices in the Interpersonal Themes as *Engagement* or *Comment Adjuncts* (see section 18.3.1.2).

*Carriers* are defined by Halliday in *SFG* (2004:242) in relational processes as “(...) typically some immobile physical feature, whereas the Actor of a ‘material’ clause of motion is typically an animate being or a mobile entity”.

Here there are some examples:

- (116) **The official view, shared by successive governments, has long been** that the UK does not negotiate and does not give in to terrorists.

(Editorial: *David Camerons mixed signals on terrorism*. Sentence 10)

- (117) But between now and mid-2014, **Mr Obama is** in a remarkably favourable position.

(Editorial: *Obama must seize the chance to make history*. Sentence 12)

*Actors* identify the one that accomplish the action. (Halliday & Matthiessen, 2004)

- (118) **Democrats won** the popular Congressional vote by some half a million nationally, but gerrymandering is a wonderful thing.

(Editorial: *Obama must seize the chance to make history*. Sentence 17)

- (119) **The International Criminal Court is investigating** allegations of war crimes including rapes and executions.

(Editorial: *Mali's Islamists are too dangerous to be ignored*. Sentence 9)

It seems that the main aim of these two roles is to present information or opinions in form of actual information and facts. This idea conflates with the definition given by Van Dijk (1989) and Alonso Belmonte (2007) on Editorials (see section 18.3)

#### 18.3.1.5 Rheme in English Editorials

The number of rhematic sentences found in the editorial corpus is of only 19 (3.24%) of the total number of complex clauses analyzed. These rhematic sentences correspond to questions and answers to rhetorical questions posed in the previous sentences, as in:

- (120) To make people responsible for their actions?

(Editorial: *We all pay the cost of the BBC's failures*. Sentence 13.)



(121) No matter.

(Editorial: *Governments should not own banks*. Sentence 15.)

### 18.3.2 Spanish Editorials

The total number of complex clauses found in the Spanish editorials corpus is of 525 clauses. Of this amount 4.95% are rhematic clauses with no thematic construction. The rest of the complex clauses 95.05% are thematic complex clauses. As in news reports the focus will be on Thematic clauses, therefore all percentages referred to the types of Thematic clauses are calculated on the total amount of these complex clauses. The following table shows the types of Thematic elements found in the editorials Spanish corpus.

**Table 18.16** Thematic Elements in Spanish editorials

English		Thematic Field		
Types	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
	18.24%	7.62%	38.88%	100%

Thematic elements are divided into Inner thematic field and Outer thematic field. The Outer thematic field includes two elements: Textual Themes (TT) and Interpersonal Themes (IT). The Inner thematic field comprises PreHeads (PH) and Head (H). Textual Themes appear in a 18.24% clause complexes. Interpersonal themes are very frequent with a 7.62% which indicates a presence of the author in the texts. In the Inner thematic field PreHead accounts for 38.88% . As dealing with thematic clauses all of them have Head (100%). This means that an amount of 35.27% of the clause complexes in Spanish editorial corpus are realized by Heads on its own, what points to few simple thematic constructions of clause complexes in English editorials. The following sections analyze this data in detail.

#### 18.3.2.1 Textual Themes in Spanish Editorials

The number of Textual themes in Spanish editorials is of 18.24% and are separated in three different types, as the following table shows:

**Table 18.17** Types of TT in Spanish editorials

Types of Textual Themes	
Binder	69.23%
Linker	28.57%
Correlative	2.20%
Total	100%

There is a statistically significant difference between the use of *Binders* and *Linkers* in Spanish editorials. *Binders* are used 69.23%, far more numerous than *Linkers* which are used 28.57%. This may indicate the introduction of counter arguments in the development of the text.

*Binders* use the adversative *Pero* [But] in 42.86% of cases, presenting a statistically significant difference with the rest of the realizations found.

- (122) **Pero** es difícil admitir que indagar sobre la existencia y naturaleza de la misión que parece haber condicionado la conducta del principal acusado sea tarea ajena al tribunal, bien directamente, bien por deducción de testimonio.

*[But it is difficult to admit that the investigation about the existence and nature of the mission that seems to have influenced the conduct of the principal accused is not a court task, either directly or by deduction of testimony.]*

(Editorial: *Yak-42: justicia a medias*. Sentence 14)

- (123) **Pero** no está claro que PP y PSOE tengan las ideas muy claras sobre la estructura y funciones de los organismos independientes.

*[But it is not clear that PP and PSOE have very clear ideas about the structure and functions of the independent bodies.]*

(Editorial: *Europa vigila*. Sentence 11)

The use of *Si* [If] in a 17.46% introduces hypothesis on the part of the author that question the account of facts of the article.

- (124) **Si** hasta ahora los presidentes de Estados Unidos debían lidiar con el conflicto palestino como principal y casi exclusivo escenario político de la región, ahora el cruce de líneas de alta tensión es más complejo debido a los errores de Bush y del anterior primer ministro israelí, Olmert.

*[If until now the U.S. president should deal with Palestinian conflict as main and*

*almost exclusive political scenario of the region, now crossing high voltage lines is more complex because of the mistakes of Bush and former israel Prime Minister Olmert.]* (Editorial: *Proyectos divergentes*. Sentence 3)

- (125) **Si** hay información política significativa, ¿por qué se oculta?  
*[If there is significant political information, why is it hidden?]*  
 (Editorial: *Escuchas y fallos*. Sentence 10)

*Linkers* on the other hand are used in only 28.57% of cases, as shown in table 18.17. The continuity in the account of the events is given by the use of *Y* [and] with a 69.23%:

- (126) **Y** sigue también sin ofrecer una explicación que reste verosimilitud a la de que trajes y regalos pagaban un trato de favor en adjudicaciones.  
*[And he is still not offering an explanation to rest authenticity to the one about costumes and gifts that paid preferential treatment in tender offers.]*  
 (Editorial: *Sigue imputado*. Sentence 11)
- (127) **Y** es que los casos de corrupción no afectan sólo a los partidos políticos.  
*[And cases of corruption do not only affect political parties.]*  
 (Editorial: *El listón de los imputados*. Sentence 8)

### 18.3.2.2 Interpersonal Themes in Spanish Editorials

The use of Interpersonal themes in Spanish editorial is of a 7.62% of clause complexes (Table 18.16). This element reflect the presence of the author in the text, revealing a subjective point of view of the matter at stake.

The analysis of these elements (the different devices to express subjectivity are explained in chapter 11.1.2) revealed the use of *Engagement* (68.42%) with a statistical significant difference over the other types of Interpersonal Themes. This use points to a strong perception of the reader from the author's perspective who recognizes and addresses to him/her the message making the reader a relevant participant in the development of the argument exposed, as in:

- (128) **Se la conoce como** ley de Acompañamiento para los Presupuestos del 2013 de la Comunidad de Madrid, fue aprobada el jueves por el pleno de la Asamblea e incluye elementos dispares, como la externalización de servicios sanitarios, la modificación de la ley del Juego para adecuarla a las exigencias de Eurovegas o la implantación del euro por receta.  
*[It is known as the Accompaniment Law for the 2013 budget Community of Madrid, it was approved Thursday by the full Assembly and includes different elements such*

*as outsourcing of health services, the modified Law Game to suit the requirements of Eurovegas or the implantation one euro per prescription.]*

(Editorial: *La sanidad pública*. Sentence 1)

- (129) **Es hora de** que, para lograr un mayor respeto de los derechos humanos en ese territorio, el contingente de la ONU (Minurso) allí desplegado se encargue de supervisar su aplicación, como lo hacen otras muchas fuerzas de paz.

*[It is time to achieve a greater respect for human rights in that territory, being the UN contingent (MINURSO) deployed there responsible for monitoring its implementation, as do many other peacekeepers.]*

(Editorial: *Juicio militar en Rabat*. Sentence 18)

In this aim of involving the reader in the text is helped by the use of *comments*, chosen as the second most preferred type of interpersonal theme after the use of *Engagement* with a 15.79%.

- (130) **Realmente**, este hecho supone un sacrificio que no le gusta a nadie, pero que han de aceptar buena parte de los trabajadores del país para poder mantener el empleo, en una Catalunya en la que hay más de 700.000 parados.

*[Actually, this fact is a sacrifice that does not like anyone, but should have to be accepted by most part of the workers to maintain their jobs, in Catalunya where there are more than 700.000 unemployed.]*

(Editorial: *Grave error en Nissan*. Sentence 3)

- (131) **En efecto**, dicha norma contempla la cesión a operadores privados de la gestión de seis hospitales y 27 centros de salud, que atienden a un total de 1,5 millones de personas.

*[In fact, this rule provides for the assignment to private operators of the management of six hospitals and 27 health centers, serving a total of 1.5 million people.]*

(Editorial: *La sanidad pública*. Sentence 18)

### 18.3.2.3 PreHead in Spanish Editorials

The number of PreHead in Spanish editorial Corpus is of 38.88%, as it was pointed out in chapter 11.2.1 serve as the expression of non-compulsory circumstances in the Thematic Field.

The following table lists the types of PreHead found in the Corpus:

**Table 18.18** Types of PreHeads

PreHead Types	%
PH-Circ	50.52%
PH-CCL	31.41%
PH-Finite	19.07%
Total	100%

Thus, the number of Circumstantial elements realized by Adverbial or Prepositional Phrases (*PH-Circ*) is of 50.52%, they present a statistically significant difference in relation to the more complex realizations of *PH-CCL* with only 30.41% of occurrences. This indicates simple syntactic structures in the expression of the circumstances proceeding the Head. These expressions are referred ( $P < 0.0001$ ) to the location in *time* of the event (31.96%) followed by expression of *contingency* (21.13%), these help the author in the development of the story.

- (132) **Ayer** mismo el líder derechista israelí aseguraba que había acordado en Washington la necesidad de conversaciones de paz con Siria y con los palestinos.

[Yesterday the Israeli right-wing leader claimed that Washington had agreed in the need for peace talks with Syria and the Palestinians.]

(Editorial: *Proyectos divergentes*. Sentence 11)

- (133) **Hoy**, lo fácil sería decir que sí.

[Today, the easy thing would be to say yes.]

(Editorial: *El listón de los imputados*. Sentence 38)

#### 18.3.2.4 Head in Spanish Editorials

All thematic sentences have thematic Head whereas Rhematic sentences, which will be seen in detail in the next section, do not have Theme, therefore, Thematic Head. In the English editorial Corpus thematic clauses are a 95.05% of the total clause complexes in the corpus. As explained in section 18.3.2.

- (134) **Lo que requiere una investigación y la correspondiente depuración de responsabilidades** es la colaboración que la empresa del expistolero presta a la lucha contra el crimen con acceso a datos confidenciales.

[What needs an investigation and the corresponding purification of responsibilities

*is the collaboration that the company of the ex-gunman's pays to the fight against crime granting them access to confidential data.]*

(Editorial: *El caso "Hellín"*. Sentence 4)

- (135) **Porque fue también la prensa la que** hace 20 años abrió los ojos a la policía dando con el paradero de Hellín en Paraguay, donde se había refugiado tras huir de la cárcel.

*[Because it was the press who 20 years ago opened the eyes of the police giving the whereabouts of Hellin in Paraguay, where he had taken refuge after fleeing from jail.]*

(Editorial: *El caso "Hellín"*. Sentence 8)

### Syntactic realizations

The following table shows the preferred syntactic realizations of Head:

**Table 18.19** Syntactic realization of Head

Syntatic forms	%
Np	46.29%
Np+App	32.46%
Verb	13.03%
Np+Np	2.40%
Hay	0.20%
Other	9.62%
Total	100%

The most frequent syntactic structure with a statistically significant difference over the rest of the realizations is that of *Noun Phrases* (46.29%) To talk about *Noun Phrases* is really broad, so to pin this information down, let's see the detailed realizations of these elements:

Both *Determiners plus Nouns* (35.50%) along with *Proper Nouns* (29.44%) are the two most frequent realizations of *Noun Phrases*.

- (136) **Dichos comicios** podrían, por tanto, saldarse con una nueva victoria para las huestes de los Hermanos Musulmanes.

*[These elections could therefore be paid off with another victory for forces of the Muslim Brotherhood.]*

(Editorial: *Egipto, de la paradoja a la gobernabilidad*. Sentence 20)

- (137) **Algunos analistas** apuntan que este declive físico se ha podido ver agravado por las intrigas sucesorias vaticanas, que afloraron con el denominado caso Vatileaks, que resultan habituales cuando el Pontífice alcanza una edad avanzada, y son de difícil gestión, incluso en plenas facultades.

*[Some analysts suggest that this physical decline could have been aggravated by the internal Vatican intrigues of succession, which prayed with the so-called Vatileaks case that are usual when the Pontiff reaches an advanced age, and are of difficult management, even in full power.]*

(Editorial: *El papado de Benedicto XVI*. Sentence 8)

#### Examples of Proper Nouns:

- (138) En total, **Benedicto XVI** habrá permanecido en su puesto poco menos de ocho años.

*[All in all, Benedict XVI would have remained in his post only eight years.]*

(Editorial: *El papado de Benedicto XVI*. Sentence 5)

- (139) **Rajoy** no viaja a Berlín en las mejores condiciones posibles, pese a las previsiones de una ligera mejora de la economía española durante el segundo semestre del 2013.

*[Rajoy is not travelling to Berlin in his best, despite expectations of a slight improvement in the Spanish economy during the second half of 2013.]*

(Editorial: *Cita en Berlín*. Sentence 4)

Acclarations on the identification of the subject are done by the use of *Appositions* following Noun Phrases, these are most frequently realized by *prepositional phrases* with a statistical significant difference over the rest (56.17%) of the elements.

- (140) **El tijeretazo al presupuesto de Van Rompuy** se da principalmente a partidas como infraestructuras de transportes (que puede afectar al corredor mediterráneo), energía y telecomunicaciones.

*[The budget snip of Van Rompuy is principally given to transport infrastructures (that may affect the mediterranean corridor) energy and telecommunications.]*

(Editorial: *España gana con el presupuesto de la UE*. Sentence 9)

- (141) Excepto en dos casos, **las condenas por “violencia sobre agentes del orden con resultado de muerte”** han resultado severas: 8 a cadena perpetua, 4 a 30 años y 10 a penas de entre 20 y 25 años.

*[Except in two cases, convictions for violence against law enforcement resulting in*

death resulted severe: 8 to life imprisonment, 4 to 30 years and 10 to prison terms of between 20 and 25.] (Editorial: *Juicio militar en Rabat*. Sentence 7)

### Semantic realizations

The experiential role most frequently used with a statistically significant difference over the rest is *Actor* (35.27%), the second most used is that of *Carrier* (26.25%). This is due to the nature of the text that mingles the account of facts through *Actor* and the presentation of opinions through the use of *Carriers*.

Examples of *Actor*:

- (142) **Pero muchos han cometido** el error de considerar sus injustificables gastos consentidos como un complemento salarial de libre disposición.

[But many have made the mistake of considering their unjustifiable expenses consented as a wage supplement of free disposal.]

(Editorial: *Sonrojo en Westminster*. Sentence 9)

- (143) **El régimen de Kim Jong-il también descartó** entonces reincorporarse a las interminables negociaciones a seis bandas para que abandone su programa bélico atómico, y que hasta ahora sólo han conducido a una ristra de promesas incumplidas.

[The regime of Kim Jong-il also rejected to rejoin the endless six-party talks to abandon its warlike atomic program, and so far have only led to a string of broken promises.]

(Editorial: *El reto de Pyongyang*. Sentence 6)

Examples of *Carriers*:

- (144) **El PP tiene** cada día más difícil mantener esa actitud, contraria a lo establecido en su propio código de conducta y que supone aplicar dos criterios diferentes en Madrid y en Valencia.

[The PP has increasing difficulties in maintaining their attitude, contrary to the provisions contained in their own code of conduct, that involve applying two different criteria one in Madrid and the other in Valencia.]

(Editorial: *Sigue imputado*. Sentence 9)

- (145) **Gobierno y PP necesitan** fortalecer la seriedad de sus actuaciones y no contribuir más a la creciente erosión de la confianza de los ciudadanos.

[The Government and PP need to strengthen the seriousness of their actions and not to contribute more to the increasing erosion of the citizens confidence.]

(Editorial: *Medidas improvisadas*. Sentence 16)



### 18.3.2.5 Rheme in Spanish Editorials

The number of rhematic clause in editorial corpus is very scarce, only a 4.95% of the total number of complex clauses. The following examples correspond to Rhematic expressions in Spanish, unlike in English these are declarative sentences.

(146) Más que las de la oposición, por cierto.

[*More than the opposite party, by the way.*]

(Editorial: *Egipto, de la paradoja a la gobernabilidad*. Sentence 19)

(147) Una perspectiva dramática para Italia y la Unión Europea.

[*A dramatic perspective for Italy and the European Union.*]

(Editorial: *Berlusconi todavía*. Sentence 7)

### 18.3.3 Contrastive Analysis of Editorials in English vs. Spanish

This section presents the results of the analysis of editorials from a language contrastive perspective.

As shown in previous sections (18.3.1 and 18.3.2) the total amount of complex clauses in English corpus is of 97.76% (568) and in Spanish of 95.05% (499), in both corpora rhematic sentences are almost nonexistent (3.24% in English and 4.95% in Spanish). Not all clause complexes in this genre are declarative sentences.

Table 18.20 shows the types of thematic elements found in the thematic field of English versus Spanish editorials. As it was pointed out in previous sections, the percentages of the different thematic elements are calculated over the total amount of the thematic clause complexes in each corpus.

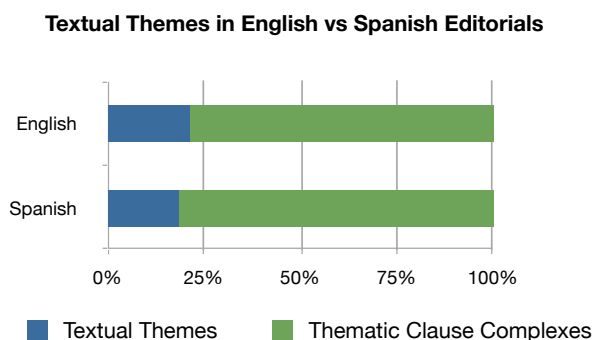
**Table 18.20** Thematic elements in English vs Spanish editorials

English vs. Spanish	Thematic Field			
	Outer Thematic Field		Inner Thematic Field	
Types	TT	IT	PH	H
English	21.12%	11.27%	31.33%	100%
Spanish	18.24%	7.62%	38.88%	100%

The following subsections will go in detail into each of these elements.

### 18.3.3.1 Textual Themes: English vs. Spanish in Editorials

Textual Themes in editorials are found in English 21.12% whereas in Spanish there are less occurrences 18.24% although this difference is not relevant.



**Figure 18.8** Textual Themes in editorials (English vs. Spanish)

As table 18.13 and 18.17 show in sections 18.3.1.1 and 18.17 for English and Spanish, English choice for *Binders* (66.67%) over *Linkers* (33.33%) was significantly relevant. The same significant difference occurred in Spanish for *Binders* (69.23%) over *Linkers* (28.57%). These choices lead to a statistically significant difference ( $P < 0.0001$ ) in the selection of *Binders* over *Linkers* or *Correlatives*. This common feature points to a genre specific characteristic, shared by both languages. The use of *Binders* as introducers of counter arguments by the author to the facts exposed in the text depict the spirit of these texts as expressed by Van Dijk (1989).

Types of Textual Themes in English vs Spanish Editorials

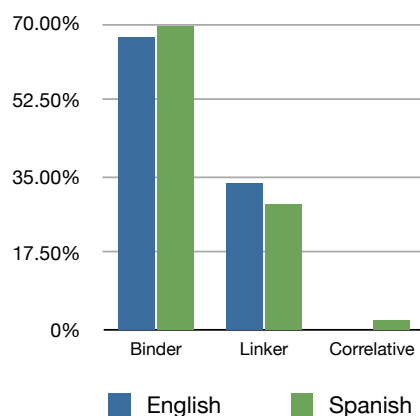


Figure 18.9 Types of Textual Themes in editorials (English vs. Spanish)

The use of *But* (41.25%) in English and *Pero* (42.86%) in Spanish is statistically preferred to the rest of *Binder* elements in both languages.

Types of Binders in English vs Spanish Editorials

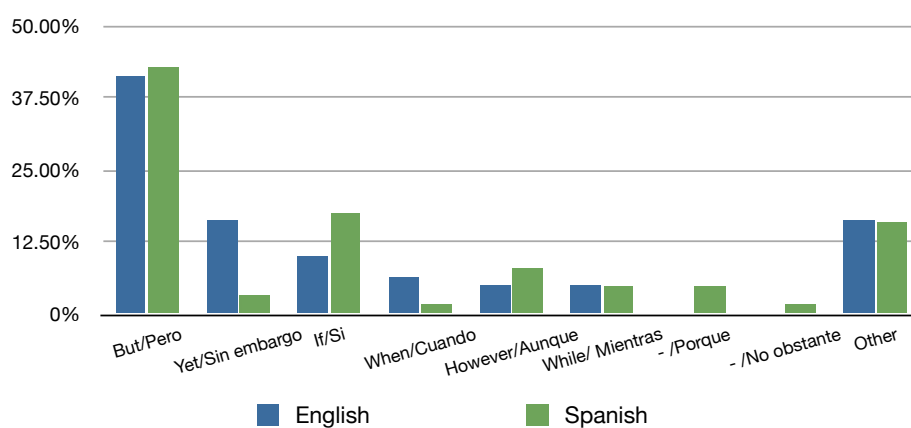


Figure 18.10 Types of binders in editorials (English vs. Spanish)

Here are some examples:

(148) **But** we in Britain should tread carefully.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 22)

- (149) **Pero** no está claro que PP y PSOE tengan las ideas muy claras sobre la estructura y funciones de los organismos independientes.

[But it is not clear that PP and PSOE have very clear ideas on the structure and functions on the independent agencies.] (Editorial: *Europa vigila*. Sentence 11)

The use of *Yet* (16.25%) is more common in English than its correspondent *Sin embargo* (3.17%) in Spanish, although both pursue the same aim of counter argumentation.

- (150) **Yet** according to Theresa May, the Home Secretary, this expression of Parliament's will is routinely being ignored by the courts.

(Editorial: *Courts and Parliament must work together*. Sentence 4)

- (151) **Sin embargo**, la Unión Europea debería haber tenido un papel mucho más activo, y no sólo ahora, sino desde el estallido de las primaveras árabes, para garantizar que las transiciones políticas en el norte de África no facilitaran el auge del terrorismo.

[However, the European Union should have taken a much more active role, not only now but since the outbreak of the Arab spring, to ensure that political transitions in North Africa do not facilitate the rise of terrorism.]

(Editorial: *Terrorismo en el Sahel*. Sentence 15)

It is interesting to highlight the use of *If* (10%) and *Si* (17.46%) in both languages as introduction of hypothetical statements by the author.

- (152) **If** Mr Cameron is serious about confronting the threat posed by Islamist terrorists, he should make sure our Armed Forces have the resources and manpower to deny them the safe havens they crave.

(Editorial: *Britain is ill-prepared for this deadly new world*. Sentence 19)

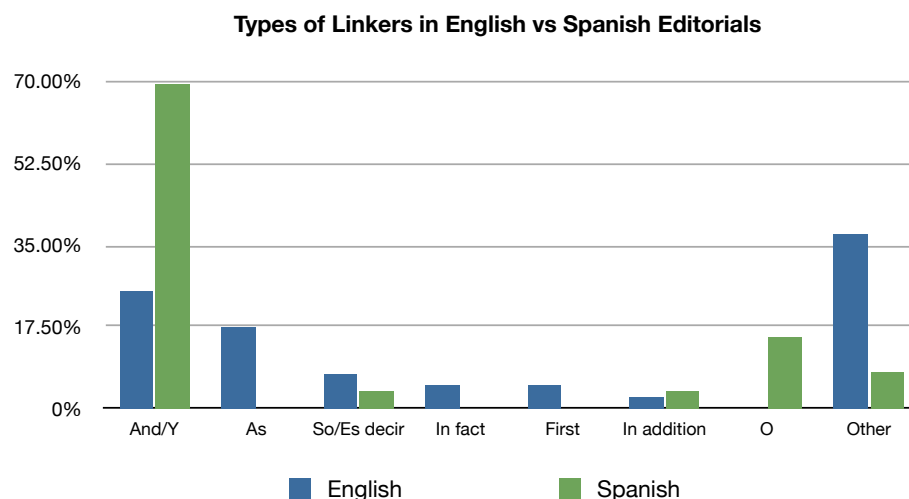
- (153) **Si** hasta ahora los presidentes de Estados Unidos debían lidiar con el conflicto palestino como principal y casi exclusivo escenario político de la región, ahora el cruce de líneas de alta tensión es más complejo debido a los errores de Bush y del anterior primer ministro israelí, Olmert.

[If until now the U.S. president should deal with Palestinian conflict as main and almost exclusive political scenario of the region, now crossing high voltage lines is more complex because of the mistakes of Bush and former israel Prime Minister Olmert.]

(Editorial: *Proyectos divergentes*. Sentence 3)

Regarding *Linkers* although neither English nor Spanish use them widely it is inter-

esting to observe the wide variety of elements used in English to perform this linking action as *and*, *as*, *so*, *in fact*, *first*, *in addition*, and 15 other elements; whereas Spanish relies on *y*, *es decir*, *dicho de otro modo*.



**Figure 18.11** Types of linkers in editorials (English vs. Spanish)

Whereas English shows this wide variety of options in the use of *Linkers* Spanish prefers fewer elements. In fact there is a statistically significant difference in the use of *y* (69%) in Spanish and its correspondent *and* in English (25%).

- (154) **And** yesterday Mark Simmonds, Foreign Office minister for Africa, spoke of providing training and support for the Malian army.

(Editorial: *France moves in*. Sentence 3)

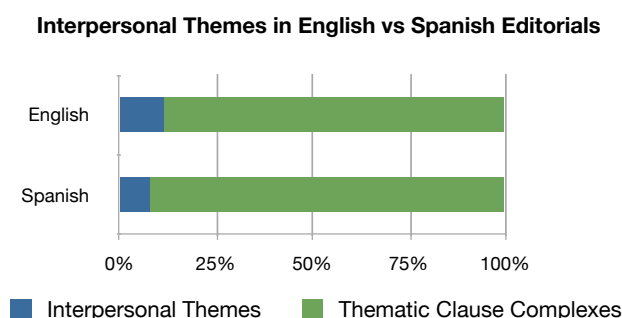
- (155) **Y** es que los casos de corrupción no afectan sólo a los partidos políticos.

[And cases of corruption do not only affect political parties.]

(Editorial: *El listón de los imputados*. Sentence 8)

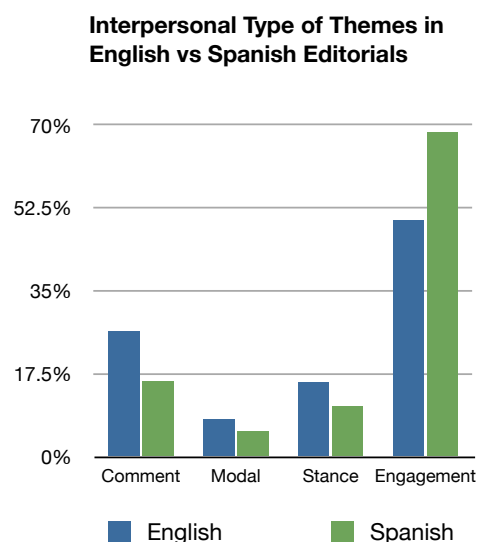
### 18.3.3.2 Interpersonal Themes: English vs. Spanish in Editorials

The occurrences of these elements in both languages does not present relevant differences, as table 18.20 shows (English 11.27% and Spanish 7.62%):



**Figure 18.12** Interpersonal Themes in editorials (English vs. Spanish)

The types of Interpersonal Themes analyzed: *Comment* and *Modal adjuncts*, *Stance* and *Engagement* strategies are shown in the following chart.



**Figure 18.13** Types of Interpersonal Themes in editorials (English vs. Spanish)

Both languages present a preference for *Engagement* elements, this suggests a genre specific characteristic of editorials common to both languages as expressions how “writers relate to their readers with respect to the positions advanced in the text” (Hyland, 2001), as shown in the following examples:

- (156) **Like it or not**, the borders will open – yet alarmingly, there is little sign that central or local government are taking the issue seriously.

(Editorial: *Heads in the sand over Eastern immigration*. Sentence 19)

- (157) **Se la conoce como** ley de Acompañamiento para los Presupuestos del 2013 de la Comunidad de Madrid, fue aprobada el jueves por el pleno de la Asamblea e incluye elementos dispares, como la externalización de servicios sanitarios, la modificación de la ley del Juego para adecuarla a las exigencias de Eurovegas o la implantación del euro por receta.

*[It is known as the Accompaniment Law for the 2013 budget Community of Madrid, it was approved Thursday by the full Assembly and includes different elements such as outsourcing of health services, the modified Law Game to suit the requirements of Eurovegas or the implantation one euro per prescription.]*

(Editorial: *La sanidad pública*. Sentence 1)

The second most preferred type of Interpersonal Theme is *Comments*, a fact that reinforces the idea of the presence of the author in the text. English uses 26.56% and Spanish 15.79%

- (158) **Doubtless**, movement is difficult in rural areas after heavy snowfall; but is it really necessary for transport and schools in towns and cities to be so badly disrupted by what is a natural seasonal occurrence, albeit one experienced less regularly today than in the past?

(Editorial: *This winter's tale*. Sentence 8)

- (159) **Realmente**, este hecho supone un sacrificio que no le gusta a nadie, pero que han de aceptar buena parte de los trabajadores del país para poder mantener el empleo, en una Catalunya en la que hay más de 700.000 parados.

*[Actually, this fact is a sacrifice that does not like anyone, but should have to be accepted by most part of the workers to maintain their jobs, in Catalunya where there are more than 700.000 unemployed.]*

(Editorial: *Grave error en Nissan*. Sentence 3)

Contrary, the use of *Stance* defined by Hyland as “the ways writers present themselves and convey their judgements, opinions, and commitments.” is scarcely used (2005:176). English does it in a 15.63% and Spanish 10.53%. This means that the author does not present him/herself outright/downright/unequivocally to the reader but the opinion he/she represents as an editorialist.

- (160) **We thought that** consumers were entitled to assume that meat – and indeed all food – was subject to random spot checks to confirm that it was what it said it was.

(Editorial: *The market needs watching*. Sentence 6)

- (161) **Esperemos** que el presidente del Gobierno ofrezca hoy alguna pista en este sentido durante el debate sobre el estado de la nación.

[Let's hope that the Prime Minister gives some hint today in the debate on the state of the nation.] (Editorial: *Exportamos bien*. Sentence 12)

As well as *Stance*, *Modal adjuncts* are reserved to very few situations. In English 7.81% and Spanish 5.26%

- (162) **Perhaps** this was because the statement in which the agency justified its action will have made even less comfortable reading for the shadow chancellor than his opposite number.

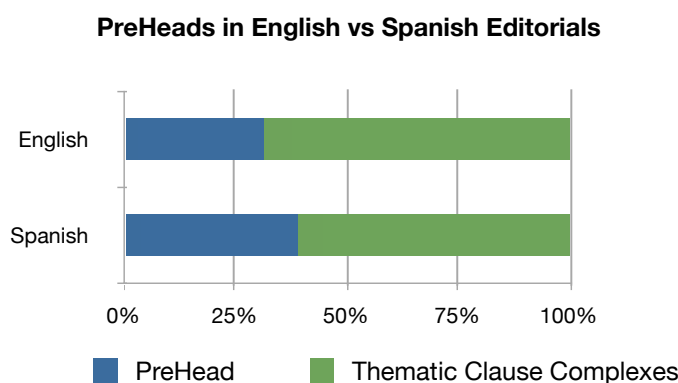
(Editorial: *Osborne needs to budget for economic growth*. Sentence 2)

- (163) Pero **quizá** beneficie al otro independentismo, el de Esquerra, que goza de una imagen más limpia.

[But maybe this benefits the other independentism, the one of Esquerra, which has a clearer look.] (Editorial: *Escuchas y fallos*. Sentence 13)

### 18.3.3.3 PreHead: English vs. Spanish in Editorials

The use of PreHeads in English and Spanish editorials present a statistically significant difference ( $P=0.0119$ ) in number of elements found in both languages, being Spanish where more elements of this kind are found.



**Figure 18.14** PreHeads in editorials (English vs. Spanish)

The reason for this difference relies in a language specific characteristic of Spanish due to the possibility that this language presents of having the verb in first position with no explicit mark of the subject. English is able to do so only in non-declarative sentences, such as directives or questions. According to explained in section 11.2.1 PreHead is the lexical part of the verb leaving its grammatical information as Head. The number of Finite elements found in English are 4.49% and in Spanish 19.07%



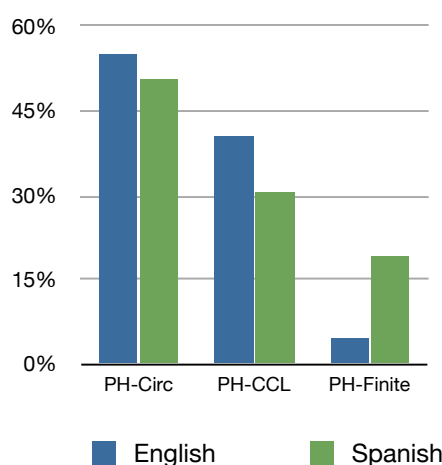
Examples:

- (164) **Agrava** la conducta de los acusados haciéndola merecedora de penas de prisión: tres años para el principal acusado, el general de Sanidad Vicente Navarro, y año y medio para el comandante y el capitán médicos José Ramón Ramírez y Miguel Ángel Sanz.

[This aggravates the conduct of the defendants making it worthy of sentences prison, three years for the main accused, General Health Vicente Navarro, year and a half to the commander and captain physician Jose Ramon Ramirez and Miguel Angel Sanz.] (Editorial: *El tribunal admite que la falsa identificación fue intencionada pero no dice los motivos*. Sentence 4)

- (165) **Chalk up** another victory for the PR departments of the global investment banks.  
(Editorial: *Blame the tax loopholes*. Sentence 1)

**PreHead Types in English vs Spanish Editorials**



**Figure 18.15** Types of PreHeads in editorials (English vs. Spanish)

The chart shows the different types of PreHeads found in the bilingual corpus. As it has already been pointed out before, the significant difference between English and Spanish relies in the realization of Finite PreHeads as a language specific characteristic of Spanish (regardless of genre).

The use of the other two types of PreHeads, *PH-Circ* and *PH-CCL* are common to both languages, with a statistical significant preference for *PH-Circumstantials* realized

by Adverbials and Prepositional Phrases, as explained in section 11.2.1. English uses a 55.06% of these elements and Spanish of 50.52%. The following examples are *Ph-Circ* realized by Adverbial (166) (167) and Prepositional phrases (168) (169).

(166) **Today** the lobby group MigrationWatch plucked from the air another dubious figure, saying there will be around 50,000 new arrivals every year once restrictions are lifted. (Editorial: *Stop all the demonising of new EU migrants*. Sentence 3)

(167) **Ayer mismo** el líder derechista israelí aseguraba que había acordado en Washington la necesidad de conversaciones de paz con Siria y con los palestinos.  
[Yesterday the Israeli right-wing leader claimed that Washington had agreed in the need for peace talks with Syria and the Palestinians.]  
(Editorial: *Proyectos divergentes*. Sentence 11)

(168) **By 2032**, the over-75s are expected to comprise 22 per cent of the population, up from 17 per cent now, so we need to act fast.  
(Editorial: *A dilemma that wont go away*. Sentence 16)

(169) **Con su decisión**, Joseph Ratzinger puede que abra paso a un nuevo aggiornamento católico.  
[With his decision, Joseph Ratzinger may open a way to a new Catholic aggiornamento.]  
(Editorial: *El papado de Benedicto XVI*. Sentence 44)

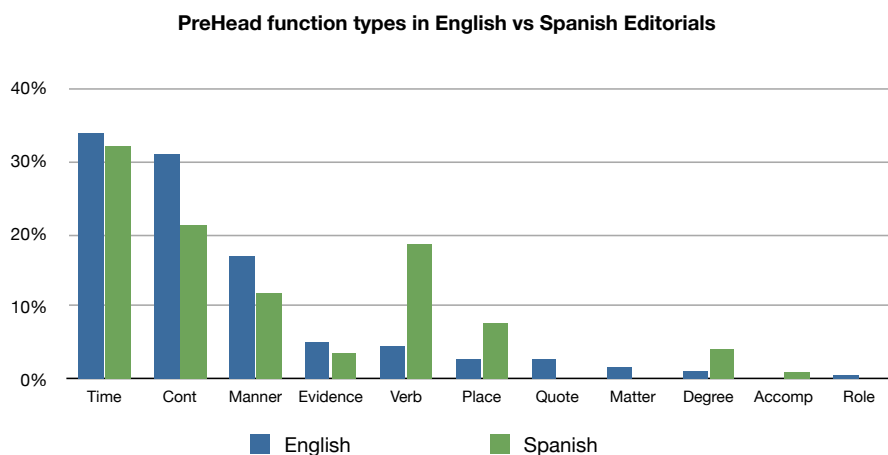
The following chart shows the main functions that these PreHeads perform in terms of Circumstance:

The setting of events in *Time* is the most frequently found in editorials. This genre specific feature pretends to accomplish the reader's agreement as main objective and to do so on recent events (Alonso Belmonte, 2007). Both English (33.71%) and Spanish (31.96%) use a similar amount of these time setting elements.

(170) But **between now and mid-2014**, Mr Obama is in a remarkably favourable position. (Editorial: *Obama must seize the chance to make history*. Sentence 12)

(171) **Hoy**, lo fácil sería decir que sí.  
[Today, the easy thing would be to say yes.]  
(Editorial: *El listón de los imputados*. Sentence 38)

Once the time framework is depicted, the arguments become the second most relevant device in use. *Contingency* is used in English in a 30.90% and Spanish 21.13% as introducers of counter argumentation or hypothesis as in the following:



**Figure 18.16** Functions of PreHeads in editorials (English vs. Spanish)

- (172) **Even if Lance Armstrong had come across as the most contrite of penitents – which he did not** – an interview with Oprah Winfrey on American TV is not sufficient to allow the cyclist to draw a line under his disgrace and “move on”.  
(Editorial: *A confession is not enough*. Sentence 1)

- (173) **Por si faltaban pruebas sobre su peligrosidad**, Teherán lanzó ayer mismo un misil en pruebas con alcance de 2.000 kilómetros.

*[In case there was lack of evidence about its dangerousness, Teheran yesterday launched a testing missile with a range of 2,000 km.]*

(Editorial: *La visita de Netanyahu a la Casa Blanca anticipa serias discrepancias sobre Oriente Próximo*. Sentence 6)

#### 18.3.3.4 Head: English vs. Spanish in Editorials

The number of thematic clauses in the English Corpus is of 97.76% and in the Spanish editorial Corpus of a 95.05% of the total amount of clause complexes found in editorials.

Interesting features that can be found in Heads are the use of Thematic Equatives and Predicated Themes, as defined in section 11.2.2, these formations are more typical of oral environments.

Thematic Equatives are the answer to a presupposed question of the reader where the prominence of the participant, attribute, circumstance or process taking part in the referred message is cognitively highlighted (Gómez-González, 2001) grouping more than one element into a single constituent (Cerban, 2008) advancing the Rhemes to the

Theme.

Examples of Thematic Equative in English and Spanish:

- (174) **What Britain needs** is a robust and unambiguous tax system that requires corporations to pay their share, without the need for negotiations with HMRC, or moral urgings from politicians, or gestures of largesse by pressurised managements.

(Editorial: *Blame the tax loopholes*. Sentence 9)

- (175) **Lo que requiere una investigación y la correspondiente depuración de responsabilidades** es la colaboración que la empresa del expistolero presta a la lucha contra el crimen con acceso a datos confidenciales.

[*What needs an investigation and the corresponding purification of responsibilities is the collaboration that the company of the ex-gunman's pays to the fight against crime granting them access to confidential data.*]

(Editorial: *El caso "Hellín"*. Sentence 4)

Predicated themes, on the other hand, in line with what was mentioned on section 11.2.2 and what Martínez Lirola (2008) states about this syntactic structure “allows the writer to be conscious that he/she is asserting or denying something in a firm way”.

The following examples show the use of Predicated Themes in English and Spanish:

- (176) **It is Mr Paterson, above all, who** has questions to answer.

(Editorial: *The market needs watching*. Sentence 27)

- (177) **Porque fue también la prensa la que** hace 20 años abrió los ojos a la policía dando con el paradero de Hellín en Paraguay, donde se había refugiado tras huir de la cárcel.

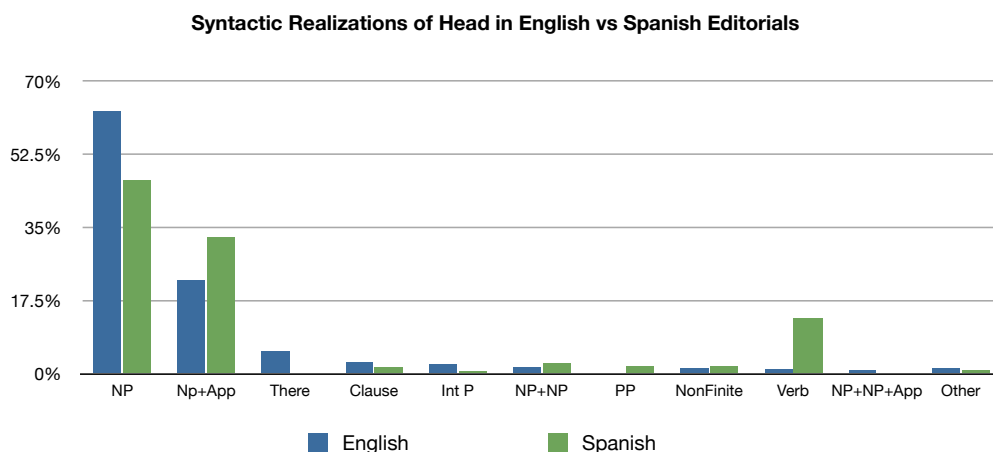
[*Because it was the press who 20 years ago opened the eyes of the police giving the whereabouts of Hellin in Paraguay, where he had taken refuge after fleeing from jail.*]

(Editorial: *El caso "Hellín"*. Sentence 8)

### Syntactic Realizations

This section studies the preferred syntactic realizations of Head in English and Spanish editorials. Both languages use *Noun Phrases* as their preferred syntactic choice, but there is a statistically significant difference between them: English 62.85% and Spanish 46.29%. As the figures show this is not a genre specific distinction considering that both languages select *Noun Phrases* as their most preferred choice. This difference may be explained as a language specific characteristic, in this case of Spanish, this language does not need to place a marked subject in front position before the verb

whereas English is forced to do so in declarative clauses. Thus looking at the number of verbs in head position (as explained in section 18.14 for English and 18.18 for Spanish) in English there are only 0.88% occurrences meanwhile Spanish has 13.03%, what reveals a statistical significant difference in the use of these structures between both languages.



**Figure 18.17** Syntactic realization of Heads in editorials (English vs. Spanish)

The realizations of *Noun Phrases* are *Pronouns* as the first option chosen for English 35.85% whereas Spanish selection is only 15.15%, as it was already mentioned, Spanish can use *Verbs* in first position and avoid the specific mark of the subject, English cannot do so, the need of an explicit mark for the subject leads to the use of a mark of any kind, in this case of *Pronouns*, whereas Spanish relies in the grammatical information of the verb to perform this action.

- (178) **They** are generally young, and have more children, which is good news, given the ageing population.

(Editorial: *Stop all the demonising of new EU migrants*. Sentence 23)

- (179) **Eso** permitiría a Monti exigir un papel dominante en la economía transalpina, a la que en buena medida ha devuelto la credibilidad.

[That would allow Monti to demand a dominant role in the transalpine economy, which largely has restored credibility.]

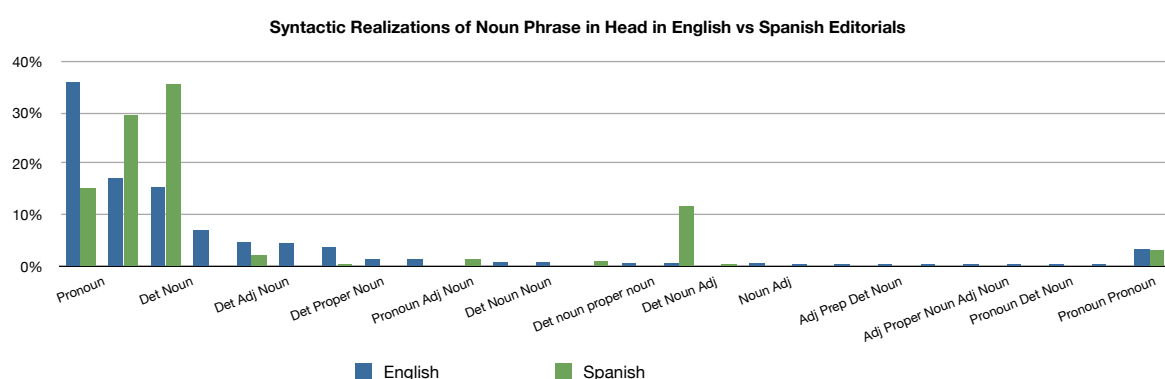
(Editorial: *Berlusconi todavía*. Sentence 11)

Example in Spanish the use of verbs in first position:

- (180) Y, desde luego, *[ellos]* **deben presentarse** una vez bien estudiadas y con la viabilidad de su aplicación asegurada.

*[And of course, [they] should introduce themselves once their applications have been well studied and with their viability secured.]*

(Editorial: *Falta elaboración política en las propuestas para salvar los sectores del automóvil e inmobiliario*. Sentence 15)



**Figure 18.18** Syntactic realization of NP Heads in editorials (English vs. Spanish)

The introduction of the participant is done through *Proper Nouns* or *Noun Phrases* and throughout the text this participant is anaphorically referred, therefore the number of *Proper Nouns* found in Head position is so scarce. In fact, the number of *Proper Nouns* used are found with a statistically significant difference in both languages (English 17.09% Spanish 29.44%) with respect to the rest of the elements in this position (except *Pronouns* treated above):

- (181) For the same reason, **David Cameron** is right to have extended logistical and other assistance to Paris.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care*. Sentence 4)

- (182) En total, **Benedicto XVI** habrá permanecido en su puesto poco menos de ocho años.

*[All in all, Benedict XVI would have remained in his post only eight years.]*

(Editorial: *El papado de Benedicto XVI*. Sentence 5)

Another strategy found presenting a statistically significant difference between English and Spanish is the use of *Determiners followed by Nouns*. English adopts this

strategy in 15.40% and Spanish in 35.50%. These are some examples:

- (183) **The crisis** is also yet another reminder that Britain's influence with allies, and its own national security, require strong military capabilities of the sort that have been drastically shrunk over the past two years.

(Editorial: *As the terror threat shifts to Africa, we must meet it with great care.* Sentence 21)

- (184) **Dichos comicios** podrían, por tanto, saldarse con una nueva victoria para las huestes de los Hermanos Musulmanes.

[*These elections could therefore be paid off with another victory for forces of the Muslim Brotherhood.*]

(Editorial: *Egipto, de la paradoja a la gobernabilidad.* Sentence 20)

It is interesting to point out that whereas Spanish tends to these formations, English compensates with the use of *Nouns* instead (English 7.00% and Spanish 0.00%). This suggests a language specific characteristic and the explanation may rely on English allowing in subject position bare plural *Noun Phrases* while Spanish and some other Romance languages do not (Pérez-Leroux et al., 2004).

- (185) **Democrats** won the popular Congressional vote by some half a million nationally, but gerrymandering is a wonderful thing.

(Editorial: *Obama must seize the chance to make history.* Sentence 17)

In Spanish a determiner would be needed to express this idea: *Los demócratas* [*Democrats*].

The second choice in both languages is *Noun Phrases followed by Appositions*, the use of these elements also reveal a statistical significant difference between English (22.36%) and Spanish (32.46%), that, as it was seen with *Noun Phrases*, cannot be attributed to a genre specific feature but to a language specific characteristic. In this case, there is no other strategy used in Spanish that justifies these differences; in fact it seems that Spanish tends to have more complex realizations of Heads due to the use of these more complex syntactic patterns.

Both select *Prepositional Phrases* as their preferred realization of *Appositions*, with a similar distribution in both languages, in English there are 50.39% and in Spanish 56.17%.

Here there are some examples of *Appositions* in the form of *Prepositional Phrases* in English and Spanish :

- (186) **A rebel victory at the town of Konna** opened the way to Mopti, Mali's second city, and thus to the whole of the south.

(Editorial: *Mali's Islamists are too dangerous to be ignored.*

Sentence 12)

- (187) **El tijeretazo al presupuesto de Van Rompuy** se da principalmente a partidas como infraestructuras de transportes (que puede afectar al corredor mediterráneo), energía y telecomunicaciones.

[*The budget snip of Van Rompuy is principally given to transport infrastructures (that may affect the mediterranean corridor) energy and telecommunications.*]

(Editorial: *España gana con el presupuesto de la UE.* Sentence 9)

After the realization of *Appositions* by *Prepositional Phrases*, *Clauses* are the second most preferred choice in both languages. English has a 15.75% and Spanish 16.67%. These resources are used as explanatory devices of the main item as the following examples show:

- (188) **The official view, shared by successive governments**, has long been that the UK does not negotiate and does not give in to terrorists.

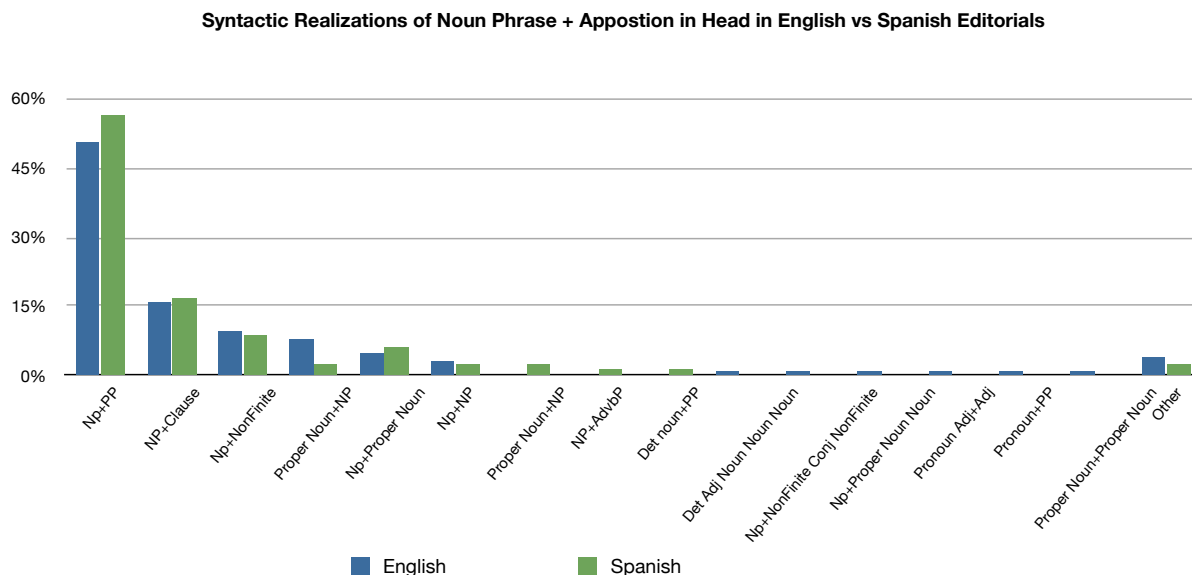
(Editorial: *David Cameron's mixed signals on terrorism.* Sentence 10)

- (189) **Las noticias que el diario conservador Daily Telegraph viene publicando** son motivo no sólo para desguazar los actuales mecanismos de control -que ni laboristas ni conservadores han tenido el menor interés en cambiar-, sino probablemente para iniciar por fin una reforma que acabe con los profundos anacronismos del sistema democrático más afianzado del mundo.

[*The news that the conservative Daily Telegraph has published recently are a reason not only to scrap the current control mechanisms - that nor Labour neither Conservatives have had the slightest interest in change - but probably to initiate a reform that will finally end with the deep anachronisms of the world's oldest democratic system.*]

(Editorial: *Sonrojo en Westminster.* Sentence 4)





**Figure 18.19** Syntactic realization of NP+App Heads in editorials (English vs. Spanish)

### Semantic Realizations

*Material processes* are selected by both languages as their first choice, in detail *Actors* are the 31.69% of English and 35.27% of Spanish experiential roles.

Some examples in of English and Spanish material processes are:

- (190) **Democrats won** the popular Congressional vote by some half a million nationally, but gerrymandering is a wonderful thing.

(Editorial: *Obama must seize the chance to make history*. Sentence 17)

- (191) **De hecho, España tendrá que aportar** 3.300 millones de euros menos a la caja común, lo cual es favorable, como mostró la caída a 375 puntos de la prima de riesgo. España gana con el presupuesto de la UE 18

[In fact , Spain will have to provide 3.300 million less to the common fund , which is favourable , as shown by the drop to 375 points of the risk premium.]

(Editorial: *España gana con el presupuesto de la UE*. Sentence 18)

*Relational processes* are the second preferred option realized by *Carriers* which are 32.30% in English and 26.25% of Spanish roles. Using this strategy allows, in my opinion, to present opinion as facts.

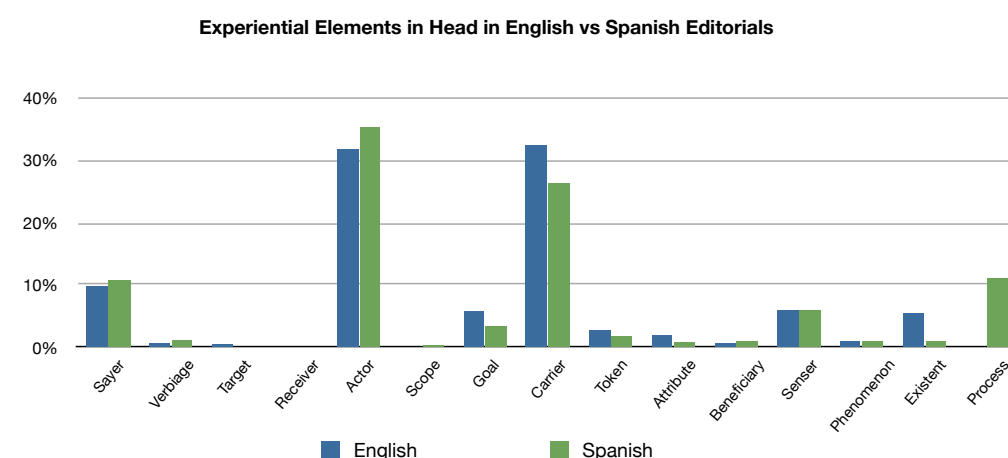
- (192) **Yet today the mood is** downbeat, the dreaded term “lame duck” is on every lip.

(Editorial: *Obama must seize the chance to make history*. Sentence 5)

- (193) **El PP tiene** cada día más difícil mantener esa actitud, contraria a lo establecido en su propio código de conducta y que supone aplicar dos criterios diferentes en Madrid y en Valencia.

[The PP is harder every day to maintain that attitude, contrary to the provisions of its own code of conduct and that involves applying two different criteria in Madrid and Valencia.]

(Editorial: *Sigue imputado*. Sentence 9)



**Figure 18.20** Semantic realization of Heads in editorials (English vs. Spanish)

The main difference in the selection of English and Spanish experiential roles is the use of processes in Spanish, as seen in sections 11.2.1.2, 18.3.2.3 this language specific characteristic allows Spanish not to convey Subjects in first position. Thus the use of processes are so frequent in Spanish 11.02%.

- (194) Viajó también a Estambul y a Jerusalén.

[He also travelled to Istambul and Jerusalem.]

(Editorial: *El papado de Benedicto XVI*. Sentence 33)

### 18.3.3.5 Rheme: English vs Spanish in Editorials

The number of rhematic sentences are similar in both languages, in English is 3.24% and in Spanish 4.95%

## 18.4 Thematic Characterization of Letters to the Editor in English and Spanish

Broadly speaking Mémet (2005) defines letters to the editor as “reactions to publications” and Pounds (2010) pins down this definition adding interesting features worth noticing, such as the author adding first-hand factual evidence or subjective personal experience, expressing his/her own conviction/perception undermining or aligning with the positive aspects of the criticised situation. Bearing these characteristics in mind the following pages analyze the LEs corpus.

Table 18.21 shows the corpus of letters to the editor in English and Spanish.

**Table 18.21** Corpus of LEs in English and Spanish

Letters to the Editor	Texts	Clauses	Words
English	50	366	8228
Spanish	50	365	9164
Total	100	731	17392

### 18.4.1 English Letters to the Editor

The total number of complex clauses found in the English letters to the editor corpus is of 366 clauses. Of this amount 1.64% are rhematic clauses with no thematic construction. The rest of the complex clauses 98.36% are thematic complex clauses (1.38% directives, 5.83% questions and the rest declaratives). As in news reports and editorials I am going to focus on Thematic clauses, therefore all percentages referred to the types of Thematic clauses are calculated over the total amount of these complex clauses. The following table shows the types of Thematic elements found in letters to the editor English corpus.

**Table 18.22** Thematic elements in English letters to the editor

English	Thematic Field			
Types	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
	13.33%	9.44%	25.83%	100%

Thematic elements are divided into Inner thematic field and Outer thematic field. The Outer thematic field includes two elements: Textual Themes (TT) and Interpersonal Themes (IT). The Inner thematic field comprises PreHeads (PH) and Head (H). Textual Themes appear in a 13.33% clause complexes. Interpersonal themes are very frequent with a 9.44% which indicates a high presence of the author in the texts. In the Inner thematic field PreHead accounts for 25.83% . As dealing with thematic clauses all of them have Head (100%). This means that an amount of 51.4% of the clause complexes in English letters to the editor corpus are realized by unaccompanied Heads, what points to fewer simple thematic constructions of clause complexes than in editorials and news reports. The following sections analyze this data in detail.

#### 18.4.1.1 Textual Themes in English Letters to the Editor

The number of Textual Themes in English letters to the editor is of 13.33%. It is important to highlight the short extension of this type of texts (average of 7.32 clause complex per text); this is something that compromises the use of Textual Themes as their cohesive function is less needed in such a short document.

**Table 18.23** Types of TT in English LEs

Types of Textual Themes	
Binder	85.42%
Linker	12.05%
Correlative	2.08%
Total	100%

Although there is a limited number of textual elements in letters to the editor there

is a statistically significant difference between the use of *Binders* and *Linkers*, this may be due to the aim of the Letters, normally as complaint of some kind to a previously identified topic. *Linkers* are used in 12.5% whereas *Binders* are used in 85.42%.

Regarding *Binders* the use of the adversative [*But*] in 29.27% of cases:

- (195) **But** we should applaud Lord McAlpine for taking action to discourage thoughtless and irresponsible tweeting.

(Letter to the Editor: *Peer brings twitter to hell*. Sentence 2)

*When* and *if* are used in 12.20% each of them, as in the examples:

- (196) **When** I was at school and university I spent most holidays doing jobs like delivering milk, working on building and road construction sites and labouring in paper mills.

(Letter to the Editor: *Free labour to stack shelves*. Sentence 2)

- (197) **If** any homosexual couple and their guests were willing to sit through that for the 45 minutes or more of my average-length sermon, I suppose they would have earned the right to the piece of paper that entitles them to the very few and steadily diminishing privileges still conferred by the legal state of marriage in this country.

(Letter to the Editor: *Gay wedding in church? Are you sure?*. Sentence 5)

The use of *Linkers* on the other hand is very varied, with an even distribution (16.6% each) between: *and*, *for*, *in addition*, *therefore*, *finally*, *in one example*.

#### 18.4.1.2 Interpersonal Themes in English Letters to the Editor

The use of Interpersonal themes in English letters to the editor is of a 9.44% in all clause complexes (Table 18.22). This element reflects the presence of the author in the text, revealing a subjective point of view of the matter at stake.

The analysis of these elements revealed that there is a clear tendency for *Engagement* (52.94%) with a statistical significant difference over the other types of Interpersonal Themes: *Stance* 23.53%; *Comment* 14.71% and *Modal* 8.82%. (see section 11.1.2 for a definition of these types of Interpersonal themes)

This use points to a strong perception of the reader from the author's perspective who recognizes and addresses to him/her the message making the reader a relevant participant in the development of the argument exposed, and part of his/her opinion group, as in:

- (198) **It is clear that** many problems would arise if the legislation as now tabled were to be implemented.

(Letter to the Editor: *Settle this Plebgate stand-off now.* Sentence 7)

- (199) **It is time** for the Director of Public Prosecutions to draw a line under the affair, by charging Mr Mitchell, based on what evidence is available, and co-charging any officer who it is believed may have lied in the course of their duty.

(Letter to the Editor: *Traditional institution of marriage is central to the welfare of society.* Sentence 6)

### 18.4.1.3 PreHead in English Letters to the Editor

The number of PreHeads found in this corpus is of 25.83% of the total number of thematic complex clauses, as shown in table 18.22

The following table shows the number of types of PreHeads in the corpus according to the definitions given in section 11.2.1

**Table 18.24** Types of PreHeads

PreHead Types	%
PH-Circ	44.57%
PH-CCL	43.48%
PH-Finite	11.96%
Total	100%

As seen there is no statistical preference for any of the three types analyzed. But it is interesting to pay attention to the *PH-Finite* formations (11.96%) found in directives such as:

- (200) **Imagine** an airline claiming that an air disaster was caused by a faulty component having been used by a rogue maintenance company.

(Letter to the Editor: *Supermarkets must take blame for horse meat.* Sentence 4)

- (201) **Fund** candidates to explain their policies, and hold the voting in May when many families have a walk out after tea to vote – instead of looking out of the window, deciding it's too cold, and staying in to watch the football instead.

(Letter to the Editor: *A disappointing day for democracy.* Sentence 7)

These expressions involve the reader in the situation, in fact (Hyland, 2005) refers to the use of directives as a way to instruct the reader to perform an action or to see things in a way determined by the writer, these actions want the reader to engage either in textual, physical or cognitive acts, being the latter the ones found in the present corpus, the reader is guided through a line of reasoning by the writer to persuade him/her of their personal point of view.

There is no difference between the other two realizations *PH-CCL* (43.48%) and *PH-Circ* (44.57%) this shows that the realizations of PreHeads tend to be evenly simple *Prepositional* and *Adverbial Phrases* as well as *Complex Clauses*, as in:

- (202) **While privacy laws may have been compromised in the telephone call,**  
the act itself is a staple of popular radio anywhere in the Western media.

(Letter to the Editor: *Bullies and victims in the hoax call row.*  
Sentence 4)

- (203) **Even if they can't get guns,** they can still mow people down with their vehicle,  
as happened here earlier this year.

(Letter to the Editor: *Limit access to guns but in wont end*  
*massacres.* Sentence 5)

*PH-Circ*:

- (204) **In 2007,** George Osborne's pledge to raise the inheritance tax threshold to £1  
million was greeted by a seven-point rise in the polls for the Tories.

(Letter to the Editor: *Tory U-turn on inheritance tax pledge puts*  
*pressure on George Osborne.* Sentence 1)

A 32.26% of these *PH-CCL* and *PH-Circ* are used to set the *Time* of the event to which the letter is related to. At the same time 25.81% are used to express *Contingency* which, in Downing (2006) words covers such meanings as cause, purpose, reason, concession and behalf. This *Contingency* expresses a counter argument or a hypothetical situation order or suggestion to be considered by the author and thought over by the reader. This can be seen in the following example:

- (205) **Whatever people thought of the merits or otherwise of that election,** the  
majority wanted to vote but knew nothing about the candidates – a particularly  
significant point when so many were concerned that the police should not become  
politicised.

(Letter to the Editor: *How PCC electors were kept in the dark.*  
Sentence 4)

#### 18.4.1.4 Head in English Letters to the Editor

The total number of complex clauses in the Corpus is of 366, of them, only 1.64% are rhematic sentences while the rest (98.36%) are thematic sentences. In these thematic sentences there are a few occurrences of *Thematic Equatives* and *Predicated Themes* with a 0.55% and 0.27% of presence respectively.

- (206) **What distinguishes mankind now from what we were then, for good or ill,** is science and technology.

(Letter to the Editor: *Intellectual vacuum*. Sentence 3)

- (207) **It is only when Gaza prospers that** those Palestinians wanting to hit Israeli targets will no longer be tolerated.

(Letter to the Editor: *Prosperity and peace for Gaza*. Sentence 6)

#### Syntactic Realizations

The analysis of the syntactic forms reveal that *Noun Phrases* (70.56%) are the preferred strategy to express Heads in English letters to the editor and only a 15% of Heads were realized by *Noun Phrase followed by Apposition*. This choice reflects the intention of the author as simple and direct (Bronia, 2005:182), it is also influenced by the shortage of space these texts are obliged to respect.

More in detail *Noun Phrases* are realized by *Pronouns* in a 51.57% of cases. Then *Determiner Noun* (13.39%) formations and in third position *Proper Nouns* (9.54%), these two of scarce use, for three main reasons: the main participant is identified either in the title or in the first line of the LE; often this Letters refer to personal experiences or opinions therefore the writer refers to him/herself using *Personal Pronouns*; and finally the shortage of space favours the use of these anaphoric forms. Let's see some examples of *Pronouns* (208), *Determiner Noun* (209) and *Proper Nouns* (210) formations:

- (208) **I** was a poll clerk at the police commissioner election in Nottinghamshire and witnessed the low turnout; it meant a much easier day for me than some recent elections, but it was a disappointing day for democracy.

(Letter to the Editor: *A disappointing day for democracy*. Sentence 1)

- (209) **This practice** is termed "split-ticketing", and it adds a further hurdle for unwary passengers within an already labyrinthine fares system.

(Letter to the Editor: *Baffling rail fares*. Sentence 4)

- (210) **Richard Garner (News, 16 November)** reports that Buckingham University has recently shown that nearly 30 per cent of trainee teachers last year had not



taken up teaching posts six months later, and that 52 per cent of teachers under the age of 60 are not within the classroom.

(Letter to the Editor: *Schools in need of discipline*. Sentence 1)

In *Noun Phrases followed by Appositions* the *Appositions* are realized by *Prepositional Phrases* in 44.44% cases, as in the following example:

- (211) **Cristina Odone's views on doctors and weekend working** ("Doctors should get off the golf course and into the wards", Comment, December 17) would have brought a wry smile to the thousands of hospital consultants who, like me, spent the weekend working in their local hospital as part of their on-call commitment.

(Letter to the Editor: *Hospital consultants already carry our work commitments at weekends*. Sentence 1)

### Semantic Realizations

*Actors* are the experiential roles most used 33.06% (examples (212) (213)), closely followed by *Carriers* with a presence of 29.17% (example (214)), these choices coincide with the nature of the text, where the importance of the argumentation of the author is of key importance.

- (212) **These multinational companies come** to Britain to enjoy the benefits of trading in a secure, politically stable environment, yet do their very best to weasel out of making any realistic contribution to the upkeep of their host nation or the education and welfare of its people.

(Letter to the Editor: *Put Starbucks out of its misery*. Sentence 3)

- (213) **But the National Rifle Association has hijacked** the Second Amendment to assert that every individual has the right to be armed with guns.

(Letter to the Editor: *Changing American's gun culture*. Sentence 5)

- (214) **This country has** no link with purple. (Letters, November 19).

(Letter to the Editor: *No purple pride*. Sentence 1)

#### 18.4.1.5 Rheme in English Letters to the Editor

There are 1.64% rhematic expressions in the corpus, as the following examples show the type of expressions are colloquial:

- (215) Good starting point.

(Letter to the Editor: *Changing American's gun culture*. Sentence 4)

(216) No change, no chance.

(Letter to the Editor: *After the EU speech and gay marriage, what next for Cameron?*. Sentence 3)

### 18.4.2 Spanish Letters to the Editor

The number of complex clauses in the Spanish letters to the editor Corpus is of 365, being 92.88% thematic clauses and 7.12% rhematic clauses. As in previous genres the focus is placed on thematic clauses and quick overview into the rhematic sentences will be given.

The following table gives an account of the elements found in the Inner Thematic field and the Outer Thematic Field, as in the rest of the analysis, each element has been contabilized over the total amount of thematic clause complexes. Thus, on the one hand, Inner Thematic Field has 47.79% of PreHeads and 100% of Heads; on the hand, the Outer Thematic Field accounts for Textual and Interpersonal elements where Textual elements sum up to 15.93% of all Thematic clause complexes and Interpersonal themes are 18.88%.

**Table 18.25** Thematic elements in Spanish letters to the editor

English	Thematic Field			
	Outer Thematic Field		Inner Thematic Field	
	TT	IT	PH	H
Types	15.93%	18.88%	47.79%	100%

#### 18.4.2.1 Textual Themes in Spanish Letters to the Editor

The number of Textual themes in Spanish letters to the editor is of 15.93%. The nature of this type of text, due to its extension, with an average number of 7.3 clauses per texts, does not need big amounts of these cohesive elements.

The following table shows the result of the analysis carried out on Textual Themes according to the types defined in section 11.1.1. *Binders* are the preferred choice in this type of text with an amount of 64.81% over *Linkers* used 35.19%.

**Table 18.26** Types of TT in Spanish Letters to the Editor

Types of Textual Themes	
Binder	64.81%
Linker	35.19%
Total	100%

The type of *Binders* selected are introducers of adversatives and contingent sentences expressing counterarguments and hypothetical thesis on the part of the author, like *Pero* (28.57%) or *Si* (34.29%). As the following examples show:

- (217) **Pero** estaría bien que a la vista de las últimas noticias sobre supuestos pagos en negro a trabajadores de sus horas extras, evitando así la correspondiente cotización a la Seguridad Social, alguien preguntara al señor Rosell si no sería más conveniente que en lugar de dejar en casa a los funcionarios supuestamente ociosos los pusiéramos a vigilar a los empresarios supuestamente tramposos, para que cumplan con sus obligaciones para con los trabajadores y para con el resto de la ciudadanía.  
*[But it would be nice to in view the latest news about alleged payments to workers on their overtime in black money, avoiding the corresponding quote Social Security, someone will ask Mr. Rosell if it would not be better instead of leaving home idle public workers, we set them to watch the supposedly cheater entrepreneurs to meet their deeds to workers and to the rest of the citizenry.]*

(Letter to the Editor: *Vigilar a los empresarios*. Sentence 2)

- (218) **Si** tenemos en cuenta que British Airways estaba quebrada y lo estará aún más gracias al desastre de su fondo de pensiones, y que por el contrario, hasta la fusión con British, Iberia estaba sana y era rica, habrá que preguntarse dónde está la caja de Iberia, quién va a pagar las pensiones de los británicos y dónde están los chorizos, porque en Londres no habrá chorizo, pero chorizos hay un montón.  
*[If we consider that British Airways was broken and will be even more thanks to the disaster in its pension fund, and instead, until the merger with British, Iberia was healthy and was rich, we must ask where is the box of Iberia, who will pay the pensions of the British and where the “chorizos” are, because in London there will not be “chorizos”, but there are plenty “chorizos”.]*

(Letter to the Editor: *Chorizos británicos*. Sentence 4)

In the case of *Linkers*, lesser in number than *Binders*, are used to add information

in a linear manner with the use of *Y [And]* (42.10%), as in :

- (219) **Y** me consta que los trámites para conseguirla no eran fáciles ni rápidos: avales, nóminas, otras propiedades, etcétera.

[And I know that the procedures to get it not easy nor fast: sureties, payroll, other properties, etcetera.] (Letter to the Editor: *Contra la amnesia*. Sentence 7)

#### 18.4.2.2 Interpersonal Themes in Spanish Letters to the Editor

The number of Interpersonal Themes found in the corpus of Spanish letters to the editor is of 18.88%

Both *Stance* (39.06%) and *Engagement* (40.63%) elements are found, almost evenly, as the two preferred choices in this position. In Hyland's (2005:186) words *Stance* "concerns writer-oriented features of interaction" and *Engagement* (2005:182) refers to "the ways writers bring readers into the discourse to anticipate their possible objections and engage them in appropriate ways". That is, the presence of the writer in the text is as relevant as the intention of the writer to involve the reader in the development of his/her argument. Therefore, the writer establishes a one to one direct relationship between him/her and the reader.

As the following examples of *Stance* show, the writer is overtly present in both cases:

- (220) **Creo desde mi modesta opinión que** es compatible aplicar medidas dirigidas a reducir el déficit público como ya se está haciendo, eso sí con un calendario más flexible para cumplirlo, con políticas de crecimiento.

[I think, from my humble opinion, that it is possible to apply measures to reduce public deficit, as it is in fact been done, but of course with more flexible schedules and accompanied by growing policies.]

(Letter to the Editor: *Otras políticas sí son posibles*. Sentence 9)

- (221) **Me alegra ver que** todavía hay personas con principios dentro del partido socialista.

[I'm glad to see that there are still people with principles inside the Socialist Party.]

(Letter to the Editor: *La conciencia por encima del cargo*. Sentence 2)

These *Engagement* examples mark how the writer is absent from the sentence but the statement is clear, this strategy engages the reader without any chance of counter-argumentation:

- (222) **Se sabe que** los ingresos hospitalarios aumentan en torno a un 2% anual en los casos de abandono de tratamiento.

[*It is known that hospital admissions increase around 2% per year in cases of treatment abandonment.*] (Letter to the Editor: *Euro por receta*. Sentence 6)

- (223) **Es preciso** bucear también otros fondos que no siempre se quieren analizar con el rigor que merecen y que tienen mucho que ver con el relativismo y el individualismo exacerbado.

[*It should also be interesting to dive in other places that are not always analyzed with the rigor they deserve and have much to do with relativism and exacerbated individualism.*] (Letter to the Editor: *Invertir en ciencia*. Sentence 4)

### 18.4.2.3 PreHead in Spanish Letters to the Editor

The number of PreHeads found in Spanish letters to the editor Corpus is 47.79%.

There is a preference for the use of *PH-Circ* realized by *Prepositional and Adverbial phrases* (40.47%) and *PH-Finite* forms (39.51%) leaving far behind in occurrences *PH-CCL* (19.75%). This reflects the use of simple and direct language, as in the next example of a *PH-Finite*:

- (224) **Soy** músico, estudiante concretamente en un conservatorio de Madrid, llevo más de 10 años tocando la trompeta (un instrumento usado habitualmente en el jazz), soy blanco, he tocado en una big band de jazz, tengo amigos músicos blancos que tocan jazz y conozco casos de músicos no negros que tocan jazz estupendamente como el trompetista Arturo Sandoval, al igual que conozco casos de músicos negros que tocan jazz y clásico de manera extraordinaria como el también trompetista Wynton Marsallis.

[*I'm a musician, a student in a Madrid conservatory, I have been more than 10-year playing trumpet (an instrument commonly used in jazz), I am white, I played in a jazz musicians big band, I have white friends who play jazz and know cases of non-black musicians who play jazz brilliantly as trumpeter Arturo Sandoval, as well as I know cases of black musicians that play jazz and classic jazz dramatically as trumpeter Wynton Marsallis.*]

(Letter to the Editor: *El arte no entiende de razas ni de sexos*.  
Sentence 3)

- (225) **Construyen** AVE minoritarios, aeropuertos sin pasajeros, falsifican empresas, muchos ayuntamientos son nidos prevaricadores.

[*They build AVE for minorities, airports without passengers, fake companies... many municipalities are transgressors nests.*]

(Letter to the Editor: *Deberían sentir miedo*. Sentence 7)

Another example of the use of simple and direct language in the use of *PH-Circ*:

- (226) **Recientemente**, el presidente de la CEOE, Juan Rosell, manifestaba que dado que “existe mucho empleo público que en este momento ni tiene competencias ni trabajo por hacer sin nada que hacer gastando bolígrafos y teléfono, saldría más barato para el país mantener a algunos funcionarios en sus casas cobrando subsidios que en sus puestos de trabajo”.

[Recently, the president of the CEOE, Juan Rosell, stated that since “there is a big amount of public jobs that right now that has not competences or not workload only spending in pens and phone, it would then be cheaper for the country to keep some public workers at home collecting subsidies in their jobs.”]

(Letter to the Editor: *Vigilar a los empresarios*. Sentence 1)

- (227) **Ahora** le toca a Rajoy y al resto de jefes de Estado y de Gobierno de los 11 países pronunciarse en este sentido.

[Now it's up to Rajoy and other heads of state and government of the 11 countries who pronounced themselves in that sense.]

(Letter to the Editor: *Tasa Robin Hood*. Sentence 8)

The functions realized by these PreHeads as Circumstantials express *Time* (24.69%, see example (228)) and *Contingency* (24.69%, i.e.: (229)). They set the action in it's timeframe and introduce counterarguments or hypothetical thesis to the subjects treated in LEs.

- (228) **Desde entonces**, la frase más escuchada era la de que alquilar era “tirar el dinero”.

[Since then, the most well known phrase was that rent was “a waste of money”.]

(Letter to the Editor: *Contra la amnesia*. Sentence 5)

- (229) **Si la crisis ya es un azote** esta se multiplica por cada miembro de una familia.

[If the crisis is already a scourge, it is multiplied by each member of a family.]

(Letter to the Editor: *Cuestión de sentido común*. Sentence 6)

#### 18.4.2.4 Head in Spanish Letters to the Editor

The number of complex clauses in Spanish letters to the editor is of 365, of them 92.88% are thematic clauses and 7.12% rhematic clauses.

A few of these Heads are realized by Thematic Equatives (0.88%, example (230)) and there are no occurrences of Predicated Themes.

- (230) **Lo que los ciudadanos no saben, y ningún grupo político parece estar interesado en contarle**, es que en la actualidad un funcionario puede ser disciplinado en un proceso que se lleva a cabo sin garantías procesales y sin la presencia de un abogado.

[What people do not know, and no political group seem to be interested in telling, is that now an official may be disciplined in a process that is performed without process and without the presence a lawyer.]

(Letter to the Editor: *Medidas preventivas Sentence. 5*)

### Syntactic realizations

Syntactic realizations of Spanish Head in letters to the editor are performed by *Noun Phrases* (41.67%) followed by *Verbs* (33.33%) with a statistical significant difference among them.

**Table 18.27** Syntactic realization of Head

Syntactic forms	%
Np	41.67%
Verb	33.33%
Np+App	15.48%
Interrogative pronouns	5.06%
Clause	2.08%
Other	2.38%
Total	100%

*Noun Phrases* are primarily realized by *Pronouns* in a 49.29% of cases and *Determiners plus Nouns* in a 27.86%, reinforcing the idea of the use of simple and direct language in LEs.

- (231) **Ellos** defienden su “neutralidad”.

[They defend their neutrality.]

(Letter to the Editor: *Las instituciones siguen copadas por los varones. Sentence 6*)

- (232) **Este gesto** ha sido el mejor tributo que ha podido ofrecer a la defensa de la vida.

[This gesture was the best tribute we could offer to the defense of life.]

(Letter to the Editor: *La conciencia por encima del cargo*. Sentence 4)

As it was explained in sections 11.2.1.2 and 11.2.2.1 the part of the verb that acts as Head is the one that carries the grammatical information of the verb as it is shown in the following examples:

- (233) Quizá **habría** menos morosos si los culpables de la crisis dejaran de mirar tanto por sus intereses y se preocuparan, aunque solo fuera un poco, por la clase trabajadora, que no es que no quiera pagar, es que no puede.

[Maybe there will be less debtors if the culprits of the crisis stop looking at their interests and worry, if only slightly, by working class, to see that it's not that they do not want to pay, is that they can not.]

(Letter to the Editor: *Quiénes son los morosos*. Sentence 7)

- (234) Al mismo tiempo, **indica** que hay indicios de la «cooperación» del Gobierno venezolano en la alianza de estos grupos de terroristas y narcoterroristas.

[At the same time, there are signs indicating that the Venezuelan government is cooperating in the alliance of these terrorists groups and narco-terrorists.]

(Letter to the Editor: *El santuario de ETA y las FARC*. Sentence 2)

### Semantic realizations

The most frequent experiential roles used are *Processes* (28.32%), *Actor* (23.30%) and *Carrier* (17.99%).

The language specific characteristic of the use of processes in Spanish is exemplified in:

- (235) **He observado** que muchos de ellos llevan vestimenta de color negro o muy oscura, lo cual hace que no destaquen contra el fondo negro del asfalto.

[I have noticed that many of them wear black clothing or of very dark colour, which does not make them stand out against the black background of the asphalt.]

(Letter to the Editor: *Ciclistas arrollados*. Sentence 3)

- (236) **Imaginemos** una situación muy frecuente, el coche tiene el parabrisas ligeramente polvoriento o sucio, la luz del sol viene de cara y, al dar en la suciedad del parabrisas, provoca que el conductor no vea bien un ciclista que va de oscuro sobre el fondo negro del asfalto o del verde oscuro de la vegetación.

[Imagine a very common situation, the car has the windshield slightly dusty or dirty, sunlight comes in the face and the dirt hit the windshield, causes the driver



does not look good a cyclist going on dark background black asphalt or dark green vegetation.] (Letter to the Editor: *Ciclistas arrollados*. Sentence 4)

Actors providing factual information appear as in:

- (237) **El PIB caerá** más de lo previsto y el número de desempleados puede llegar a superar los seis millones.

[GDP will fall more than expected and the number of unemployed may reach more than six million.]

(Letter to the Editor: *Otras políticas sí son posibles*. Sentence 5)

- (238) **El pasado 30 de noviembre ASPEPC-SPS inició** en el Illustre Col·legi de Doctors i Llicenciats de Catalunya las I Jornadas de Secundaria sobre el fracaso escolar.

[Last November 30 SPS ASPEPC started in the Illustre Col·legi Doctors Catalunya i Llicenciats the First Conference on Secondary school failure.]

(Letter to the Editor: *Las causas del fracaso escolar*. Sentence 3)

And finally, the use of Carriers conflating with admonishing and belief expressions conveyed by the authors:

- (239) **Gaza es** la representación de un odio perpetuo alimentado a lo largo de los siglos, que solo llega a silenciarse al momento de tomar un respiro para, luego, seguir adelante.

[Gaza is the representation of a perpetual hatred fed along centuries, silenced only when to take a breath to, afterwards, go forward.]

(Letter to the Editor: *La violencia en Gaza*. Sentence 1)

- (240) **Guillermo Toledo, tan buen cómico como incoherente intolerante, debía tener** la conciencia intranquila por no haberse manifestado en las últimas horas.

[Guillermo Toledo, such a good comic as an incoherent intolerant, should have uneasy conscience for not having manifested in the last hours.]

(Letter to the Editor: *Willy Toledo y la muerte de Zapata*. Sentence 1)

#### 18.4.2.5 Rheme in Spanish Letters to the Editor

The number of rhematic sentence found Spanish letters to the editor are of 7.12% As in English letters to the editor the cases found are colloquial or rhetorical responses to the development of the argument, as in the following examples:

(241) Y de lo más importante, del incremento del riesgo de los pacientes.

[And of the most important, the increase in patients risk.]

(Letter to the Editor: *Euro por receta*. Sentence 8)

(242) No solo desde un punto de vista académico, sino también desde un punto de vista formativo.

[Not only from an academic point of view but also from the formative side.]

(Letter to the Editor: *Por una educación para los alumnos*. Sentence 6)

### 18.4.3 Contrastive Analysis in Letters to the Editor English vs. Spanish

The following graph shows the number of thematic and rhematic clause complexes in the corpus of letters to the editor in English and Spanish. As in the other two corpora of news reports and editorials, the focus is on thematic clauses.

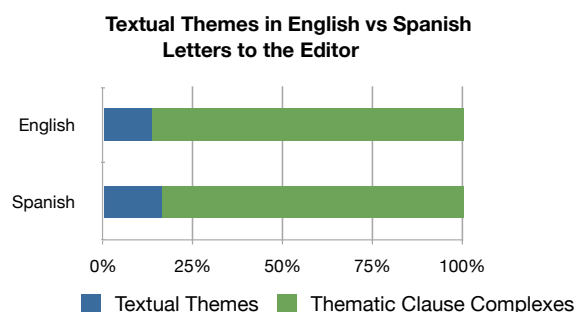
**Table 18.28** Thematic elements in English vs Spanish letters to the editor

L.E.	Thematic Field			
	Outer Thematic Field		Inner Thematic Field	
Types	TT	IT	PH	H
English	13.33%	9.44%	25.83%	100%
Spanish	15.93%	18.88%	47.79%	100%

There is a statistically significant difference in the use of Interpersonal themes and PreHeads among English and Spanish letters to the editor, being in both cases Spanish the one that accounts for more occurrences of these elements.

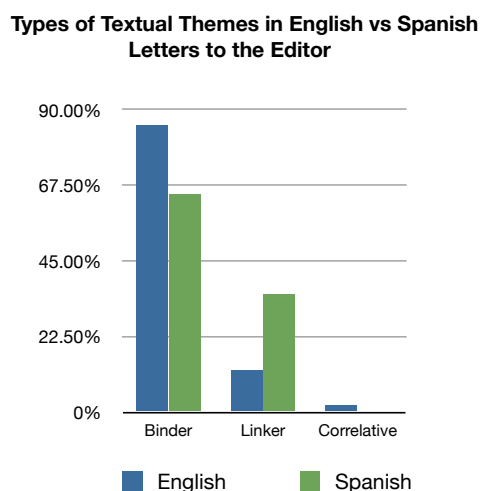
#### 18.4.3.1 Textual Themes: English vs. Spanish in Letters to the Editor

It is relevant to highlight the extension of this type of texts; this is something that affects the use of textual themes. In English the average number of clause complexes per text is of 7.32, and in Spanish of 7.3. Both cases point to short texts where the use of cohesive elements is not very relevant.



**Figure 18.21** Textual Themes in letters to the editor(English vs. Spanish)

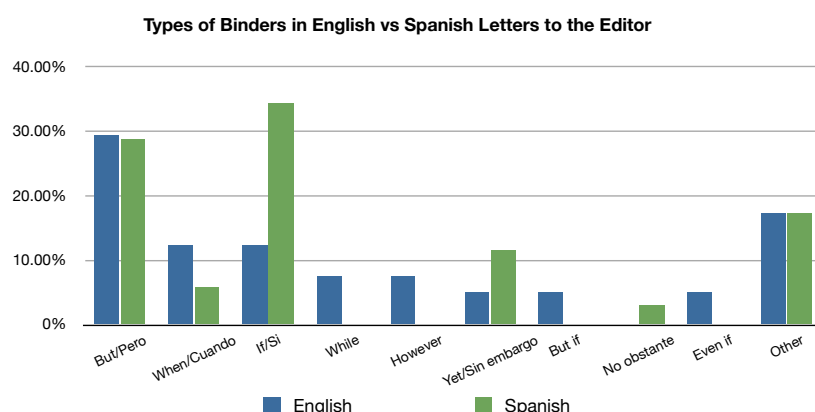
Of these Textual Themes the study of its types reveal that the use of *Binders* and *Linkers* present a statistically significant difference between English and Spanish letters to the editors being *Binders* more used in English (85.42%) than Spanish (64.81%) and *Linkers* more used in Spanish (35.19%) than English (12.5%). (The percentages are calculated over the total number of Textual Themes of each language)



**Figure 18.22** Types of Textual Themes in letters to the editor (English vs. Spanish)

Although this limited number of textual elements in letters to the editor a statistically significant difference between the use of *Binders* (85.42% in English and 64.81% in Spanish) and *Linkers* (12.5% in English and 35.19% in Spanish) is found in both languages, being *Binders* more frequent than *Linkers* may be due to the aim of the Letters, normally containing a complaint of some kind to a previously identified/referred topic.

The following chart shows a comparison of the type of *Binders* used in English and Spanish:



**Figure 18.23** Binders in letters to the editor (English vs. Spanish)

Even though both languages select the use of *Binders* as their preferred Textual Theme, there is a statistically significant difference between both languages, being more abundant its use in English than in Spanish. The choices of the preferred *Binder* elements are *But* (29.27%) in English and *Si* [*If*] in Spanish (34.29%), being *Pero* [*But*] the second preferred choice (28.57%). This contingent vision introduced in Spanish by *Si* is performed in English through the use of a wide range of elements such as *If* 12.20%, *While* (7.32%), *However* (7.32%) and *When* (12.20%).

- (243) **But** we should applaud Lord McAlpine for taking action to discourage thoughtless and irresponsible tweeting.

(Letter to the Editor: *Peer brings twitter to hell*. Sentence 2)

- (244) **Pero** estaría bien que a la vista de las últimas noticias sobre supuestos pagos en negro a trabajadores de sus horas extras, evitando así la correspondiente cotización a la Seguridad Social, alguien preguntara al señor Rosell si no sería más conveniente que en lugar de dejar en casa a los funcionarios supuestamente ociosos los pusiéramos a vigilar a los empresarios supuestamente tramposos, para que cumplan con sus obligaciones para con los trabajadores y para con el resto de la ciudadanía.
- [*But it would be nice to in view the latest news about alleged payments to workers on their overtime in black money, avoiding the corresponding quote Social Security, someone will ask Mr. Rosell if it would not be better instead of leaving home idle public workers, we set them to watch the supposedly cheater entrepreneurs to meet their deeds to workers and to the rest of the citizenry.*]

(Letter to the Editor: *Vigilar a los empresarios*. Sentence 2)

- (245) **Si** hacemos un pequeño resumen de todo lo acontecido últimamente, veremos que aquí de miedo, nada de nada, o no hay viña o lo tienen todo atado y bien atado.

[If we make a small summary of everything that has happened lately, we see that here there is no fear here, or there is no vineyard or he has all well buttoned up.]

(Letter to the Editor: *Deberían sentir miedo*. Sentence 6)

- (246) **While** the fungus that causes ash dieback may be windblown up to a point, the greatest threat to our native broadleaves comes from lax quarantine rules in garden centres and tree nurseries. (Letter to the Editor: *Native ash tree hope*. Sentence 1)

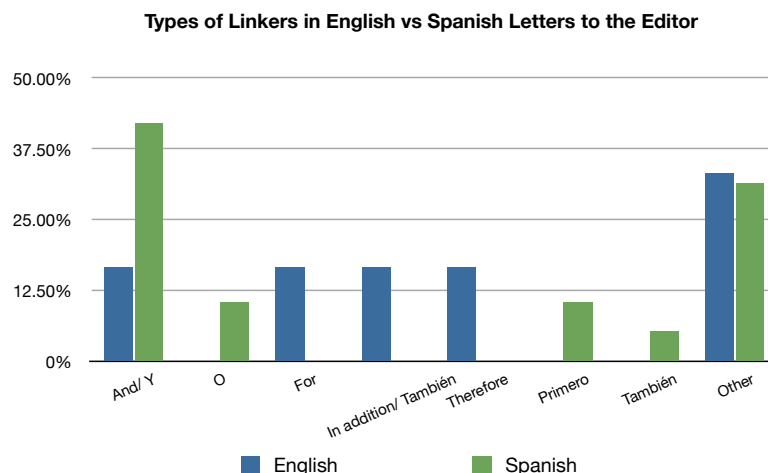
- (247) **If** any homosexual couple and their guests were willing to sit through that for the 45 minutes or more of my average-length sermon, I suppose they would have earned the right to the piece of paper that entitles them to the very few and steadily diminishing privileges still conferred by the legal state of marriage in this country.

(Letter to the Editor: *Gay wedding in church? Are you sure?*. Sentence 5)

- (248) **However**, most road users accept that these crossings are the safest way for a cyclist to cross a busy road, and if dismounted cyclists cross with the care exercised by most pedestrians, they are unlikely to bother anyone.

(Letter to the Editor: *Correct etiquette for cyclists crossing the road*. Sentence 11)

The scarce use of *Linkers* in both languages is grouped in Spanish under *Y* [And] (42.10%) whereas English uses *and*, *in addition*, *therefore*... (65.8%)



**Figure 18.24** Linkers in letters to the editor (English vs. Spanish)

- (249) **Y** me consta que los trámites para conseguirla no eran fáciles ni rápidos: avales, nóminas, otras propiedades, etcétera.

[*And I know that the procedures to get it not easy nor fast: sureties, payroll, other properties, etcetera.*] (Letter to the Editor: *Contra la amnesia*. Sentence 7)

- (250) **And** for me this all goes far beyond the tax-avoidance issue, with economic life now seeming to have become a farcical game of monopoly, with the finance industry and corporations calling the shots and with both Brussels and national governments happily dancing to their tune.

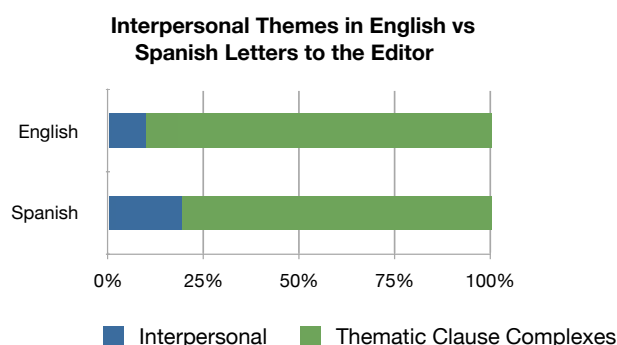
(Letter to the Editor: *Water companies call the shots*. Sentence 10)

- (251) **Therefore**, the problem, to which the Chief of the Defence Staff alludes, is not so much the size of the Army but how, in an expeditionary age, other than in very small packets, expeditionary forces can safely and speedily be transported to where they are needed, and provided with adequate air and sea power support in-theatre.

(Letter to the Editor: *Overburdened forces*. Sentence 3)

#### 18.4.3.2 Interpersonal: English vs. Spanish in Letters to the Editor

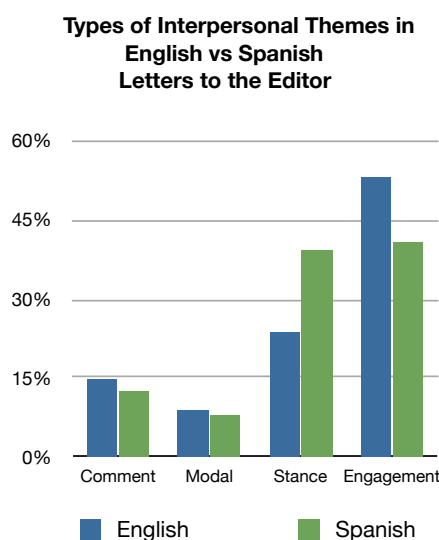
There are 9.44% occurrences of Interpersonal Themes in English and 18.88% in Spanish. There is a statistically significant difference between the number of Interpersonal Themes used in Spanish and the number of these elements used in English.



**Figure 18.25** Interpersonal Themes in letters to the editor (English vs. Spanish)

Regarding the types of Interpersonal themes used in English, as seen in section 18.4.1.2, there is a clear tendency towards the use of *Engagement* strategies with a statistical significant difference over the rest of the categories: 52.94%; *Stance* 23.53%; *Comment* 14.71% and *Modal* 8.82%.

Spanish types, on the contrary, share the weight of Interpersonal Themes types evenly between *Stance* and *Engagement* (39.06% and 40.63% respectively). Thus Spanish writers seem to show themselves in the text more than English writers do, who tend to address the reader more rather than showing themselves in the texts.



**Figure 18.26** Types of Interpersonal Themes in LEs (English vs. Spanish)

English *Engagement* strategies:

- (252) **It is clear that** many problems would arise if the legislation as now tabled were to be implemented.

(Letter to the Editor: *Settle this Plebgate stand-off now.* Sentence 7)

- (253) **It is time for** the Director of Public Prosecutions to draw a line under the affair, by charging Mr Mitchell, based on what evidence is available, and co-charging any officer who it is believed may have lied in the course of their duty.

(Letter to the Editor: *Traditional institution of marriage is central to the welfare of society.* Sentence 6)

Spanish *Stance* strategies:

- (254) **Creo desde mi modesta opinión que** es compatible aplicar medidas dirigidas a reducir el déficit público como ya se está haciendo, eso sí con un calendario más flexible para cumplirlo, con políticas de crecimiento.

*[I think, from my humble opinion, that it is possible to apply measures to reduce public deficit, as it is in fact been done, but of course with more flexible schedules and accompanied by growing policies.]*

(Letter to the Editor: *Otras políticas sí son posibles.* Sentence 9)

- (255) **Me alegra ver que** todavía hay personas con principios dentro del partido socialista.

*[I'm glad to see that there are still people with principles inside the Socialist Party.]*

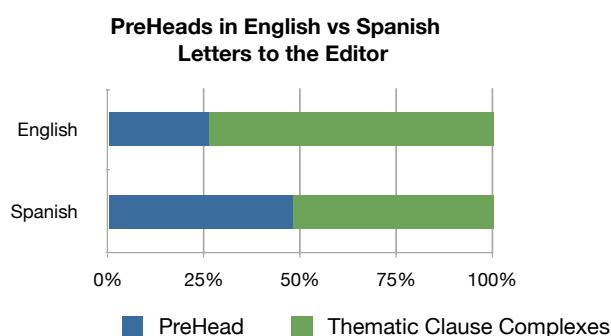
(Letter to the Editor: *La conciencia por encima del cargo.* Sentence 2)

Biber and Finegan (1989) suggest as Pounds (2010) comments in her work *Writer's argumentative attitude: A contrastive analysis of 'Letters to the Editor' in English and Italian* that English “stance” may be more often “integrated into text rather than overtly marked”. This aligns with the results found in the present study along with the idea pointed out in Pound's work in relation to the Italian culture, that can be applied here also to Spanish, and that is that writers argumentation retain more traces of oral language being thus the speaker's or writer's presence more visible in the text.

#### 18.4.3.3 PreHead: English vs. Spanish in Letters to the Editor

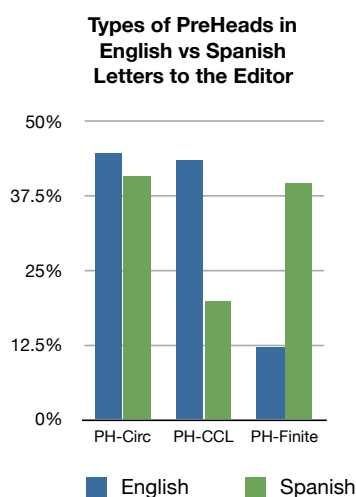
The number of PreHead used in the letters to the editor English Corpus is of 25.83% and in Spanish 47.79%. The difference in the use of these elements is statistically significant.





**Figure 18.27** PreHead in letters to the editor (English vs. Spanish)

The distribution of PreHeads in the types analyzed are displayed in the following graph, according to the labels explained in section 11.2.1.



**Figure 18.28** Types of PreHeads in letters to the editor (English vs. Spanish)

In English there is no statistical preference for *PH-CCL* (43.48%) or *PH-Circ* (44.57%) whereas Spanish shows a preference for *PH-Circ* (40.74%) and *PH-Finite* (39.51%) PreHead types instead of *PH-CCL* (19.75%). This means that Spanish relies on simple syntactic structures as *Prepositional Phrases*, *Adverbial Phrases* or *Verbs* in PreHead position, while English uses more complex elements in the expression of PreHeads, as in:

- (256) **Whatever people thought of the merits or otherwise of that election**, the majority wanted to vote but knew nothing about the candidates – a particularly significant point when so many were concerned that the police should not become

politicised. (Letter to the Editor: *How PCC electors were kept in the dark.*  
Sentence 4)

(257) **While privacy laws may have been compromised in the telephone call,**  
the act itself is a staple of popular radio anywhere in the Western media.

(Letter to the Editor: *Bullies and victims in the hoax call row.*  
Sentence 4)

Contrasting with the simplicity of Spanish *PH-Circ* and *PH-Finite* in the following examples:

(258) **Desde entonces,** la frase más escuchada era la de que alquilar era “tirar el dinero”.  
[*Since then, the most well known phrase was that rent was “a waste of money”.*]

(Letter to the Editor: *Contra la amnesia.* Sentence 5)

(259) **Construyen AVE** minoritarios, aeropuertos sin pasajeros, falsifican empresas, muchos ayuntamientos son nidos prevaricadores.

[*They build AVE for minorities, airports without passengers, fake companies... many municipalities are transgressors nests.*]

(Letter to the Editor: *Deberían sentir miedo.* Sentence 7)

Most of these *PH-CCL* and *PH-Circ* are used to set the timeframe of the event related or to express contingency which, the used to expresses counter argument or a hypothetical situations to be considered and thought over by the reader.

#### 18.4.3.4 Head: English vs. Spanish in Letters to the Editor

The corpus of English letters to the editor is of 366 complex clauses and 365 in Spanish. Of these clause complexes 98.36% in English are Thematic clauses and 92.88% in Spanish. The number of Rhematic clauses in each corpus is of 1.64% in English and 7.12% in Spanish.

The use of Thematic Equatives and Predicated Themes are very few, Thematic Equatives are used in 0.55% in English (261) and 0.88% in Spanish (262). Regarding Predicated Themes there are 0.27% occurrences in the English corpus (260) but not in the Spanish one.

(260) **It is only when Gaza prospers that** those Palestinians wanting to hit Israeli targets will no longer be tolerated.

(Letter to the Editor: *Prosperity and peace for Gaza.* Sentence 4)

Thematic Equatives in English and Spanish:

- (261) **What distinguishes mankind now from what we were then, for good or ill,** is science and technology.

(Letter to the Editor: *Intellectual vacuum*. Sentence 3)

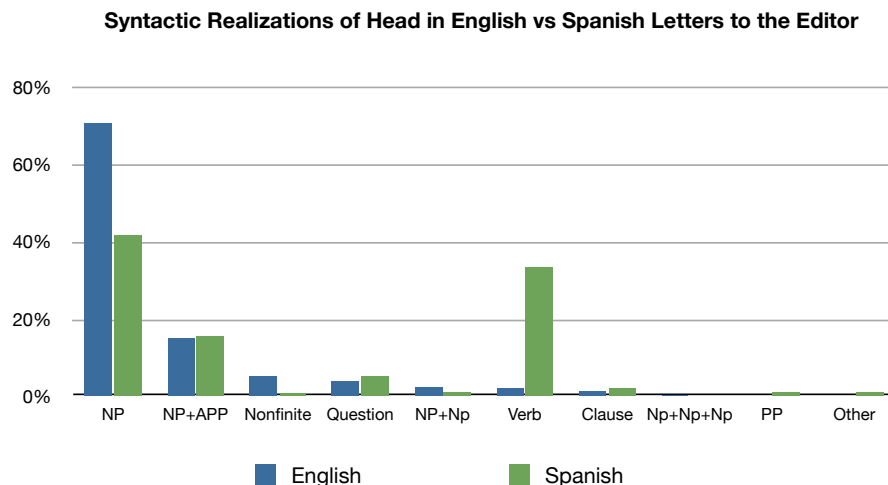
- (262) **Lo que los ciudadanos no saben, y ningún grupo político parece estar interesado en contarle,** es que en la actualidad un funcionario puede ser disciplinado en un proceso que se lleva a cabo sin garantías procesales y sin la presencia de un abogado.

[What people do not know, and no political group seem to be interested in telling, is that now an official may be disciplined in a process that is performed without process and without the presence a lawyer.]

(Letter to the Editor: *Medidas preventivas*. Sentence 5)

### Syntactic realization

The following chart shows the most frequent syntactic realization of Head in English and Spanish letters to the editor.



**Figure 18.29** Syntactic realisation of Heads in letters to the editor (English vs. Spanish)

Both English and Spanish select as more frequent the usage of *Noun Phrases* (English 70.56% and Spanish 41.67%) but they present a statistically significant difference between them.

The second choice in Spanish is the use of *Verbs* with 33.33%, on the contrary English uses *Noun Phrases followed by Appositions* 15.00%.

Taking a look into these data the use of *Verbs* in Spanish is realized by the grammatical part of the verb, that is, person and tense.

The most frequently used tense is the *Simple Present* in the *3rd person of the singular* (42.30%):

As in:

- (263) Nunca **podrá quedar** tan en evidencia la falta de sentido común como en este caso.

[It could never be so apparent meaninglessness the lack of common sense as in this case.]

(Letter to the Editor: *Un cierre incomprensible*. Sentence 2)

The next preferred person is the *1st person singular* (30.76%). In it, the author reveals him/herself with a direct voice.

- (264) **Creo** que por fin la mayoría de los catalanes está desenmascarando al independentismo.

[I think at last most Catalans are unmasking the independence movement.]

(Letter to the Editor: *El fracaso del independentismo*. Sentence 2)

The last most relevant form of the verb selected in Spanish is the *3rd person plural of Simple Present* (15.38%):

- (265) Recientemente **han desfilado** por los platós un ex director general de la Guardia Civil y un ex alcalde de Marbella corruptos; un bailar flamenco que atropelló mortalmente a un peatón y se dio a la fuga; incluso maltratadores, famosos o sin serlo, que a los pocos días de ser entrevistados por morbosos presentadores han seguido acosando a sus víctimas.

[Recently there have paraded by the TV sets a former general director general of the Guardia Civil and a former corrupt mayor of Marbella, a flamenco dancer who ran mortally a pedestrian and fled the scene, even abusive, famous or not, that a few days of being interviewed by morbid presenters have continued to harass their victims.]

(Letter to the Editor: *Las televisiones siempre un paso por detrás*. Sentence 2)

Let's take a closer look at the Thematic Head realized by the part of the verb carrying person and tense information of the verb in English, to do so, it is important to take into account that when verbs appear in English functioning as Heads they do not do it in Declarative sentences but in Directives, as in:

- (266) **Fund** candidates to explain their policies, and hold the voting in May when many families have a walk out after tea to vote – instead of looking out of the window, deciding it’s too cold, and staying in to watch the football instead.

(Letter to the Editor: *A disappointing day for democracy*. Sentence 7)

Or in *Let* constructions:

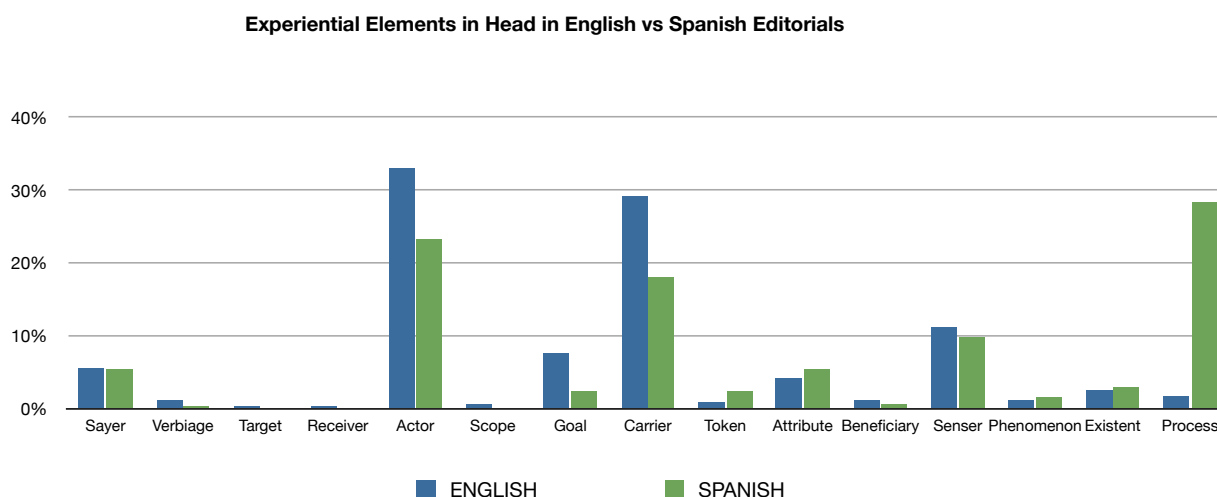
- (267) **Let** us hope that, for once, lessons really have been learnt by officialdom.

(Letter to the Editor: *How PCC electors were kept in the dark*. Sentence 7)

The type of sentences where the verb appears in first position in English is not a Declarative clause, so its function does not correspond between one language and the other, the function that Spanish performs with the verb in first position in Declarative sentences is the one that English does with the use of *Pronouns* as Head, in fact the use of *Pronouns* is of a 44.44% of the total number of *Noun Phrases* functioning as Heads.

### Semantic realization

The following graph shows the Experiential roles used in English and Spanish letters to the editor Corpus:



**Figure 18.30** Semantic realization of Heads in letters to the editor (English vs. Spanish)

As explained in the previous section, processes are used in Spanish Processes:

- (268) **Tell** that to my parents as they huddled in the kitchen of their cramped council flat in north London in the 1950s and 1960s.

(Letter to the Editor: *BBC inspires all classes*. Sentence 2)

- (269) **He observado** que muchos de ellos llevan vestimenta de color negro o muy oscura, lo cual hace que no destaquen contra el fondo negro del asfalto.

[*I have noticed that many of them wear black clothing or of very dark colour, which does not make them stand out against the black background of the asphalt.*]

(Letter to the Editor: *Ciclistas arrollados*. Sentence 3)

Actors:

- (270) **These multinational companies come** to Britain to enjoy the benefits of trading in a secure, politically stable environment, yet do their very best to weasel out of making any realistic contribution to the upkeep of their host nation or the education and welfare of its people.

(Letter to the Editor: *Put Starbucks out of its misery*. Sentence 3)

- (271) **El pasado 30 de noviembre ASPEPC-SPS inició** en el Illustre Col·legi de Doctors i Llicenciats de Catalunya las I Jornadas de Secundaria sobre el fracaso escolar.

[*Last November 30 SPS ASPEPC started in the Illustre Col·legi Doctors Catalunya i Llicenciats the First Conference on Secondary school failure.*]

(Letter to the Editor: *Las causas del fracaso escolar*. Sentence 3)

Carriers:

- (272) **This country has** no link with purple. (Letters, November 19).

(Letter to the Editor: *No purple pride*. Sentence 1)

- (273) **Gaza es** la representación de un odio perpetuo alimentado a lo largo de los siglos, que solo llega a silenciarse al momento de tomar un respiro para, luego, seguir adelante.

[*Gaza is the representation of a perpetual hatred, fed along centuries, silenced only when to take a breath, to, afterwards, go forward.*]

(Letter to the Editor: *La violencia en Gaza*. Sentence 1)

#### 18.4.3.5 Rheme: English vs Spanish Letters to the Editor

Both English and Spanish letters to the editor have more occurrences of rheme than the other types of texts, due to the more colloquial style developed in the Letters.

Although Spanish accounts for more apparitions of rhematic expressions (7.12%) than English (1.64%) there is no significant difference between them. Both types of rhemes respond to the same instances of colloquialism and rhetoric responses in the development or their arguments, as in can be seen in the following examples:

(274) No change, no chance.

(Letter to the Editor: *After the EU speech and gay marriage, what next for Cameron?*. Sentence 3)

(275) No solo desde un punto de vista académico, sino también desde un punto de vista formativo.

[*Not only form an academic point of view but also from the formative side.*]

(Letter to the Editor: *Por una educación para los alumnos*. Sentence 6)

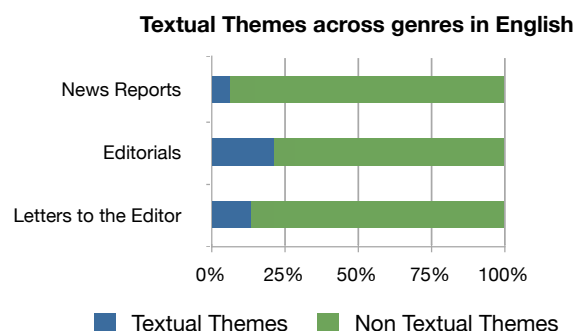
## 18.5 Thematic Characterization from a Genre Perspective: news reports, editorials and Letters to the Editor

In this section the analysis focuses on news reports, editorial and letters to the editor from a contrastive genre perspective. The section is organized following the same structure as the previous ones treating each type of thematic element separately. For the sake of clarity a further division according to language will be made.

The Corpus of study as described in chapter 16 is composed of three different genres: news reports, editorials and letters to the editor in English and Spanish.

### 18.5.1 English and Spanish Textual Themes Across Genres

**In English** Textual Themes are displayed in the different genres as shown in the following chart:



**Figure 18.31** Textual Themes across genres in English

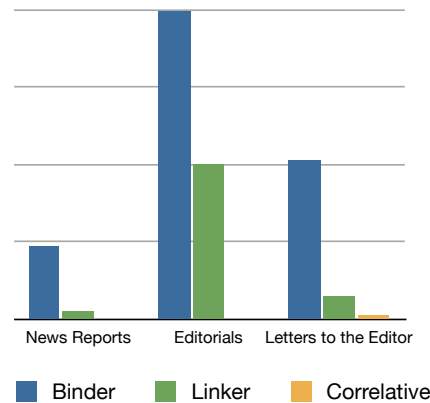
As shown in the chart above the number of Textual Themes found in News reports is of 6.21% of the total number of thematic clause complexes in the corpus, in editorials this amount increases up to 21.13%, and letters to the editor present 13.33%. The data show that there is a statistical significant difference ( $P < 0.0001$ ) in the distribution of these elements among the three types of texts, this rejection of the null hypothesis points to a relation between the use of these elements and the genre of the text in which they are used. The amount of elements used in editorials is the biggest of the three, the nature of these elements and the extension of the texts (average extension of editorials is of 19.56 complex clauses per text, News reports 11.63 and letters to the editor 7.32) makes the distribution of textual themes organized as the chart above shows.

As McCabe & Heilman (2007) correctly comment in their study on Textual Themes, *Binders* are not “innocent” regarding interpersonality, and, as they say, these elements are not mere linkers of ideas but in them the writer inserts his/her opinion in the text, performing a blending function of conjuncts and subjective markers.

The results depicted in the chart 18.32 reflect this view of editorials and letters to the editor corresponding with a higher number of Textual Themes of the type *Binders* used, whereas in news reports the occurrences of these elements is less.



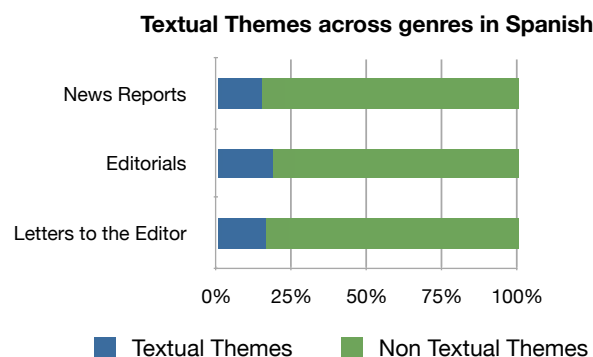
Types of Textual Themes across genres in English

**Figure 18.32** Types of Textual Themes across genres in English

In the two relevant types: *Binders* and *Linkers* the null hypothesis has been rejected with a P value of less than 0.0001. The distribution of *Binders* in number in all genres is always greater than *Linkers* increasing proportionally in relation to the extent of the text, from news reports to letters to the editor. Again the fact that editorials are longer than the rest of the texts seem to be a determinant factor for this numeric distribution of elements and in terms of percentages and in relation to the use of both types of Textual Themes, letters to the editor are the ones that use in a smaller amount of text (7.32 clauses per text) a greater number of Textual elements, and what is more, of *Binders*, this strategy allows the writer to introduce counter arguments and hypothesis.

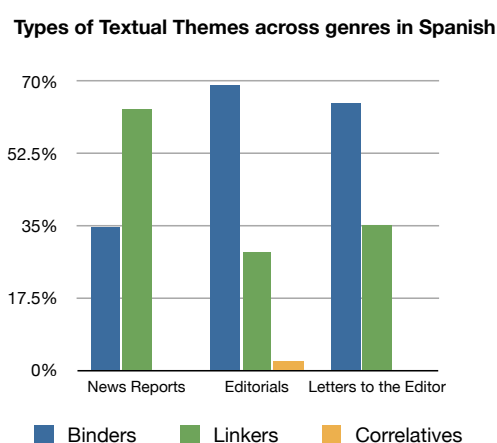
Another factor that is relevant and worth mentioning is the high amount of *Binders* in news reports, this fact does not seem to respond to a any sensible explanation in relation to genre but could rely on a language characteristic, as it will be study in the following lines.

**In Spanish** the distribution of Textual Themes among the three genres does not present a statistical difference according to the results of the chi-square test. In fact the percentages are almost evenly distributed among news reports (14.69%), editorials (18.23%) and letters to the editor (15.92%). Although the number of Textual Themes does not present a significant difference the distribution of its types does, as the P value is less than 0.0001.



**Figure 18.33** Textual Themes across genres in Spanish

The use of *Linkers* in Spanish is necessary to the development of the text, in fact its use is relevant in all three genres. Of the total number of Textual Themes used in news reports 63.46% were *Linkers*, in editorials 28.57% and letters to the editor account for 35.19%. There is always a relevant number of these elements to mark the continuative flow of the text. This feeling is most relevant in news reports where a 63.46% of Textual Themes are dedicated to this matter, helping in the account of facts in this genre. Editorials are, on the other hand, where less amount of *Linkers* are used highlighting its nature in the preference shown for *Binders*. Letters to the editor is a special case where the extension of the text is the least of the three (avg # of clauses per text: letters to the editor 7.30, editorials 17.5 and news reports 11.69) and, as news reports, the number of *Linkers* is relevant in the overall amount of Textual Themes in the texts 35.19%.

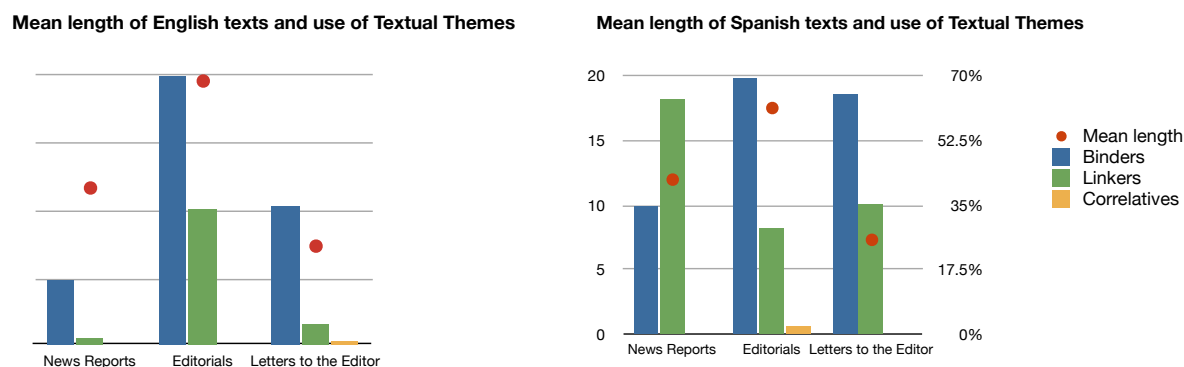


**Figure 18.34** Types of Textual Themes across genres in Spanish

*Binders* show also a statistical difference among the three genres. Editorials

(69.23%) and letters to the editor (64.81%) prefer the use of these elements to news reports (34.62%). These cohesive elements introduce the counter arguments representative of editorials and letters to the editor.

**English and Spanish Textual Themes** The following charts show a comparison between the use of English and Spanish types of Textual Themes. The red dots indicate the average length of each type of text in the aim of displaying the relation of the extent of the texts with the proportional distribution of these types of Textual Themes.



**Figure 18.35** Mean length of English and Spanish texts and use of Textual Themes

The use of *Linkers* in Spanish is greater than in English in all cases, whereas English seems to avoid the use of these elements, there is no substitution strategy for them just the continuity of sentences as in example (276), the idea of evolution of the text is taken for granted or is not signalled through any other visible mean. As many studies have shown Vogel (2008), Shaw (2009), Tse & Hyland (2006), Hyland & Bondi (2006), Charles et al. (2011), the use of *Linkers* by native English speakers is scarce.

The clearer example on this aspect is shown in news reports where the amount of *Linkers* used in English is of 9.52% and in Spanish 63.46%.

- (276) **The decision** was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remodel and regulate capitalism in the wake of the financial crisis.

**French commentators** hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is

now behind us. We love each other as much as we did on the first day.”

(Report: *El FMI anuncia ayudas financieras para Ucrania y Hungría.*  
Sentences 10,11)

(277) **Además**, el director gerente del organismo, Dominique Strauss-Kahn, ha prometido una “importante ayuda financiera” para Hungría con el objetivo de estabilizar la economía de este país centroeuropeo.

[In addition, the managing director of the agency, Dominique Strauss-Kahn, has promised a “major financial aid” for Hungary in order to stabilize the economy of this Central European country.]

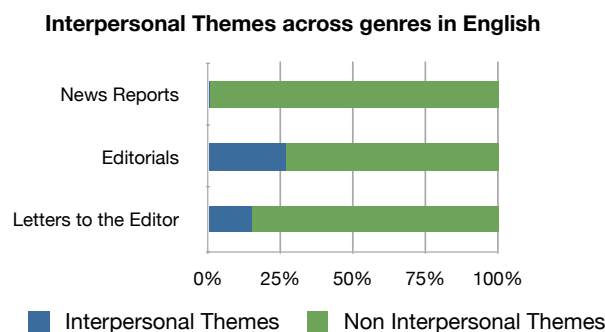
(Report: *El FMI anuncia ayudas financieras para Ucrania y Hungría.* Sentence 2)

The strategies adopted by both languages in the use of these elements is quite different. English tends to use *Binders* leaving aside the use of *Linkers* as the examples above show, probably because they do not add any extra information but rather present the continuative flow of information, meanwhile the introduction of *Binders* contrast the previous information, therefore these elements do add something new about the relationship among clauses.

### 18.5.2 English and Spanish Interpersonal Themes Across Genres

**English** does present an extremely significant difference in the use of Interpersonal Themes across genres ( $P < 0.0001$ ).

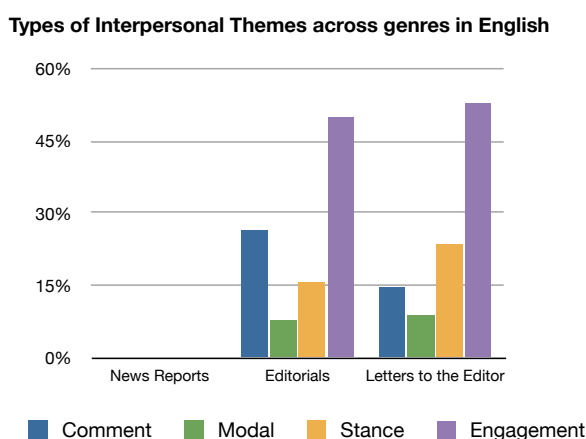
Francis (1990) in his study also found a higher frequency of Interpersonal Themes in editorials and Letters than in news reports, as the first two genres are more “personal and evaluative, as well as more concerned with persuasion than information, therefore it is likely to be more direct interaction between writer and reader”



**Figure 18.36** Interpersonal Themes across genres in English

News reports, as it was seen on section 18.2.1.2, does present only one occurrence of Interpersonal Theme whereas editorials and letters to the editor increase the use of these numbers up to 14.71% in letters to the editor and 26.56% in editorials.

In relation to the types of Interpersonal Themes, the following chart presents their distribution in the different genres:



**Figure 18.37** Types of Interpersonal Themes across genres in English

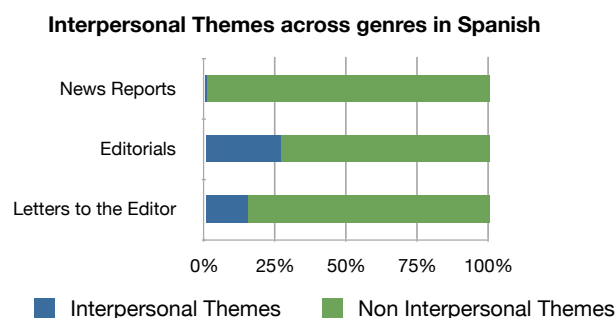
According to the definition given on section 11.1.2, *Comments*:

Express the speaker or writer's judgement or his/her attitude to the content of the message. (Halliday & Matthiessen, 2004:81)

These elements present a significant relevant difference in its use through the three genres, being more frequent in editorials than Letters to the Editor, this responds to the nature of these texts as "...opinionated genres intended to persuade the reader" (Biber, 1988)

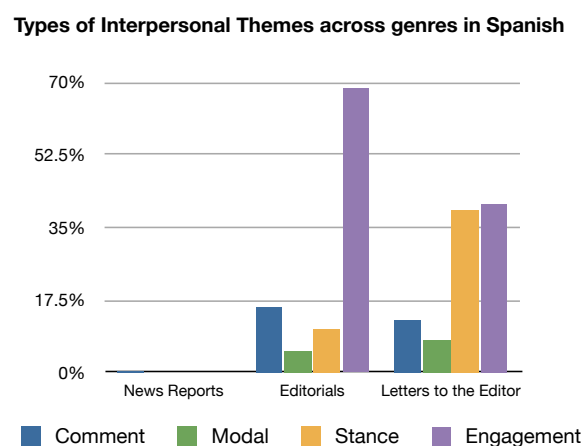
On the other hand, the use of *Engagement Strategies* in both editorials and letters to the editor is evenly high (editorials 50.00% and letters to the editor 52.94%) with no difference in its use among them. The same happens with the use of *Stance* (editorials 15.63% and letters to the editor 23.53%) and *Modal* (editorials 7.81% and letters to the editor 8.82%).

**Spanish** does also present a significant difference in the use of Interpersonal Themes across genres ( $P < 0.0001$ ).



**Figure 18.38** Interpersonal Themes across genres in Spanish

News reports only have one occurrence of this type of Interpersonal elements. But the use of *Engagement* (editorials 68.42% and letters to the editor 40.63% ) and *Stance* (editorials 10.53% and letters to the editor 39.06%) present a extremely significant difference and a very significant difference respectively, amongst the three genres. On the other hand the use of *Modals* and *Comments* does not present a significant difference in its distribution.



**Figure 18.39** Types of Interpersonal Themes across genres in Spanish

Letters to the editor are characterized as carriers of *Stance* elements, described by Hyland as “adopting a point of view in relation to both the issues discussed in the text and to others who hold points of view on those issues.” This means that the author shows off in the text clearly, presenting him/herself as the responsible for the opinions in the text.

On the other hand, editorials as “the voice of the newspaper” are characterized

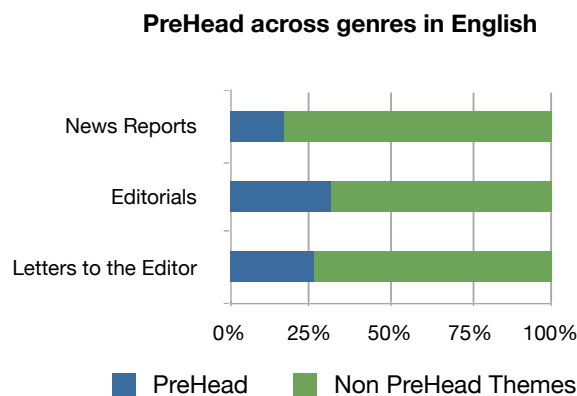
by the use of *Engagement* strategies conveying opinions but not directly, in Hyland's words (2005:176):

This is an alignment dimension where writers acknowledge and connect to others, recognizing the presence of their readers, pulling them along with their argument, focusing their attention, acknowledging their uncertainties, including them as discourse participants, and guiding them to interpretations.

**English vs. Spanish** Spanish presents more variation in terms of statistical differences in the distribution of the types of Interpersonal elements. In fact the variation in English sticks to the use of *Comments* whereas Spanish shows distributional discrepancies in *Engagement* and *Stance* strategies. Expressions of interpersonality are more thematic in Spanish than in English, that means that English does not present in the theme the marks of interpersonality as much as Spanish does.

### 18.5.3 English and Spanish PreHead Across Genres

**English:** There is a statistical significance difference in the use of PreHeads across genres.

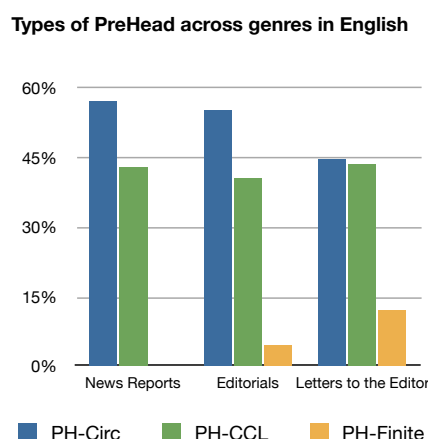


**Figure 18.40** PreHead across genres in English

As the figure above shows there is an increase in the use of PreHeads from the formal writing of news reports (16.57%) to letters to the editor (25.83%) and editorials (31.33%). Although there seems to be a direct relation between the subjectivity of the genre with the use of these elements, the reason may be related to the need of these two types of texts to set a frame to the story accounted giving more information to help

the reader understanding the argumentation whereas in news reports these elements are used to set time and place as pure informative elements.

In the following chart there is a detail of the types of PreHeads found:

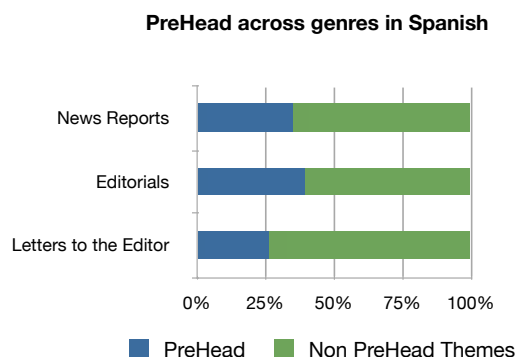


**Figure 18.41** Types of PreHead across genres in English

The difference relies mostly in the PH-Finite type of elements that in English are represented by directive and imperative sentences rather than declaratives, these sentences are more typical of non-formal language thus the use of them increases from its scarce use in news reports (0%) then editorials (4.49%) and finally letters to the editor (11.96%). This is a distinguishing factor among the three genres, something that does not happen with the use of PH-Circ and PH-CCL. Its use is evenly distributed between news reports (PH-Circ 57.14% and PH-CCL 42.86%) editorials (PH-Circ 55.06% and PH-CCL 40.45%) and letters to the editor (PH-Circ 44.57% and PH-CCL 43.48%).

**Spanish:** presents more variation in the choices of these elements, with a statistically significant difference between PreHeads among genres (fig. 18.42) as well as an extremely statistical difference amidst the three types of PH: PH-Circ, PH-CCL and PH-Finite (fig. 18.43)



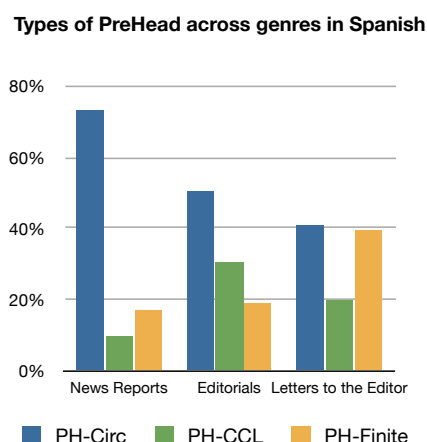


**Figure 18.42** PreHead across genres in Spanish

PH-Circ set the frame of the text through the use of simple syntactic forms, as mentioned in section 11.2.1, these forms present a statistically significant difference among the three genres. news reports (73.17%) are the ones that use them more frequently, setting Time and Place of the event, editorials follow them with an amount of 50.52% and letters to the editor use up to 40.74% (see sections 18.2.2.3, these two types of texts use PH-Circ for other purposes 18.3.2.3 and 18.4.2.3 for more details).

PH-CCLs reflect a higher syntactic complexity, its use increases in editorials (30.41%) and in letters to the editor (19.75%) whereas in news reports there is not much presence of these elements (9.76%). It is interesting to note that although the length of letters to the editor is lower than editorials (letters to the editor 7.30 and editorials 17.5) the use of PH-CCL in letters to the editor is in proportion very high revealing syntactic density in this genre.

A special case that differs from English is the use of PH-Finite forms realized by the lexical part of the verb, the use of imperatives and questions. These elements present a statistically significant difference in their realisation across genres. The use of these forms increases as the informality of the language used rises in relation to the type of texts, thus: news reports (17.07%) editorials (19.07%) and letters to the editor (39.51%)



**Figure 18.43** Types of PreHead across genres in Spanish

As the graph shows it can be said that the syntactic complexity of PreHeads in Spanish is related to genre, realized by PH-CCL and PH-Finite, using more objective language implies the use of less number of PreHeads and less complexity of the ones used, at the same time PH-Circ decreases in the opposite direction: less used when the language tends to be less formal.

**English vs. Spanish:** The contrast of English vs Spanish on the distribution of these elements differ only in the different uses of PH-Finite, and this, as explained above is a language specific characteristic.

The most relevant function that PreHead perform in both languages is the expression of *Time* in all genres and both languages, the idea of immediateness is present in the three genres as Newspapers are related to the account of recent facts from different perspectives, news reports take care of facts that have just happened (Clark, 1975), editorials present the opinion of the Newspaper on recent news events (Van Dijk, 1989) and letters to the editor express the reader's opinion on hot topics (Vázquez y del Árbol, 2005).

There is a contrast ( $P < 0.0001$ ) in the use of *Place* and *Contingency* expressions in both languages ranging, in the case of *Place* to a relevant use in news reports 14.29% in English and 7.32% in Spanish to an almost non existent appearance in letters to the editors (English 3.23% Spanish 0.62%) (in editorials they express *Place* 2.81% in English and 7.73% in Spanish). This fails within the nature of this type of texts as news reports give account of issues that have to be set in *Place* and *Time*.

*Contingency* on the contrary introduces counter arguments or hypothesis, more

used in editorials (English 30.90% Spanish 21.13%) and letters to the editor (English 25.81% Spanish 16.67%) than in news reports (English 5.36% Spanish 8.94%). New Reports for this same purpose use *Quotations* as the justification of their reasoning.

*Evidence* (English 10.71% Spanish 13.82%) and *Quotations* (English 19.64% Spanish 0.81%) appear only in a remarkable amount in news reports and are almost inexistent in the other two genres. Through these strategies news reports approach reliable sources to present objective and identifiable points of view (more occurrences are found in English than in Spanish as explained on sections 18.2.1.3, 18.2.2.3 and 18.3).

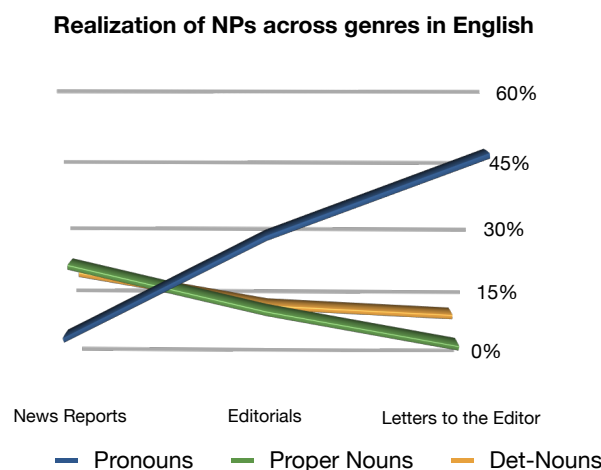
#### 18.5.4 English and Spanish Head Across Genres

As it was pointed out in several occasions along this presentation of results, all thematic clauses analyzed have *Head*, therefore the focus will fall upon its syntactic and semantic realizations.

##### Syntactic realizations

**English** presents a statistically significant difference in the distribution of the use of Noun Phrases and Noun Phrases followed by Appositions across the three genres.

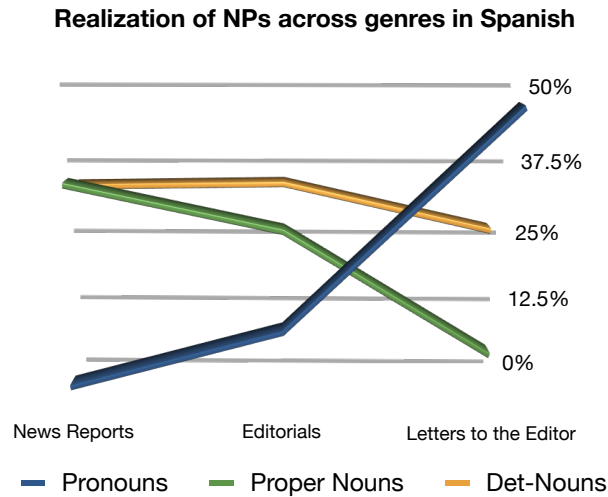
Noun Phrases are realized mainly by *Pronouns*, *Proper Nouns* and *Determiner plus Noun* formations. These are found with a statistically significant distribution among the three genres. As the following figure (18.44) shows the use of *Pronouns* increases from news reports (14.47%), to editorials (35.48%) and letters to the editor (51.57%) being this tendency contrary to what happens with *Proper Nouns* (news reports 27.19%, editorials 17.19% and Letters to the Editor 9.54%) and *Determiner plus Nouns* formations (news reports 22.81%, editorials 15.40% and Letters to the Editor 13.39%). This seems to be related to the type of language used in the different languages. The objectivity of news reports makes direct reference through the use of *Proper Nouns* to the subject of the news. Editorials, being longer in extension use both less *Proper Nouns* and more *Pronouns* to refer to an already mentioned entity. Finally, letters to the editor use more *Pronouns* than the other two texts because they refer to an already known person or fact, that has been been stated with a high probability in the headline. This type of texts, as its extension is very short, does not introduce new entities in the text, but refer all the time to the same known subject.



**Figure 18.44** NPs realization across genres in English

*Noun Phrases followed by Appositions (Np-App)* are widely used in editorials (22.36%) and news reports (23.96%) being less common in letters to the editor (15.00%). This seems to confirm the use of *Nouns followed by Appositions* as adders of information to help in the identification of the subject. This can be done in longer texts, being this strategy less necessary in shorter texts.

**Spanish** *Noun Phrases*, as in English, present a statistically significant difference in its distribution among the three genres. *Pronouns* and *Proper Nouns* are found in the three types of texts with a statistical significant difference, whereas the use of *Determiner plus Noun* formations are evenly distributed without relevant disparities among genres (news reports 35.00%, editorials 35.50% and letters to the editor 27.86%). As the graph shows the increase in the use of *Pronouns* from news reports to letters to the editor is very significant. The same movement but on the opposite direction is the one experienced by the use of *Proper Nouns* that decrease also significantly from its use in news reports to letters to the editor. This behaviour is determined by the type of genre. One thing that has to be taken into account is the use in Spanish of verbs in first position, diminishing the use of *Pronouns* as a language specific characteristic, this fact affects the three genres in the same extent.



**Figure 18.45** NPs realization across genres in Spanish

*Noun Phrases followed by Appositions* uses in Spanish are significantly different distributed among genres: news reports use 28.25%, similar amount than editorials with 32.46%, letters to the editor, on the other hand, use only 15.48%.

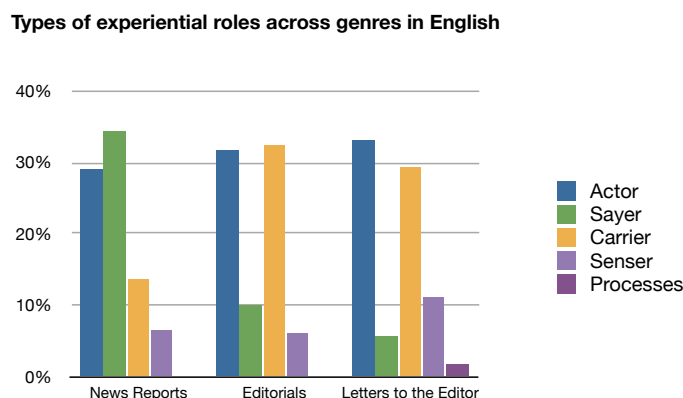
**English and Spanish** The use of *Noun Phrases* and its most common realizations: *Pronouns*, *Proper Nouns* and *Determiner plus Nouns*, present a similar distribution in both genres, what points to a genre related difference in the use of these elements.

*Pronouns* are more used in letters to the editor than in the other two genres, showing a decrease from editorials to news reports. *Proper Nouns* experiment the same behaviour but in the opposite sense than *Pronouns* being more numerous in news reports than in letters to the editor. *Determiner plus Nouns* formation are the ones that present in both languages less variation, in fact Spanish does not even present a significant difference in the use of these elements in the three genres.

Finally *Noun Phrases followed by Apposition* also reveal the same pattern according to each genre in both languages: news reports and editorials make a relevant use of these elements whereas letters to the editor do not.

### Semantic realizations

**English** presents a statistical difference in the use of *Sayer*, *Carriers*, *Senser* and *Processes* among genres.



**Figure 18.46** Experiential roles across genres in English

The figure above (fig. 18.46) shows these data graphically.

The only experiential role that is found evenly in all three types of text is *Actors*.

Francis (1990) also noted a higher frequency of *Sayers* (people quoted or whose words were reported on) in Theme position in News Reports. *Sayers* are the preferred choice of news reports (34.32%), while editorials and letters to the editor use them only in a 9.86% and 5.56% respectively. To explain these choices it is necessary to refer to the definition of news reports as White (1998) puts it the striving for objectivity makes it necessary to go to authoritative voices and present their opinion in direct quotations, as in the examples:

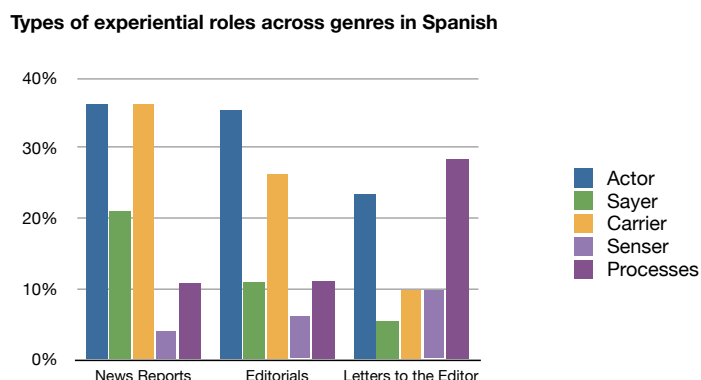
*Carriers* are used to present opinions as facts, going away from the objectivity of news reports (13.61%), they are more used in editorials (32.3%) and letters to the editor (29.17%).

The use of *Sensers* is not very high in any of the three types of texts but it is more relevant in letters to the editor (11.11%) where the author gives an opinion of a person or fact based on his/her feelings and beliefs. (New Reports and editorials only use *Senser* in 6.51% and 5.99% respectively). As *Sensers*, *Processes* are found in letters to the editor (1.67%) being in the other two types of texts inexistent. This experiential role comes from non-declarative sentences in English, which more common in a less formal writing as letters to the editor could be.

**Spanish** experiential roles analyzed are all distributed with a statistically significant difference among the three genres.

The use of *Actors* among the three genres is almost as high as the use of *Carriers* in news reports (36.16% and 36.16%) and in editorials (35.27% and 26.25%). The difference in the use of these elements is found in letters to the editor where *Actors* are

used in 23.30% and *Carriers* a 9.73%. There is not a clear cut in the use *Carriers* and *Actors* among genres except for letters to the editor.



**Figure 18.47** Experiential roles across genres in Spanish

The use of *Sensers* is not very numerous in any of the three genres, although an increase can be noted from its scarce use in news reports (3.95%) to the most numerous in letters to the editor 9.73%, going through a mid step with editorials using 6.01%.

Finally *Processes* as the other two genres have an increase in its use from the more formal news reports with the lowest use of this element 10.73%, editorials 11.02% and letters to the editor 28.32% where the number of *Processes* used is the highest.

It can be said that the use of experiential roles in Spanish show more relevant differences in the use of both news reports and editorials in front of letters to the editor.

**English and Spanish.** Experiential roles are elements that show a statistical significant difference in almost all their types and both languages. In English the use of *Actors* is the only element that remains constant in all three genres (news reports 28.99% editorials 31.69% letters to the editor 33.06%) whereas in Spanish they present a statistically significant difference among genres (news reports 36.16%, editorials 35.27% letters to the editor 23.30%).

*Carriers* in both languages present a statistical difference in its use, but English is the one that uses this device in less objective type of texts, on the contrary Spanish does not present a clear difference in its use among news reports and editorials.

*Processes* are influenced by the language specific peculiarity of Spanish (see section 11.2.1.2) but do present the same pattern both in English and Spanish being more numerous in letters to the editor than in any of the other two genres. This is due to the use of questions and directives in English in this type of texts, less objective

than the language used in news reports where there are only declarative sentences. In Spanish the findings present the same pattern but not because of sentences others than declaratives but because of the use of verbs in first position as a more colloquial strategy proper of editorials and letters to the editor.



# 19

## Annotating Thematic Progression Patterns

### 19.1 Corpus

The study on Thematic Progression patterns within Moves is carried out on a sample corpus of fifty two texts evenly extracted from the three genres of study of the larger corpus where the thematization has been analyzed (16). The texts are distributed as follows: sixteen news reports (eight in English and eight in Spanish), sixteen editorials (eight in English and eight in Spanish) and twenty letters to the editor (ten in English and ten in Spanish) (part of the results of this study are presented in Moratón & Lavid (2013), Moratón et al. (2012), Lavid et al. (2013c).

### 19.2 Models for the Study of Moves and Thematic Progression

This chapter reviews the theoretical basis under which the analysis of moves in the first place and thematic progression in the second has been carried out.

The approach to the study of moves and thematic progression has taken into consideration both aspects at the same time, considering that thematic progression patterns map into moves. Therefore the study, first identifies moves as understood by the litera-

ture in each genre, and once these moves are located in the texts, then is when the study of thematic progression is developed, relating the idea of moves to its performance in terms of thematic progression.

We have made previous studies on this topic in Lavid et al. (2011), Moratón et al. (2012) and Moratón & Lavid (2013). Not many studies have been carried out in this very same line but other authors have similarly studied rhetorical structures and thematic progression pattern in Research Articles as Khany & Nejad (2010) or in oral speech Taboada & Lavid (2003) or in second language learning as Gruber & Muntigl (2005). Vázquez y del Árbol (2005) studies editorials and letters to the editor in biomedical magazines.

In the following sections I give a brief account of the methodology and main theoretical models used for the analysis (results can be found in section 19.3)

### 19.2.1 Moves

For the identification of moves the concept applied is the one contained in Generic Structure Potential (GPS from now on) coined by Halliday & Hasan (1989). The three main tenants of the theory can be broadly summarised in the the following three ideas:

1. Halliday & Hasan (1989:5) take into consideration the “context of situation” as “the total environment in which a text unfolds”, this context is understood as the sum of field, tenor and mode, that is, the who, what is taking place and what role the language plays. Collectively they constitute the “register” of the text. (Halliday & Hasan, 1989:12)

The surface structure underlying texts, context, can be used for making predictions about text structure and the sequence and order of elements in which they are organized. This configuration consists of: field of discourse, tenor of discourse, and mode of discourse, these elements can be thought of as defining variables.

**Field:** what is happening, refers to the social situation that is taking place.

**Tenor:** who is participating, refers to the status and roles of the participants.

**Mode:** the organization of the text, the communicative channel in which the communicative event takes place.

2. The structure underlying texts can be used for making predictions about its deeps and surface structure, thus every genre has a Generic Structure Potential (GSP), consisting of obligatory, optional, and recursive elements, in Hasan words:

(...) associated with each genre of text - i.e. type of discourse - is a generalized structural formula, which permits an array of actual structures. Each complete text must be a realization of a structure from such an array. The generic membership of the text is determined by reference to the structural formula to which the actual structure can be shown to belong. A text will be perceived as incomplete if only a part of some recognizable actual structure is realized in it; and the generic provenance of a text will remain undetermined, if the part so realized is not even recognizable as belonging to some distinct actual structure.

(Hasan, 1978:229)

3. The definition of Genre is given by obligatory elements and their sequence, the function of optional elements may account for variation of the main type of text.

In the following sections I will give a detail account of the move structure for news reports, editorials and letters to the editor.

#### 19.2.1.1 News Reports

As White (1997) argues, news reports present an orbital structure in which there is a core realized by the *headline* plus the *lead* and a posterior development of the event.

This simple structure is represented by the merge of the *headline* and *lead* (these are separated in the present study to be comparable to the other genres where these elements appear in separated stages) being the core information of the News followed by a posterior development stage where the qualification, elaboration and explanation of the first move takes place.

Bronia (2005) explains the so called 'orbital' structure as allowing the writer to change the display of Move 2, adding or deleting details, without affecting the rest of the coherence and functionality of the text

The structure presents a simple display as shown in the following table:

**Table 19.1** GSP news reports moves according to White (1997)

GSP News Report moves		
Move 0	(H) Headline	It is a summary of the lead. The headline in the vast majority of cases exactly repeats part of the informational content of the lead
Move 1	(PI) The Lead: Presenting the issue	The core subject matter of the report
Move 2	(FI) Facts about the issue	Does not introduce new information but elaborates, contextualises or justifies the lead

#### 19.2.1.2 Editorials

As Fowler points out as cited by Bronia (2005) in his work is that in editorials what is distinctive is “that they employ textual strategies which foreground the speech act of offering values and beliefs”. This inner aim of editorials has its reflection on the moves structure of the genre.

In the next table it is how Ansary & Babaii (2005) describe the GSP of editorials result to a study carried out on 30 editorials of the *Washington Times*:

**Table 19.2** GSP editorial moves according to Ansary & Babaii (2005)

GSP Editorial moves		
<b>Move 0</b>	<b>(H)</b> Headline	Establishes the theme around which the editorial is centred
<b>Move 1</b>	<b>(BI)</b> Providing background information	Setting the scene to give background information on the major issue addressed in the text to later develop it
	<b>(AI)</b> Addressing an issue	This is the reason for the existence of the editorial, this is the motivation for the editorial itself and the subsequent presentation of the debate
<b>Move 2</b>	<b>(A)</b> Argumentation	This element refers to the process of argumentation (A - 1 and A - 2) and the arguments produced. It is in this part where the editorial displays its full power of persuasion
	<b>(A -1)</b> Developing of the argumentation	Exposition of facts
	<b>(A -2)</b> Closing the argumentation	Evaluation
<b>Move 3</b>	<b>(AP)</b> Articulating a position	This element specifies the newspaper position in relation to the topic at stake, describing the angle from which it looked at the discussed issue

Ansary & Babaii (2005) found that some of these elements appeared always in the editorials studied being some others not always found. as such they defined compulsory and optional elements. Obligatory elements are: *Headline*, *Addressing an Issue*, *Argumentation*, and *Articulating a Position* and optional elements *Background Information*, *Initiation of Argumentation* and *Closing the argumentation*, these elements do not always appear in this order but may vary in the scheme of apparition. A general structure may well be the following:

$$(H^{\wedge}[(BI).AI]^{\wedge}\{(IA)^{\wedge}A - 1^{\wedge}A - 2^{\wedge}...(CA)\}_n)$$

### 19.2.1.3 Letters to the Editor

In Ghadessy (1983) study of letters to the editor, where he analyzes 43 letters from the *Daily Mail*, proposes the Situation, Problem, Solution and Evaluation structure. According to it, and as Bronia (2005) puts it each of these elements do not correspond in a one to one relationship to a paragraph identification, in fact, in one single paragraph more than one move can be identified, indeed “it is not uncommon to find one-paragraph letters including all the four categories” (2005:32).

Ghadessy (1983) explains that these elements do not have to follow a fixed linear pattern of appearance but rather each of them can take any place, being interchangeable in order, this is what he called the “loop” approach.

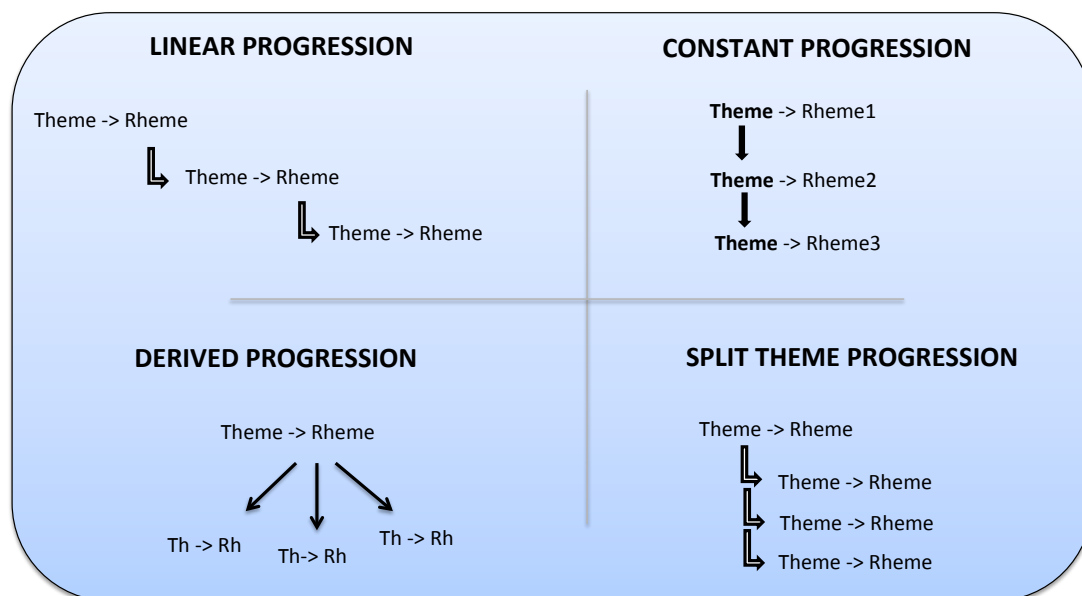
Based on Ghadessy, Bronia (2005:169) proposes the following Generic Structure Potential of letters to the editor in the three parts structure depicted in the following table.

**Table 19.3** GSP letters to the editor moves according to Ghadessy (1995)

GSP Letters to the Editor moves		
Move 0	(H) Headline	Establishes the theme around which the editorial is centred
Move 1	(I) Introducing the issue	Settling the issue: refer to a published article, pointing to the problem, stating one’s position
Move 2	(A) Outlining the arguments	Evaluates the causes and effects, proposes causes of action or provides background information
Move 3	(C) Closing	Restating one’s position, makes a pledge an appeal for cooperation, etc...

### 19.2.2 Thematic Progression

For Thematic progression patterns the same theoretical tenants as the ones exposed in Moratón & Lavid (2013) will be followed. As Eggins (2004:324) defines it, the texts’ method of development is “how Thematic elements succeed each other” and those are defined by Downing & Locke (2006:246) in four main types: linear, constant, derived and split, as it is summarize in the following figure:



**Figure 19.1** Types of Thematic Progression patterns

### 1. Simple linear progression

Downing adapts from Danes the definition as “something introduced as new information in the Rheme of the first clause is taken up to be the Theme of the second. The wording need not be identical.”

- (1) El producto interior bruto (PIB) registró una contracción del 0,1% en el tercer trimestre del año (equivalente a un ritmo anual de caída del 0,3%), el peor dato desde la última crisis económica, en 2001.

*[Gross domestic product (GDP) shrank by 0.1% in the third quarter of the year (equivalent to an annual rate of decrease of 0.3%), the worst figure since the last economic crisis in 2001.]*

- (2) El motivo de este frenazo reside en la mayor caída del consumo -motor de la primera economía del mundo- de los últimos 28 años y el pesado lastre de la devaluación de la vivienda, en medio de la peor crisis financiera desde la Gran Depresión.

*[The reason for this slowdown lies in the greater drop in consumption-engine of the world's leading economy of the last 28 years and the heavy burden of the devaluation of house pricing, in the midst of the worst financial crisis since the Great Depression.]*

(Report: “La crisis financiera lleva a Estados Unidos al borde de la

*recesión.”)*

## 2. Continuous progression (constant Theme)

The same Theme is maintained in the following clauses, creating a continuous reference chain.

- (3) For the full-year ending June 2009, Microsoft forecasts earnings per share to range between \$2.00 to \$2.10 on revenue from \$64.9 billion to \$66.4 billion.
- (4) Microsoft’s previous estimates called for earnings per share from \$2.12 to \$2.18 and revenue ranging from \$67.3 billion to \$68.1 billion.
- (5) Microsoft’s guidance for the current quarter was marginally lower than Wall Street was expecting.

(Report: “*Microsoft results offer hope to tech sector.*”)

## 3. Derived progression

In this third type, the different themes of a number of Theme–Rheme structures are all relate to a ‘hypertheme’ or ‘global topic’. This type turned out to be one of the most relevant patterns in the study.

- (6) Sigue imputado  
[*Still impeachment*]
- (7) El presidente de la Generalitat valenciana, Francisco Camps, compareció ayer como imputado ante el Tribunal Superior de Justicia de la Comunidad Valenciana, del que salió, (...)  
[*The President of the Generalitat Valenciana, Francisco Camps, appeared yesterday under impeachment in the High Court of Justice of the Valencian Community, from where he came out (...)*]
- (8) El tribunal facilitó una escueta nota en la que afirma que “el proceso sigue su curso”.  
[*The court provided a brief note in which it was stated that “the process is ongoing.”*]  
(Editorial: “*Sigue Imputado.*”)

## 4. Split progression

It is a mixture of Linear progression (type 1) and Continuous progression (type 2), this is explained in the following figure from Downing & Locke (2006:249):



Clause 1      T1 – R1. The Rheme implies two items, A and B:  
 Clauses 2–3    In one (A) \_\_\_\_\_ (T2 ----- R2)  
 Clauses 4–5    In the other (B) \_\_\_\_\_ (T2'' ----- R2'')

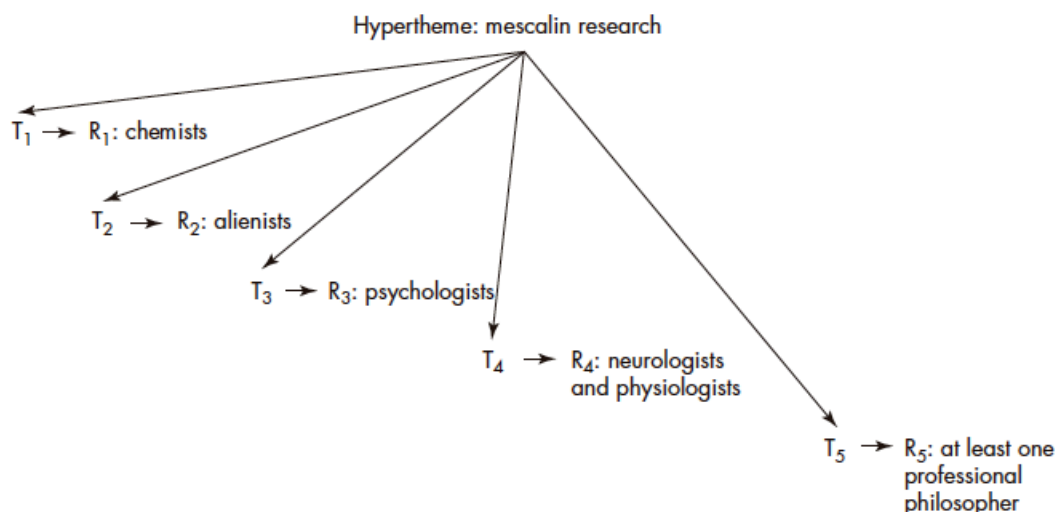


Figure 19.2 Split Theme

## 19.3 Analysis & Results of the Study of Moves and Thematic Progression

This chapter presents the analysis and results of the study of the thematic progressions patterns found within moves. The analytical model followed for the study is described in sections 19.2.1 and 19.2.2, where a detailed account of the concept of *Generic Structure Potential* given by Halliday & Hasan (1989) and its extension by Ansary & Babaii (2005) are given, along with organizational charts with the GSP of news reports 19.1, editorials 19.2 and letters to the editor 19.3. An explanation of *Thematic Progression* given by Danes (1974), Eggins (2004) and Downing & Locke (2006) and followed in this analysis are also given.

The chapter is organized as follows: a presentation of the Corpus of study, then an individual analysis of each genre in terms of Thematic Progression patterns within Moves in English and Spanish and, finally, a comparison among the three genres under study.

### 19.3.1 News Reports

This section presents the analysis and results obtained in the study of Moves and the Thematic Progression patterns that characterizes them. news reports Move structure, as table 19.1 shows consists of three moves.

**Move 0** Headline

**Move 1** The lead: presenting the issue

**Move 2** Facts about the issue

**Move 0** corresponds to the headline, that in Bell's (1991:20) words is "an abstract of the abstract", that in Spanish tends to be longer and more explanatory than the English one that are shorter and direct with a "distinctive telegraphic syntax" (1991:160), as the following examples show:

(9) Opec hawks want to cut oil production to keep up price.

(10) Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown.

[Barclays receives 9200 millions of state funds to avoid calling Browns' funds.]

**Move 1** for Bell (1991:20) "the lead paragraph is the journalists primary abstract of a story" and it is devoted to the clarification of the previous headline, in most cases in no more than 4 or 5 sentences, in this move there is often an exact repetition of what it has already been stated in the headline adding more information to contextualize the issue. In terms of the amount of complex clauses that perform this function are only of 10.61% in English and 16.21% in Spanish of the total amount of clause complexes that conform the texts. From a thematic point of view the transition from this Move 0 to Move 1 is done through a constant thematic pattern, this strategy reinforces the idea of repeating the idea stated on the headline. Move 1 also revealed to be the move where more information is added to the same topic, using predominantly this constant progression in both languages English 75% and Spanish 45%.

(11) **Move 0** Opec hawks want to cut oil production to keep up Price.

**Move 1** Opec's mettle will be tested next week when the cartel assembles in Vienna at a hastily convened meeting to discuss the economic crisis and whether the exporters' club should take action to arrest the fall in the price of crude oil.

(News Report: *Opec hawks want to cut oil production to keep up Price.*)

- (12) **Move 0** Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown.

*[Barclays receives 9200 millions of state funds to avoid calling Browns' funds.]*

**Move 1** El Barclays recurrirá en su mayor parte a inversores del Golfo Pérsico para llevar a cabo una ampliación de capital de hasta 7.300 millones de libras (9.305 millones de euros), ha comunicado hoy el banco británico al regulador de la Bolsa de Londres.

*[Barclays will use mostly Gulf investors to carry out a capital increase of up to 7.300 billion pounds (9.305 million euros), the British bank informed today the London Stock Exchange bank regulator.]*

La entidad emitirá 3.000 millones de libras (3.822 millones de euros) en opciones sobre acciones, destinadas a esos inversores árabes, (...)

*[The company will issue 3.000 billion pounds (3.822 million euros) in stock options, aimed at Arab investors (...)]*

(News Report: *Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown.*)

**Move 2** presents the facts about the issue, this move gives the details on the piece of news introduced by Move 0 and 1, in fact the extension of this move in both languages is longer than any other move both in English (89.38%) and in Spanish (83.78%). In this second move similar patterns are found in thematic progression in English and Spanish, each new move is introduced by a derived theme that is developed by either constant or linear themes. The number of English constant themes is 20.79% and linear themes 28.71%. In Spanish the amount of constant themes is of 21.51% and linear themes are 34.41%. Although Spanish shows a preference for the selection of linear themes instead of constant themes it does not present a statistically significant difference.

Example (13) shows the behaviour of constant themes in English and example (14) presents a linear pattern in Spanish news reports.

- (13) For the full-year ending June 2009, Microsoft forecasts earnings per share to range between 2.00to2.10 on revenue from 64.9billionto66.4 billion.

Microsoft's previous estimates called for earnings per share from 2.12to2.18 and revenue ranging from 67.3billionto68.1 billion.

Microsoft's guidance for the current quarter was marginally lower than Wall Street was expecting.

(News Report: *Microsoft results offer hope to tech sector.*).

- (14) El producto interior bruto (PIB) registró una contracción del 0,1% en el tercer trimestre del año (equivalente a un ritmo anual de caída del 0,3%), el peor dato desde la última crisis económica, en 2001.

*[Gross domestic product (GDP) shrank by 0.1% in the third quarter of the year (equivalent to an annual rate of decrease of 0.3%), the worst figure since the last economic crisis in 2001.]*

El motivo de este frenazo reside en la mayor caída del consumo -motor de la primera economía del mundo- de los últimos 28 años y el pesado lastre de la devaluación de la vivienda, en medio de la peor crisis financiera desde la Gran Depresión.

*[The reason for this slowdown lies in the greater drop in consumption-engine of the world's leading economy of the last 28 years and the heavy burden of the devaluation of house pricing, in the midst of the worst financial crisis since the Great Depression.]*

(News Report: *La crisis financiera lleva a Estados Unidos al borde de la recesión.*)

To sum up, the only difference that can be observed in English vs Spanish news reports is the tendency in Move 2 for linear themes in Spanish instead of constant themes that are the ones preferred in English.

### 19.3.2 Editorials

Editorials move structure is longer than news reports due to the nature of these texts that account for a mean extension in terms of sentences per text of 19.56 for English and 17.5 for Spanish. This genre accounts for four different main moves:

**Move 0** Headline

**Move 1** Addressing the issue

**Move 2** Argumentation

**Move 2.a** (op) Providing background information of developing of the argument

**Move 2.b** (op) Evaluation or Closing argumentation

**Move 3** Articulating a position

**Move 0** corresponds to the headline, in this case of a very different nature than news reports, this tends to be short and with a metaphorical sense that does not let the reader grasp at first sight what the content of the article is about.

(15) A confession is not enough.

- (16) Sigue imputado.

*[Still impeachment]*

**Move 1** develops and clarifies the topic of the editorial as the headline is not often enough. This move is not very long although there is a striking difference in its extension between English 5.40% and Spanish 16.51%. The thematic transition of Move 0 to Move 1 is done through derived thematic progression, and this is something found in all editorials and both languages, what makes sense regarding its explanatory nature of the metaphorical headline, what is more in 50% of texts in English and 45% of texts in Spanish the only constituent of this move is performed by this derived theme on its own. The second most preferred pattern is the derived-linear pattern (English 30% Spanish 30%) that is an introduction of the issue to the reader and its posterior development through the linear pattern.

- (17) **Move 0** A confession is not enough.

**Move 1** Even if Lance Armstrong had come across as the most contrite of penitents – which he did not – an interview with Oprah Winfrey on American TV is not sufficient to allow the cyclist to draw a line under his disgrace and “move on”.

(Editorial: *A confession is not enough.*)

- (18) **Move 0** Sigue imputado.

*[Still impeachment]*

**Move 1** El presidente de la Generalitat valenciana, Francisco Camps, compareció ayer como imputado ante el Tribunal Superior de Justicia de la Comunidad Valenciana, del que salió, (...)

*[The President of the Generalitat Valenciana, Francisco Camps, appeared yesterday under impeachment in the High Court of Justice of the Valencian Community, from where he came out (...)]*

El tribunal facilitó una escueta nota en la que afirma que “el proceso sigue su curso”.

*[The court provided a brief note in which it was stated that “the process is ongoing”.]*

(Editorial: *Sigue imputado.*)

**Move 2** is the core of the editorial where the author outlines the arguments and it also takes the longer extension of the text (English 51.26% and Spanish 58.71%). This longer span of text allows even for the appearance of submoves as Ansary & Babaii (2005) explain that corresponds to the developing of the argument, evaluation or exposing facts and the closing of the argumentation.

What the analysis revealed is that these sub-moves do not appear in a sequence pattern but rather in a leap frogging fashion where there is an exposition submove followed by an evaluative one and then back again to exposing some arguments and evaluating them. This long moves starts with the use of a derived theme, again acting as a starter and then a development that in English is done through the use of split themes (16.08%) linear (20.28%) and constant progression patterns (16.78%). On the other hand, Spanish uses the same starting point with a derived theme and develops the argument mainly through the use of linear patterns (29.69%). The distribution of these elements in submoves reveal that in the second move the number of submoves found in English are 24 explanatory and 25 evaluative realized by 82 and 65 sentences respectively. This points to a more extensive explanatory move in terms of sentences (82) grouped into 24 submoves followed by a shorter (65 sentences) its evaluative parts grouped into 25 submoves. This structure shows the complexity of this move which is reinforced by the finding of elements of move 3. The following table exemplifies this idea:

**Table 19.4** GSP example of editorial Move 2

<b>GSP Editorial moves 2</b>
<b>Move 2.a Exposition of facts - Providing background information:</b>
(1) David Cameron gave a succinct explanation in the House of Commons yesterday for his determination to strike a new deal with the European Union.
(2) He warned that the most dangerous position this country could take in the light of changes being forced by the eurozone sovereign debt crisis would be to stand back and say we are going to do nothing about it.
(3) The Prime Minister said that Britain had to fight for the new deal it wants “so then we can ask for the consent of the British people to settle this issue once and for all”.
<b>Move 2.b Evaluation:</b>
(4) This will be the broad thrust of the much-anticipated speech on Europe that he will deliver in the Netherlands tomorrow.
(5) Mr Cameron hopes it will quell the growing restiveness in his own ranks over the issue, though that is probably a forlorn hope.
<b>(Back to) Move 2.a Exposition of facts - Providing background information:</b>
(12) Yesterday, the Fresh Start group of more than 100 Tory MPs published a manifesto for change, but not withdrawal.
(13) This vigorous debate includes those who advocate withdrawal and those, led by Michael Heseltine and Kenneth Clarke, who caution against any serious attempt to challenge the status quo.
<b>Move 3 Articulating a position - Recommendation:</b>
(16) If Mr Miliband deems this worthy only of his sneers, he is making a serious mistake.
(17) But his response is characteristic of the internationalist Left, who are quick to champion every country’s point of view except their own.
<b>(Back to) Move 2.b Evaluation:</b>
(20) They do not want Britain to rock the boat.
(21) Sir Nigel Sheinwald, formerly our ambassador to both the EU and Washington, spoke in the authentic voice of the Brussels elites yesterday when he warned that “a Britain on the sidelines of Europe, even more out of Europe, would not be in the American interest and would not be in the interest of other major European partners”.
(22) Such threats are par for the course, but are essentially empty.

In Spanish these submoves are shorter than in English, probably due to the shorter extension of the move itself. The explanatory submove (2.a) is realized by 29 clauses and the evaluative submove (2.b) by 35. These sentences are grouped into 13 submoves each. The structure shows the same leap frogging movement than in English but with a longer extension of evaluative moves what suggests a more subjective treatment of editorials in Spanish, whereas English seems to be more balanced between the explanatory side and the evaluative one.

**Move 3** articulates the position of the author after finishing with the exposition and evaluation of facts. This move presents a significant difference in extension between English (43.32%) and Spanish (24.77%). It is possible that the longer evaluative part of Spanish Move 2 integrates the evaluative function that this Move 3 does in English. The extension of this third move is much longer than moves 0 and 1 but shorter than 2. In the analysis of this move it is found that, as in move 2, there were elements similar to the submoves described by Ansary & Babaii (2005), in this case they have been called informational units, these units develop one idea. Each of these pieces start with a derived theme, which in many cases is its only constituent (English 47.36% and 75% Spanish). The next preferred patterns are derived-linear in English (26.31%) and derived-constant in Spanish (15%).

### 19.3.3 Letters to the Editor

In letters to the editor the analysis of the Thematic progression patterns are not as straight forward as in the other two genres, this is due to the short extension of these texts, that, as said in section 19.2.1.3 citing Bronia (2005:169) “in one single paragraph more than one move can be identified”. The GSP proposed by Ghadessy:

**Move 0** Headline

**Move 1** Introducing the issue

**Move 2** Outlining the arguments

**Move 3** Closing

Regarding the mentioned extension, the mean length of these texts is of 7.32 clauses per text in English letters to the editor and 7.30 clauses per text in Spanish. Despite this fact, in the present analysis there were no move changes inside the same paragraph, and the results confirm the structure proposed by Ghadessy (above). Almost all texts in both languages respond to the structure of the four moves explained above. There



is one single case in which the last move (Move 3 Closing) does not appear in the text, which is not representative.

Headlines in **Move 0**, both in English and Spanish, are in mid way between the characteristics of editorials and news reports headlines, that is, they are not as metaphorical as editorials' nor so self explanatory as news reports', they introduce the topic of the Letter giving the reader a hint on what it is going to be about.

(19) Schools in need of discipline.

(20) Euro por receta.

[Euro per prescription]

**Move 0** as López Maestre (1997:163) observes in her study on headlines, these of letters to the editor “are mostly formed by independent nominal phrases and can be said to have a nominal style defined by the absence of verbs”. In fact in the sample selected for this study the use of verbs in the headlines if any, appear in its non finite form: *Dangerous ways to raise money*, *Baffling rail fares*, *Un cierre incomprensible* [An incomprehensible close], *La violencia en Gaza* [Violence in Gaza]

**Move 1** introduces the issues, referring directly to the matter that will be given the opinion upon. This move does take a mean of one to three sentences long, that is 21.57% of the clauses of the English texts and 18.09% of in the Spanish corpus. This move takes a clear and explanatory style where the presentation of the matter at stakes is plainly stated. The transition between Move 0 to Move 1 is done through a derived theme in 80% of texts in English and in the same amount in Spanish:

(21) **Move 0** Schools in need of discipline.

**Move 1** Richard Garner (News, 16 November) reports that Buckingham University has recently shown that nearly 30 per cent of trainee teachers last year had not taken up teaching posts six months later, and that 52 per cent of teachers under the age of 60 are not within the classroom. For Alan Smithers, one of the authors, to argue that the Government's Teaching Agency should withdraw funding for places from teacher training providers who perform poorly in supplying recruits to the profession is wholly inappropriate.

(Letter to the Editor: *Schools in need of discipline.*)

(22) **Move 0** Euro por receta.

[Euro per prescription]

**Move 1** El Gobierno de la Comunidad de Madrid sigue defendiendo la necesidad de poner el copago de un euro por receta.

[The Comunidad de Madrid Government continues to advocate the need to charge an euro copayment per prescription.] (Letter to the Editor: *Euro por receta.*)

Within this move, and due to its short extension there are not long patterns reflected but rather what it is seen is that in English Move 1 is realized by isolated sentences in 90% of cases and of those 80% are realized by a derived theme. In Spanish only 50% of the moves are realized by isolated clause complexes, being 60% realized by derived themes and 40% by Linear themes. This difference in the realization of Move 1 by isolated sentences in English and more than one in Spanish is considered to be statistically significant. The elements realized by more than one sentence in Spanish show the following distributional pattern: DDD, DD, DL, DLD, DD (being D derived, and L linear Thematic Progression). This is not relevant enough as to establish an alternative pattern rather than the consideration of derived as the preferred Thematic progression pattern both for English and Spanish in Move 1.

**Move 2** outlines the writer's arguments in the main body of the text. English devotes to this move 54.90% of the clause complexes and Spanish 58.51%.

Again the extension of the text determines the nature of this move in letters to the editor being quite simple in comparison to the display of patterns variety of editorials and news reports.

- (23) **Move 2** The biggest problem facing trainee teachers, and most teachers, is the poor behaviour that teachers experience day in, day out.

(Letter to the Editor: *Schools in need of discipline.*)

- (24) **Move 2** Supongo que cuando se toman este tipo de medidas hay profesionales que estudian todos los impactos que ello puede producir.

[I guess when such measures are taken is because there are professionals who study all impacts this may cause.]

Pero en este caso considero, desde mi profesión de médico, que están dejando fuera aspectos tan importantes como no valorar que muchas personas, y muy preferentemente los parados, las personas de bajos salarios y los jubilados, puedan dejar de comprar su medicación en cuanto los síntomas disminuyan; o peor aún, desde el primer momento.

[But in this case I think, from my medical profession, that they are leaving out relevant aspects such as not taking into account that many people mostly unemployed, those on low wages and retirees who can stop buying their medication as soon as symptoms subside; or worse, from the first moment.]

(Letter to the Editor: *Euro por receta.*)

In English the thematic types used are derived 60.17%, linear 21.43%, constant 10.71% and split 7.14%, choosing with a statistically significant difference the use of derived rather than linear, constant or split themes.

In Spanish the thematic types used are derived 58.18%, constant 18.18%, linear 12.73% and split 10.91%, as in English, the preference are derived themes.

Regarding the distributional patterns present in both languages are the pattern DD being this structure obvious since the amount of these elements found are far more numerous than the other types. DL or LD in English are the second preferred pattern but these are found in such scarce number that there is no relevance in its number of appearances. Spanish on the other hand presents a wider variety of occurrences, being as in English DD the preferred one, the most common varieties point to the use of constant themes rather than linear themes as English does, finding the following patterns DC (four occasions), in three texts, being the more complex DLC.

**Move 3** Closes the Letter with a recommendation or a final thought of the writer wrapping up the Letter. This last move acts as a final conclusion to the previous stated arguments. This move is parallel in extension to the first move (Move 1) closing in circular structure the Letter. The amount of clauses devoted to this last move in both languages is quite similar, of 23.53% in English and 23.40% in Spanish. This short move is realized in 80% of cases by isolated sentences in English and 40% of cases in Spanish. Spanish closing structures are longer than English but this difference is not considered statistically significant. In English the realization of this move is done through 66.67% of derived themes, 25% of constant themes and 8.33% of linear themes. It is not a statistically significant difference, although it shows a tendency for the realization of this move by derived themes. In the case of Spanish Closing move is performed by derived themes in 90.9% and 9.1% by linear themes. In Spanish the difference is statistically significant showing a clear preference for the use of this type of thematic progression.

- (25) **Move 3** As a teacher, the depressing statistics do not surprise me.

(Letter to the Editor: *Schools in need of discipline.*)

- (26) **Move 3** Se sabe que los ingresos hospitalarios aumentan en torno a un 2% anual en los casos de abandono de tratamiento.

*[It is known that hospital admissions increased by around 2% annually for treatment abandonment.]*

¿Nadie es consciente de esto?.

*[Is there anyone aware of this?]*

Y de lo más importante, del incremento del riesgo de los pacientes.

*[And of the most important thing, the increase risk of patients.]*

(Letter to the Editor: *Euro por receta.*)

# 20

## Summary and Discussion

The present part summarizes and discusses the results of the manual corpus annotation in terms of thematic characterization (chapter 18) and moves and thematic progression (chapter 19).

In chapter 11, I focussed in detail the model of thematization empirically validated in the previous part of the study and applied to the analysis of the bilingual corpus (described in section 16). The analysis and results (in chapter 18.1) are presented from a language and a genre perspective. From the language point of view, each genre is analyzed in Spanish and English and the results compared and commented. In the second perspective, from a genre point of view, all thematic elements are compared based on the type of texts to which they belong. The analysis yields interesting results in the selection of the thematic elements both from a language and genre viewpoint.

From the generic point of view, the following results were obtained:

In **News reports**, although both languages present a direct and informative style, they achieve these goals through different procedures. The use of Textual elements is scarce in both languages, the little occurrences presented in English are chosen to be of the type Binders whereas Spanish tends to use more Textual elements of the type Linkers seeming to add in the concatenation of ideas. The

position held by PreHeads are covered by the use of expressions of time following what Bell (1991:155) highlights about the overturning of temporal sequences in news reports, these expressions are built on adverbial groups and prepositional phrases (PH-CCL) in English. Spanish opts for simpler expressions (PH-Circ). Another interesting feature of PreHeads is that it is revealed as the Spanish preferred slot, being a pro-drop language, for the lexical part of the verb when it appears in first position. In English this only occurs when the type of sentences are questions or directives. The position of quotations in news reports is another key difference among English and Spanish. English up fronts its appositions to the PreHeads whereas Spanish delays the apposition of quotations in the rheme that can be traced by the use of a verb with a communicative meaning. The realization of Heads are balanced between English and Spanish in the use of Noun Phrases and Noun Phrases followed by Appositions, both elements more frequently used in English where Spanish uses verbs, in this case, its grammatical information. These verbs both in English and Spanish, in its semantic role, act, as Francis (1990) pointed out as Actors and Sayers.

In **editorials** the distributional map of thematic elements changes drastically. Due to the longer and explanatory nature of the texts, the use of Textual Themes increases in both languages. There is also an increase in the use of Interpersonal Themes that reveal the strong presence of the author in the text, in most cases in the form of Engagement where the writer relates to the readers in respect to his/her position in the text (Hyland, 2001). In PreHeads the forms are evenly distributed preferring both languages the use of the more simple forms of PH-Circ for the expression of Time and more complex (and less frequent) for the expression of Contingency, introducing counterarguments or hypothesis. A language characteristic of Spanish is the apposition of the lexical part of the verb in this position. In Head position appears the use, although in not many occasions, of Thematic Equatives and Predicated Themes, symbol again of the presence of the writer who through these means “is conscious that (he/she) is asserting or denying something in a firm way” (Martínez Lirola, 2008). The realizations of Noun Phrases are Pronouns as the first option chosen for English, instead of this strategy Spanish uses Verbs in first position avoiding the explicit mark of the subject, English cannot do so, the need of a subject mark leads to the use of, in this case, Pronouns, whereas Spanish relies in the grammatical information of the verb. Another language characteristic is the use in English of common nouns

---

(noun not preceded by determiner) in subject position whereas Spanish does not present any occurrence of this kind, this, as Pérez-Leroux et al. (2004) notes, "...English allows in subject position bare plural Noun Phrases while Spanish and some other Romance languages do not". The semantic realizations belong to Actors and Relational processes in this type of genre.

In **Letters to the editor** it is important to highlight the short extension of this type of texts that will influence certain thematic choices. It happens with Textual Themes that reflect the nature of these texts in the use of Binders both in English and Spanish, Linkers, although less used in both cases, are preferred by Spanish rather than English. Interpersonal Themes are expected to appear more due to the nature of the texts, and they do and also reveal a language specific strategy to reveal the presence of the author in the text: English who selects the use of Engagement approach to show the author in the texts, does so in less number than Spanish who selects in a varied form both Stance and Engagement forms in higher numbers than English. Does this mean that English authors do now show themselves in the texts? Rather the explanation points to the reason given by Pounds (2010) who says that English "stance" may be more often "integrated into text rather than overtly marked". In PreHead position the selection of simple syntactic constructions in Spanish rather than English marks the tendency of this language to express in this slot Time and Contingency through the more complex formations of PH-CCL. In Head the apparition of Thematic Equatives and Predicated Themes point to the more colloquial feature of this type of texts. The preferred syntactic realization of Head is that of Noun Phrases in both languages, although English presents a statistical significant difference in its number, Spanish balances this selection with the use of verbs in first position. In this type of texts, contrary to the other genres analyzed there are other types of clauses than declaratives, therefore it can be also studied the use of verbal forms in Head position in English directives and questions. The choice of experiential elements is more varied appearing in relevant numbers: Actors, Carriers, Senser, Sayers and Goal (and the special mention of processes in Spanish as a language characteristic).

From the language-specific point of view, the thematic study yielded the following results:

**Textual Themes:** the use of *Linkers* in Spanish is greater than in English in all cases, the idea of evolution of the text is taken for granted and it is not signaled

through any other visible mean. As many studies have shown Vogel (2008), Shaw (2009), Tse & Hyland (2006), Hyland & Bondi (2006), Charles et al. (2011), the use of *Linkers* by native English speakers is scarce. The use of *Binders* are used to contrast the previous information, therefore these elements do add something new about the relationship among clauses.

**Interpersonal Themes:** In his study Francis (1990) also found a higher frequency of Interpersonal Themes in editorials and letters to the editor than in news reports, as the first two genres are more “personal and evaluative, as well as more concerned with persuasion than information, therefore it is likely to be more direct interaction between writer and reader”. Comments in English present a significant relevant difference in its use through the three genres, being more frequent in editorials than Letters to the Editor, this responds to the nature of these texts as “...opinionated genres intended to persuade the reader” (Biber, 1988). On the other hand, the use of *Engagement Strategies* in both editorials and letters to the editor is evenly high (editorials 50.00% and letters to the editor 52.94%) with no difference in its use among them. The same happens with the use of *Stance* (editorials 15.63% and letters to the editor 23.53%) and *Modal* (editorials 7.81% and letters to the editor 8.82%).

Spanish news reports only have one occurrence of Interpersonal elements. But the use of *Engagement* (editorials 68.42% and letters to the editor 40.63%) and *Stance* (editorials 10.53% and letters to the editor 39.06%) present a extremely significant difference and a very significant difference respectively, amongst the three genres. On the other hand the use of *Modals* and *Comments* does not present a significant difference in its distribution. Letters to the editor are characterized as carriers of *Stance* elements, described by Hyland as “adopting a point of view in relation to both the issues discussed in the text and to others who hold points of view on those issues.” This means that the author shows off in the text clearly, presenting him/herself as the responsible for the opinions in the text.

On the other hand, editorials as “the voice of the newspaper” are characterized by the use of *Engagement* strategies conveying opinions but not directly.

Spanish presents more variation in terms of statistical differences in the distribution of the types of Interpersonal elements. In fact the variation in English sticks to the use of *Comments* whereas Spanish shows distributional discrepancies in *Engagement* and *Stance* strategies. Expressions of interpersonality are more thematic in Spanish than in English, that means that English does not present in



---

the theme the marks of interpersonality as much as Spanish does.

**PreHead:** the most relevant function that PreHead perform in both languages is the expression of *Time* in all genres and both languages, the idea of immediateness is present in the three genres as Newspapers are related to the account of recent facts from different perspectives, news reports take care of facts that have just happened (Clark, 1975), editorials present the opinion of the Newspaper on recent news events (Van Dijk, 1989) and letters to the editor express the reader's opinion on hot topics (Vázquez y del Árbol, 2005).

There is a contrast ( $P < 0.0001$ ) is in the use of *Place* and *Contingency* expressions in both languages ranging, in the case of *Place* to a relevant use in news reports 14.29% in English and 7.32% in Spanish to an almost non existent appearance in letters to the editors (English 3.23% Spanish 0.62%) (in editorials they express *Place* 2.81% in English and 7.73% in Spanish). This fails within the nature of this type of texts as news reports give account of issues that have to be set in *Place* and *Time*.

*Contingency* on the contrary introduces counter arguments or hypothesis, more used in editorials (English 30.90% Spanish 21.13%) and letters to the editor (English 25.81% Spanish 16.67%) than in news reports (English 5.36% Spanish 8.94%). New Reports for this same purpose use *Quotations* as the justification of their reasoning.

*Evidence* (English 10.71% Spanish 13.82%) and *Quotations* (English 19.64% Spanish 0.81%) appear only in a remarkable amount in news reports and are almost inexistent in the other two genres. Through these strategies news reports approach reliable sources to present objective and identifiable points of view (more occurrences are found in English than in Spanish as explained on sections 18.2.1.3, 18.2.2.3 and 18.3).

**Head:** the use of *Noun Phrases* and its most common realizations: *Pronouns*, *Proper Nouns* and *Determiner plus Nouns*, present a similar distribution in both genres, what points to a genre related difference in the use of these elements.

*Pronouns* are more used in letters to the editor than in the other two genres, showing a decrease from editorials to news reports. *Proper Nouns* experiment the same behaviour but in the opposite sense than *Pronouns* being more numerous in news reports than in letters to the editor. *Determiner plus Nouns* formation are the ones that present in both languages less variation, in fact Spanish does

not even present a significant difference in the use of these elements in the three genres.

Finally *Noun Phrases followed by Apposition* also reveal the same pattern according to each genre in both languages: news reports and editorials make a relevant use of these elements whereas letters to the editor do not.

Experiential roles are elements that show a statistical significant difference in almost all their types and both languages. In English the use of *Actors* is the only element that remains constant in all three genres (news reports 28.99% editorials 31.69% letters to the editor 33.06%) whereas in Spanish they present a statistically significant difference among genres (news reports 36.16%, editorials 35.27% letters to the editor 23.30%).

*Carriers* in both languages present a statistical difference in its use, but English is the one that uses this device in less objective type of texts, on the contrary Spanish does not present a clear difference in its use among news reports and editorials.

*Processes* are influenced by the language specific peculiarity of Spanish (see section 11.2.1.2) but do present the same pattern in English and Spanish being more numerous in letters to the editor than in any of the other two genres. This is due to the use of questions and directives in English in this type of texts, less objective than the language used in news reports where there are only declarative sentences. In Spanish the findings present the same pattern but not because of sentences others than declaratives but because of the use of verbs in first position as a more colloquial strategy proper of editorials and letters to the editor.

### **Moves & Thematic progression**

In chapter 19 the analysis of thematic progression patterns and moves (GSP) yields the following results, summarized in the following figure:

**Table 20.1** Contrastive table of moves and thematic progression patterns

Moves	News Reports	Editorial	Letter to the Editor
<b>M. 0</b>	Explicative headline	Metaphorical headline	Summarizing headline
<b>M. 1</b>	Presenting the issue	Addressing the issue	Introducing the issue
ThP	Constant	Derived	Derived
<b>M. 2</b>	Facts about the issue	Argumentation: eval, facts	Outlining the arguments
ThP	Hypertheme Constant Eng Hypertheme Linear Spa	Hypertheme Derived	Derived
<b>M. 3</b>		Articulating a position	Closing
ThP		Hypertheme Derived	Derived

The table above presents a summary of the study carried out on GSP and their thematic progression patterns within them. It reveals that the organization of moves (GSP) in both languages and genres is quite similar, finding in editorials the main differences. Although both languages display a complicated organization of information in submoves with an intermingling of evaluative submoves and expositions of facts, Move 2 in Spanish editorials acquires a special relevance taking from Move 3 part of its function and extension. The thematic progression patterns selected in English and Spanish show a common preference for the Derived theme as a trigger of moves, submoves and units of information. This finding gives an important clue for the identification of these elements. In letters to the editor, due to their short extension moves are realized by isolated clause complexes or they are so short in extension that is difficult to identify a hypertheme and its posterior development takes the headline as constant hypertheme that is developed according to each move in the letter itself.

In the other two genres, news reports and editorials present a difference among English and Spanish in the development of these moves, being English prone to use constant themes where Spanish uses linear thematic progression.

Regarding genres, editorials and letters to the editor use more derived themes whereas news reports use constant or linear progression patterns. News reports and letters to the editor stick to the theme introduced in the headline and add information to it while editorials state related hyperthemes that are further evolved, evaluated or opinionated. All these particular features lead us to reach a general conclusion: the use of thematic progression patterns is relevant for the identification of moves and units of

## *20. Summary and Discussion*

---

information, and therefore for the identification of genres.

## **E. TOWARDS AUTOMATIC THEME ANNOTATION**



# 21

## Introduction

In this Part I will focus on the steps taken towards automating the annotation of the thematic features which were manually annotated by the author of this dissertation, as described in Part D above (Annotating a Larger Corpus). As can be inferred from the reading of Part D, the manual annotation process is a lengthy and time-consuming task that relies on human annotation to ensure that the created dataset is the ‘gold standard’ and is of high-quality. Therefore, after this lengthy process I decided to investigate the possibility to automate it and compare the results obtained using each method: human-coded or automatic annotation. The remainder of this Part is organized as follows: First, I will describe the tool used to annotate the English and the Spanish datasets. This is the GATE annotation platform, whose interface is easily accessible and it offers a wide range of available resources as well as the integration of our own APIs containing rules for automatic annotation of thematic features. In English the bootstrapping allowed for the creation of an automatic application (API) for the annotation of Theme (chapter 24). A ‘translation’ process began where the guidelines are transformed into JAPE rules. In Spanish, on the other hand, it was not possible to develop a full automatic annotation process within GATE due to difficulties in integrating existing processing resources into the platform, as it will be shown below (see chapter 25). Finally, once the automatic annotation (for English) and the

semi-automatic (for Spanish) were completed, a comparison of the results obtained was carried out (chapters 26.1 and 26.2). Two main results are derived from this work: (i) the reliable automation of the annotation process for thematic features in the English dataset ; and (ii) the creation of a bilingual (English-Spanish) corpus of newspaper texts from different genres, fully annotated with thematic features, and which can be used for a number of linguistic and NLP applications.



# 22

## The GATE Tool

For the current investigation, I decided to use GATE Tool *General Architecture for Text Engineering*, (Cunningham et al., 2011). This is a multipurpose tool that supports computer scientists with the possibility of developing in JAPE annotation applications; and a general graphical interface: GATE Developer that provides “a rich set of graphical interactive tools for the creation, measurement and maintenance of software components for processing human language” (Cunningham et al., 2011).

The reason for the selection of this tool can be summarized in the requirements specified by Orasan (2003):

1. **Easy to use** with a minimum time required to learn how it works.
2. **Friendly interface** hiding unnecessary details, as XML tagging, from the annotator.
3. **Multitasking** so the annotators do not need to learn how to use a new tool for different tasks.
4. **Multi language** ensures consistency between annotators that can use a language-independent tool.

Some relevant concepts cited by GATE's authors (Cunningham et al., 2011), are essential for the understanding of the tool, such as **annotation sets**, that are groups of annotations; or **processing resources**: plugins that create annotations on documents, these plugins automatically annotate the document or corpus and can be arranged in sequences, allowing for consecutive annotations of various types in different layers.

The following GATE components have been used throughout this process:

1. **Processing Resources (PRs)**. Some of the PRs used for the annotation of the corpus are provided by GATE. In the following lines I refer to the ones used in this study. These are arranged in different layers: first the three layers for the pre-processing or bootstrapping, namely: Morphosyntactic, syntactic and semantic layer; and a fourth layer, called Functional Layer which is the one developed for the automatic annotation of Thematic Features for the English corpus. This application is developed in JAPE and added to GATE as a new layer of annotation, that is, as another Processing Resource.

(a) Morphosyntactic Layer

The morphosyntactic layer makes use of ANNIE resources. ANNIE is A Nearly-New IE system (Cunningham et al., 2011) its components are included in GATE. Among them the following have been used:

- Tokenizer: ANNIE English and Spanish Tokeniser splits the text into very simple tokens as numbers, punctuation and words of different types.
- Segmentation: ANNIE Sentence Splitter segments the text into sentences, our units of analysis. This resource is the same for both languages
- Part of Speech Tagger: ANNIE PoS Tagger produces a part-of-speech tag as an annotation on each word or symbol, this is used for English. Spanish makes use of Tree Tagger, although not part of ANNIE resources, it is also available in GATE.

(b) Syntactic Layer

This layer of annotation has only been used for English, since at the moment when this study was carried out there were no reliable and available syntactic parsers for Spanish<sup>1</sup>. English makes use of the Stanford Parser. The tag-set used by Stanford in its annotations comes from the Penn Treebank project (Mladová et al., 2008). For the description of the different types of sentences

---

<sup>1</sup>The use of Freeling or Meaning Cloud was not available as a downloadable PR for GATE

---

it is necessary to identify them in the Penn Treebank tag-set, to be able to construe rules appropriately.

(c) Semantic Layer

These annotations are not strictly needed for the identification of the Thematic Field or its components but it will be of use in further exploitation of the annotated corpus.

- Number annotation: Numbers Tagger is a Processing Resource provided by ANNIE that annotates numbers.
- Date annotation: Date Tagger is a Processing Resource provided by ANNIE that identifies and annotates dates.

(d) Functional layer

This final layer is a new Processing Resource that has been created in JAPE (Java Annotation Patterns Engine) to annotate the Thematic Field and its components. The prior condition needed for its proper functioning is that the documents have been previously annotated with the morphosyntactic, syntactic and semantic parsers, described above, and at token level.

2. **ANNIC**. The other component of GATE used in this study is ANNIC (ANNo-tations In Context) (Cunningham et al., 2010) a tool for indexing and retrieving information integrated within the data stores. This tool allows for querying and studying the context where the annotations occur. The queries can be very complex as its language is based on JAPE. Statistical analysis can be carried out over the results obtained and these export to external files. This tool has been used to extract data from the automatically annotated corpora to compare it with the manual annotation.



# 23

## Corpus and Metadata

The corpus used for the automatic annotation was the same that was manually annotated and stored in Excel files, as described in Part D above. For the automation task, the corpus was subdivided into three subcorpora, each corresponding to the three genres analyzed in this study, as shown on table 23.1:

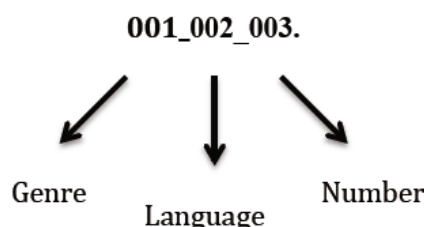
**Table 23.1** Corpus for automatic annotation

Corpus	English Texts	Spanish Texts
Reports	30	30
Editorials	30	30
Letters to the Editor	50	50
Total	110	110

As to corpus metadata, Bird & Simons (2001) explain that linguistic resources are increasingly expanding but the accessibility to these resources is not always easy and web indexation to find these resources is not always sufficient. They propose and encourage linguists to keep a proper documentation of their work in metadata.

Metadata is said to be “data about data” and it serves for different purposes that ranges from the proper storage and exploitation of the data to a posterior retrieval of the annotated corpora, if accessible through the internet. The Open Language Archives Community (OLAC) is probably the best system to organize information contained in corpus metadata for inclusion in the OLAC search engine (Hughes, 2005). However, the progress of my automatic annotated corpora still needs further development until it can be included in OLAC. For that reason, corpus metadata in this study was organized by creating a correlation between the corpus file names and a XML archive data storage, as shown in figure 23.1 below:

The naming of the files will ensure the unique identification of each text under the following structure:



**Figure 23.1** Metadata file format

The names of the archives could have in 001 either NR (news reports), Ed (editorials) or LE (letters to the editor). In 002 the available languages Eng (English) or Spa (Spanish) and in 003 the number of the text ranging from 1 to 50. The corresponding file extension is added after the dot, depending on whether it is a preprocessed text (.TXT) or a processed text in XML format.

**NR\_Spa\_07.**

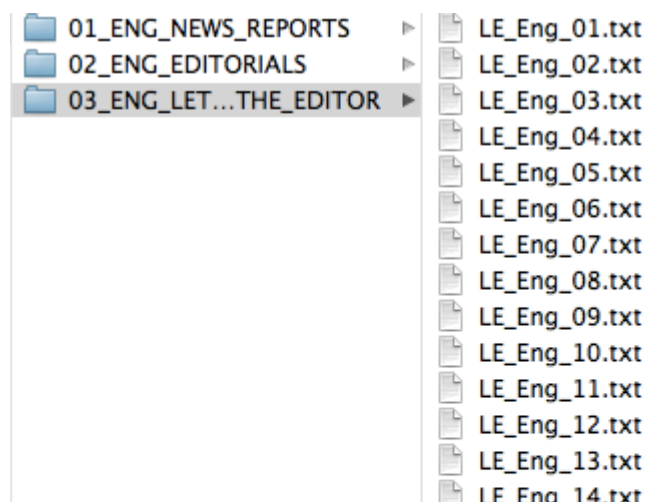
**Figure 23.2** Example of a file name

In order to preprocess the corpus in GATE, all texts were transformed from the original file (normally HTML or XML as they are gathered from the Internet) to .TXT (UTF-8) each in a separate file, and named with an ID, which identifies each text in an Excel file that contains all metadata. The Excel file has the ID for the text in each row and the following fields: Title, Genre, Newspaper, Language, Date and Author. An example of the metadata page is shown in Figure 23.3 (the full metadata are presented in appendix E):

ID	Title	Genre	Newspaper	Language	Date	Author
NR_Eng_01	Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair	News Report	Times online	English	2008	Adam Sage
NR_Eng_02	Opec hawks want to cut oil production to keep up price	News Report	Times online	English	2008	Carl Mortished
NR_Eng_03	Microsoft results offer hope to tech sector	News Report	Times online	English	2008	Suzy Jagger and Mike Harvey
NR_Eng_04	Dutch Government to inject €10bn of fresh capital into ING savings bank	News Report	Times online	English	2008	Lilly Peel
NR_Eng_05	G7 'preparing to drive down the yen'	News Report	Times online	English	2008	Leo Lewis
NR_Eng_06	Ukraine agrees terms of \$16.5bn rescue by IMF	News Report	Times online	English	2008	Lilly Peel
NR_Eng_07	Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi	News Report	Times online	English	2008	Grainne Gilmore
NR_Eng_08	US economy officially on brink of recession	News Report	Times online	English	2008	Gary Duncan
NR_Eng_09	Narrow win for 'Merkel's Mac' in Lower Saxony	News Report	The Independent online	English	2013	Tony Patterson
NR_Eng_10	Georgian protesters want President Saakashvili out	News Report	The Independent online	English	2013	Reuters
NR_Eng_11	Top executive suppressed report damning Barclays	News Report	The Independent online	English	2013	James Moore
NR_Eng_12	Think-tank calls for new strategy to fire up the economy	News Report	The Independent online	English	2013	Ben Chu
NR_Eng_13	Santander eyes a double bank buy with £2bn tag	News Report	The Independent online	English	2013	James Thompson
NR_Eng_14	Tesco opens fifth dotcom store	News Report	The Independent online	English	2013	James Thompson
NR_Eng_15	UK firms to bid for Japan's nuclear clean-up	News Report	The Independent online	English	2013	Mark Leftly
NR_Eng_16	Hilco UK emerges among favourites to give HMV a spin	News Report	The Independent online	English	2013	James Thompson
NR_Eng_17	UK seems likely to lose EU battle over limits on bonuses	News Report	The Independent online	English	2013	Nick Goodway

**Figure 23.3** Metadata example

As shown in Figure 23.3, the first column contains the unique ID for each text, the second column lists the title, the third one the genre, the fourth column shows the source of the text, the fifth the language, the sixth the date and the seventh the author. Thus, the list of the preprocessed texts in the archive looks like this (the full preprocessed corpus is available in appendix E):



**Figure 23.4** Folder with preprocessed files





# 24

## Automatic Annotation of Theme in English

### 24.1 Corpus Preprocessing

The preprocessing or ‘bootstrapping’ of the English corpus with GATE consisted of the following steps: segmentation, tokenization, part-of-speech tagging, and parsing. This means that certain categories will be automatically pre-annotated, these annotations will be used as a basis for the annotation of Theme, a higher level annotation. Each of these pre-annotations are described in detail below.

#### 1. Segmentation

In this process the string of characters which form what it is known as a ‘text’ is segmented into sentences, the unit of analysis. The tool used for this process is the ‘Sentence Splitter’, which is domain-and-application-independent (Cunningham et al., 2011). It works on a cascade of finite-state transducers that segment text into sentences. It uses a list of abbreviations to distinguish full-stops of sentences from other possible dots. The result is the following:

- the unit of analysis appears as ‘Sentence’
- the space in the sentence break as ‘Split’ annotation with two possible values ‘Internal’ for exclamations or question marks and ‘External’ for a newline

## 24. Automatic Annotation of Theme in English

The following figure shows the Split marks (highlighted in purple) that serve as delimiters to identify sentences.

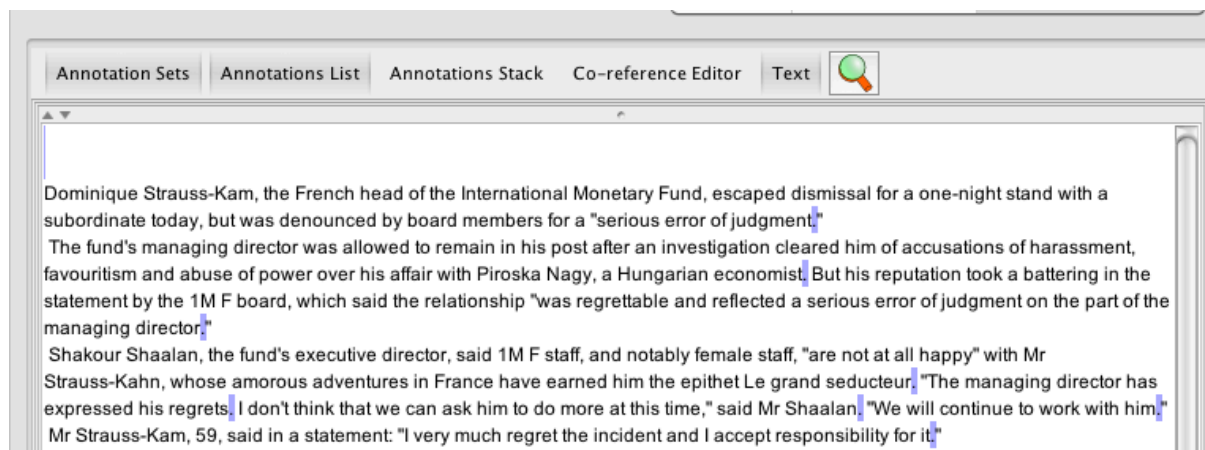


Figure 24.1 English Sentence splitting

### 2. Tokenization

In this process the text is split into simple tokens as words of different types, numbers, symbols, punctuation and space tokens. Words are distinguished in upper-case, lowercase and mix caps, and between end and start punctuation. The tool used is called 'Tokenizer' (Cunningham et al., 2011). Image 24.2 shows the identification of tokens in a text (highlighted in blue).

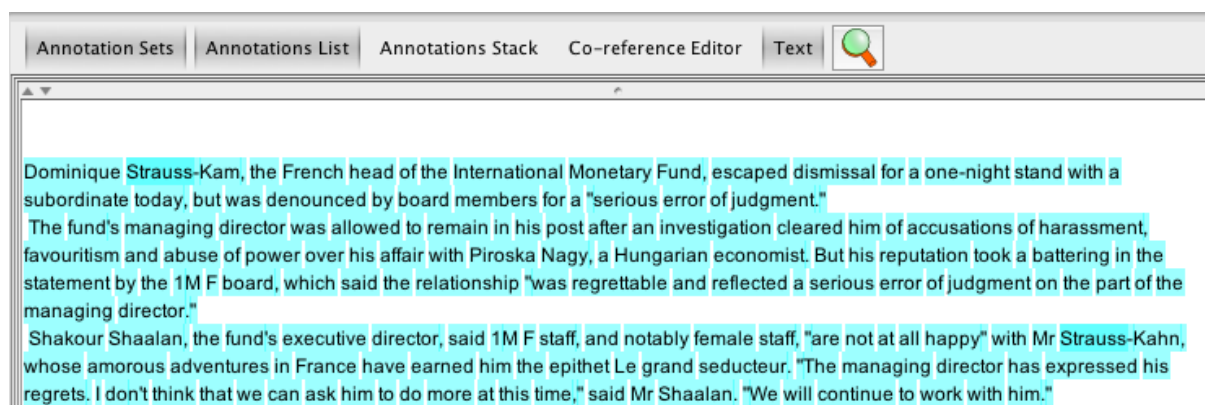
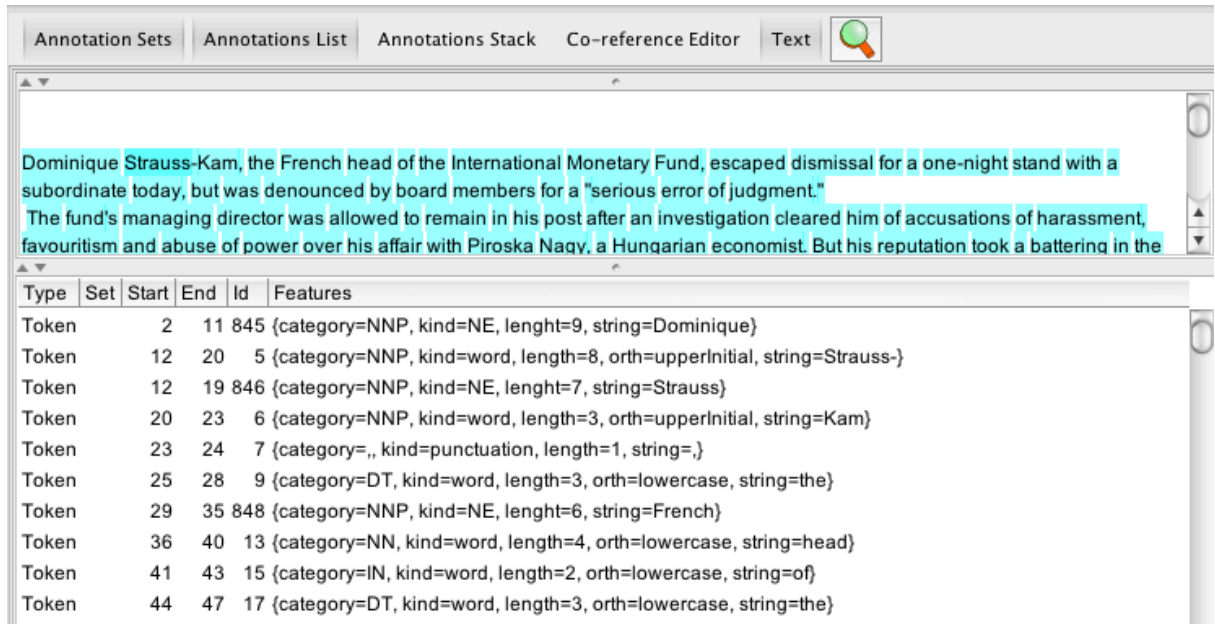


Figure 24.2 English Tokenizer

3. **Part of Speech Tagging** The ANNIE Part-of-Speech Tagger used is based on the Brill tagger, it produces a part-of-speech tag as an annotation on each word

or symbol. The following figure shows an example of some of the tags used in the PosTagger<sup>1</sup>:



The screenshot shows the Stanford POS Tagger interface. The top bar includes tabs for 'Annotation Sets', 'Annotations List', 'Annotations Stack', 'Co-reference Editor', and 'Text'. The main text area displays a paragraph about Dominique Strauss-Kam. Below the text, a table lists the tokens and their annotations.

Type	Set	Start	End	Id	Features
Token		2	11	845	{category=NNP, kind=NE, length=9, string=Dominique}
Token		12	20	5	{category=NNP, kind=word, length=8, orth=upperInitial, string=Strauss-}
Token		12	19	846	{category=NNP, kind=NE, length=7, string=Strauss}
Token		20	23	6	{category=NNP, kind=word, length=3, orth=upperInitial, string=Kam}
Token		23	24	7	{category=., kind=punctuation, length=1, string=,}
Token		25	28	9	{category=DT, kind=word, length=3, orth=lowercase, string=the}
Token		29	35	848	{category=NNP, kind=NE, length=6, string=French}
Token		36	40	13	{category=NN, kind=word, length=4, orth=lowercase, string=head}
Token		41	43	15	{category=IN, kind=word, length=2, orth=lowercase, string=of}
Token		44	47	17	{category=DT, kind=word, length=3, orth=lowercase, string=the}

Figure 24.3 English PoS Tagging

#### 4. Parsing

In this study the Stanford Parser Application (SPA) was chosen for the parsing process. It is a probabilistic parsing system implemented in JAVA by Stanford University's Natural Language Processing Group, it is available for parsing Arabic, Chinese, English, and German. Documents to be processed by the Parser PR must be already segmented and tokenized.

The Stanford Parser annotates:

**ConstituentAnnotations** annotates syntactic constituents.

**SyntaxTreeNode annotations** compatible with the Syntax Tree Viewer.

**DependencyAnnotations** indicate the dependency relations in the sentence.

**DependencyFeatures dependencies** are Token annotations to indicate the dependency relations in the sentence.

<sup>1</sup>A full list of the tags used can be found at <https://gate.ac.uk/sale/tao/splitap7.html> (Cunningham et al., 2011).

## 24. Automatic Annotation of Theme in English

Other ready-made applications, as reliable as the SPA were first tried on the bilingual corpus (e.g. LingPipe, ANNIE and OpenNLP), but they are oriented to Information Extraction and Semantic Annotation whereas the aims of this study are closer to syntax, and therefore the SP was preferred.

The following figure shows a display of the bootstrapping processed performed in GATE with the SPA at the end of the processing pipeline. the different tools mentioned above are listed as a pipeline in the right top window.

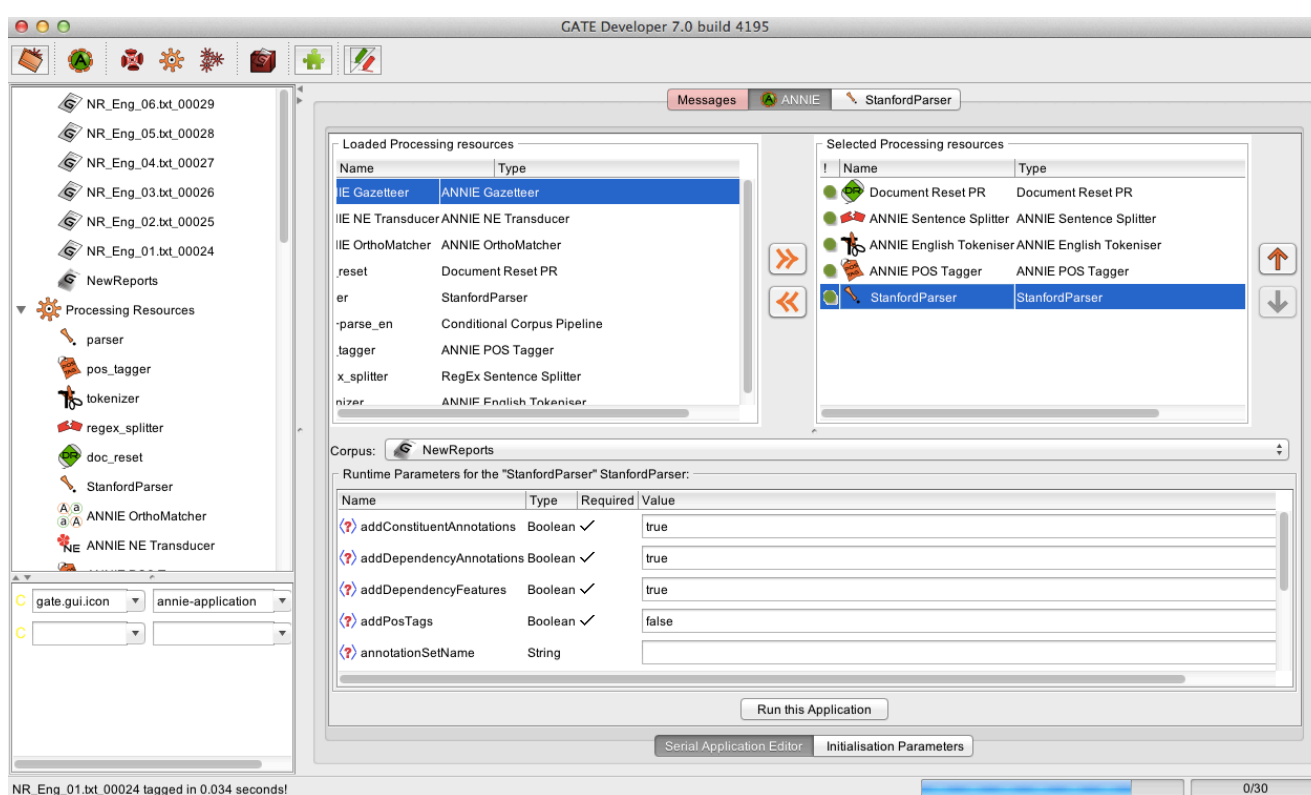


Figure 24.4 English bootstrapping pipeline

## 24.2 Rule Based Automatic Annotation of Theme in English

Once the preprocessing or bootstrapping was carried out, the high level annotation of Theme takes place. For this step an *ad hoc* tool has been developed: an application (API) that runs in GATE to automatically add new layers of annotation containing the

studied features. The API was fully programmed in JAPE<sup>1 2</sup> from scratch based on the analysis and the work presented in Parts B and C of this dissertation. The definition of JAPE rules to automate the annotation of the Thematic Field and its elements, was a ‘translation’ of the annotation guidelines to a format which could be understood by a programming language such as JAPE.

This proved to be a hard task since the “translation” procedure required a great deal of clarification and explanation from the point of view of human annotators. In the following, I will describe the rules created for the annotation of the features of the Theme model which were empirically validated in Part C.

### 24.2.1 Rule for Thematic Field in English

The definition given in the guidelines is the following: “Initiating clause span of varying length up to and including the first nuclear constituent in main clause.” (see appendix C.1). This means that the main verb of the sentence has to be identified first and then marked as a reference for the identification of the Thematic Field span.

The identification of the main verb is one of the factors that have to be carefully pinned down according to the previous bootstrapping. The information that has to be stated to the system is *what* is the main verb, previously identifying the types of sentences annotated by the Stanford parser. The tag-set used by Stanford in its annotations is the Penn Treebank project (Marcus et al., 1993), therefore for the description of the different types of sentences it is necessary to identify them in the Penn Treebank tag-set. The tags included in this tag-set are listed below for convenience:

S	Simple declarative clause
SBAR	Clause introduced by subordinating conjunction or 0
SBARQ	Direct question introduced by wh-word or wh-phrase
SINV	Declarative sentence with subject-aux inversion

Once the types of sentences are identified, the identification of the main verb is also done on the basis of the tags for the Verb categories in the Penn Treebank tag-set, as specified below:

---

<sup>1</sup>JAPE: Java Annotation Patterns Engine

<sup>2</sup>The actual writing of the JAPE rules has been carried out by Sara Lana Serrano, lecturer at the Telematic Engineering and Electronics Department of the Escuela Técnica Superior de Ingeniería y Sistemas de Telecomunicación (Universidad Politécnica de Madrid. Spain)

VB	Verb, base form
VBD	Verb, past tense
VBP	Verb, non-3rd person singular present
VBZ	Verb, 3rd person singular present

Figure 24.5 shows an example of the information provided by the Stanford Parser in an annotated sentence:

**Parse**

```
(ROOT
  (S
    (NP
      (NP (NNP Dominique) (NNP Strauss-Kam))
      (, ,)
      (NP
        (NP (DT the) (JJ French) (NN head))
        (PP (IN of)
          (NP (DT the) (NNP International) (NNP Monetary) (NNP Fund))))
      (, ,))
    (VP
      (VP (VBD escaped)
        (NP (NN dismissal))
        (PP (IN for)
          (NP (DT a) (JJ one-night) (NN stand))))
        (PP (IN with)
          (NP (DT a) (JJ subordinate) (NN today))))
      (, ,)
      (CC but)
      (VP (VBD was)
        (VP (VBN denounced)
          (PP (IN by)
            (NP
              (NP (NN board) (NNS members))
              (PP (IN for)
                (NP
                  (NP (DT a) (`` ``) (JJ serious) (NN error))
                  (PP (IN of)
                    (NP (NN judgment))))))))))
      (. .) (`` '')))
```

**Figure 24.5** English Stanford parsing example

An important issue that has to be considered is that the creation of a rule depends on the accuracy of the previous preprocessing steps, i.e., segmentation, tokenization, pos-tagging and parsing. On the basis of this preprocessing, the following rule for thematic field identification was created:

For Sentences of the type S and SINV: given a token of the type VP,  
search for all VP in which this VP token may be found and annotate it  
with the label TF CONTEXT VP when it is not part of a subordinate type

For Sentences of the type SBARQ and SBAR: Thematic Field are the group of  
tokens from the beginning of the sentence until TF CONTEXT VP

The step by step command include:

Find the first TF CONTEXT VP (the one with the smallest offset)

Search TF SINV

If any mark until the first VP

If the att type of the verb is impersonal label including the first VP  
found

...

All steps taken into the identification of the relevant elements for the annotation  
have to be explicitly stated, as well as what to do if these elements are not found.

### 24.2.2 Rule for Textual Theme in English

The Textual Theme of the SFGS model is defined in the annotation guidelines as  
follows: “Elements which are instrumental in the creation of the logical connections in  
the texts, such as linkers, binders or correlatives.” (see appendix C.3). The features  
which were considered when creating the rules for annotating Textual Theme in English  
are the following:

It is not an obligatory element

It is possible to find only one or several consecutively found

It is contained in the Thematic Field

It is found at the beginning of the sentence

It can be preceded by an Interpersonal Theme

It is not part of a NP

It is not part of a IT Gazetteer

These features have to be clearly stated and it is important to mark them as part  
of the Thematic Field. Also, the elements which are candidates to be Textual Themes  
are previously annotated in the Tokenizer as:

**EVIDENCES** elements that can realized Textual Theme are marked as:

CC	Conjunctions
LS	List item marker (first, second...)
IN	Preposition of conjunction

When these rules were first tested on the documents many elements were not properly annotated. In order to improve the results, a Gazetteer was introduced. This “consists of a set of lists containing names of entities such as cities, organizations, days of the week, etc. These lists are used to find occurrences of these names in text, e.g. for the task of named entity recognition. (...) Gazetteers usually do not depend on Tokens or on any other annotation and instead find matches based on the textual content of the document” (Cunningham et al., 2011:chapter 13). Since the elements prone to be Textual Themes are based on finite lists of elements (conjunctions, prepositions..) I decided to include a Gazetteer to solve possible mistakes in the annotation of this elements.

In a subsequent step, a classification of the annotation rules was added to annotated the subtypes of Textual Theme (Binders and Linkers). For this purpose the aid of the Gazetteer helped in the distinction between linkers and binders. Also, since the list of elements is limited, it is possible that some might be left out. To avoid this problem, the type attribute is left blank but it is identified with the label ‘Textual Theme’.

### 24.2.3 Rule for Interpersonal Theme in English

The definition given in the guidelines for Interpersonal Themes and its subtypes is: “These are elements which express the attitude and the evaluation of the speaker with respect to his/her message. These include: Vocatives, comment adjuncts, modal adjuncts, engagement and stance expressions” (see appendix C.3)

Due to the complexity of this element and its subtypes, the transformation of the rules was arranged in two different steps. The first step took care of the types vocatives, comment adjuncts and modal adjuncts. The second dealt with the types engagement and stance.

The features which were considered when creating the rules for annotating Interpersonal Theme in English are the following:



It is not an obligatory element, a sentence may not have any

There can be only one or more in sequence

It is contained in a Thematic Field

It is found at the beginning of the sentence

It can be preceded by a Textual Theme

It is not part of a text quote

It is not part of a NP

It is not part of a TT Gazetteer

In this phase the identification of the three first types of Interpersonal themes posed difficulties, as the differentiation between Comment adjuncts and Modal adjuncts is purely semantic. The decision was to unify or ‘neuter’ both subtypes into only one (Comment and Modal into ComModal), having in this phase identified Vocatives and ComModal types.

From the preprocessing of these elements, I used the following evidences:

**EVIDENCES** elements that can realize Interpersonal Themes are marked as:

UH Interjections

RB Adverbs

NNP! Vocatives

A Gazetteer was also used for the disambiguation of overlapping occurrences of Adverbs.

For the identification of the types Stance and Engagement, whose disambiguation depends on semantic information, I decided to ‘neuter’ these two categories into a unified one which I labelled StaEng. This follows the suggestion made by Hovy & Lavid (2010), who propose to ‘neuter’ certain theoretical categories when low agreement is found between annotators but still one would want to annotate them.

In a posterior classification the identification of the types Stance and Engagement was addressed. As with Comment and Modal adjuncts the differences between Stance and Engagement are also purely based on semantics, and what is more, the results obtained through inter-annotator agreement were the lowest of the studied categories 13.3.2.4. Therefore the decision was to treat these elements as one single category: ComModal.

#### 24.2.4 Rule for PreHead in English

The annotation guidelines based on the Theme model of SFGS defined the PreHead as “any circumstantial and/or finite element preceding the Thematic Head” (see figure C.3).

The features which were considered when creating the rules for annotating PreHead are the following:

It is not an obligatory element, a sentence may not have any

There can be only one

It is contained in a Thematic Field

It is found at the beginning of the sentence

It can be preceded by a Textual Theme(s) and/or Interpersonal Theme(s)

The complexity of this element did not make it possible to take direct advantage of the automatically annotated elements as such, but further work needed to be done on them to benefit from the coded information. For this purpose a rule was created to identify adverbs and subordinate sentences that fulfill the requirements these components require (as stated above). Two labels were added:

TF RB TOKEN            adverbs

TF SUB SENTENCES    subordinate sentences

This new layer of annotation helps in the identification of adverbs, annotated at token level, and subordinate sentences, at parsing level, susceptible of being annotated as PreHeads. For instance, the latter rule comprises elements such as those included in the PH Gazetteer (the aim of the Gazetteer is to disambiguate as in Interpersonal Themes), those labelled by the syntax tree node as ADVB or SBAR or PP, or are of the type TIMEX (expressions of time).

#### 24.2.5 Rule for Thematic Head in English

The guidelines account for Head as the “... first nuclear constituent (not circumstantial) element in the clause, this can be a Participant or a Process. The types of Thematic Head are There, Thematic Equative and Predicated Theme.” (see appendix C.3)

The main features used for the annotation of Thematic Head in English are the following:

It is an obligatory element, all sentences have to have one Thematic Head

There is only one Thematic Head per sentence

It is contained in the Thematic Field

Thematic Head and Thematic Field can be both realized by the same elements

The Thematic Head can never be larger than the Thematic Field

It is not part of a text quote

It is not part of a PreHead

It can appear preceded by an Interpersonal Theme(s) or/and Textual Theme(s) or/and PreHead

It is the subject in declarative sentences

It is the wh- element in wh- questions

It is the verb in questions

It is the verb in imperative sentences

As shown above, the features to be considered when annotating Head are far more complex than the rest of the elements in the Thematic Field. This complexity is reflected in the preprocessing steps which use both token level annotation and syntax level annotations. For instance, in declarative clauses, the Subject can be Thematic Head when:

Token NNP	Proper Names
Token NN	Common Nouns
SyntaxTreeNode NP	Noun Phrases
SyntaxTreeNode NP + SBAR	Noun Phrases + Appositions

In wh- questions:

Token WRB    wh adverb

Token WP\$    Wh pronoun what

In yes/no questions, the first set of verbs can be Head on its own and combined with the second set of verbs:

Token VBD	Verb, past tense
Token VBP	Verb, non-3rd person singular present
Token VBZ	Verb, 3rd person singular present
Token VBG	Verb, gerund or present participle
Token VBN	Verb, past participle
Token VB	Verb, base form
Token MD	Modal auxiliary

Another particular feature of Head (that will be mirrored on the Thematic Field) is the treatment of *quoted text* or *reporting clauses* (see chapter 11.2.2.8). Quotations are considered Head when the next element is a verb (TF CONTEXT VP), but if followed by the subject then they will be noted as PreHead.

### 24.2.5.1 Rule for There Theme in English

Existential clauses (see chapter 11.2.2.4) are identified with the Head *There*. That is, when Head in declarative clauses is of this type, the type of Head is noted as *There*.

### 24.2.5.2 Rule for Predicated Theme in English

Predicated Theme (see chapter 11.2.2.3) presents a particular syntactic structure:

- (1) [It is Mr Paterson, above all, who] has questions to answer
- (2) [It is you who] is to blame

In this rule the normal annotation pattern has to be altered: the default annotation of *It* as Thematic Head is not correct in this particular case. The complex structure has to be identified, to do so the second part of the structure is considered as the beginning of an *SBAR* (*who has questions to answer* and *who is to blame*). Thus, the structures that match with *It* and *SBAR* are selected and the span between them labelled as Predicated themes.

### 24.2.5.3 Rule for Thematic Equative in English

The third type of Head (see chapter 11.2.2.2) is enclosed inside another syntactically difficult structure that is not straight forward identified by Stanford. The solution

found to identify these elements is the following:

**Transformation of a PreHead in Thematic Equative if:**

SBAR category

Begins by *what*

The container Thematic Field does not have any other identified Head



# 25

## Semi-Automatic Annotation of Theme in Spanish

While for English, the GATE platform offered a wide range of tools for corpus pre-processing, this was not the case for Spanish. The lack of availability of accurate pre-processing tools has conditioned the development of automatic annotation of Theme in this language. In the following sections, I will describe the steps taken towards this goal.

### 25.1 Corpus Preprocessing in Spanish

The bootstrapping pipeline for Spanish was the following:

1. **Segmentation**

*Sentence Splitter*, as said in English Bootstrapping section 24.1, is domain and application-independent (Cunningham et al., 2011). Therefore it was also used for the segmentation into units of analysis in the Spanish corpus.

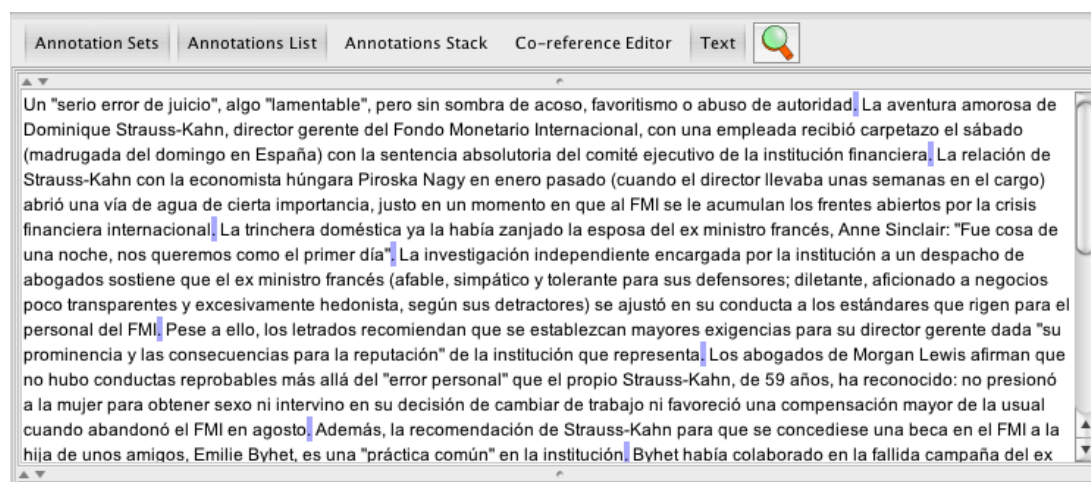


Figure 25.1 Spanish Sentence splitting

## 2. Tokenizer

GATE Unicode Tokenizer splits the text into simple tokens as words of different types, numbers, symbols, punctuation and space tokens. Words are distinguished in uppercase, lowercase and mix caps, and between end and start punctuation. (Cunningham et al., 2011)

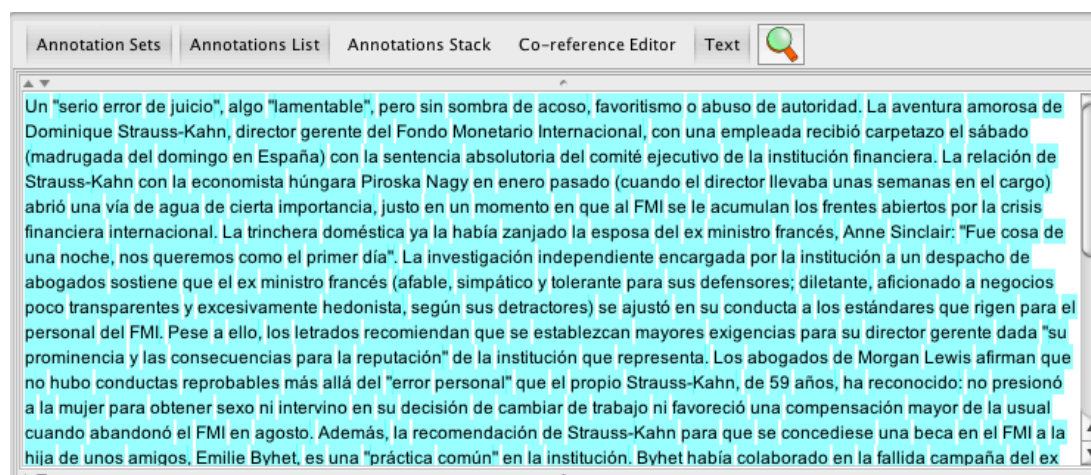


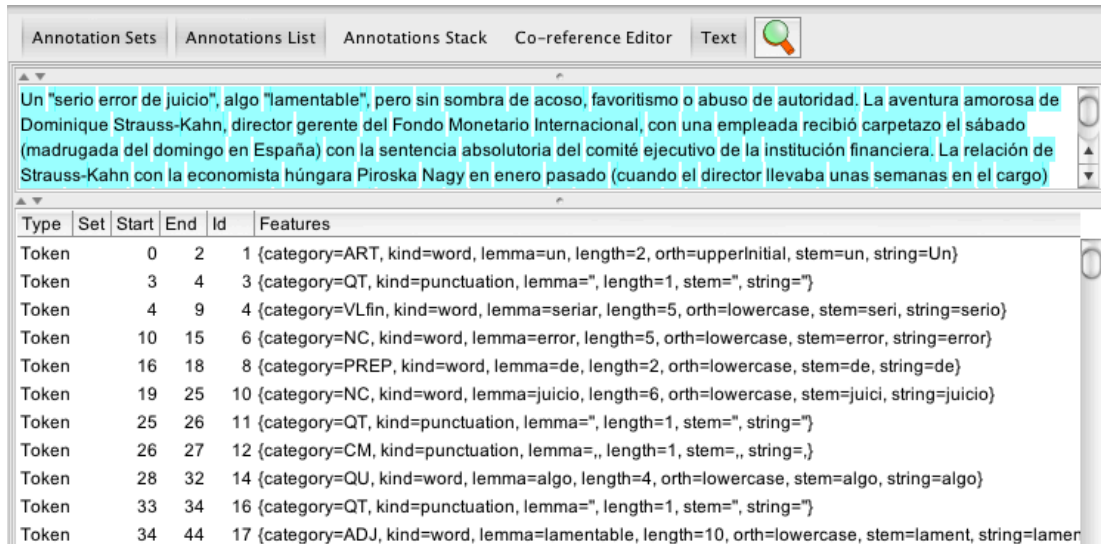
Figure 25.2 Spanish Tokenizer

## 3. Part of Speech Tagger

The tool used for the tokenization of the text was the Tree Tagger. This plugin is a language independent part-of-speech tagger that supports several languages, among them Spanish. Being an external program, GATE calls it through its



wrapper and adds gate tokens two new features: the postagger tag and the lemma of the token (Cunningham et al., 2011). The installation and running in the GATE pipeline was not a trivial task, causing a number of difficulties.

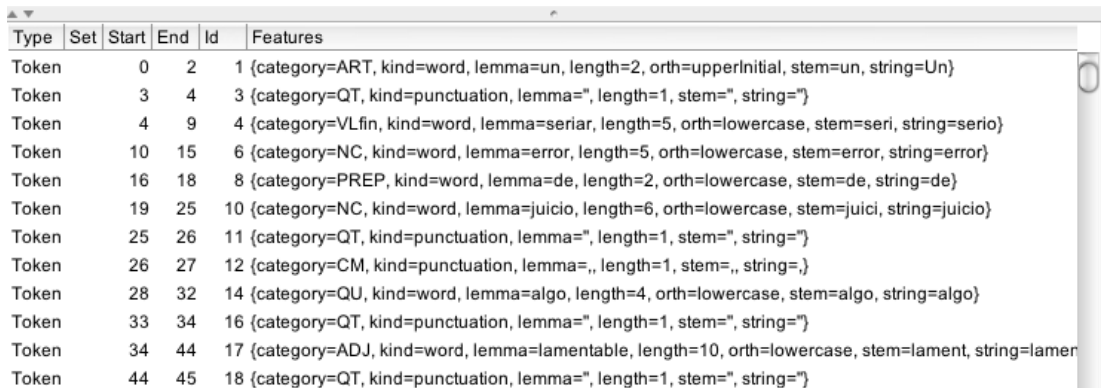


Type	Set	Start	End	Id	Features
Token		0	2	1	{category=ART, kind=word, lemma=un, length=2, orth=upperInitial, stem=un, string=Un}
Token		3	4	3	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		4	9	4	{category=VLfin, kind=word, lemma=seriar, length=5, orth=lowercase, stem=seri, string=serio}
Token		10	15	6	{category=NC, kind=word, lemma=error, length=5, orth=lowercase, stem=error, string=error}
Token		16	18	8	{category=PREP, kind=word, lemma=de, length=2, orth=lowercase, stem=de, string=de}
Token		19	25	10	{category=NC, kind=word, lemma=juicio, length=6, orth=lowercase, stem=juici, string=juicio}
Token		25	26	11	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		26	27	12	{category=CM, kind=punctuation, lemma=., length=1, stem=., string=.,}
Token		28	32	14	{category=QU, kind=word, lemma=algo, length=4, orth=lowercase, stem=algo, string=algo}
Token		33	34	16	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		34	44	17	{category=ADJ, kind=word, lemma=lamentable, length=10, orth=lowercase, stem=lament, string=lamen

Figure 25.3 Spanish PoS Tagging

#### 4. Stemmer

The stemmer plugin ‘Stemmer Snowball’, consists of a set of PRs for 11 European languages. Each Token is annotated with a new feature ‘stem’, with the stem for that word as its value. This feature is important for Spanish for the annotation of the lexical and grammatical form of the verb as PreHead and Head in the Theme.



Type	Set	Start	End	Id	Features
Token		0	2	1	{category=ART, kind=word, lemma=un, length=2, orth=upperInitial, stem=un, string=Un}
Token		3	4	3	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		4	9	4	{category=VLfin, kind=word, lemma=seriar, length=5, orth=lowercase, stem=seri, string=serio}
Token		10	15	6	{category=NC, kind=word, lemma=error, length=5, orth=lowercase, stem=error, string=error}
Token		16	18	8	{category=PREP, kind=word, lemma=de, length=2, orth=lowercase, stem=de, string=de}
Token		19	25	10	{category=NC, kind=word, lemma=juicio, length=6, orth=lowercase, stem=juici, string=juicio}
Token		25	26	11	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		26	27	12	{category=CM, kind=punctuation, lemma=., length=1, stem=., string=.,}
Token		28	32	14	{category=QU, kind=word, lemma=algo, length=4, orth=lowercase, stem=algo, string=algo}
Token		33	34	16	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}
Token		34	44	17	{category=ADJ, kind=word, lemma=lamentable, length=10, orth=lowercase, stem=lament, string=lamen
Token		44	45	18	{category=QT, kind=punctuation, lemma="", length=1, stem="", string=","}

Figure 25.4 Spanish Stemmer

#### 5. Parser

At this point the pipeline for Spanish is incomplete since there is no available tool that can be integrated into GATE and parse the Spanish texts in a similar way to what the Stanford Parser does for English. The lack of this element has made it impossible to develop the rules for automatic annotation.

Figure 25.5 below shows a screenshot of the preprocessing for Spanish texts with the GATE Tool. The described tools are listed as a pipeline in the right top window.

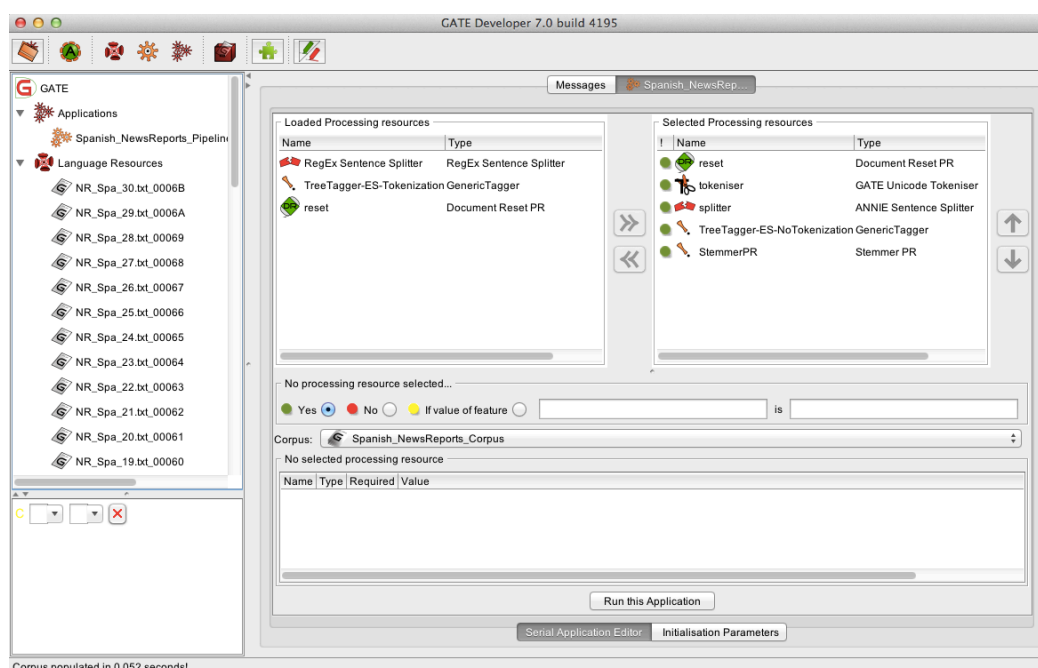


Figure 25.5 Spanish bootstrapping pipeline

## 25.2 Semi-automatic Annotation of the Spanish Corpus

Since automatic parsing was not possible for the Spanish corpus, the annotation of the Thematic features has been performed in a semi-automatic manner. The process was carried out in two steps: first, annotating the scope of the Thematic Field, and second, marking the elements of the Thematic Field, i.e.: Textual Theme, Interpersonal Theme, PreHead and Head. For this purpose it was necessary to create as XML a schema to define the tags to annotate.

Figure 25.6 shows the schema for the manual annotation of Thematic Field and its elements:

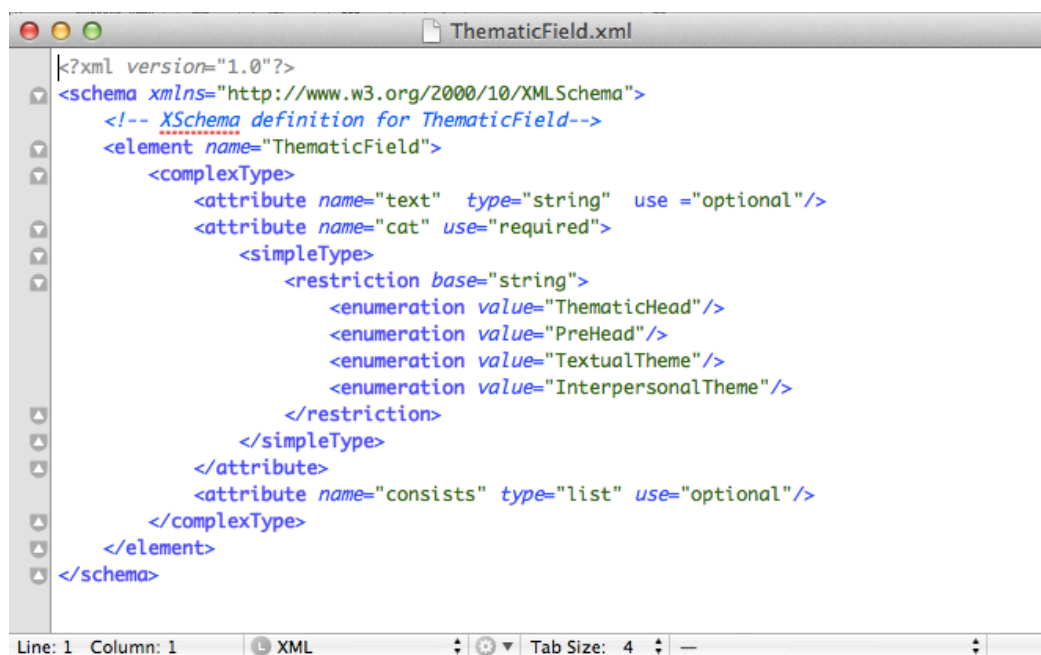
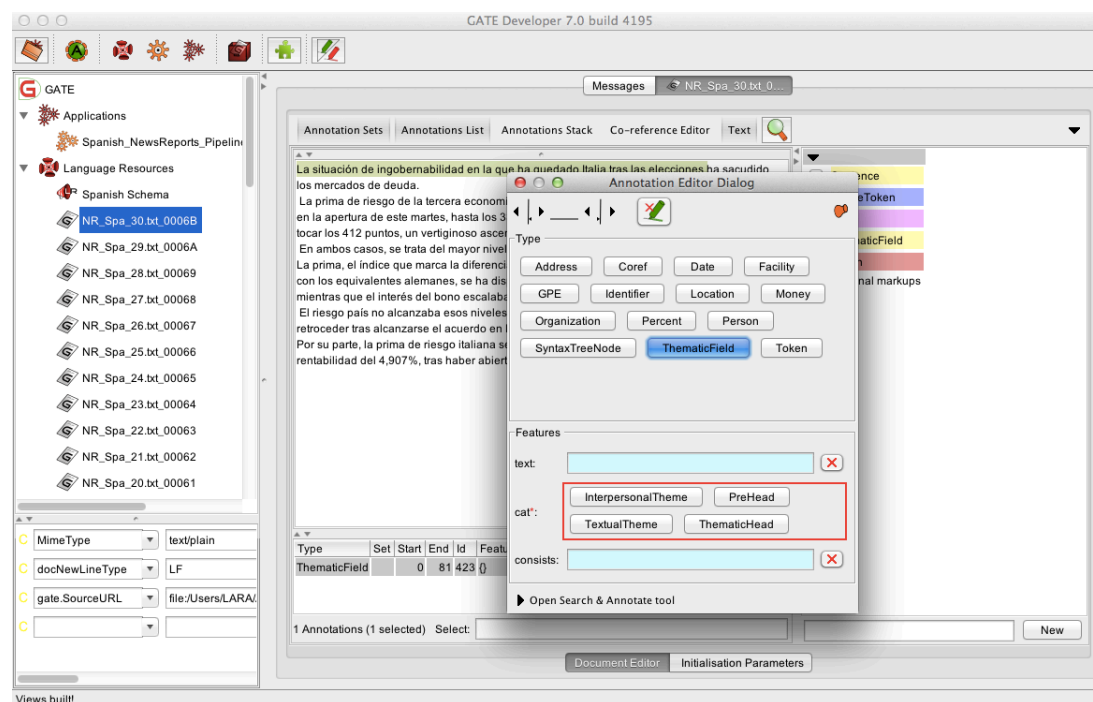


Figure 25.6 Schema for Thematic Field annotation and its elements

Once this schema is loaded in GATE the process of annotation begins. The benefits of using of this type of schemas is that it reduces the chances to misspell the name of a label, thus avoiding further problems in the annotation process.

## 25. Semi-Automatic Annotation of Theme in Spanish



**Figure 25.7** Schema annotation process

One new label is added and the annotations of the Thematic Field are kept within this new category of annotations. They can thus be exploited and indexed as any other automatic category having the possibility to run statistical analysis on them (see 27).

# 26

## Comparing the Results of Hand-coded versus Automatic Annotation

This section is dedicated to provide a quantitative and qualitative comparison of the results obtained through manual (or hand-coded) annotation, as carried out in Part D, and the results obtained through the automatic annotation for English and automatic and semi-automatic annotation for Spanish carried out with the GATE platform, as described above. The basis for comparison is the number of annotations obtained through hand-coded methods versus the number of annotations obtained through automatic procedures. The procedure analyzes whether there are statistically-significant<sup>1</sup> differences between these two types of annotations and study the qualitative relevance of the annotations at thematic level.

### 26.1 Comparing the Results of Hand-coded versus Automatic Annotation in English

In the following subsections I will present the results of comparing the manual with the automated methods of annotation. I will deal first with the sentence splitting and will

---

<sup>1</sup>Chi-square test was used according to Butler (1985)

later analyze the different thematic features studied in the English corpus.

### 26.1.1 Sentence Splitting in English

The comparative analysis examines whether there are statistically-significant differences between the number of clause complexes split automatically versus those which were manually segmented, as shown in table 26.1 below:

**Table 26.1** Sentence splitting automatic/manual annotation

Corpora	Manual	Automatic
Reports	349	395
Editorials	587	614
Letters to the Editor	366	393
Total	1302	1402

The distribution shows what seems to be a larger amount of clause complexes automatically split in comparison with those manually segmented. However, when the data was analyzed statistically, the differences were not found to be statistically-significant ( $P < 0.05$ ). This means that the hand-coded elements selected in one corpus as units of analysis (hand-coded) do not differ from the selection of these elements done automatically.

### 26.1.2 Thematic Field in English

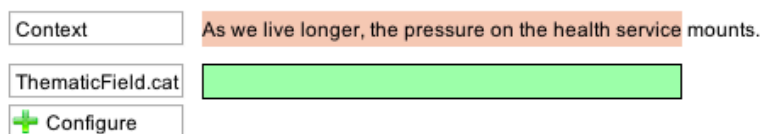
When comparing the number of sentences automatically annotated as thematic -and thus containing a Thematic Field- and those annotated manually, the results did not show any statistically-significant differences between both methods ( $P < 0.05$ ), as shown on table 26.2:

**Table 26.2** Thematic Field

Corpora	Manual	Automatic
Reports	338	309
Editorials	559	562
Letters to the Editor	357	325
Total	1254	1196

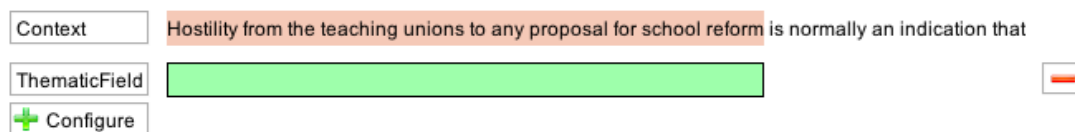
The following figures show two examples of Thematic Field automatically annotated. The green box underlines the span of the ThematicField:

- (1) As we live longer, the pressure on the health service mounts.



Thematic Field automatic annotation in GATE  
(ED-Eng-02. Sentence Id 787)

- (2) Hostility from the teaching unions to any proposal for school reform is normally an indication that the plan has merit.



Thematic Field automatic annotation in GATE  
(ED-Eng-11. Sentence Id 677)

It has to be pointed out that the accuracy of the annotations is quite high, the examples selected are not simple Thematic fields but constitute complex occurrences of the field.

### 26.1.3 Textual Themes in English

The comparison of manual and automatic annotations of Textual Themes in the English dataset is shown in the following table. The first line called Textual Theme in each type of text contains the proportions of Textual Themes found on the total number clauses in the corpus. The types Binder and Linkers are calculated on the number of Textual Themes.

## 26. Comparing the Results of Hand-coded versus Automatic Annotation

**Table 26.3** Textual Themes in automatic vs manual annotation

Corpus	Thematic Element	Manual	Automatic
News Reports	Textual Theme	6.21%	7.12%
	Binder	90.48%	63.64%
	Linker	9.52%	36.36%
Editorials	Textual Theme	21.12%	25.62%
	Binder	66.67%	59.03%
	Linker	33.33%	40.97%
Letters to the Editor	Textual Theme	13.33%	16.18%
	Binder	85.42%	74%
	Linker	12.05%	26%
	Correlative	2.08%	-

The differences are not statistically-significant which means that the automatic method achieves the same results as the manual one does. It also means that the JAPE rules created for this automatic annotation are a useful procedure for automating the annotation process. The same happens with the subtypes of Textual Themes, such as linkers and binders.

The following examples show the annotation of an occurrence of Binder and Linker in the corpus:

- (3) However, Hilco faces competition from other suitors for the UK business and the structure of any deal is likely to be complex, partly because HMV has many loss-making stores that a new owner is likely to jettison.

Context

entertainment group in 2011. However, Hilco faces competition from

TFC\_Textual\_Theme

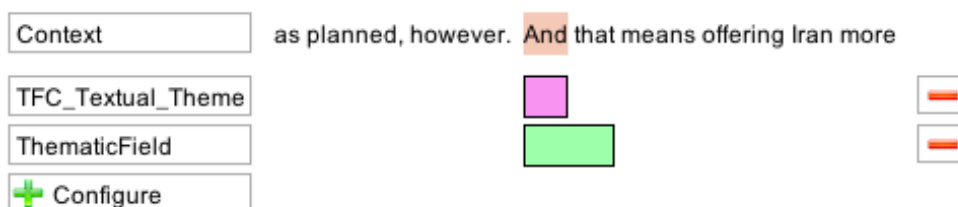
ThematicField

+ Configure

Textual Theme automatic annotation. Binder  
(NR-Eng-16. Sentence Id 2720)

- (4) And that means offering Iran more tempting incentives to scale back its enrichment programme.





Textual Theme automatic annotation. Linker  
(ED-Eng-29. Sentence Id 813)

The examples are taken from ANNIC, the tool that indexes all annotations and allows the user to query on them. It is through this means that these examples have been obtained, and as it can be noted in the examples above, different layers can be retrieved, not only the queried Textual Theme but also the Thematic Field annotations.

#### 26.1.4 Interpersonal Themes in English

Interpersonal Theme is one of the most difficult categories to annotate. In fact, the validation of the annotation categories 8 showed the lowest agreement (see chapter 13.3.2.4) among annotators. This was reflected in the creation of the rules. The first step was to merge two categories into one, as explained in section 24.2.3, the categories Comment and Modal considered as one (ComMod), and Engagement and Stance noted as EngSta. The results are shown in the column on the right account for these merged types.

**Table 26.4** Clause complex automatic annotation of Interpersonal Themes

Corpus	Thematic Element	Manual	Automatic
News Reports	Interpersonal Theme	0.30%	-
	ComMod	-	-
	EngSta	-	-
Editorials	Interpersonal Theme	11.27%	6.05%
	ComMod	34..37%	26.46%
	EngSta	65.63%	73.53%
Letters to the Editor	Interpersonal Theme	9.44%	5.23%
	ComMod	23.53%	29.41%
	EngSta	76.47%	70.59%



As shown above, the tendency of the annotated categories goes in the right direction, having found more elements in EngSta than in ComMod. Despite this fact, the number of identified

## 26. Comparing the Results of Hand-coded versus Automatic Annotation

---



elements is scarce in relation to the manually annotated categories. Regarding the accuracy of the annotations and taken into consideration the aforementioned difficulty in its proper marking, it has to be highlighted the level of complexity achieved. Some examples of EngSta (figures (5) and (6)) and ComMod (figures (7) and (8)) are shown below:

- (5) It is not hard to imagine what the BBC itself would say about any other organization that behaved in this manner, particularly one that had promised to be entirely open.

Context	not want to hear. It is not hard to imagine what the BBC itself would say about any other organisation that behaved in this manner,	
TFC_Interpersonal_Theme.type	EngSta	
 Configure		



Interpersonal Theme automatic annotation. EngSta  
(ED-Eng-28. Sentence Id 1365)

- (6) It is worth remembering too that Americans as a whole did not want a Republican House.

Context	as the bubonic plague. It is worth remembering too that Americans as a whole did	
TFC_Interpersonal_Theme.type	EngSta	
 Configure		

Interpersonal Theme automatic annotation. EngSta  
(ED-Eng-02. Sentence Id 1371)

- (7) Sadly, one suspects that if the banks were capable of being shamed out of this kind of antisocial behaviour, it would have happened long ago.

Context	" about their reputations. Sadly , one suspects that if	
TFC_Interpersonal_Theme.type	ComMod	
 Configure		

Interpersonal Theme automatic annotation. ComMod  
(ED-Eng-08. Sentence Id 736)

- (8) Clearly, those words of reproach were either not intended to be taken that seriously, or we just move on fast from such squabbles because, as this newspaper reveals, the UK sold several million pounds worth of arms last year to a regime we castigated only a few years ago.

Context

demanding a separate state. Clearly , those words of reproach

TFC\_Interpersonal\_Theme.type

ComMod

Configure

Interpersonal Theme automatic annotation. ComMod  
(ED-Eng-21. Sentence Id 686)

Although the annotation does not respond to the recall expected regarding the number of elements, it does so in terms of precision where the elements are annotated according to the rule.

### 26.1.5 PreHead in English

The following table shows the results obtained in the annotation of PreHead. As in previous sections the column Manual shows the results of the human-coded corpus and the column on the right Automatic, the results obtained in the automatic annotation of the corpus. As the table shows, there is no significant difference in the number of PreHeads found in news reports and letters to the editor, while there is a significant difference between manual and automatically annotated elements in editorials that will be explained below in detail.

**Table 26.5** Clause complex automatic annotation of PreHeads

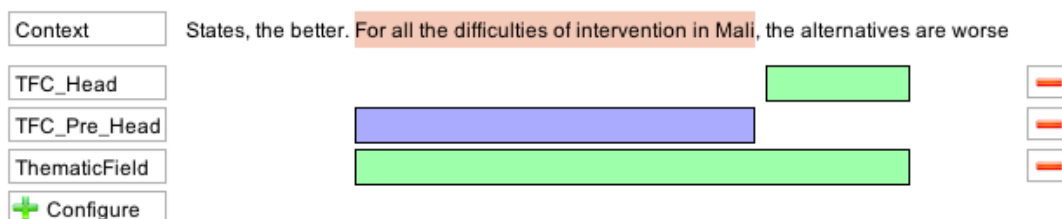
Corpus	Thematic Element	Manual	Automatic
News Reports	PreHead	16.57%	17.80%
	Circumstantials	57.14%	81.81%
	CCL	42.86%	18.18%
	Finite	-	-
Editorials	PreHead	31.33%	40.93%
	Circumstantials	55.06%	76.52%
	CCL	40.45%	23.47%
	Finite	4.49%	-
Letters to the Editor	PreHead	25.83%	32.31%
	Circumstantials	44.57%	70.10%
	CCL	43.48%	29.80%
	Finite	11.96%	-

Figures (10) and (9) show the complete annotation of the clause complex, the annotation of PreHead, Head and the span of the Thematic Field. The figure shows the complete annotation

## 26. Comparing the Results of Hand-coded versus Automatic Annotation

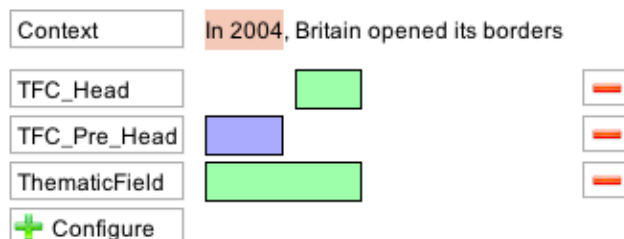
of the clause complex, the annotation of PreHead, Head and the span of the Thematic Field.

- (9) For all the difficulties of intervention in Mali, the alternatives are worse.



PreHead automatic annotation. Circ  
(ED-Eng-06. Sentence Id 1397)

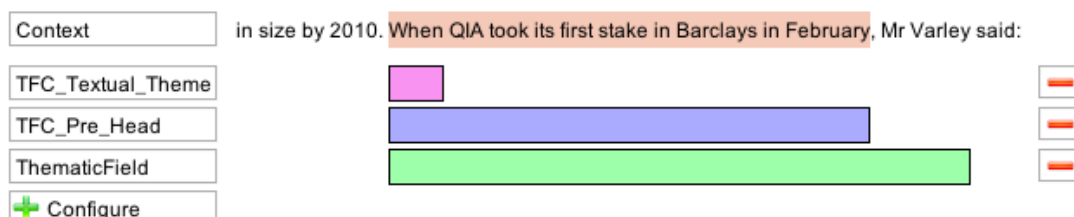
- (10) In 2004, Britain opened its borders to the new Eastern European members of the EU.



PreHead automatic annotation. Circ  
(ED-Eng-14. Sentence Id 1033)

The following two examples (11) and (12) show the annotation of the type of PreHead CCL realized by clause complexes. Again the figures are of the full annotation of the clause complex, Textual Theme, PreHead and the Thematic Field span:

- (11) When QIA took its first stake in Barclays in February, Mr Varley said: “The participation of players such as these is one of the most positive manifestations of globalisation. It’s highly positive that there are pools of money that can be put to work in this way. That’s very different to the world of five or ten years ago.”

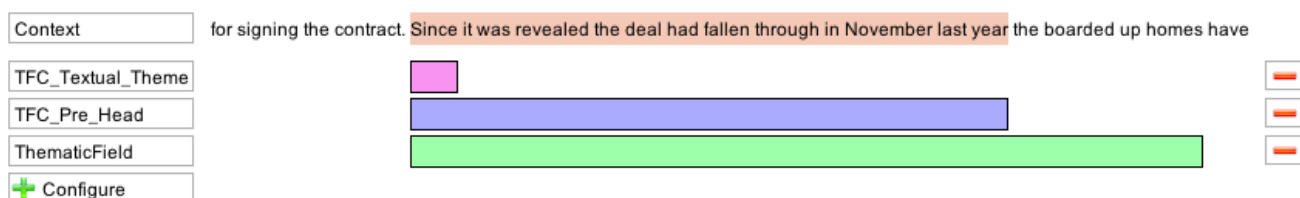


PreHead automatic annotation. CCL

(NR-Eng-07. Sentence Id 1362)

In both examples there is a Textual Theme, above *When* and below *Since*, as shown it is marked under the label Textual Theme and also included inside the span of PreHead. This has been done as such for practical reasons and does not affect the account for Textual Themes. In further analysis if PreHeads are to be studied in isolation, the use of Textual Themes can be used for the identification of the message.

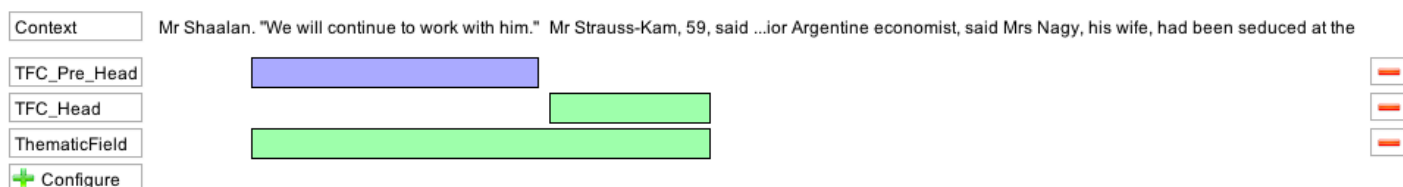
- (12) Since it was revealed the deal had fallen through in November last year the boarded up homes have remained empty.



PreHead automatic annotation. CCLc  
(NR-Eng-24. Sentence Id 956)

Finally, the last example (13) shows the use of quotations as PreHeads, commonly used in English news reports. The annotation is done as follows:

- (13) "We will continue to work with him." Mr Strauss-Kam, 59, said in a statement.



PreHead automatic annotation. Quotations  
(NR-Eng-01. Sentence Id 809)

The Quotation is marked as PreHead and *Mr Strauss-kan, 59*, as Head, and both elements together as Thematic Field

As it was pointed out in the analysis of the table presented above (table 26.6) the number of annotated elements in PreHead position present a statistically significant difference, being always bigger the number of elements of this kind in the automatic part than in the manual one. A closer look at the annotations revealed the following examples:

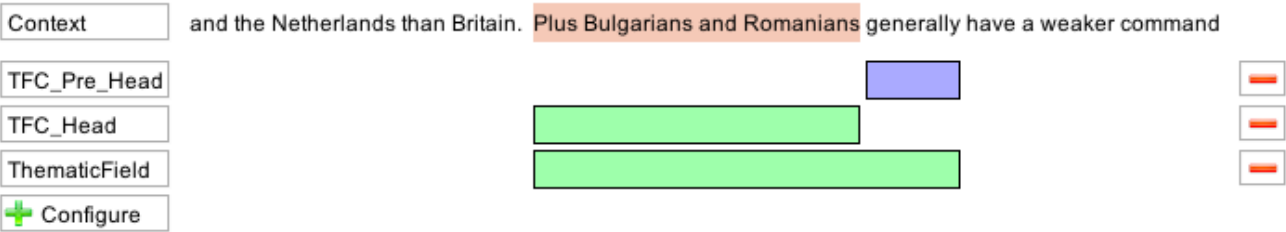
- (14) Barclays also issued a trading update saying group profit in the first nine months of this year was "slightly ahead" of the same level a year earlier.

26. Comparing the Results of Hand-coded versus Automatic Annotation



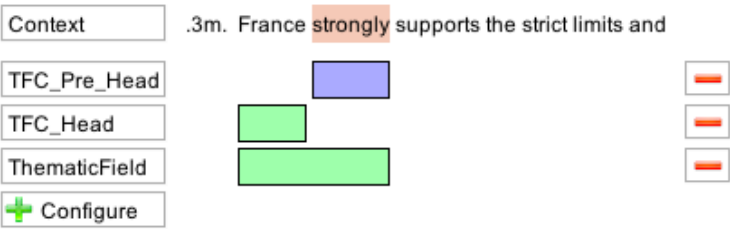
PreHead automatic annotation.  
(NR-Eng-07. Sentence Id 1354)

- (15) Plus Bulgarians and Romanians generally have a weaker command of English than Poles.



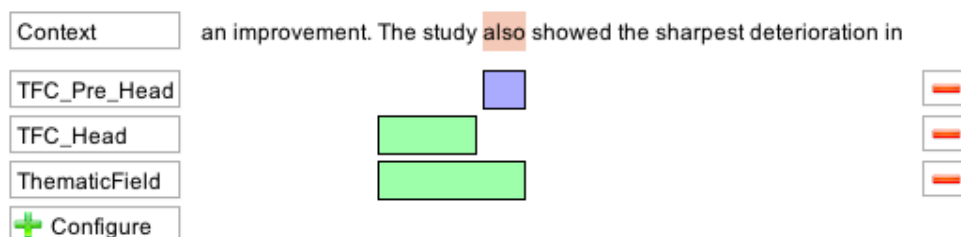
PreHead automatic annotation.  
(ED-Eng-04. Sentence Id 1384)

- (16) France strongly supports the strict limits and Germany has moved towards accepting them as a compromise.



PreHead automatic annotation.  
(NR-Eng-17. Sentence Id 350)

- (17) The study also showed the sharpest deterioration in families' cash availability since June last year, with 35pc of people reporting a decline.



PreHead automatic annotation.  
(NR-Eng-23. Sentence Id 985)

The elements *also* in example (14), *generally* in (15), *strongly* (16) and *also* in (17) are found after the Head and automatically annotated as PreHeads. According to the definition given in the guidelines these elements are correctly annotated even though they do not occupy the expected position before the Head. In the manual annotation these elements are not annotated as PreHeads, therefore the difference in the amount of elements found as PreHead come from this disparity in the annotation (as mentioned above). This means that the definition provided for PreHead should probably be refined and made more specific so that only the elements before the Head are annotated as PreHeads.

### 26.1.6 Thematic Head in English

The last part of the rules deals with the identification of the Thematic Head, which includes subtypes such as Thematic Equatives, Predicated Themes and There themes as special Heads. Table 26.6 shows the distribution of each type in the corpus. Although there is some disparity in the results there is no statistical difference between the manually and the automatically annotated occurrences.

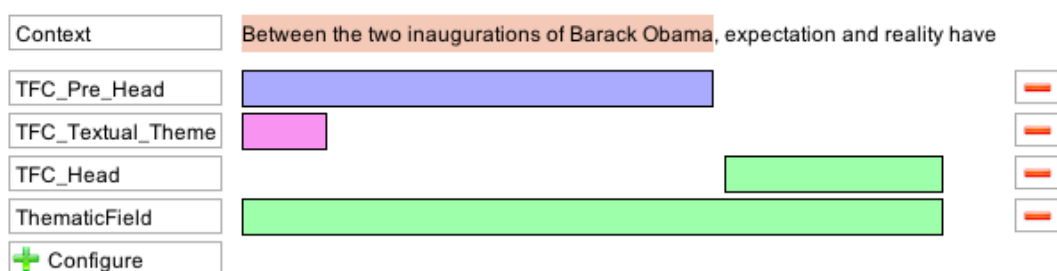
## 26. Comparing the Results of Hand-coded versus Automatic Annotation

**Table 26.6** Clause complex automatic annotation of Heads

Corpus	Thematic Element	Manual	Automatic
New Reports	Head	96.85%	78.22%
	Thematic Equatives	-	-
	Predicated Themes	-	0.32%
	There	0.30%	0.32%
Editorials	Head	95.23%	92.83%
	Thematic Equatives	0.88%	0.89%
	Predicated Themes	0.70%	1.60%
	There	5.19%	4.63%
Letters to the Editor	Head	97.54%	82.44%
	Thematic Equatives	0.55%	0.31%
	Predicated Themes	0.27%	2.15%
	There	2.84%	1.54%

The following two examples (figures (18) and (19)) show the annotation of two sentences with PreHead and Head. These element do not only show the annotation of Head but they also overcome the problem of multiple theme annotation, something which was not solved in previous attempts (c.f. Schwarz et al., 2008).







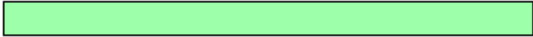


- (18) Between the two inaugurations of Barack Obama, expectation and reality have been turned on their head.



Head automatic annotation.  
(ED-Eng-03. Sentence Id 1356)

- (19) For all the difficulties of intervention in Mali, the alternatives are worse.







Context	States, the better. For all the difficulties of intervention in Mali, the alternatives are worse	
TFC_Pre_Head		
TFC_Textual_Theme		
TFC_Head		
ThematicField		
 Configure		

Head automatic annotation.  
(ED-Eng-06. Sentence Id 1397)








The following examples illustrate the annotation of special types of Head: Thematic Equative (example (20)), Existential clauses (there) (example (21)) and Predicated Theme (example (22)). As explained in chapter 11.2.2 these elements have a different notation form and had to be treated separately creating rules for their identification (see sections 24.2.5.3, 24.2.5.2 and 24.2.5.1):

- (20) What distinguishes mankind now from what we were then, for good or ill, is science and technology.

Context	,000 years ago. What distinguishes mankind now from what we were then, for good or ill	
TFC_Textual_Theme		
TFC_Head.type	Thematic Equative	
ThematicField		
 Configure		

Head automatic annotation. Thematic Equative  
(LE-Eng-30. Sentence Id 209)

- (21) On the other hand, there was hardly any sport for two months, so maybe we are not such wimps today after all.

Context	On the other hand, there was hardly any sport for	
TFC_Textual_Theme		
TFC_Head		
ThematicField		
 Configure		


Head automatic annotation. There  
(ED-Eng-16. Sentence Id 556)


- (22) It is only when Gaza prospers that those Palestinians wanting to hit Israeli targets will no longer be tolerated.


Context

land access with Egypt. It is only when Gaza prospers that those Palestinians wanting to hit

ThematicField





 Configure

Head automatic annotation. Predicated Theme  
(LE-Eng-08. Sentence Id 514)

## 26.2 Comparing the Results of Hand-coded versus Semi-automatic Annotation in Spanish

As with the English corpus, I will first present the results of comparing the manual vs the semi-automatic annotation of the Spanish corpus, focusing first on the splitting of sentences and secondly on the thematic features, annotated according to explained in chapter 25.

### 26.2.1 Sentence Splitting

Comparing the Results of Sentence Splitting in the Spanish Dataset

The first elements that can be compared in this distribution are the number of clause complexes identified in English and Spanish. Table 26.1 shows the number of clause complexes identified manually in the Spanish corpus in each genre and in the following column the number of clause complexes automatically found in GATE (Sentence Splitter) in each genre.

**Table 26.7** Clause complex bootstrapping Spanish

Corpora	Manual	Bootstrapping
Reports	359	356
Editorials	525	563
Letters to the Editor	365	413
Total	1249	1332

As shown there are apparent differences but they are not statistically-significant.

### 26.2.2 Comparing Thematic Elements in Spanish

The following table shows the results of the manual versus the semi-automatic annotation of the Spanish corpus. (M stands for the results obtained in the first manual annotation and Semi for the results obtained with the schema annotation in GATE). No statistically-significant differences were found ( $P > 0.05$ ).

**Table 26.8** Manual vs semi-automatic annotation in Spanish

Corpora	M TT	Semi TT	M IT	Semi IT	MPH	Semi PH	M H	Semi H
Reports	52	50	1	1	102	103	354	352
Editorials	91	102	38	39	158	169	499	531
Letters	55	52	64	62	98	103	336	338
Total	198	204	103	102	358	375	1189	1221

Both Textual and Interpersonal themes do not present a statistical significant difference. In the case of PreHeads it must be pointed out that the figures in the table do not account for the realization of Verbs, the reason for this is that the semi-automatic annotation of these elements would need a special label to be able to annotate two different categories in the same word. In order not to biased the results, the ones presented rely in the PreHeads without verbal realization both in manual and semi-automatic annotation.

In conclusion, semi-automatic analysis results do not differ from the ones obtained through pure manual annotation. What has been achieved is a reliable annotated corpus with thematic features that, when resources are available, could be further annotated adding a new layer of features (syntax and dependency) increasing its exploitation possibilities.



# 27

## Summary and Discussion

The present Part has focused on the steps undertaken towards automating the annotation of thematic features in the English and in the Spanish datasets. The final goal is to speed up the annotation process -which is laborious and time consuming when done manually- and to be able to annotate larger quantities of data in a more efficient way.

The computer platform chosen to carry out this process is GATE as presented in chapter 22 due to its easy of use, friendly interface, multitasking and multi language features (Orasan, 2003), along with the possibility to create your own processing resources to add to the pipeline of study of your corpus.

Once the corpus is prepared in the appropriate format (see section 23), it is uploaded to the platform and the bootstrapping is carried out. This process takes advantage of already made processing resources to automatically annotate the first layers of knowledge needed for the posterior automatization: in this case, the processing resources used were a tokenizer, a part-of-speech tagger and a parser (Moratón & Lavid, 2012).

The preprocessing or bootstrapping of the English and the Spanish dataset differ. The quality of the bootstrapping process is key to develop the automatization processes, therefore its annotations should be of the best possible quality. In English, once the bootstrapping has been carried out (see section 24.1), the translation process of the linguistic rules previously validated (Empirical Validation of the Thematization Model Part) starts, and a set of rules are created (see section 24.2) and included in an API that add a new layer of annotation to the preprocessed corpus. The resulting new annotated features are (chapter 26):

**Head:** and its subtypes: Thematic Equatives, Predicated Themes and Existential clauses

**PreHead:** with Circ and CCL types

**Interpersonal Themes:** EngSta and ComMod types

**Textual Themes:** Linker and Binders

**Thematic Field:** the full span of Thematic Field

The comparison of these automatic results with the manually annotated corpus showed that the number of automatically annotated items under each category does not present a statistically-significant difference with the manually annotated one. The only category that presented difficulties is Interpersonal Theme, something that may be expected as the intra-annotation agreements measures on this feature presented the lowest rate (see chapter 13.3.2.4).

In the case of Spanish, the available resources to be implemented in GATE at the time of the study did not include a part-of-speech tagger and a parser which can be easily loaded into the GATE platform. In order to overcome this problem, semi-automatic annotation of the corpus with thematic features was carried out and the results were compared to the manual annotation carried out in chapter 18 (manual annotation). The lack of statistically-significant differences between the number of features annotated according to both methods indicates that the semi-automatic method can be used with confidence to annotate thematic features in the Spanish corpus, although it would be highly desirable to count on reliable, open-code and easy-to-use processing software for Spanish which can be easily integrated in the GATE platform, as it is the case of the Stanford Parser for English.

The automatic annotation of both corpora allows for further and deeper study of the annotated data. The computer supported format admits the use of the platform ANNIC (ANNotations-In-Context) which is a “full-featured annotation indexing and retrieval system” built on top of Apache Lucene – a high performance full-featured search engine implemented in JAVE, which supports indexing, searching and querying of large document collections (Cunningham et al., 2011). This makes possible the exploitation of the relation of the new thematic layers of annotation in relation to any other annotated feature or to annotate.

For example: `ThematicField.cat=="ThematicHead",Token.category=="VLfin"` retrieves the number of Thematic Heads realized by verbs in finite form in the selected database. This information is accompanied by statistical accounts for the occurrences of this match in context, in the datastore or in the database.

Figure 27.1 shows an example in which the query `TFC-Head, SyntaxTreeNode.cat=="NP"` obtain the number of elements annotated as Head that are realized by NP:

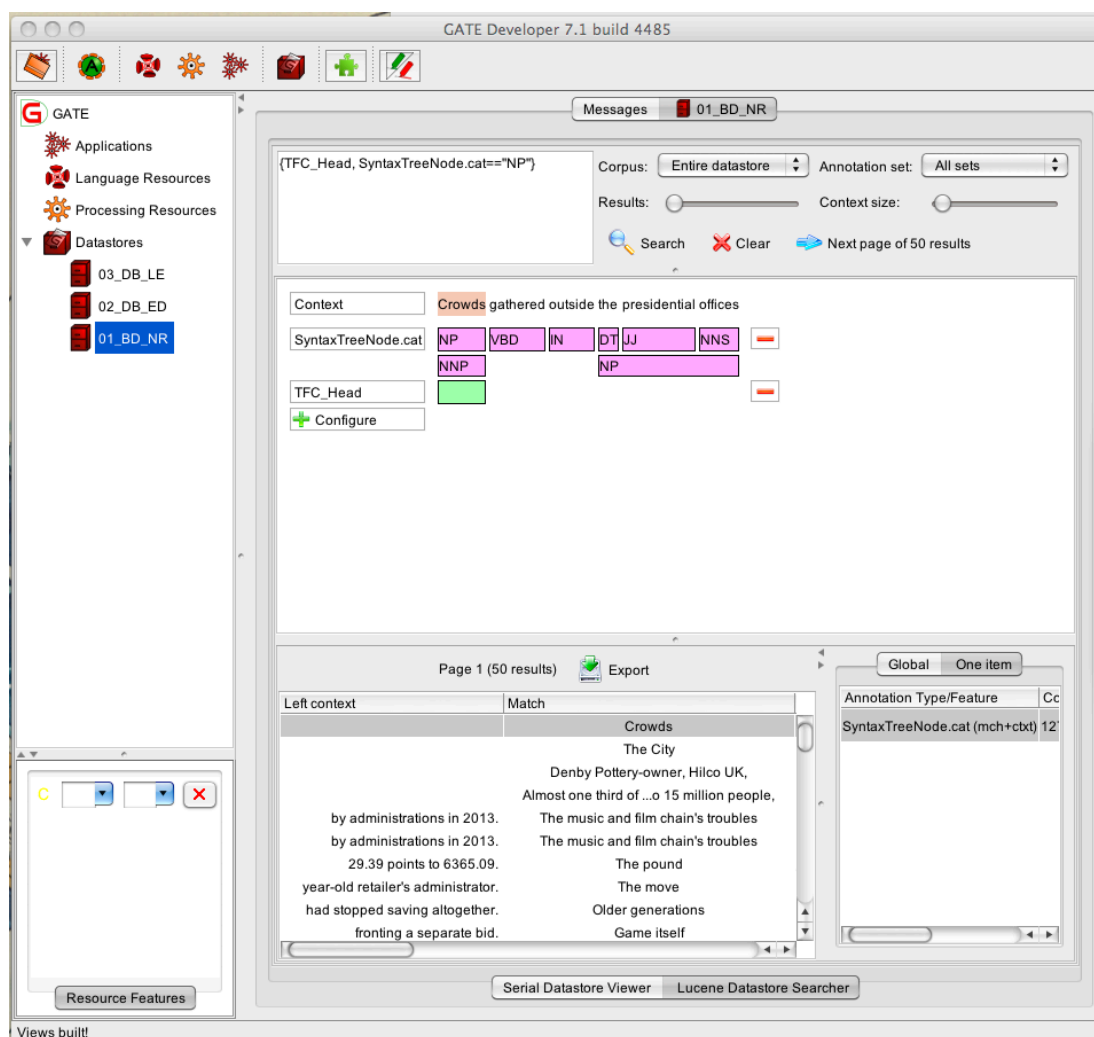


Figure 27.1 ANNIC English corpus screenshot

This tool expands the exploitation potential of the corpus to a wide range of possibilities, ensuring its reusability to which more layers of annotations (manual or automatic) can be added.





## **F. FINAL REMARKS**



# 28

## Conclusions

As explained in the introductory chapter, the work reported in this dissertation has investigated three main research questions related to more general aims framed in the fields of Corpus Linguistics, Contrastive Functional Linguistics and Natural Language Processing.

The first research question was framed in the area of Corpus Linguistics and addressed the issue of whether the theoretical phenomenon of thematization can be empirically tested through corpus annotation. In order to answer this question a number of agreement studies were designed to ascertain the reliability of certain thematic features extracted from Lavid et al.'s model of thematization (2010d:chapter 5). There were two types of agreement studies: the first one focused on the identification of the Thematic Field and thematic elements, and the second one focused on the labeling of thematic elements (chapter 13). These tasks were accompanied with detailed guidelines including explanations of each assignment and theoretical descriptions to guide the coder for the best quality annotations (chapter 12). The measurement of the results was carried out according to the characteristics of the annotations: for agreement study 1, where the coders had to select spans of text, AGR metrics was used; for agreement 2, focused on labeling markables, Kappa was used, obtaining thus interpretable data from the different type of annotations (chapter 13.3). The results showed that the level of agreement in the identification of the thematic span was significant in both languages and the annotations could be considered reliable, although agreement levels were lower in Spanish. The fact that Spanish is a pro-drop language affects the decisions of the coders in the selection of the thematic span. The use of verbs as starting point in the sentence compel the coders to select the lexical and

grammatical part of the verb as Thematic Field, but in many occasions they included only one of them. Although this is addressed by the Theme model and in the guidelines, it represents a problem for the coders. In the second task, the coders had to select the span belonging to the following thematic elements: PreHead (PH), Thematic Head (TH), Textual Theme (TT), Interpersonal Theme (IT), Predicated Theme (PT) and There/Hay-Theme. The results obtained a high level of agreement in the coding task in both languages, except in the English Interpersonal Theme, showing that the annotations were reliable. Most disagreements were due to the selection of Textual Theme instead of Interpersonal Theme, and vice versa. Interestingly enough, and contrary to expected, the choice of Head and PreHead in Spanish, did not pose any problem for the coders in this second task.

The second agreement study consisted of the labeling of marked items from a given drop-down list; the first task included thematic elements: PreHead (PH), Thematic Head (TH), Textual Theme (TT), Interpersonal Theme (IT) or None (None). The rest of the tasks addressed more fine grained categories, i.e, subtypes of thematic elements. The first task was accomplished with high agreement in both languages. The second task dealing with the selection of experiential elements, attained poorer results in Spanish, although still substantial for the validation. The relational type presented problems for the coders in Spanish but not in English. As such, a redefinition of this category in Spanish guidelines may be needed. The third task included the annotation of Textual Theme types: binder, linker and correlative. In English, the rate of agreement was slightly lower than in Spanish but it was not relevant. Task 4 was related to the types of Interpersonal Theme and both languages revealed a very low level of agreement in this assignment. The reason for this measurement is the significant subjectivity of this category. In fact, and according to the results obtained, this category can not be considered as validated. In task 5 related with the subtypes of PreHead, a high level of agreement was obtained between annotators.

All in all, it can be stated that, all thematic categories of Lavid et al.'s model of thematization can be considered validated except for the types of Interpersonal Theme, both in English and Spanish. That is, the categories of Thematic Field span, PreHead, Thematic Head, Textual Theme, Interpersonal Theme, Predicated Theme and There/Hay, on the one hand, and the subtypes of Thematic Head (experiential roles), subtypes of Textual Theme and subtypes of PreHead obtained a significant level of agreement among annotators. This means that the corresponding annotations schemes and guidelines which were based on Lavid et al.'s model of Theme are reliable and stable.

The second research question was framed in the area of Contrastive Functional Linguistics and focused on the investigation of language-specific thematic selection and progression patterns in the three newspaper genres studied: news reports, editorials and letters to the editor. It also focused on possible correlations between thematic patterns and generic moves in these three newspaper genres. According to Leech "corpora are useful only if we can extract knowledge or

---

information from them” (1997:4). Therefore, once the thematic categories were validated, they were used for the annotation of a larger corpus. Due to the impossibility of annotating a larger corpus by more than one annotator, I decided to follow Dligach et al. (2010) who claim that the most efficient and reliable corpora is obtained through more single manual or handed-coded annotations of an expanded corpus. The larger corpus consisted of a large collection of news reports, editorials and letters to the editor, in English and Spanish, summing up to 220 texts, and a total of 2551 clause complexes. The manual annotation of the thematic features of this corpus was presented in chapter 18, including the moves and Thematic Progression patterns in the three genres in chapter 19.

The results of this manual annotation yielded interesting thematic features of each of the three newspaper genres studied in this thesis, which can be explained by the different communicative goals that each genre fulfills. Thus, news reports are defined as communicatively-based rather than event-based (White, 1997) and strive to remain objective and use neutral language while presenting a diversity of opinions, voices, and perspectives of the event, incident, or issue under discussion. Similarly, Thompson (1994) explains that news reports should “give a reasonably objective “true” picture of the event and of the evidence, reactions, and opinions of the people involved.” These definitions are resembled in the thematic choices of news reports: they are characterized by the use of quotations in PreHead position. English up fronts quotations to PreHead whereas Spanish delays its occurrence to the Rheme, although these quotations can be traced by a verb of communicative meaning. Also, a common feature shared by both languages is the use of PreHead to “highlight about the overturning of temporal sequences” (Bell, 1991), as a temporal setting, through more complex devices in English (PH-CCL), and simpler expressions in Spanish (PH-Circ). Another interesting feature of PreHead in Spanish news reports is that it is revealed as the preferred slot for the lexical part of the verb when it appears in first position. In addition, the Thematic Head is realized by Noun Phrases and Noun Phrases followed by Appositions, both elements more frequently used in English. In its place, Spanish uses verbs, particularly its grammatical information. These verbs, both in English and Spanish, act as *actors* and *sayers*, as Francis (1990) pointed out. The composition of the Thematic Field in news reports is a clear reflection of the nature of these texts. The use of elements of the Inner thematic field and the almost non-existent occurrences of elements of the Outer thematic field is indicative of the genre of news reports.

Regarding *Generic Structure Potential* (GSP) and Thematic Progression patterns, as explained in chapter 19, news reports are the only genre of the three analyzed that has three moves instead of four. Its generic organization is simple. The results revealed that after the explicative headline, the next move, in which the issue is presented, is characterized by the use of constant theme, where the same Thematic Head is repeated and some extra information is added to it. The development of the piece of news, in the last move, where the facts about the issue are presented, is deployed by a Hypertheme followed by Constant Thematic progression in English, and linear patterns in Spanish.

Editorials are conceptualized as articles giving “the opinion of the editor or publisher on a topic or item of news” (Sinclair (1995) in Ansary & Babaii, 2005). The distributional map of thematic elements changes drastically in this genre in comparison to news reports. To begin with, and due to the longer and explanatory nature of the texts, the use of Textual Themes increases in both languages. It is worth noticing the contrast in the use of binders in English and in Spanish. These elements are the preferred textual strategy for English in long and explanatory texts, as it is the case; whereas Linkers are seldom used nor on editorials neither on other type of texts. Spanish, on the other hand, follows this general pattern but accounts for a wider use of Linkers in the three genres (see chapter 18.5). The occurrence of Interpersonal Themes reveals the strong presence of the author in the text, in most cases in the form of Engagement where the writer relates to the readers in respect to his or her position in the text (Hyland, 2001), “moulding opinion and encouraging critical thinking” (Roy, 2011). In PreHead position, both languages prefer the use of simple forms of the type PH-Circ for expression of Time and more complex (and less frequent) to express contingency, introducing counterarguments or hypothesis. A characteristic feature of Spanish is the use of the lexical part of the verb in this position. Thematic Heads reflect the presence of the writer in the text through the use of Thematic Equatives and Predicated Themes. Head is realized by Noun Phrases, particularly pronouns in English, as its most frequent strategy. Spanish uses Verbs in first position instead, avoiding the explicit marking of the subject. English also uses common nouns, not preceded by determiners whereas Spanish does not present any occurrence of this kind, as Pérez-Leroux et al. (2004) notes: “(...) English allows in subject position bare plural Noun Phrases while Spanish and some other Romance languages do not.” The semantic realizations which typically realize Thematic Heads in this genre are *actors* and *relational processes* that serve to present facts to the reader. Also, the use of both Inner and Outer thematic fields reflex the complexity of theme in these type of texts.

As to the relationship between *Generic Structure Potential* (GSP) and Thematic Progression patterns, editorials present a four move structure, as explained in chapter 19. After the metaphorical headline, the next move is introduced by a derived theme that addresses the issue. The argumentation of facts and its evaluation is organized by *units of information*: an introductory Hypertheme and a set of derived themes attached to it. The same arrangement is followed in the last move, where the articulation of a position on the part of the writer takes place.

In letters to the editor “the writers take position by expressing their views on current events and states of affairs and by appealing for action. In so doing they also reveal their feelings towards such events and provide their evaluation of the people and circumstances involved” (Pounds, 2010). The short extension of these texts limits the choice of certain elements in the Thematic Field.

The selection of Textual Themes reflects the nature of the texts in the use of binders both in English and Spanish. Linkers, as mentioned above, are preferred by Spanish rather than

---

English.

As to Interpersonal Themes, English selects the use of Engagement to highlight the presence of the author in the texts, but does so less frequently than Spanish. Spanish evenly prefers Stance and Engagement. Does this mean that English authors do not appear in the texts? Rather, the explanation points to the reason given by Pounds (2010), who says that English “stance” may be more often “integrated into text rather than overtly marked.”

In PreHead position Spanish selects simple syntactic constructions and English more complex formations to express time and contingency.

As in editorials, Thematic Heads are, in some cases, realized by Thematic Equatives and Predicated Themes. The preferred syntactic realization of Head is that of Noun Phrases in both languages, although English presents a statistical significant difference, Spanish balances this selection with the use of verbs in first position. In this type of texts, contrary to other genres, clauses are not only declaratives, but it is also possible to find verbal forms in Head position in English directives and questions. The choice of experiential elements is more varied, appearing in relevant numbers: *actors*, *carriers*, *senser*, *sayers* and *goal*, and the special mention of *process* in Spanish as a feature of this language.

Regarding the *Generic Structure Potential* (GSP) and Thematic Progression patterns, as explained in chapter 19, letters to the editor are characterized by a four move structure. The headline summarizes the issue at stake. Although the text extension is very short, it allows for an introduction of the issue, outlining the argumentation and closing. All of them are derived themes, and in this particular case the short extension of the text does not allow to develop more complex or elaborated patterns, being in most cases a single clause complex the only component of the move.

The different composition of the thematic field varies according to the genre under study: News reports present a simple structure achieved by the presence of the Inner thematic field, consisting of a PreHead mainly used for temporal setting or quotations, simpler in English and of a higher syntactic complexity in Spanish. Head acquires the form of Noun Phrase in some cases followed by Appositions in English whereas Spanish uses verbs. These are repeated through the text, as shown by the Constatain Thematic Progression patterns in English and the linear one in Spanish.

Editorials increase the complexity of the Thematic Field making use of both outer and the inner elements. Textual and Interpersonal Themes in the Outer Thematic Field account, in the first case, for the use of hypotactic relations in the text; and, in the second, for the presence of the author in the text. Inner thematic theme reports the use of PreHead as a time settler. Head appears realized by pronouns, a substitution strategy for an already mentioned entity. These strategies are reflected in the selection of Thematic Progression patterns, using *units of information* introduced by a Hypertheme and developed by derived themes.

Letters to the editor make use of elements of both the Inner and Outer Thematic Fields. The Interpersonal Theme presents a strong presence of the author in the text although less relevant than the one showed in editorials. In the Inner thematic field, PreHead presents time

settings and contingency circumstances. Head is occupied by verbs also in English as in this type of texts there are declaratives and other types of clauses. Thematic structure patterns find in the presentation of the issue at the beginning of the text, the only Hypertheme to which adds information through derived themes.

The third research question was framed in the area of Natural Language Processing and focused on the possibility of automating the manual annotation of thematic features in the bilingual English-Spanish corpus, and on comparing the results obtained using each method: human-coded or automatic annotation. The corpus selected was the same as the one manually annotated and discussed in the previous part. The guidelines that were validated in the first part of the study, and tested in the second part in manual annotation, took now a salient role, not only helping in the writing of automatic rules, but also in the documentation of the study for the benefit of future users of the corpus (Dipper et al., 2004).

The tool to support the process had to fulfill a number of minimum requirements: easy of use, friendly interface, multitasking and multi language (Orasan, 2003). These features were key for the study to be usable, reproducible, flexible and interactive. With all these points in mind, the selection of the tool to annotate was GATE -General Architecture for Text Engineering- (Cunningham et al., 2011).

Corpus data was prepared to be automatically treated (format, metadata (see chapter 23)) and uploaded to GATE, and submitted to a bootstrapping process, involving tokenization, segmentation, part-of-speech tagging and parsing, as well as the identification of the units of analysis, since these steps are all necessary for the definition of rules to automatically annotate the corpus with thematic features. It is at this moment where the processes followed for each language diverged, due to availability of the different language processing tools for each language: while English had accessible tools to perform these tasks accurately (sentence splitter, tagger, part-of-speech tagger and parser), similar tools were not available for Spanish, at the time of writing this dissertation. For this reason, the automatization process differed for these two languages: while in English a set of JAPE rules were created to automatically annotate the Thematic Field span, Textual Theme, Interpersonal Theme, PreHead and Thematic Head, as well as their subtypes, in Spanish the process was done semi-automatically. For English, the “translation” of the set of guidelines to JAPE proved to be a hard work, involving the transfer of much linguistic knowledge to the computer, that is a blank canvas. However, it was worth it, since the API that run on the English pipeline proved to be able to automatically annotate the same number of thematic features as the ones done manually, thus saving work in the long run.

In the case of Spanish, the available resources to be implemented in GATE, at the time of the study, did not allow to find a part-of-speech tagger and a parser that provided codings accurate codings and could be easily integrated within the GATE environment. The solution given to this situation was to perform a thematic semi-automatic annotation (with the aid of



---

a schema in GATE platform) of the corpus where the results were compared to the manual annotation carried out in chapter 18 (manual annotation) obtaining results that ensure that the corpus was reliably annotated in the following categories: Thematic Head, PreHead, Interpersonal Theme, Textual Theme and Thematic Field span. Both in the English and in the Spanish case, the automatization procedure proved to be less time- consuming and equally reliable as manual annotation.

To conclude, the current work has shed light into a number of interesting issues related to the phenomenon of Theme in a contrastive perspective which might be useful in the fields of Corpus Linguistics, Contrastive Functional Linguistics and Natural Language Processing.

From the Corpus Linguistics perspective, the use of corpus annotation has proved to be a useful mechanism for the empirical validation of aspects of the thematization model proposed by Lavid et al. (2010d) through the design of reliable annotation schemas which can be tested through agreement studies. As those agreement studies showed, the annotation schemas designed so far, including a core and an extended tag-set, proved to be reliable and stable for the thematic types which can occur as part of the Thematic Field in English and Spanish declarative clauses, although the definition for the the Interpersonal Theme tag needs to be refined or reformulated in order to be considered as reliable for annotation purposes.

From the contrastive perspective, the current dissertation has provided a thorough characterization of the thematic features of the three newspaper genres studied (news reports, editorials and letters to the editor), both in terms of their thematic selection choices, but also in terms of the Thematic Progression patterns which correlate with their generic structures. The thematic characterization has been achieved through the manual annotation of a bilingual comparable corpus of two-hundred and twenty texts with the core and the extended Theme tag-sets previously validated in the agreement studies. From the NLP perspective, the current dissertation has achieved two main goals: (i) the automation of the annotation process for thematic features in the English dataset, and (ii) the creation of a bilingual (English-Spanish) corpus of newspaper texts from different genres, fully annotated with thematic features, and which can be used for a number of linguistic and NLP applications. The first goal proved to be a difficult one, requiring the creation of JAVA rules which used linguistic knowledge extracted from the previous preprocessing steps. This was possible and turned out to be cost-effective for the English dataset, since preprocessing components were available and integrated in the GATE platform for English. For Spanish it was not possible to fully automate the annotation process due to the lack of availability of these preprocessing resources. However, I am confident that when these resources are available, it will also be possible to develop similar JAVA rules for Spanish which will allow the automatic annotation of Spanish data.

In the meantime, it is expected that the current work will be useful for the NLP community, due to the need for corpora in different languages with discourse-level annotations as the ones provided in this dissertation.

## 28. *Conclusions*

---

Future work will be focused on extending the current work to other genres and registers in both English and Spanish, whose thematic features may also contribute to their generic characterization, as has been shown for the three newspaper genres studied in this dissertation.

## REFERENCES

- Adam, J.-M. (1992). *Les textes: types et prototypes-Récit, description, argumentation et dialogue*. Paris, Nathan. 30
- Agno, A., Bufí, N., Civit, M., Martí, M., & Navarro, B. (2003). Análisis cuantitativo y cualitativo del acuerdo entre anotadores en el desarrollo de corpus interpretados lingüísticamente. *Revista SEPLN*, (31), 201.
- Albornoz, L. A. (2005). *Los diarios online de información general. El caso de los grandes periódicos en español*. Ph.D. thesis, Servicio de Publicaciones, Vicerrectorado de Cultura, Deporte y Política Social, Universidad Complutense de Madrid.
- Alishahi, A. (2010). Computational Modeling of Human Language Acquisition. *Synthesis Lectures on Human Language Technologies*, 3(1), 1–107.
- Alonso Belmonte, I. (2007). Newspaper Editorials and Comment Articles: A “Cinderella” Genre? *Revista Electrónica de Lingüística Aplicada (RAEL), Different Approaches to Newspaper Opinion Discourse*, 1–9. 169, 180, 198
- Andersen, G. (2012). *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian*, vol. 49. John Benjamins Publishing. 35
- Ansary, H., & Babaii, E. (2005). The Generic Integrity of Newspaper Editorials: A Systemic Functional Perspective. *RELC Journal*, 36, 271–295. xviii, 37, 256, 257, 261, 265, 268, 338, 381, 390
- Artstein, R., & Poesio, M. (2008). Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, 34(4), 555–596. 53, 55, 56, 57, 65
- Arús, J., Lavid, J., & Moratón, L. (2012). Annotating Thematic Features in English and Spanish: A Contrastive Corpus-Based Study. *Linguistic and the Human Science*, 6(1-3), 173–192. 3, 22, 66, 69, 134, 378, 386
- Arús, J., Lavid, J., & Moratón, L. (2013). Annotating Thematic Features in English and Spanish: A Contrastive Corpus-Based Study. In Maite Taboada and Susana Doval Suárez and Elsa González Álvarez (Ed.) *Contrastive Discourse Analysis: Functional and Corpus*

## REFERENCES

---

- Perspectives*, chap. 8, (pp. 164–182). Sheffield, Bristol. UK: Equinox. 3, 22, 69, 94, 134, 378, 386
- Ballesteros Calvo, A., Cárcamo Escorza, G., & Duobert Collazos, E. (2013). *Recubrimiento y normalización de recursos lingüísticos: aplicación a las anotaciones morfosintácticas [sic] y sintácticas de Freeling*. Master's thesis, Universidad Complutense de Madrid, Spain.
- Basile, V., & Bos, J. (2011). Towards Generating Text from Discourse Representation Structures. In *Proceedings of the 13th European Workshop on Natural Language Generation*, (pp. 145–150). Association for Computational Linguistics. 378, 387
- Bateman, J. A. (1990). Upper Modeling: Organizing Knowledge for Natural Language Processing. Tech. rep., DTIC Document.
- Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An Introduction to History, Theory, Research, and Pedagogy*. West Lafayette, IN: Parlor Press. 32, 33
- Bayerl, P. S., & Paul, K. I. (2011). What Determines Inter-Coder Agreement in Manual Annotations? A Meta-Analytic Investigation. *Computational Linguistics*, 37(4), 699–725. 57
- Bazerman, C., Bonini, A., & Figueiredo, D. d. C. (2009). *Genre in a Changing World*. Fort Collins, Colorado.; West Lafayette, IND: WAC Clearinghouse; Parlor Press.
- Bazerman, C., et al. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.
- Bednarek, M. (2006). *Evaluation in Media Discourse: Analysis of a Newspaper Corpus*. London; New York, NY: Continuum.
- Beigman Klebanov, B., & Beigman, E. (2009). From Annotator Agreement to Noise Models. *Computational Linguistics*, 35(4), 495–503. 55
- Bell, A. (1991). *The Language of News Media*. Blackwell Oxford. 36, 144, 164, 262, 274, 337, 380, 390
- Bello, A. (1860). *Gramática de la lengua castellana destinada al uso de los americanos*. Alicante: Biblioteca Virtual Miguel de Cervantes. 15
- Bhatia, V. K. (1993). *Analysing Genre: Language Use in Professional Settings*. London: Longman. 32
- Bhatia, V. K., Flowerdew, J., & Jones, R. H. (2008). *Advances in Discourse Studies*. Routledge. 4
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press. 30, 241, 276

- Biber, D. (1995). *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge University Press. 30
- Biber, D., & Finegan, E. (1989). Styles of Stance in English: Lexical and Grammatical Marking of Evidentiality and Affect. *Text-interdisciplinary Journal for the Study of Discourse*, 9(1), 93–124. 228
- Biemann, C., & Mehler, A. (Eds.) (2014). *Text Mining: From Ontology Learning to Automated Text Processing Applications*, vol. Festschrift in Honor of Gerhard Heyer of *Theory and Applications of Natural Language Processing Edited volumes*. Springer.
- Bird, S., & Simons, G. (2001). The Olac Metadata Set and Controlled Vocabularies. In *Proceedings of the ACL 2001 Workshop on Sharing Tools and Resources-Volume 15*, (pp. 7–18). Association for Computational Linguistics. 289
- Bloor, K. N. (1991). Thematic Progression in Professional and Popular Medical Texts. *Functional and Systemic Linguistics: Approaches and Uses*, 55, 369. 5
- Bock, K., & Levelt, W. (2002). Language Production. *Psycholinguistics: Critical Concepts in Psychology*, 5, 405.
- Bontcheva, K., Cunningham, H., Roberts, I., & Tablan, V. (2010). Web-Based Collaborative Corpus Annotation: Requirements and a Framework Implementation. *New Challenges for NLP Frameworks*, (pp. 20–27). 48
- Bosque, I., & Demonte, V. (2000). *Gramatica descriptiva de la lengua española*. Colección Nebrija y Bello. Real Academia Española. Madrid: Editorial Espasa Calpe, 3 ed. 71, 82, 83
- Boye, K. (2014). Propositions, States of Affairs and Verbal Mood. A Danish Case Study. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 103
- Brill, E. (1995). Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part-of-Speech Tagging. *Computational Linguistics*, 21(4), 543–565.
- Bronia, S. P. C. (2005). *Analysing Newspaper Genres with a View to Developing a Genre-Based Pedagogy for the Teaching of Writing in Hong Kong Schools*. Ph.D. thesis, City University of Hong Kong. 37, 212, 255, 256, 258, 268
- Buráňová, E., Hajičová, E., & Sgall, P. (2000). Tagging of very Large Corpora: Topic-Focus Articulation. In *Proceedings of the 18th conference on Computational Linguistics*, vol. 1, (pp. 139–144). Association for Computational Linguistics. 4
- Butler, C. S. (1985). *Statistics in Linguistics*. Basil Blackwell. 140, 313

## REFERENCES

---

- Butler, C. S. (2003). *Structure and Function a Guide to Three Major Structural-Functional Theories: Part 2: From Clause to Discourse and beyond*, vol. 64. John Benjamins Publishing.
- Caffarel, A., Martin, J. R., & Matthiessen, C. M. (2004). *Language Typology: A Functional Perspective*, vol. 253. John Benjamins Publishing. 71
- Caporaletti, S. (1997). The Thematization of Time in E. M. Forster's "The Eternal Moment" and Joyce's "The Dead". *Twentieth Century Literature*, 43(4), 406–419. J1: Twentieth Century Literature; M3: Article.
- Carletta, J. (1996). Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics*, 22(2), 249–254. 53, 55, 56
- Carletta, J., McKelvie, D., & Isard, A. (2002). Supporting Linguistic Annotation Using XML and Stylesheets. *Readings in Corpus Linguistics*. 44, 50
- Carreras, X., & Padró, L. (2002). A Flexible Distributed Architecture for Natural Language Analyzers. In *Proceedings of the LREC*, vol. 2.
- Cerban, M. (2008). The Textual Metafunction of the Message: Identifying Themes in Special Thematic Structures. *Analele Universității din Craiova, Științe filologice, Lingvistică, anul XXX(1-2)*, 157–164. 199
- Chafe, W. L. (1994). *Discourse, Consciousness and Time: Flow and Displacement of Conscious Experience in Speaking and Writing*. University of Chicago Press. 135
- Chafe, W. L., & Li, C. N. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View in Subject and Topic. *Subject and Topic*, (pp. 26–56). 19
- Charles, M., Hunston, S., & Pecorari, D. (2011). *Academic Writing: At the Interface of Corpus and Discourse*. Bloomsbury Publishing. 239, 276
- Clark, P. P. (1975). Newspapers and Novels: Some Common Functions and Themes. *Studies in the Novel*, 7(2), 166–180. 246, 277
- Cohen, J., et al. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46. 52
- Cooley, L., Trigueros, M., & Baker, B. (2007). Schema Thematization: A Framework and an Example. *Journal for Research in Mathematics*, 38(4), 370–392.
- Córdova Jiménez, A. (2012). Las cartas al director como género periodístico. *Revista de Estudios de Comunicación*, 16(30).
- Cucatto, A. (2003). Los 'Temas' lingüísticos: la construcción figurativa del agente y su relación con la tematización. *Sintagma: Revista de Lingüística*, 15, 43–60.

- Cunningham, H. (2000). *Software Architecture for Language Engineering*. Ph.D. thesis, University of Sheffield.
- Cunningham, H., Hanbury, A., & Rüger, S. (2010). Scaling up High-Value Retrieval to Medium-Volume Data. In H. Cunningham, A. Hanbury, & S. Rüger (Eds.) *Advances in Multidisciplinary Retrieval*, Lecture Notes in Computer Science, (pp. 1–5). Vienna, Austria: Springer. 287
- Cunningham, H., Maynard, D., & Bontcheva, K. (2011). *Text Processing with GATE (Version 6)*. Stanford, CA. 6, 50, 285, 286, 293, 294, 295, 300, 307, 308, 309, 330, 340, 383, 384, 392, 393
- Cunningham, H., Maynard, D., Bontcheva, K., & Tablan, V. (2002). GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. In *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02)*.
- Dafouz-Milne, E. (2008). The Pragmatic Role of Textual and Interpersonal Metadiscourse Markers in the Construction and Attainment of Persuasion: A Cross-Linguistic Study of Newspaper Discourse. *Journal of Pragmatics*, 40(1), 95–113.
- Danes, F. (1974). Functional Sentence Perspective and the Organization of the Text. *Papers on Functional Sentence Perspective*, 106, 128. 17, 261
- de Beugrande, R., & Dessler, W. (1981). *Introduction to Text Linguistics*. Harlow: Longman. 30
- Dipper, S., Götze, M., & Skopeteas, S. (2004). Towards User-Adaptive Annotation Guidelines. In *Proceedings of the COLING 2004 5th International Workshop on Linguistically Interpreted Corpora*, (pp. 23–30). 48, 340, 383, 392
- Dligach, D., Nielsen, R. D., & Palmer, M. (2010). To Annotate More Accurately or to Annotate More. In *Proceedings of the Fourth Linguistic Annotation Workshop*, (pp. 64–72). 55, 66, 132, 337, 380, 389
- Downing, A. (1991). An Alternative Approach to Theme: A Systemic Functional Perspective. *Word*, 42(2), 119–143. 3, 19, 21, 76, 143, 377, 386
- Downing, A., & Locke, P. (2006). *English Grammar: A University Course*. Routledge, 2 ed. 76, 78, 82, 83, 152, 176, 211, 258, 260, 261
- Drummond, C. (2009). Replicability is Not Reproducibility: Nor is it Good Science. In *Proceedings of the Twenty-Sixth International Conference on Machine Learning: Workshop on Evaluation Methods for Machine Learning IV*. Montreal, Quebec, Canada.
- Drummond, C. (2012). Reproducible Research: A Dissenting Opinion. Retrieved from <http://cogprints.org/8675/1/ReproducibleResearch.pdf> Consulted January 2015.

## REFERENCES

---

- Duszak, A. (1994). *On Thematic Configurations in Texts: Orientation and Goals*. Amsterdam, Philadelphia: John Benjamins. 378, 387
- Eckart de Castilho, R. (2014). *Natural Language Processing: Integration of Automatic and Manual Analysis*. Ph.D. thesis, Technical University of Darmstadt. xiii, 45, 48, 50, 59, 383, 392
- Eckart de Castilho, R., Holtz, M., & Teich, E. (2009). Computational Support for Corpus Analysis Work Flows: The Case of Integrating Automatic and Manual Annotations. In *Book of Abstracts GSCL Workshop: Linguistic Processing Pipelines*, (p. 19). 49
- Eggs, S. (1994). *Introduction to Systemic Functional Linguistics*. Printer London, 1st ed. 32
- Eggs, S. (2004). *Introduction to Systemic Functional Linguistics*. Bloomsbury Publishing, 2nd ed. 258, 261
- Eiler, M. (1986). Thematic Distribution as a Heuristic for Written Discourse Function. *Functional Approaches to Writing: Research Perspectives*, (pp. 49–68). 5, 40
- Enkvist, N. E. (1985). A Parametric View of Word Order. *Text Connexity, Text Coherence*, (pp. 320–336).
- Faigley, L., & Meyer, P. (1983). Rhetorical Theory and Readers' Classifications of Text Types. *Text*, 3(4), 305–325. 30
- Fawcett, R. P. (2007). The Many Types of 'Theme' in English: Their Semantic Systems and their Functional Syntax. *Research Papers in the Humanities*, (pp. 1–105). 20
- Fernández Lorences, T. (2010). *Gramática de la tematización en español*. Universidad de Oviedo. 21
- Ferrucci, D., & Lally, A. (2004). UIMA: An Architectural Approach to Unstructured Information Processing in the Corporate Research Environment. *Natural Language Engineering*, 10(3-4), 327–348. 51
- Firbas, J. (1964). On Defining the Theme in Functional Sentence Analysis. *Travaux linguistiques de Prague*, 1, 267–280.
- Firbas, J. (1972). On the Interplay of Prosodic and Non-Prosodic Means of Functional Sentence Perspective. *The Prague School of Linguistics and Language Teaching*, (pp. 77–94). 16
- Firbas, J. (1987). On the Delimitation of the Theme in Functional Sentence Perspective. *Functionalism in Linguistics*, 20, 137–156. 16, 17
- Firth, J. R. (1957). *Papers in Linguistics, 1934-1951*. Oxford University Press.



- Francis, G. (1989). Thematic Selection and Distribution in Written Discourse. *Word*, 40(1-2), 201–221. 5, 40
- Francis, G. (1990). Theme in the Daily Press. *Occasional Papers in Systemic Linguistics*, 4. 5, 169, 240, 250, 274, 276, 337, 381, 390
- Fries, P. H. (1983). On the Status of Theme in English: Arguments from Discourse. In J. Petöfi, & E. Sözer (Eds.) *Micro and Macro Connexity of Texts*, vol. 45, (pp. 116–152). Hamburg: Buske Helmut Verlag. 16, 39, 40
- Fries, P. H. (1994). On Theme, Rheme and Discourse Goals. In M. Coulthard (Ed.) *Advances in Written Text Analysis*, (pp. 229–249). London and New York: Routledge. 39
- Fries, P. H. (1995a). A Personal View of Theme. In M. Ghadessy (Ed.) *Thematic Development in English texts*, vol. 1, (p. 19). London: Pinter London. 39
- Fries, P. H. (1995b). Themes, Methods of Development, and Texts. In R. Hasan, & P. H. Fries (Eds.) *On Subject and Theme: A Discourse Functional Perspective*, vol. 118 of *Current Issues in Linguistic Theory*, (p. 317). John Benjamins Publishing. 39
- Fries, P. H., & Francis, G. (1992). Exploring Theme: Problems for Research. *Occasional Papers in Systemic Linguistics*, 6(1992), 45–60. 5
- García Riaza, B. (2012). *Attribution and Thematization Patterns in Science Popularization Articles of the Guardian Newspaper*. Ph.D. thesis, Departamento de Filología Inglesa, Universidad de Salamanca. 84
- Gelbukh, A. (Ed.) (2009). *Computational Linguistics and Intelligent Text Processing*. Mexico City, Mexico: Springer.
- Gelbukh, A. (Ed.) (2013). *Computational Linguistics and Intelligent Text Processing. Part I*. Samos, Greece: Springer.
- Ghadessy, M. (1983). Information Structure in Letters to the Editor. *IRAL. International Review of Applied Linguistics in Language Teaching*, 21(1), 46–56. 37, 40, 258
- Ghadessy, M. (1995). Thematic Development and Its Relationship to Register and Genres. In M. Ghadessy (Ed.) *Thematic Development in English Texts*, (pp. 129–146). London: Pinter. xviii, 3, 5, 40, 258, 377, 386
- Ghadessy, M. (1997). Thematic Organisation as a Criterion for Establishing English Text-types. *Estudios Ingleses de la Universidad Complutense*, (5), 129–146. 40
- Givón, T. (1995). *Functionalism and Grammar*. John Benjamins Publishing.

## REFERENCES

---

- Glynn, S., & Krawczak, K. (2014). Operationalisation and Robust Manual Annotation of Non-observable Usage-features. An Exploratory Study in English and Polish. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 103
- Gómez-González, M. A. (1994). The Relevance of Theme in the Textual Organization of BBC News. *WORD - Journal of the International Linguistic Association*, 45(3), 293–305. 19
- Gómez-González, M. A. (2001). *The Theme-Topic Interface: Evidence from English*. John Benjamins Publishing. 80, 81, 199
- Görlach, M. (2004). *Text Types and the History of English*, vol. 139. Walter de Gruyter. 30
- Gosden, H. (1992). Discourse Functions of Marked Theme in Scientific Research Articles. *English for Specific Purposes*, 11(3), 207–224.
- Gries, S. T. (2010). Useful Statistics for Corpus Linguistics. In A. Sánchez, & M. Almela (Eds.) *A Mosaic of Corpus Linguistics: Selected Approaches*, vol. 66, (pp. 269–291). Peter Lang.
- Gries, S. T., & Newman, J. (2013). Creating and Using Corpora. *Research Methods in Linguistics*, (pp. 257–287). 44
- Gruber, H., & Muntigl, P. (2005). Generic and Rhetorical Structures of Texts: Two Sides of the Same Coin? *Folia Linguistica*, 39(1-2), 75–113. 254
- Grzegorek, M. (1984). Thematisation in English and Polish. A Study in Word Order. *Uniwersytet Imienia Adama Mickiewicza w Poznaniu. Seria Filologia Angielska Poznan*, (18), 1–148. 3, 377, 386
- Gunter, B. (2003). *News and the Net*. Routledge.
- Hajič, J., Votrubec, J., Krbeč, P., Květoň, P., et al. (2007). The Best of Two Worlds: Cooperation of Statistical and Rule-Based Taggers for Czech. In *Proceedings of the Workshop on Balto-Slavonic Natural Language Processing: Information Extraction and Enabling Technologies*, (pp. 67–74). Association for Computational Linguistics.
- Halliday, M. A. K. (1967). Notes on Transitivity and Theme in English Part I. *Journal of Linguistics*, 3(01), 37–81. 17, 18
- Halliday, M. A. K. (1985). Dimensions of Discourse Analysis: Grammar. In T. A. Van Dijk (Ed.) *Handbook of Discourse Analysis*, vol. 2, (pp. 29–56). London: Academic Press. 18
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold. 17
- Halliday, M. A. K., & Hasan, R. (1989). *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford University Press. 32, 33, 254, 261

- Halliday, M. A. K., & Matthiessen, C. (2004). *An Introduction to Functional Grammar*. Routledge, 3rd ed. 3, 18, 19, 70, 72, 73, 80, 84, 85, 86, 87, 88, 89, 90, 139, 171, 173, 180, 241, 377, 386
- Halliday, M. A. K., & Webster, J. (2009). *Continuum Companion to Systemic Functional Linguistics*. London; New York, NY: Continuum. 135
- Hasan, R. (1978). Text in the Systemic-Functional Model. In W. Dreesler (Ed.) *Current Trends in Text Linguistics*, (pp. 228–246). Walter de Gruyter. 255
- Hasselgård, H. (1996). *Where and When?: Positional and Functional Conventions for Sequences of Time and Space Adverbials in Present-Day English*. Universitetet i Oslo.
- Hasselgård, H. (2004). Thematic Choice in English and Norwegian. *Functions of Language*, 11(2), 187–212. 3, 377, 386
- Hausser, R. (2001). *Foundations of Computational Linguistics: Human-Computer Communication in Natural Language*. Springer, 2nd rev. and ext. ed. 133
- Hawes, T., & Thomas, S. (1996). Rhetorical Uses of Theme in Newspaper Editorials. *World Englishes*, 15(2).
- Hayes, A. F., Glynn, C. J., & Shanahan, J. (2005). Willingness to Self-censor: A Construct and Measurement Tool for Public Opinion Research. *International Journal of Public Opinion Research*. 52
- Hayes, A. F., & Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, 1(1), 77–89. 52
- Hidalgo Downing, R. (2003). *La Tematización en el español hablado: estudio discursivo sobre el español peninsular*, vol. 429. Madrid: Gredos. 3, 21, 377, 386
- Ho-Dac, L.-M. (2008). *La Position Initiale Dans L’Organisation Du Discours: Une Exploration En Corpus*. Ph.D. thesis, Département de Sciences du Langage de l’Université Toulouse le Mirail.
- Honnibal, M., & Curran, J. R. (2007). Creating a Systemic Functional Grammar Corpus from the Penn Treebank. In *Proceedings of the Workshop on Deep Linguistic Processing*, (pp. 89–96). Association for Computational Linguistics. 4
- Hovy, E. H., & Lavid, J. (2007). Classifying Clause-Initial Phenomena in English: Insights for Microplanning in NLG. In *Proceedings of the 7th International Symposium on Natural Language Processing (SNLP)*. 3, 378, 387
- Hovy, E. H., & Lavid, J. (2010). Towards a ‘Science’ of Corpus Annotation: A New Methodological Challenge for Corpus Linguistics. *International Journal of Translation*, 22(1), 13–36. xiii, 44, 45, 46, 47, 52, 55, 59, 65, 69, 301

## REFERENCES

---

- Huddleston, R. (1988). Constituency, Multi-Functionality and Grammaticalization in Halliday's Functional Grammar. *Journal of Linguistics*, 24(01), 137–174. 19
- Huddleston, R. (1991). Further Remarks on Halliday's Functional Grammar: A Reply to Matthiessen and Martin. *Occasional Papers in Systemic Linguistics*, 5, 75–129. 19
- Hughes, B. (2005). Metadata Quality Evaluation: Experience from the Open Language Archives Community. In *Digital Libraries: International Collaboration and Cross-Fertilization*, (pp. 320–329). Springer. 290
- Hyland, K. (2001). Bringing in the Reader Addressee Features in Academic Articles. *Written Communication*, 18(4), 549–574. 194, 274, 338, 381, 390
- Hyland, K. (2005). Stance and Engagement: A Model of Interaction in Academic Discourse. *Discourse Studies*, 7(2), 173–192. 73, 74, 173, 174, 195, 211, 216, 243
- Hyland, K. (2010). Metadiscourse: Mapping Interactions in Academic Writing. *Nordic Journal of English Studies*, 9(2), 125–143.
- Hyland, K., & Bondi, M. (2006). *Academic Discourse across Disciplines*, vol. 42. Peter Lang. 239, 276
- Ide, N., & Romary, L. (2004). International Standard for a Linguistic Annotation Framework. *Natural Language Engineering*, 10(3-4), 211–225.
- Ide, N., & Suderman, K. (2009). Bridging the Gaps: Interoperability for GrAF, GATE, and UIMA. In *Proceedings of the Third Linguistic Annotation Workshop*, (pp. 27–34). Association for Computational Linguistics.
- Jenset, G. B. (2013). Basic Statistics for Corpus Linguistics.
- Jiménez Juliá, T. (1986). *Aproximación al estudio de las funciones informativas*. Málaga: Librería Ágora. 15, 16
- Jordan, M. P. (1985). Non-thematic Re-entry: An Introduction to and Extension of the System of Nominal Group Reference/Substitution in Everyday English Use. *Systemic Perspectives on Discourse*, 1, 322–32. 40
- Kano, Y., Baumgartner, W. A., McCrohon, L., Ananiadou, S., Cohen, K. B., Hunter, L., & Tsujii, J. (2009). U-Compare: Share and Compare Text Mining Tools with UIMA. *Bioinformatics*, 25(15), 1997–1998. 52
- Karlgren, J. (2000). *Stylistic Experiments for Information Retrieval*. Ph.D. thesis, Stockholm University, Department of Linguistics. 30
- Khany, R., & Nejad, A. M. (2010). The Interaction between Rhetorical Structure and Thematization in Academic Research Articles. 41, 254

- Kilgarrriff, A., & Grefenstette, G. (2003). Introduction to the Special Issue on the Web as Corpus. *Computational Linguistics*, 29(3), 333–347. 30
- Krippendorff, K. (2004). Reliability in Content Analysis. *Human Communication Research*, 30(3), 411–433. 52, 54, 56, 57, 383, 392
- Kruijff-Korbayová, I., & Kruijff, G.-J. M. (2004). Discourse-Level Annotation for Investigating Information Structure. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation*, (pp. 41–48). Association for Computational Linguistics. 4
- Lavid, J. (1998). The Role of Corpora for Linguistic and Computational Studies: Thematization as an Example. In M. T. Turell, & E. Vallduví (Eds.) *IV Jornades de corpus lingüístics: els corpus en la recerca semàntica i pragmàtica*, Publicaciones del Instituto de Lingüística Aplicada (IULA). Universidad Pompeu Fabra. 40
- Lavid, J. (2000a). Contextual Constraints on Thematization in Written Discourse: An Empirical Study. *Formal Aspects of Context*, 20, 37. 26, 40
- Lavid, J. (2000b). Text types, Chaining Strategies and Theme in a Multilingual Corpus: A Cross-linguistic Comparison for Text Generation. In *Estudios de Filología Inglesa: Homenaje a Jack White*, (pp. 107–122). 40
- Lavid, J. (2010). Contrasting Choices in Clause-Initial Position in English and Spanish: A Corpus-Based Analysis. In E. Swain (Ed.) *Thresholds and Potentialities of Systemic Functional Linguistics: Multilingual, Multimodal and Other Specialised Discourses*, (pp. 49–68). Trieste: EUT. 22
- Lavid, J. (2011). Corpus Analysis and Annotation in CONTRANOT: Linguistic and Methodological Challenges. In I. Moskowich (Ed.) *Encoding the Past, Decoding the Future: Corpora in the 21st Century*. Cambridge Scholars Publishing. 52
- Lavid, J. (2014). Compiling and Annotating a Register-Controlled Bilingual (English-Spanish) Corpus: Linguistic and Computational Issues. In Universidad de Las Palmas de Gran Canaria (Ed.) *6th International Conference on Corpus Linguistics: Input a Word, Analyze the World: Corpus Linguistics and Society*. 103, 104
- Lavid, J., Arús, J., Carretero, M., Moratón, L., & Zamorano, J. R. (2014). Contrastive Corpus Annotation in the CONTRANOT Project: Issues and Challenges. In M. Gómez González, F. Ruiz de Mendoza Ibáñez, & F. González García (Eds.) *Form and Function in Language: Functional, Cognitive and Applied Perspectives. Essays in Honour of Christopher S. Butler (FFL)*. Amsterdam: John Benjamins. 75, 94
- Lavid, J., Arús, J., & Moratón, L. (2009a). Comparison and Translation: Towards a Combined Methodology for Contrastive Corpus Studies. *IJES, International Journal of English Studies, Special Issue on Recent and Applied Corpus-based Studies*(1), 159–173. 3, 378, 386

## REFERENCES

---

- Lavid, J., Arús, J., & Moratón, L. (2009b). Thematic Choices in English and Spanish: An SFL Analysis of Two Newspaper Genres. In *21st European Systemic Functional Linguistics Conference and Workshop*. Cardiff University, Wales. 134
- Lavid, J., Arús, J., & Moratón, L. (2010a). Annotating Thematic Features in English and Spanish. In M. Taboada, S. Doval Suárez, & E. González Álvarez (Eds.) *Contrastive Discourse Analysis*. London, England.: Equinox. 22, 94, 134
- Lavid, J., Arús, J., & Moratón, L. (2010b). Signalling Genre through Theme: The Case of News Reports and Commentaries. In *8th MAD: Signalling Text Organisation*, (pp. 82–92). Moissac, France: University of Toulouse. 3, 84, 378, 386
- Lavid, J., Arús, J., & Moratón, L. (2010c). Towards an Annotated English-Spanish Corpus with SFL-based Textual Features. In *37th International Systemic-Functional Congress*. Vancouver, Canada: University of British Columbia. 80
- Lavid, J., Arús, J., & Moratón, L. (2011). Thematic Variation in English and Spanish Newspaper Genres: A Contrastive Corpus-Based Study. In *ICAME 32 (Workshop on Corpus-based Contrastive Analysis)*. Norway: Oslo University. 134, 254
- Lavid, J., Arús, J., & Moratón, L. (2012). Genre Realized in Theme: The Case of News Reports and Commentaries. *Discours: Special Issue on Multidisciplinary Approaches to Signalling Text Organisation, Special Issue on Multidisciplinary Approaches to Signalling Text Organisation*,(10). 3, 26, 40, 69, 78, 134, 378, 386
- Lavid, J., Arús, J., & Moratón, L. (2013a). Investigating Thematic Choices in Two Newspaper Genres: An SFL-based Study Analysis. In G. O'Grady, & L. Fontaine (Eds.) *Choices in Language: Applications in Text Analysis*, (pp. 187–209). London: Equinox. 3, 69, 90, 94, 103, 134, 377, 386
- Lavid, J., Arús, J., & Moratón, L. (2013b). Investigating Thematic Choices in Two Newspaper Genres: An SFL-based Study Analysis. In G. O'Grady, & L. Fontaine (Eds.) *Choices in Language: Applications in Text Analysis*, (pp. 187–209). London: Equinox. 65, 71, 84
- Lavid, J., Arús, J., & Moratón, L. (2013c). Thematic Variation in English and Spanish Newspaper Genres: A Contrastive Corpus Based Study. In *Advances in Corpus Based Contrastive Linguistics. Studies in Honor of Stig Johansson*, no. 54 in Studies in Corpus Linguistics, (pp. 261–286). Amsterdam: John Benjamins. 26, 27, 36, 37, 38, 69, 134, 253
- Lavid, J., Arús, J., & Zamorano-Mansilla, J. R. (2010d). *Systemic Functional Grammar of Spanish: A Contrastive Study with English*, vol. 1. A&C Black. 3, 4, 5, 6, 13, 20, 22, 23, 24, 25, 26, 69, 70, 72, 74, 76, 77, 78, 93, 144, 153, 163, 335, 341, 377, 384, 386, 394
- Lavid, J., & Moratón, L. (2013a). Intersubjective Positioning and Thematisation in English and Spanish: A Contrastive Analysis of Letters to the Editor. In *Paper presented at ICAME*

- 34: *English corpus linguistics on the move: Applications and implications. Pre-Conference Workshop "Cross linguistics studies at the interface between lexis and grammar"*. Spain: Universidad Santiago de Compostela.
- Lavid, J., & Moratón, L. (2013b). La anotación automática de rasgos temáticos en inglés y español. In *13th International Symposium on Social Communication*, vol. 1 of *Lingüística Aplicada. Computational Linguistics*. Centro de Lingüística Aplicada. Actualizaciones en Comunicación Social. Santiago de Cuba, Santiago de Cuba: Universidad de Santiago de Cuba. 60
- Lavid, J., & Moratón, L. (2014). Writer's Stance and Engagement in English and Spanish Newspaper Genres: A Contrastive Corpus Based Study. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 38, 103
- Lavid, J., & Moratón, L. (2015). Intersubjective Positioning and Thematisation in English and Spanish: A Contrastive Analysis of Letters to the Editor. *Nordic Journal of English Studies*, 14(1), 289–319.
- Leach, G., & Wilson, A. (1996). Recommendations for the Morphosyntactic Annotation of Corpora. *Relatório técnico, Expert Advisory Group on Language Engineering Standard (EAGLES)*.
- Leech, G. (1997). Introducing Corpus Annotation. In R. Garside, G. Leech, & A. McEnery (Eds.) *Corpus Annotation: Linguistic Information from Computer Text Corpora*, (p. 281). London: Longman. 43, 60, 93, 337, 380, 389
- Leech, G. (2004). Developing Linguistic Corpora: A Guide to Good Practice. Adding Linguistic Annotation. Retrieved from <http://www.ahds.ac.uk/guides/linguistic-corpora/chapter2.htm> Consulted March 2015.
- López Maestre, M. D. (1997). Is There Stylistic Variation within the Headlines of a Single Newspaper?: A Study of the Intranewspaper Variation of the Grammar of the Headlines in Six Sections of the Times (1970-1990). A Quantitative Approach. In Universidad de Murcia. Departamento de Filología Inglesa (Ed.) *Estudios de lingüística aplicada y literatura: homenaje póstumo al profesor Juan Conesa Sánchez*, (pp. 155–166). EDITUM. 269
- Lorés, R. (2004). On RA Abstracts: From Rhetorical Structure to Thematic Organisation. *English for Specific Purposes*, 23(3), 280–302. 41
- Lotfipour-Saedi, K., & Rezai-Tajani, F. (1996). Exploration in Thematization Strategies and Their Discoursal Values in English. *Text*, 16(2). 3, 377, 386
- Marcus, M. P., Marcinkiewicz, M. A., & Santorini, B. (1993). Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2), 313–330. 297

## REFERENCES

---

- Martin, J. R. (1992). *English Text: System and Structure*. John Benjamins Publishing.
- Martin, J. R. (1997). Analysing Genre: Functional Parameters. In F. Christie, & J. R. Martin (Eds.) *Genre and Institutions: Social Processes in the Workplace and School*, (pp. 3–39). A&C Black. 31
- Martin, J. R., Christie, F., & Rothery, J. (1987). Social Processes in Education: A Reply to Sawyer and Watson (and Others). *The Place of Genre in Learning: Current Debates*, (pp. 35–45). 31
- Martin, J. R., et al. (1984). Language, Register and Genre. *Children Writing: Reader, 1*, 984. 40
- Martínez Lirola, M. (2008). Exploring Predicated Themes from a Systemic Functional Point of View in Alan Paton's Novels. *Journal of Literary Studies*, 24(1), 101–128. 200, 274
- Martínez Lirola, M., & Smith, B. (2007). Predicated Themes as textual resource: An analysis of Alan Paton's *Cry, the Beloved Country*. In L. Barbara, & T. Berber Sardinha (Eds.) *Proceedings of the 33rd International Systemic Functional Congress*, (pp. 236–259).
- Mathesius, V. (1975). *A Functional Analysis of Present Day English on a General Linguistic Basis*. 208. Walter de Gruyter. 15
- Matthiessen, C. (1995). *Lexicogrammatical Cartography: English Systems*. International Language Sciences Publications. 72
- Maxwell-Reid, C. R. (2011). *Effect of Bilingual Education on Students First Language Written Discourse: A Contrastive Spanish-English Study Using Systemic Functional Linguistics*. Ph.D. thesis, The University of Edinburgh. 22
- McCabe, A., & Heilman, K. (2007). Textual and Interpersonal Differences between a News Report and an Editorial. *Revista Alicantina de Estudios Ingleses*, 20, 139–156. 40, 171, 236
- McCabe, A. M. (1999). *Theme and Thematic Patterns in Spanish and English History Texts*. Ph.D. thesis, Aston University. 21, 40, 82, 83
- McCabe, A. M., & Alonso Belmonte, M. I. (2001). Theme, Transitivity and Cognitive Representation in Spanish and English Written Texts. *CLAC*. 3, 21, 377, 386
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh University Press. 60
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based Language Studies: An Advanced Resource Book*. Taylor & Francis. 44
- McNaught, J., & Calzolari, N. (1996). EAGLES Morphosyntactic Annotation. Retrieved from <http://www.ilc.cnr.it/EAGLES/annotate/annotate.html> Consulted June 2015. 60



- Mémet, M. (2005). Letters to the Editor. *European Journal of English Studies*, 9(1), 75–90. 37, 207
- Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction*. Cambridge University Press.
- Mitkov, R. (2005). *The Oxford Handbook of Computational Linguistics*. Oxford University Press.
- Mladová, L., Zikanova, S., & Hajicová, E. (2008). From Sentence to Discourse: Building an Annotation Scheme for Discourse Based on Prague Dependency Treebank. In *LREC*. 286
- Moratón, L. (2009). *Comparing Thematic Features in English and Spanish: A Functional Analysis of Two Newspaper Genres*. Master's thesis, Universidad Complutense de Madrid, Spain. 3, 134, 160, 378, 386
- Moratón, L., Arús, J., & Lavid, J. (2012). Themes, Hyperthemes and Generic Structure in English and Spanish Newspaper Genres. In *Invited paper presented at 11th Conference of the European Society for the Study of English*. Turkey: Bogazici University of Istanbul. 3, 41, 253, 254, 378, 386
- Moratón, L., & Lavid, J. (2012). Towards the Automatic Annotation of Thematic Patterns in English and Spanish: A Contrastive Study for Computational Purposes. In *Paper Presented at Interdisciplinary Conference of AHLIST (Association of History, Literature, Science and Technology)*. Spain: Universidad Complutense de Madrid. 3, 134, 329, 378, 386
- Moratón, L., & Lavid, J. (2013). Thematic Progression Patterns in English and Spanish Newspaper Genres. In *COLD OC. La question des genres à l'écrit et à l'oral*. Université de Paris-Ouest, France. 41, 253, 254, 258
- Munday, J. (1998). Problems of Applying Thematic Analysis to Translation between Spanish and English. *Cadernos de tradução*, 1(3), 183–213. 21
- Munday, J. (2000). Using Systemic Functional Linguistics as an Aid to Translation between Spanish and English: Maintaining the Thematic Development of the ST. *Revista Canaria de Estudios Ingleses*, 40, 37–58. 21
- Noda, S. (1994). Thematization: A Contrastive Study of Japanese and Spanish. *Gengo Kenkyu. The Linguistic Society of Japan*, 105, 32–53.
- Nuyts, J. (1995). Functionalism vs. Formalism. In *Handbook of Pragmatics. Manual*, (pp. 293–300). Amsterdam.
- Orasan, C. (2003). Palinka: A Highly Customisable Tool for Discourse Annotation. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialog*, (pp. 39–43). 285, 329, 340, 383, 392

## REFERENCES

---

- Östman, J.-O., & Virtanen, T. (1999). Theme, Comment, and Newness as Figures in Information Structuring. *Amsterdam Studies in the Theory and History of Linguistic science*, (pp. 91–110).
- Pareja-Lora, A. (2012). *Ontotag- a Linguistic and Ontological Annotation Model Suitable for the Semantic Web*. Ph.D. thesis, Facultad de Informática. Universidad Politécnica de Madrid, Spain. 60, 61
- Pareja-Lora, A., & Aguado de Cea, G. (2010). Modelling Discourse-related Terminology in OntoLingAnnots Ontologies. In *Workshop on Establishing and Using Ontologies as a Basis for Terminological and Knowledge Engineering Resource*. Dublin, Ireland.
- Pérez Blanco, M. (2014). From Evidentiality to Expectedness: The Functions of English Evidential Adverbs in Newspaper Opinion Discourse. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 103
- Pérez-Leroux, A. T., Munn, A., Schmitt, C., & DeIrish, M. (2004). Learning Definite Determiners: Genericity and Definiteness in English and Spanish. In *BUCLD 28 Proceedings Supplement*. 203, 275, 338, 381, 391
- Pérez Veneros, M., & Elorza, I. (2014). The Flow of Authorized and Journalists' Voices in English and Spanish: A Cross-linguistic Study of Science Popularizations. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 103
- Perfetti, C. A., & Goldman, S. R. (1974). Thematization and Sentence Retrieval. *Journal of Verbal Learning and Verbal Behavior*, 13(1), 70–79. 4
- Perfetti, C. A., & Goldman, S. R. (1975). Discourse Functions of Thematization and Topicalization. *Journal of Psycholinguistic Research*, 4(3), 257–271. 4
- Petrillo, M., & Baycroft, J. (2010). *Introduction to Manual Annotation*. 47
- Philip, R. P., & Kathirvalavakumar, T. (Eds.) (2013). *Mining Intelligence and Knowledge Exploration*, vol. Lecture Notes in Artificial Intelligence of *Subseries of Lecture Notes in Computer Science*. Tamil Nadu, India: Springer.
- Ping, A. L. (2004a). Delimiting the Theme of the English Clause—An Inference-Boundary Account. *SKY Journal of Linguistics*, 17, 167–187.
- Ping, A. L. (2004b). *Theme and Rheme: An Alternative Account*, vol. 266. Peter Lang.
- Ping, A. L. (2005). Talking Themes: The Thematic Structure of Talk. *Discourse Studies*, 7(6), 701–732.

- Poesio, M. (2004). Discourse Annotation and Semantic Annotation in the Gnome Corpus. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation*, (pp. 72–79). Association for Computational Linguistics.
- Pounds, G. (2010). Writer’s Argumentative Attitude: A Contrastive Analysis of ‘Letters to the Editor’ in English and Italian. *Pragmatics*, 15(1). 37, 207, 228, 275, 338, 339, 382, 391
- Quirk, R., Greenbaum, S., Leech, G. N., Svartvik, J., et al. (1972). *A Grammar of Contemporary English*. Essex, England: Oxford University Press. 17
- Rasinger, S. M. (2008). *Quantitative Research in Linguistics: An Introduction*. Research Methods in Linguistics. Continuum International Publishing Group Ltd. 131
- Rehm, G. (2002). Towards Automatic Web Genre Identification: A Corpus-Based Approach in the Domain of Academia by Example of the Academic’s Personal Homepage. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*, (pp. 1143–1152).
- Romary, L., & Ide, N. (2003). Outline of the International Standard Linguistic Annotation Framework. In *Workshop Linguistic annotation: Getting the model right. ACL 2003*.
- Roy, B. (2011). *Beginner’s Guide to Journalism and Mass Communication*. V&S Publishers. 37, 338, 381, 390
- Schwarz, L., Eckart de Castilho, R., & Teich, E. (2008). Exploring Automatic Theme Identification: A Rule-Based Approach. In *Text Resources and Lexical Knowledge. Selected Papers from the 9th Conference on Natural Language Processing KONVENS 2008*, (pp. 15–26). Mouton de Gruyter. 3, 324, 384, 393
- Shaw, P. (2009). Linking Adverbials in Student and Professional Writing in Literary Studies: What Makes Writing Mature. *Academic Writing: At the Interface of Corpus and Discourse. London: Continuum*, (pp. 215–235). 239, 276
- Stodden, V. C. (2010). Reproducible Research: Addressing the Need for Data and Code Sharing in Computational Science. *Computing in Science & Engineering*, 12(5), 8–12.
- Stubbs, M. (1996). *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Blackwell Oxford. 30
- Subirats Rüggeberg, C. (1998). Automatic Extraction of Textual Information in Spanish. *Language Design: Journal of Theoretical and Experimental Linguistics*, 1, 001–13.
- Sullivan, A. D. (2014). *Determining an Inter-Rater Agreement Metric for Researchers Evaluating Student Pathways in Problem Solving*. Ph.D. thesis, Iowa State University. 53, 54, 56

## REFERENCES

---

- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press. 32, 147
- Taboada, M. (1995). *Theme Markedness in English and Spanish: A Systemic-Functional Approach*. Master's thesis, Universidad Complutense de Madrid, Spain. 21, 81
- Taboada, M. (2013). Contrasting Opinions: Appraisal in a Multilingual Corpus. In *International Conference on Genre- and Register-Related Text and Discourse Features in Multilingual Corpora*. Brussels, Belgium. 4
- Taboada, M., & Lavid, J. (2003). Rhetorical and Thematic Patterns in Scheduling Dialogues: A Generic Characterization. *Functions of Language*, 10(2), 147–178. 254
- Taboada, M., & Zabala, L. H. (2008). Deciding on Units of Analysis within Centering Theory. *Corpus Linguistics and Linguistic Theory*, 4(1). 135
- Taylor, C. V. (1983). Structure and Theme in Printed School Text. *Text-Interdisciplinary Journal for the Study of Discourse*, 3(2), 197–228.
- Thompson, G. (1994). *Reporting: Collins Cobuild English Guides 5*. London: Harper Collins Publishers. 36, 84, 337, 380, 389
- Tinsley, H. E., & Brown, S. D. (2000). *Handbook of Applied Multivariate Statistics and Mathematical Modeling*. Academic Press. 56
- Tse, P., & Hyland, K. (2006). Gender and Discipline: Exploring Metadiscourse Variation in Academic Book Reviews. *Academic Discourse across Disciplines*, (pp. 177–202). 239, 276
- Uebersax, J. (2001). Statistical Methods for Rater Agreement. Retrieved from <http://john-uebersax.com/stat/agree.htm> Consulted June 2015.
- Van Dijk, T. A. (1985). Structures of News in the Press. *Discourse and Communication: New Approaches to the Analysis of Mass Media Discourse and Communication*, 10, 69.
- Van Dijk, T. A. (1987). *News Analysis: Case Studies of International and National News in the Press*. L. Erlbaum Associates.
- Van Dijk, T. A. (1989). Race, Riots and the Press. An Analysis of Editorials in the British Press about the 1985 Disorders. *Gazette*, 43, 229–253. 169, 180, 190, 246, 277
- Van Noord, G., Schuurman, I., & Vandeghinste, V. (2006). Syntactic Annotation of Large Corpora in STEVIN. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*. Genoa, Italy.
- Van Valin, R. D. (1997). *Syntax: Structure, Meaning, and Function*. Cambridge University Press.

- Vande Kopple, W. J. (1991). Themes, Thematic Progressions, and some Implications for Understanding Discourse. *Written Communication*, 8(3), 311–347. 39
- Vázquez y del Árbol, E. (2005). A Genre-Based Study of Biomedical Editorials and Letters to the Editor: A Contrastive Analysis. *Ibérica*, 10, 145–160. 37, 246, 254, 277
- Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Fam Med*, 37(5), 360–363. 53, 54, 104, 126, 380, 389
- Virtanen, T. (1992). *Discourse Functions of Adverbial Placement in English. Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions*. ERIC.
- Virtanen, T. (2004). Point of Departure: Cognitive Aspects of Sentence-Initial Adverbials. *Trends in Linguistics Studies and Monographs*, 147, 79.
- Vogel, R. (2008). Sentence Linkers in Essays and Papers by Native vs. Non-Native Writers. *Discourse and Interaction*, 1. 239, 276
- Weikum, G. (2007). Information Retrieval and Data Mining. In *Computer Science Lecture at University of Saarland, Winter Term*, vol. 2008.
- Weil, H. (1978). *The Order of Words in the Ancient Languages Compared with that of the Modern Languages*, vol. 14. John Benjamins Publishing. 14
- Werlich, E. (1976). *A Text Grammar of English*. Quelle & Meyer Heidelberg. 30
- White, P. R. (1997). *Death, Disruption and the Moral Order: The Narrative Impulse in Mass-Media 'Hard News' Reporting*. London: Continuum. xviii, 36, 142, 143, 255, 256, 337, 380, 389
- White, P. R. (1998). *Telling Media Tales: The News Story as Rhetoric*. Ph.D. thesis, The University of Sydney, Australia. 137, 143, 144, 147, 165, 250
- Wiebe, J., Wilson, T., & Cardie, C. (2005). Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation*, (pp. 165–210). 52, 54, 104
- Wierzbicka, A. (1988). *The Semantics of Grammar*, vol. 18. John Benjamins Publishing.
- Wilcock, G. (2009). Introduction to Linguistic Annotation and Text Analytics. *Synthesis Lectures on Human Language Technologies*, 2(1), 1–159.
- Zamorano-Mansilla, J. R., & Carretero, M. (2014). The Annotation of Modality and Evidentiality in English-Spanish Comparable and Parallel Texts. In *International Conference on Evidentiality and Modality in European Languages (EMEL)*. Spain: Universidad Complutense de Madrid. 103



# APPENDIXES







## Curriculum Vitae

## *A. Curriculum Vitae*

---

# Education

## **PhD IN ENGLISH LINGUISTICS**

June 2010 - 2015

**Universidad Complutense de Madrid (Spain)**

*Thesis: Thematic Patterning in English and Spanish:  
Contrastive Annotation of a Bilingual Newspaper Corpus  
for Linguistic and Computational Applications.*

## **MA IN ENGLISH LINGUISTICS:**

October 2007 - June 2009

**New Applications and International Communication**

**Universidad Complutense de Madrid (Spain)**

*Thesis: Comparing thematic features in English and Spanish:  
A functional analysis of two newspaper genres.*

## **MA IN LINGUISTICS**

October 2005 - September 2007

**Universidad Complutense de Madrid (Spain)**

Major in Computational Linguistics.

## **MA IN ENGLISH STUDIES**

October 2002 - June 2007

**Universidad Complutense de Madrid (Spain)**

Major in Applied Linguistics.

Major in Theoretical Linguistics.

## *A. Curriculum Vitae*

---

# Professional Experience

## **VOICE TEAM MANAGER**

April 2013 - to date

### **Samsung Electronics Iberia**

S Voice team manager: Manager of a testing group that develops and updates all steps of the recognition process.

S Voice Natural Language Understanding european coordinator: Coordinator of all european teams for SVoice in the area of NLU.

## **LINGUISTIC CONSULTANT**

February 2012 - April 2013

### **Appen Butler Hill Inc.**

Project LIVINGSTONE: Developing oral expressions in natural language for Spanish.

Web Data Evaluator.

## **RESEARCHER in**

March 2012 - December 2012

### **CORPUS & COMPUTATIONAL LINGUISTICS**

#### **Universidad Complutense de Madrid**

Project CONTRANOT: Creation and validation of contrastive functional descriptions (English-Spanish) through analysis and corpus annotation: linguistic, computational and methodological aspects. (Project of Ministry of Science and Innovation.)

## **COMPUTATIONAL LINGUIST**

March 2011 - January 2012

### **Nimbus Systems**

Research Team coordinator for grammar-based voice recognition engines (May 2010 - March 2011).

Information structure and dialogue flow design for  
Interactive Voice Response Applications (IVR) for Orange, Yoigo, Simyo.  
Design of voice recognition grammars (ASR, DTMF).  
Fine Voice Tuning (Nuance 8.5).  
Refining and improving TTS dictionaries.  
Prompt writing for marketing campaigns.

**COMPUTATIONAL LINGUIST**

October 2009 - May 2010

**Fonetic Solutions, S.A.**

Voice Speech Mining: Call Center speech analysis  
for Telefónica, La Caixa, BBVA, etc.  
using Speech Analytics for the implementation of  
marketing campaigns and business analysis.

**LINGUISTIC CONSULTANT**

June 2009 - October 2009

**Molino de Ideas, S.A.**

Project ONOMA: Natural Language Processing:  
analysis of the Spanish verb to build  
an on-line verbal conjugator (<http://www.onoma.es/>)

**COMPUTATIONAL LINGUIST**

February 2008 - June 2009

**Infospeech, S.A.**

Call Center analysis and content structure  
using Speech recognition tools (Speech Analytics) Telefónica, BBVA.  
Dialogue flow design for Telefonica IVR.  
Usability studies for Bankinter Help Desk systems.  
Fine Voice Coaching.  
Project I3Media. Creation of a voice synthesizer Spanish-Catalan.

**LINGUIST**

January 2008 - December 2008

**Universidad Complutense de Madrid**

---

Collaborator in the Project: Linguistic annotation of  
English and Spanish texts with computational and didactic aims.  
Comunidad de Madrid and Complutense University.

**LINGUIST**

March 2006 - December 2006

**Universidad Complutense de Madrid**

Collaborator in the Project: Text compilation and exploitation for  
translation and contrastive learning (English- Spanish). Grant by European Higher  
Education Area.





# Publications

## Book Chapters

Lavid, J., Arús, J., Carretero, M., Moratón, L. & JR Zamorano (2014) Contrastive Corpus Annotation in the CONTRANOT project: Issues and Problems. In M.A. Gómez González, F. Ruiz de Mendoza Ibáñez, F. González García & A. Downing (Eds.), *The Functional Perspective on Language and Discours: Applications and Implications*. Amsterdam, Holland: John Benjamins.

Lavid, J., Arús J., & Moratón, L. (2013) Investigating Thematic Choices in Two Newspaper Genres: An SFL-based Study Analysis. In Gerard O'Grady & Lise Fontaine (Eds.), *Choices in Language: Applications in Text Analysis*. London, England: Equinox.

Lavid, J., Arús J., & Moratón, L. (2013) Thematic Variation in English and Spanish Newspaper Genres: A Contrastive Corpus Based Study. In K. Ajimer & B. Altemberg (Eds.), *Advances in Corpus Based Contrastive Linguistics. Studies in Honor of Stig Johansson*. Amsterdam, Holland: John Benjamins.

Lavid, J., Arús J., & Moratón, L. (2012). Annotating Thematic Features in English and Spanish: A Contrastive Corpus Based Study. In Maite Taboada, Susana Doval Suárez & Elsa González Álvarez (Eds.), *Contrastive Discourse Analysis*. London, England: Equinox.

## Scientific Journals and Master's Thesis

Lavid J., & L. Moratón, L. (2015) Intersubjective Positioning and Thematisation in English and Spanish: A Contrastive Analysis of Letters to the Editor. *Nordic Journal of English Studies*, 14(1):289–319.

Moratón, L., & Lavid, J. (2013) Thematic Progression Patterns in English and Spanish Newspaper Genres. *COLDOC La question des genres à l'écrit et à l'oral*. Université de Paris-Ouest (France). Text available at: [https://coldoc2013fr.files.wordpress.com/2014/11/livret\\_actescoldoc\\_version2.pdf](https://coldoc2013fr.files.wordpress.com/2014/11/livret_actescoldoc_version2.pdf)

Lavid, J., Arús J., & Moratón, L. (2012). Genre Realized in Theme: The Case of News Reports and Commentaries. *Discours: Special Issue on Multidisciplinary Approaches to Signalling Text Organisation*, 10. Text available at: <http://discours.revues.org/8623>

Arús J., Lavid, J., & Moratón, L. (2012). Annotating Thematic Features in English and Spanish: A Contrastive Corpus Based Study. *Linguistics and the Human Sciences*, 6(1-3),173-192.

Arús J., Lavid, J., & Moratón, L. (2009) Comparison and Translation: Towards a Combined Methodology for Contrastive Corpus Studies. In Pascual Cantos and Aquilino Sánchez (Eds.), *International Journal of English Studies Special Issue on Recent and Applied Corpus Studies*. Publicaciones de la Universidad de Murcia, 159-173.

Moratón, L. (2009) *Comparing Thematic Features in English and Spanish: A Functional Analysis of two Newspaper Genres* (Master's thesis). Available at E-Prints Complutense: [http://eprints.ucm.es/17388/1/Moraton\\_Lara.pdf](http://eprints.ucm.es/17388/1/Moraton_Lara.pdf)

## Presentations at Scientific Congresses and Proceedings

Lavid, J., & Moratón, L. (2014) *Writer's Stance and Engagement in English and Spanish Newspaper Genres: A Contrastive Corpus Based Study*. Paper presented at EMEL (International Conference on Evidentiality and Modality in European Languages). Universidad Complutense de Madrid, Spain, 6-8 October 2014

Lavid, J., & Moratón, L. (2014, April 30-May 5) *Global Discourse Structures, Rhetorical Relations and Thematic Patterns in English and Spanish Journalistic Texts: A Comparative Study*. Paper presented at ICAME 35 (Corpus Linguistics, Context and Culture). University of Nottingham, Great Britain, April 30-May 5 2014

Lavid, J., & Moratón, L. (2013) La anotación automática de rasgos temáticos en inglés y español. *13th International Symposium on Social Communication*. Centro de Lingüística Aplicada. Actualizaciones en Comunicación Social. Santiago de Cuba. Text available at: [//www.santiago.cu/hosting/linguistica/actas.php?Simposios=XIII&id=en&Actas=1&numpage=16&total=99](http://www.santiago.cu/hosting/linguistica/actas.php?Simposios=XIII&id=en&Actas=1&numpage=16&total=99)

Lavid, J., & Moratón, L. (2013) *Investigating the Automated Annotation of Thematic Features in English and Spanish*. Paper presented at ICLC7 - UCCTS 3. Ghent University, Belgium, 11-13 July 2013

---

Lavid, J., & Moratón, L. (2013) *Intersubjective Positioning and Thematisation in English and Spanish: A Contrastive Analysis of Letters to the Editor*. Paper presented at ICAME 34: English Corpus Linguistics on the Move: Applications and Implications. Pre-Conference Workshop “Cross linguistics studies at the interface between lexis and grammar”. Universidad de Santiago de Compostela, Spain, 22-26 May 2013

Lavid, J., Arús, J., Carretero, M., Moratón, L., & Zamorano, JR. (2012) *Annotating Thematic and Modal Categories in English and Spanish: Issues and Challenges*. Invited paper presented at 11th Conference of the European Society for the Study of English. Bogazici University of Istanbul, Turkey, 4-8 September 2012

Moratón, L., Arús, J., & Lavid, J. (2012) *Themes, Hyperthemes and Generic Structure in English and Spanish Newspaper Genres*. Paper presented at 11th Conference of the European Society for the Study of English. Bogazici University of Istanbul, Turkey, 4-8 September 2012

Moratón, L., & Lavid, J. (2012) *Towards the Automatic Annotation of Thematic Patterns in English and Spanish: A Contrastive Study for Computational Purposes*. Paper presented at Interdisciplinary Conference of AHLiST (Association of History, Literature, Science and Technology). Universidad Complutense de Madrid, Spain, 27-29 June 2012

Lavid, J., Arús J., & Moratón, L. (2011) *Thematic Variation in English and Spanish Newspaper Genres: A Contrastive Corpus Based Study*. Paper presented at ICAME 32. Workshop on Corpus-based Contrative Analysis. Oslo University, Norway, 1 June 2011

Lavid, J., Arús J., & Moratón, L. (2011) *What Thematic Annotation can Reveal about Interpersonal Issues in Cross-linguistic Research*. Paper presented at 38th International Systemic-Functional Congress. University of Lisbon, Portugal, July 2011

Lavid, J., Arús J., & Moratón, L. (2010) *Annotating Thematic Features in English and Spanish*. Paper presented at 6th International Contrastive Linguistics Conference (ICLC 6). Freie Universität Berlin, Germany, September 30-October 3 2010

Lavid, J., Moratón, L., & Arús J. (2010). *Towards an Annotated English-Spanish Corpus with SFL-based Textual Features*. Paper presented at 37th International Systemic Functional Congress. University of British Columbia, Canada, 18-23 April 2010

Lavid, J., Arús J., & Moratón, L. (2010). *Investigating Thematic Meaning in English and Spanish: A Methodological Proposal*. Paper presented at 22nd European Systemic Functional Linguistics Conference and Workshop. Koper University, Slovenia, 7-10 July 2010

Lavid, J., Arús J., & Moratón, L. (2010). *Corpus Annotation in the CONTRANOT Project*. Paper presented at Interdisciplinary Conference of AHLiST (Association of History, Literature, Science and Technology). Universidad Complutense de Madrid, Spain, 6 June 2010

Lavid, J., Arús J., & Moratón, L. (2010). Signalling Genre Through Theme: The Case of News Reports and Commentaries. *Proceedings of the 8th MAD Multidisciplinary Perspectives on Signalling Text Organisation*. Mossaic, France, 17-20 March, 2010. Text available at: <http://w3.workshop-mad2010.univ-tlse2.fr/>

Moratón, L. & Lavid, J., Arús J. (2009). Using Comparable and Translated Texts in Contrastive Corpus Analysis: Thematisation as an Example. *Proceedings of the Corpus Linguistics Conference 2009 (CL 2009)*. University of Liverpool: Great Britain. Text available at: <http://ucrel.lancs.ac.uk/publications/CL2009/#papers>.

Lavid, J., Arús J., & Moratón, L. (2009). *Using Comparable and Translated Texts in Contrastive Corpus Analysis: Thematisation as an Example*. Poster presented at International Corpus Linguistics Conference. University of Liverpool, United Kingdom, 20-10 July, 2009

Lavid, J., Arús J., & Moratón, L. (2009). *Thematic Choices in English and Spanish: An SFL Analysis of Two Newspaper Genres*. Paper presented at 21st European Systemic Functional Linguistics Conference and Workshop. University of Cardiff, Wales, 8-10 July, 2009

# B

## English and Spanish Summary

### B.1 English Summary

The motivation for the work reported in this dissertation is rooted in two main research areas: On the one hand, Systemic-Functional Linguistics and its application to language comparison and corpus analysis; and Natural Language Processing (NLP), on the other. In Systemic-Functional Linguistics, the phenomenon of thematization has been recognized as a fundamental mechanism in the construction of messages and texts, and several scholars have pointed out its discourse role as a signal through which writers guide their readers in the interpretation of messages and in the characterization of texts (see Halliday & Matthiessen, 2004; Downing, 1991; Lotfipour-Saedi & Rezai-Tajani, 1996).

A productive line of research has focused on the relationship between thematic selection, registers, and genres (see Ghadessy, 1995), focusing almost exclusively on English texts. More recently, several scholars with an interest in comparing English to other languages have pointed out the potential of thematization for both clausal and discourse structuring (see Grzegorek, 1984; Hasselgård, 2004; Hidalgo Downing, 2003; McCabe & Alonso Belmonte, 2001; Hidalgo Downing, 2003; Lavid et al., 2013a).

The current dissertation relies specifically on the thematization model developed by Lavid, Arús, & Zamorano-Mansilla (2010d) in *Systemic Functional Grammar of Spanish: A Contrastive Study with English*, and developed under the scope of the projects CONTRANOT<sup>1</sup>

---

<sup>1</sup>The CONTRANOT project was financed by the Spanish Ministry of Science and Innovation

and MULTINOT<sup>1</sup>, led by Professor Julia Lavid. It is also framed within the range of activities developed by the FUNCAP (Functional Linguistics English-Spanish and its Applications) research group at Universidad Complutense to which I belong. This research group has produced fruitful results in the area of contrastive annotation, as illustrated by the work of Lavid, Arús, & Moratón (2012, 2010b, 2009a); Arús, Lavid, & Moratón (2013, 2012); Moratón (2009); Moratón, Arús, & Lavid (2012); Moratón & Lavid (2012).

However, in spite of thematization for message and textual structure, to date there are no linguistic studies that have undertaken the task of validating its aspects in a contrastive manner, either for linguistic or computational purposes. This study, therefore, seeks to fill a research gap implementing a methodology based on contrastive corpus annotation, which will allow to empirically validate thematization in English and Spanish.

This study attempts to advance knowledge in this area by focusing on three newspaper genres—news reports, editorials and letters to the editor—in English and Spanish, whose thematic features have seldom been studied contrastively before, as indicative of their generic characterization. It analyses in detail Thematic Field and its elements: Textual Theme, Interpersonal Theme, PreHead and Head. It gives account of its organization in the text studying the thematic progression of the three genres in both languages.

In the field of Natural Language Processing (NLP), the phenomenon of thematization has also attracted the attention of researchers working in Discourse Parsing and in Natural language generation (NLG). In NLG, for example, thematization plays a fundamental role in micro planning tasks (see Hovy & Lavid, 2007; Duszak, 1994; Basile & Bos, 2011).

The following research questions have been investigated throughout the different sections of the manuscript.

1. Can the theoretical phenomenon of thematization in English and Spanish be empirically tested through corpus annotation?
2. Are there any language-specific thematic selection and progression patterns in the three newspaper genres studied, i. e., news reports, editorials and letters to the editor? Are there any correlations between thematic patterns and generic moves in these three newspaper genres?
3. Is it possible to automate the manual annotation of thematic features in English and Spanish and how can this be achieved? If possible, is the result obtained reliable?

The organization of the thesis resembles the methodology followed and is structured in three main parts:

---

(MICINN) for the period 2009-2012.

<sup>1</sup>The MULTINOT project is currently under way and it is financed by the Spanish Ministry of Economy and Competitiveness for the period 2013-2015.

## **Empirical Validation of the Model of ‘Theme’**

The empirical validation of the thematization model consists of the preparation of a sample corpus on which agreement studies will be carried out. The aim of these studies is to achieve a measure of agreement that renders the categories reliable and valid enough to be annotated on a larger corpus. The first step after the selection of the sample corpus is to determine the concepts that the coders are going to annotate. These are described in the guidelines that lead the annotators in their tasks. The second step is to design the studies and the scope to cover, as well as their ability to be statistically measured. Once these elements are ready, the work of the two coders start.

The two agreement studies carried out for the validation of thematic categories revealed interesting aspects of the annotation of thematic features. The first study, deals with the identification of spans, one referred to identifying the Thematic Field and the other to delimit its elements,. This task proved to be more difficult than the second task where the labeling of already selected elements was requested. It was more difficult for Spanish than for English, due to the complex morphology of the Spanish verbal group.

**Agreement Study 1.** In task 1 of this Agreement Study the annotators had to identify the span corresponding to the Thematic Field, a reliable result in AGR agreement (0.97 and 0.74) was achieved in both languages. The Spanish result is lower than the English, as mentioned earlier, for the more complex morphology of the Spanish verbal group.

Task 2 also obtained a high level of agreement in the selection of the spans corresponding to Thematic Head, PreHead, Interpersonal Theme, Textual Theme, Predicated Theme and There/Hay Theme. Thematic Head in English is at 0.93 and Spanish at 0.91, although there is a slight difference in the values reflecting a better agreement in English, both AGR results shows the relevance of the result. PreHead shows a better agreement in Spanish (0.89) than in English (0.78). Interpersonal Themes selection also presents a higher value in Spanish (0.94) than in English (0.37). Textual Themes show an almost even value in both languages (English 0.96 and Spanish 0.97). Predicated Themes and There show almost a perfect agreement in the selection of their spans.

**Agreement Study 2** consisted of five tasks in which the coders had to label the spans marked in the complex clauses according to the categories delivered in a drop down list.

In Task 1, the coders had to select the type of Thematic element highlighted in the dataset. The Kappa value obtained was 0.915 for English and 0.839 for Spanish.

The second Task, in which the type of experiential roles of the Head were marked, the kappa value obtained in English was 0.875 and for Spanish 0.475. The moderate level of agreement is due to discrepancies presented by the coders in the identification of relational processes where Token and Value types were mistaken.

Task 3 consisted in the annotation of types of Textual Themes. In this task, the Kappa values obtained were 0.664 in English and 0.927 in Spanish, although both are relevant enough as to consider this category validated there is a clear difference of lower agreement in English than in Spanish. In this case the problem yields in the selection by one of the coders of one

category where the other marked it differently. This may indicate the need for further definition of these elements in the guidelines.

In Task 4, coders had to select the type of Interpersonal Theme. In this task both languages present a very low Kappa measurement (0.405 in English and 0.291 in Spanish). Although these values are considered “moderate” and “fair” respectively, according to Viera & Garrett (2005), their reliability is compromised. The subjectivity of these concepts makes it advisable to be especially cautious in their application. Regarding automatic annotation rules, it may not be possible to use these types of Interpersonal Themes as they are defined in the guidelines.

Task, 5, revealed a high level of agreement in the selection of the types of PreHead where the Kappa values were 0.927 in English and 0.934 in Spanish.

To sum up, the categories subject to the empirical validation process: Thematic Field and its elements; in the light of the statistical results obtained these categories are considered to be validated from the point of view of its identification in a clause complex (span selection of the first agreement study) and also from the point of view of its further classification (labeling in the second agreement study, except for the types of Interpersonal Theme).

## **Annotating a Larger Corpus**

As Leech (1997) says “corpora are useful only if we can extract knowledge or information from them.” Now that the categories have been validated, they can be applied to the annotation of a larger corpus. Due to the impossibility of annotating a larger corpus by more than one annotator, and according to Dligach et al. (2010) who claim that the most efficient and reliable corpora is obtained through more single handed coded annotations of an expanded corpus; a large corpus is collected. The corpus consists of 220 texts news reports, editorials, and letters to the editor, both in English and Spanish, which make up to 2551 clause complexes for analysis.

After the analysis, the genres for study can be thematically characterized as follows:

**News reports** are defined by White (1997) as communicatively-based rather than event-based, striving to remain neutral and objective while presenting a diversity of opinions, voices, and perspectives of the event, incident, or issue under discussion. This picture may be completed with what Thompson (1994) expects when turning to a newspaper report, that “(...) it will give a reasonably objective “true” picture of the event and of the evidence, reactions, and opinions of the people involved.” These definitions are resembled in the thematic choices of news reports.

The “diversity of opinions, voices...” depicted in news reports are represented by the use of quotations in PreHead position. English up fronts quotations to PreHead whereas Spanish delays its apparition to the rheme, although these quotations can be traced by a verb of communicative meaning. A common feature shared by both languages is the use of PreHead to “highlight about the overturning of temporal sequences” (Bell, 1991), that is, as a temporal setting, using more complex devices in English (PH-CCL), and simpler expressions in Spanish (PH-Circ). Another interesting feature of PreHead in Spanish is that it is revealed as the preferred slot, being a pro-drop language, for the lexical part of the verb when it appears in first



position.

The realization of Head is accomplished by the use of Noun Phrases and Noun Phrases followed by Appositions, both elements more frequently used in English. In its place, Spanish uses verbs, particularly its grammatical information. These verbs both in English and Spanish, in its semantic role, act, as Francis (1990) pointed out as *actors* and *sayers*.

The composition of the thematic field in news reports is a clear reflection of the nature of these texts. The use of elements of the inner thematic field and the almost non existent apparition of outer thematic components reveal a clear picture of this type of texts.

Regarding *Generic Structure Potential* (GSP) and thematic progression patterns, news reports are the only genre of the analyzed that has three moves instead of four. It is, in its structural organization, simple. The results revealed that after the explicative headline, the next move, in which the issue is presented, is characterized by the use of constant theme. That is, repeating the same Head and adding information to it. The development of the piece of news in the last move, where the facts about the issue are presented, is deployed by a hypertheme followed by constant thematic progression in English, and linear patterns in Spanish.

**Editorials** considered as “an article in a newspaper that gives the opinion of the editor or publisher on a topic or item of news” (Sinclair (1995) in Ansary & Babaii, 2005), “(…) seeking to encourage critical thinking, to mould opinion, and to promote action.” (Roy, 2011:109). The distributional map of thematic elements change drastically.

Due to the longer and explanatory nature of the texts, the use of Textual Themes increases in both languages. It is worth noticing the contrast in the use of Binders in English and in Spanish. These elements are the preferred textual strategy for English in long and explanatory texts; whereas Linkers are seldom used on Editorials or on other types of texts. Spanish, on the other hand, follows this general pattern but accounts for a wider use of Linkers in the three genres.

The apparition of Interpersonal Theme reveals the presence of the author in the text, in most cases in the form of Engagement where the writer conveys his or her position in the argument (Hyland, 2001), “moulding opinion and encouraging critical thinking” (Roy, 2011).

In PreHead position, both languages prefer the use of simple forms of the type PH-Circ for expression of Time and more complex (and less frequent) to express Contingency, introducing counterarguments or hypothesis. A language characteristic of Spanish is the use of the lexical part of the verb in this position.

Head reflects the presence of the writer in the text through the use of Thematic Equatives and Predicated Themes. Head is realized by Noun Phrases, particularly pronouns in English, as its most frequent strategy. Spanish instead, uses Verbs in first position avoiding the explicit mark of the subject. Another language characteristic in this position in English is the use of common nouns, not preceded by determiner whereas Spanish does not present any occurrence of this kind, this, as Pérez-Leroux et al. (2004) notes: “(…) English allows in subject position

bare plural Noun Phrases while Spanish and some other Romance languages do not.” The semantic realizations belong to *actors* and *relational processes* in this type of genre, that serve to present facts to the reader.

The composition of thematic field in editorials is a clear reflection of the nature of these texts. The use of both inner and outer thematic field reflex the complexity of theme in these type of texts.

Regarding the *Generic Structure Potential* (GSP) and thematic progression patterns, editorials present a four move structure. After the metaphorical headline, the next move is introduced by a derived theme that addresses the issue. The argumentation of facts and its evaluation is organized by *units of information* that follow the same pattern of an introductory hypertheme and a set of derived themes attached to it. The same arrangement is followed in the last move, where the articulation of a position on the part of the writer takes place.

In **letters to the editor** “the writers take position by expressing their views on current events and states of affairs and by appealing for action. In so doing they also reveal their feelings towards such events and provide their evaluation of the people and circumstances involved” Pounds (2010). The short extension of these texts compromise the choice of certain elements in the thematic field.

The selection of Textual Themes reflect the nature of the texts in the use of binders both in English and Spanish, linkers, as mentioned above, and although less used in both cases, are preferred by Spanish rather than English.

Contrary to expected, Interpersonal Themes although present in the texts, were expected in higher numbers. The subjectivity is exposed in the texts but it is revealed through language specific strategies. English selects the use of Engagement to highlight the presence of the author in the texts, but does so in less number than Spanish. Spanish evenly prefers Stance and Engagement. Does this mean that English authors do now appear in the texts? Rather, the explanation points to the reason given by Pounds (2010) who says that English “stance” may be more often “integrated into text rather than overtly marked.”

In PreHead position Spanish selects simple syntactic constructions and English more complex formations to express Time and Contingency.

As in editorials, Head is, in some cases, realized by Thematic Equatives and Predicated Themes. The preferred syntactic realization of Head is that of NP in both languages, although English presents a statistical significant difference, Spanish balances this selection with the use of verbs in first position.

In this type of texts, contrary to other genres analyzed there are other types of clauses than declaratives, therefore it can be also studied the use of verbal forms in Head position in English directives and questions. The choice of experiential elements is more varied, appearing in relevant numbers: *actors*, *carriers*, *senser*, *sayers* and *goal* (and the special mention of *process* in Spanish as a language characteristic).

Regarding the *Generic Structure Potential* (GSP) and thematic progression patterns,

letters to the editor are structures in four moves. The headline summarizes the issue at stake. Although the text extension is very short, it allows for an introduction of the issue, outlining the argumentation and closing. All of them are derived themes, in this particular case the short extension of the text does not allow to develop more complex or elaborated patterns, being in most cases a single clause complex the only component of the move.

## **Towards Automatic Theme Annotation**

The last step of this process is devoted to the exploration of thematic annotation automatization. The aim of this effort is to make available a reliable annotated corpus for further research, and to try to achieve this goal through automatic annotation.

The corpus selected is the same as the manually annotated corpus and discussed above. This allows for the comparison of the results of both manual and automatic annotation, and to test its reliability. The guidelines that have been validated in the first part of the study, and tested in the second part in manual annotation, now acquire a salient role, helping in the writing of automatic rules, and in the documentation of the study for the benefit of future users (Dipper et al., 2004).

The tool to support the process had to fulfill a number of minimum requirements: ease of use, friendly interface, its ability to multitask and multilanguage (Orasan, 2003). These claims are key to making the study usable, reproducible, flexible and interactive. It was important that the tool be useful and accessible to all users, not only to skilled computer users (Eckart de Castilho, 2014); reproducible and therefore reliable (Krippendorff, 2004); flexible, to allow changes in the use of the tool components; and interactive enough to be used by others. With all these points in mind, the selection of the annotation tool was GATE (General Architecture for Text Engineering) (Cunningham et al., 2011).

Data was prepared to be automatically processed (format, metadata (see chapter 23)) and uploaded to GATE. The preprocessing of texts consists of preparing the corpus prior to its annotation. In this case, the final annotation is the thematic annotation, but before this task takes place, relevant data has to be obtained from the corpus, and this is achieved through bootstrapping. To be able to automatically annotate theme, it is mandatory to identify the units of analysis, its tokens, part-of-speech tagging and parsing. The identification of these data serve as the basis for the formulation of rules that inform the automatic annotation of the corpus. At this moment, the processes for each language diverges, due to availability of these annotation tools. English had tools available to perform these tasks accurately, but Spanish tools were not at my disposal at the time of this study.

English bootstrapping annotate the basis on which the automatic rules are based, that is, tokens and part of speech tagging. Here begun the “translation” of the guidelines to JAVA. After this process the set of JAVA rules ordered in an API is run after the bootstrapping pipeline (sentence splitter, tagger, part-of-speech tagger and parser) and thus annotate the span of thematic field, Textual Theme (binder and linkers), Interpersonal Theme (EngSta and ComMod),

PreHead (Circ and CCL) and Head (Thematic Equatives, Predicated Themes and Existential clauses). The “translation” of the set of guidelines to JAVA proved to be a hard work<sup>1</sup> (see chapter 24.2). The high stage annotation at discourse level implies a wide background linguistic knowledge that has to be transferred to the computer, that is a blank canvas. This iterative process resulted in an API that run on the English pipeline gave the following results when compared to the manually annotated corpus: the number of automatically annotated items under each category does not present a statistically significant difference in relation to the same elements manually annotated. The only category that presented difficulties is Interpersonal Theme, something that may be expected as the intra-annotation agreements measures on this feature presented the lowest rate. This success overcomes previous drawbacks in the automatic annotation of theme (see Schwarz et al., 2008).

In the case of Spanish the available resources to be implemented in GATE, at the time of the study, did not allow to find a part-of-speech tagger and a parser that provided codings accurate enough. The solution given to this situation is to perform a thematic semi-automatic annotation of the corpus where the results are compared to the manual annotation carried out in chapter 18 (manual annotation) obtaining results that ensure that the corpus has been reliably annotated in the following thematic elements: Head, PreHead, Interpersonal Theme, Textual Theme and Thematic Field. Thus delivering an authoritative thematic annotated corpus, as one of the main objectives of the study.

The automatic annotation of both corpus allows for further and deeper study of the annotated data. The computer-supported format allows the use of the ANNIC (ANNotations-In-Context) platform which is a “full-featured annotation indexing and retrieval system” built on top of Apache Lucene—a high performance search engine implemented in JAVA, which supports indexing, searching, and querying of large document collections (Cunningham et al., 2011). This makes it possible to understand the relationship of new thematic layers of annotation to any other annotated feature.

The use of computer-supported devices and automatic annotation should be further explored in linguistic studies. It is, in the light of the results of this work, reliable and accessible, ensuring the replicability of the experiments and the availability for the research community. This procedure is less time consuming than manual annotation and efficacy is not compromised.

This study accomplishes the task of validating aspects of the thematization model proposed by Lavid et al. (2010d) empirically by implementing a methodology based on contrastive corpus annotation in English and Spanish. The bilingual newspaper corpus has been thematically characterized relating its features to the different genres of study, namely, news reports, editorials, and letters to the editor. The same corpus is automatically annotated with thematic

---

<sup>1</sup>JAVA code has been carried out by Sara Lana, in a common effort to share linguistic and computer knowledge

features at clausal and discourse levels, proving its reliability through the comparison of the results obtained through manual and automatic annotation.

Future work may overcome the limitations found in the lack of availability of processing resources for automatic annotation in Spanish, and enhance the developed guidelines. Future work may also extend the study to different genres, and thus explore its automatic identification through thematic characterization.

## B.2 Resumen en castellano

### La estructuración temática en inglés y español: anotación contrastiva de un corpus bilingüe para aplicaciones lingüísticas y computacionales

El trabajo presentado en esta tesis se entronca en dos áreas principales de investigación. Por un lado, en el ámbito de la Lingüística Sistémico-Funcional, y concretamente en su aspecto contrastivo y de análisis de corpus; y por otro lado, en el de Procesamiento del Lenguaje Natural (PLN). En Lingüística Sistémico-Funcional el fenómeno de tematización se ha reconocido como un mecanismo fundamental en la construcción de mensajes y textos, varios investigadores han apuntado su papel discursivo como el modo a través del que los autores guían a sus lectores en la interpretación de mensajes y caracterización de textos (véase Halliday & Matthiessen, 2004; Downing, 1991; Lotfipour-Saedi & Rezai-Tajani, 1996).

Una productiva línea de investigación se ha concentrado en la relación que existe entre la selección temática y el género (Ghadessy, 1995) centrada sobre todo en un único idioma, el inglés. Más recientemente varios investigadores se han interesado en la comparación del inglés con otros idiomas, insistiendo en la idea de Tema como elemento clave en la estructuración de la frase y el discurso (Grzegorek, 1984; Hasselgård, 2004; Hidalgo Downing, 2003; McCabe & Alonso Belmonte, 2001; Hidalgo Downing, 2003; Lavid et al., 2013a).

Es en este punto en el que se sitúa este trabajo, apoyándose en el modelo de tematización desarrollado por Lavid, Arús, & Zamorano-Mansilla (2010d) en *Systemic Functional Grammar of Spanish: A Contrastive Study with English* y, bajo el amparo de los proyectos CONTRA-NOT<sup>1</sup> y MULTINOT<sup>2</sup>, liderados por la profesora Dra. Julia Lavid. El grupo de investigación FUNCAP (Functional Linguistics English-Spanish and its Applications) bajo su dirección, ha contribuido prolíficamente a esta línea de investigación, con los siguientes trabajos, entre otros: Lavid, Arús, & Moratón (2012, 2010b, 2009a); Arús, Lavid, & Moratón (2013, 2012); Moratón (2009); Moratón, Arús, & Lavid (2012); Moratón & Lavid (2012).

A pesar de la importancia del estudio de la tematización en la estructuración del mensaje y del texto, hasta ahora, no ha habido estudios que hayan llevado a cabo la tarea de validar aspectos de este fenómeno desde un punto de vista contrastivo (inglés-castellano), relacionando además propósitos lingüísticos y computacionales. Este estudio pretende rellenar un hueco en este área, desarrollando una metodología basada en la anotación contrastiva de un corpus que valide empíricamente aspectos temáticos en castellano e inglés.

Se pretende además incidir en el ámbito del género periodístico, en concreto en: noticias, editoriales y cartas al director, en inglés y castellano. La tesis aborda detalladamente el campo temático y sus elementos: Textual Theme, Interpersonal Theme, PreHead y Head. También

---

<sup>1</sup>El proyecto CONTRANOT fue financiado por el Ministerio de Ciencia e Innovación durante los años 2009-2012.

<sup>2</sup>El proyecto MULTINOT está siendo financiado por el Ministerio de Economía y Competitividad en el período 2013-2015.

estudia su organización en el texto de acuerdo a los patrones de progresión temática en los tres géneros y ambas lenguas.

En el campo del Procesamiento del Lenguaje Natural (PLN), el fenómeno de la tematización ha atraído la atención de investigadores de las áreas de Análisis de discurso y Generación de Lenguaje Natural (GLN). En GLN, por ejemplo, la tematización juega un papel fundamental en las tareas de micro planning (Hovy & Lavid, 2007; Duszak, 1994; Basile & Bos, 2011).

El presente trabajo pretende responder a las siguientes preguntas de investigación:

1. ¿Se puede comprobar empíricamente el concepto teórico de tematización en inglés y castellano a través de la anotación de corpus?
2. ¿Existe alguna selección en la elección temática y su progresión que se pueda asociar a los tres géneros estudiados (noticias, editoriales y cartas al director)? ¿Existe alguna relación entre los patrones temáticos y los *moves* en alguno de los géneros estudiados?
3. ¿Se puede automatizar la anotación de categorías temáticas en inglés y castellano?, si esto es así, ¿de qué modo?, ¿son los resultados obtenidos fiables?

La estructuración de la tesis se asemeja a la metodología escogida y se organiza en las siguientes partes: una primera parte introductoria, una segunda dedicada al marco teórico en el que se presentan las nociones utilizadas a lo largo del trabajo. La tercera, en la que se desarrolla la validación empírica de las categorías temáticas que luego serán anotadas en un corpus más extenso. Una cuarta parte que aborda la anotación de rasgos temáticos y su progresión en el corpus más extenso. El quinto bloque está dedicado a explorar la anotación automática de este corpus mediante la creación de una herramienta en JAVA a partir de las guías utilizadas en la validación empírica; y, finalmente, la última parte recoge a las conclusiones.

A continuación se resume el análisis y los resultados más destacables del estudio en las partes de: validación empírica, anotación del corpus extenso y anotación automática.

## Validación empírica

La validación empírica del modelo de tematización consiste en la preparación de una muestra del corpus sobre la que se van a llevar a cabo los experimentos de acuerdo entre anotadores. El propósito de estos estudios es el de obtener medidas de acuerdo entre los anotadores suficientemente relevantes como para considerar las categorías fiables y válidas para ser anotadas en un corpus más extenso.

El primer paso tras la selección de la muestra del corpus es definir las categorías para dárselas a conocer a los anotadores. Estos conceptos se describirán en las guías de anotación que servirán de ayuda en la tarea de codificación.

El segundo paso es el diseño de los experimentos y su ámbito de aplicación, así como su estructura para poder ser sometidos a medición estadística.

Una vez que este trabajo está hecho comienza el de los dos anotadores.

Los dos experimentos de anotación llevados a cabo para la validación empírica pusieron al descubierto aspectos interesantes de las categorías temáticas. El primer experimento, en el que tenía que identificarse el ámbito de aplicación del campo temático (thematic field) y el de sus elementos, resultó ser más difícil que el segundo en el que una sección de texto previamente identificado tenía que ser clasificado de acuerdo a una lista de opciones definidas. También resultó más complicada en castellano que en inglés, debido a la complejidad morfológica del idioma.

En las siguientes líneas detallamos los resultados de estos experimentos de anotación:

El **Experimento 1**, en la primera tarea, en la que los anotadores tenían que identificar el ámbito del campo temático (thematic field), dio como resultado en la medición de AGR 0.97 y 0.74 en inglés y castellano respectivamente, ambos considerados una medida de acuerdo suficientemente fiable, aunque en castellano son algo menores que en inglés debido a la complejidad morfológica de este idioma.

En la segunda tarea de este primer experimento se obtuvo un alto nivel de acuerdo entre los anotadores en la selección del texto correspondiente a las categorías de Thematic Head, PreHead, Interpersonal Theme, Textual Theme, Predicated Theme y There/Hay Theme. Thematic Head en inglés alcanza un nivel de acuerdo de 0.93 y en castellano de 0.91, a pesar de esta pequeña diferencia entre ambos idiomas, los dos resultados son suficientemente relevantes. En el PreHead el nivel de acuerdo es más alto en castellano de 0.89, mientras que en inglés es de 0.78. La selección de los Interpersonal Theme también muestra una tasa de acuerdo más alta en castellano (0.94) que en inglés (0.37). Los valores de Textual Theme son casi idénticos en los dos idiomas inglés 0.96 y castellano 0.97. Tanto Predicated Themes como There/Hay tienen un resultado de acuerdo casi total.

El **Experimento 2** consistió en la realización de cinco tareas en las que los codificadores tenían que etiquetar una selección de texto previamente subrayado de acuerdo a las categorías disponibles en una lista desplegable.

En la primera tarea, los anotadores tenían que elegir el elemento temático al que correspondía la parte subrayada del texto. Se obtiene un resultado Kappa de 0.915 en inglés y 0.839 en castellano.

La segunda tarea consistía en escoger el tipo de papel semántico del Head (*actor*, *sayer*, *value*, *goal*, *etc.*...) , los resultados muestran un desigual acuerdo entre inglés, con un valor de Kappa de 0.875 y 0.475 en castellano. El nivel moderado del acuerdo obtenido en castellano es debido a las discrepancias entre los anotadores a la hora de clasificar los procesos relacionales en sus tipos de *value* y *token* que fueron sistemáticamente confundidos.

La tercera tarea trataba de la anotación del Textual Theme. En ella el valor de kappa fue de 0.664 en inglés y 0.927 en castellano, a pesar de que ambos resultados muestran un nivel de acuerdo lo suficientemente alto como para ser considerado relevante y la categoría validada, hay una clara diferencia entre los dos resultados. En este caso el problema se encuentra en que uno de los anotadores eligió un elemento textual donde el otro escogió otra categoría distinta.



Este problema puede resolverse con una clarificación de la definición de estos elementos en las guías de anotación.

En la cuarta tarea los anotadores seleccionan el tipo de Interpersonal Theme. En esta tarea ambos idiomas presentan una tasa de Kappa muy baja (0.405 en inglés y 0.291 en castellano) y, a pesar de que Viera & Garrett (2005) lo consideran un valor de acuerdo moderado, para los fines de este estudio su fiabilidad queda comprometida. La subjetividad de esta categoría recomienda que se sea especialmente precavido con respecto a ella. En cuanto a su reflejo en las reglas de anotación automática, puede que no sea posible usar los tipos de elementos interpersonales tal y como quedan definidos en la guías de anotación.

Finalmente, la última tarea, número 5, revela un alto nivel de acuerdo en la selección de los tipos de PreHead, siendo el valor de Kappa 0.927 en inglés y 0.934 en castellano.

En resumen, las categorías sometidas al proceso de validación empírica: el campo temático y sus elementos, y, a la luz de los resultados estadísticos obtenidos se pueden considerar validadas desde el punto de vista de su identificación en una frase, (ámbito de ocurrencia de los elementos en el primer ejercicio) y también sus tipos dentro de una clasificación más fina (clasificación en el segundo experimento, excepto los tipos de elementos interpersonales)

## Anotación de un corpus más extenso

En palabras de Leech (1997) “corpora are useful only if we can extract knowledge or information from them” [Los corpus son útiles solo si se puede extraer conocimiento o información de ellos]

Debido a la imposibilidad de anotar el corpus por más de una anotador, se toma la decisión, siguiendo a Dligach et al. (2010), de que los corpus más eficientes y fiables son aquellos anotados por un único anotador pero más extensos. De este modo se compila un corpus de noticias, editoriales y cartas al director, en inglés y en castellano, que suman un total de 220 textos, lo que implica un total de 2551 frases analizadas.

Tras el análisis, los géneros estudiados se pueden caracterizar temáticamente de la siguiente manera:

Los artículos, tal y como los define White (1997) están basados en actos comunicativos, de carácter objetivo, utilizando un lenguaje neutral en su presentación de distintas opiniones, voces y perspectivas del incidente o situación que se esté tratando. A esta definición pueden añadirse las palabras de Thompson (1994) sobre lo que él espera cuando se acerca a un periódico “(...) it will give a reasonably objective true picture of the event and of the evidence, reactions, and opinions of the people involved” [obtener una imagen razonablemente objetiva de un hecho o situación, así como reacciones y opiniones de la gente involucrada]. Estas definiciones se reflejan en la selección de los temas en los artículos de noticias.

La *diversidad de opiniones y voces...* aparecen representadas a través de las citas en los PreHead. En inglés estas aparecen principalmente en el PreHead, mientras que en castellano se retrasan hasta el rema, usando en verbos de tipo comunicativo que adelantan la presencia de estos elementos. Una característica común compartida por ambos idiomas es el uso de los PreHead para “highlight about the overturning of temporal sequences” [resaltar aquello que

tiene que ver con el desarrollo temporal de los eventos] (Bell, 1991), usando expresiones más complejas en inglés (PH-CCL), y más sencillas en castellano (PH-Circ). Otra característica interesante de los PreHead en castellano es que en él tiene lugar la parte léxica de los verbos que se aparecen en primera posición de frase.

Head son realizados por sintagmas nominales de forma aislada y también por sintagmas nominales seguidos de una aposición, ambos elementos se usan con más frecuencia en inglés que en castellano. En su lugar el castellano usa verbos, particularmente su información gramatical. Estos verbos, tanto en inglés como en castellano en su papel semántico actúa como dice Francis (1990) como *actors* y *sayers*.

La composición del campo temático en los textos periodísticos es un claro reflejo de la naturaleza de los mismos. El uso de estos elementos en el *inner thematic field* y la casi inexistente aparición del *outer thematic field* revela un claro dibujo de este tipo de textos.

En cuanto a la *Generic Structure Potential* (GSP) y los patrones de progresión temática, los artículos de noticias son el único género de los tres analizados que está organizado en tres movimientos en vez de en cuatro. Es, en cuanto a su estructura, simple. Los resultados muestran que tras un encabezado explicativo, el siguiente movimiento, en el que se presenta los hechos, está caracterizado por el uso de un tema constante, esto es, repite el mismo encabezado añadiendo información sobre este en el rema. El desarrollo de la noticia está formado por un hipertema seguido de temas constantes en inglés y lineares en castellano.

Los editoriales son considerados como “an article in a newspaper that gives the opinion of the editor or publisher on a topic or item of news” (Sinclair (1995) in Ansary & Babaii, 2005) [el artículo del periódico en el que se da la opinión del editor sobre un tema noticioso] “(...) seeking to encourage critical thinking, to mould opinion, and to promote action.” (Roy, 2011:109) [con la intención de animar el pensamiento crítico, modular opinión o promover la acción.] La distribución de los elementos temáticos cambia radicalmente. Para comenzar, y debido a la mayor extensión y naturaleza expositiva de estos textos, el uso de Textual Theme se incrementa en los dos idiomas. Es importante destacar el contraste en el uso de los Binder en inglés y en castellano. Estos elementos son la herramienta textual preferida para el inglés en textos largos y explicativos, como en este caso; en cambio, los Linkers son utilizados raramente en los editoriales y en el resto de los tipos de textos. El castellano, en cambio, sigue este patrón general pero utiliza un uso más extenso de Linkers en los tres géneros.

La aparición de elementos interpersonales revela la presencia del autor en el texto, en la mayoría de los casos en la forma de *Engagement* donde el escritor se refiere al lector en cuanto a su postura en el texto (Hyland, 2001), “moulding opinion and encouraging critical thinking (Roy, 2011) [modelando opinión y alentando el pensamiento crítico].

En el lugar del PreHead, ambas lenguas prefieren el uso de formas simples del tipo PH-Circ para expresar temporalidad y más complejas (y menos frecuentes) para expresar contingencia, introduciendo argumentos o hipótesis. Una característica del castellano es el uso de la parte léxica del verbo en este lugar.

Head refleja la presencia del escritor en el texto mediante el uso de *Thematic Equative* y

*Predicated Theme*. El resto de realizaciones de Head se encuentran sobre todo en forma de sintagmas nominales, en concreto pronombres, como la estrategia más frecuente en inglés. En cambio, en castellano se usan verbos en esta posición sin la marca explícita del sujeto. Otra característica propia del inglés es el uso de nombres comunes, sin ir precedidos de determinantes, mientras que en castellano no aparece ninguno de estos casos como Pérez-Leroux et al. (2004) observa: “(...) English allows in subject position bare plural Noun Phrases while Spanish and some other Romance languages do not”. [En inglés se permite usar en posición de sujeto los sintagmas nominales puros mientras que en castellano y otras lenguas romances esto no es posible]. Las realizaciones semánticas son *actors* y *relational processes*.

La composición del campo temático en editoriales es un reflejo claro de la naturaleza de este tipo de textos. El uso de ambos *inner* y *outer thematic field* reflejan la complejidad temática de estos textos.

En cuanto al *Generic Structure Potential* (GSP) y los patrones de progresión temática en los editoriales tienen una estructura en cuatro partes. Tras el título metafórico, el siguiente movimiento es introducido por un tema derivado que aborda el tema a tratar. La argumentación de los hechos y su evaluación se organiza en *unidades de información* que siguen el mismo patrón: una introducción mediante un hipertema, seguido de un grupo de temas derivados dependientes de este. La misma estructura se refleja en el último movimiento donde se plasma la postura del autor.

En las cartas al director “the writers take position by expressing their views on current events and states of affairs and by appealing for action. In so doing they also reveal their feelings towards such events and provide their evaluation of the people and circumstances involved” Pounds (2010) [el autor toma postura expresando sus puntos de vista respecto a los acontecimientos actuales y los hechos que acontecen llamando a la acción. Haciendo esto revelan sus pensamientos respecto a los hechos relatados y evalúan a las personas y las circunstancias implicadas]. La corta extensión de estos textos condiciona la selección de ciertos elementos del campo temático.

La selección del Textual Theme refleja la naturaleza de los textos en el uso de *Binders*, tanto en inglés como en castellano. Los *Linkers* como ya se mencionaba, y a pesar de que son menos usados en ambos casos, son más numerosos en castellano que en inglés.

Al contrario de lo que pudiera parecer, los elementos interpersonales, a pesar de estar presentes en el texto, no son tan numerosos como cabría esperar. Encontramos subjetividad en los textos pero se revela a través de estrategias propias de cada idioma. En inglés el uso de *Engagement* señala la presencia del autor en el texto pero lo hace en menor medida que en castellano. En español se usan por igual estrategias de *Stance* tanto como de *Engagement*. ¿Significa esto que los autores ingleses no aparecen en el texto? Todo lo contrario, como explica Pounds (2010) “stance may be more often integrated into text rather than overtly marked” [stance suele estar integrado en el texto más que abiertamente marcado].

En la posición de PreHead en castellano se suelen seleccionar estructuras sintácticas simples mientras que en inglés se escogen elementos más complejos para expresar temporalidad y

contingencia.

Como en los editoriales, el Head es, en algunas ocasiones realizado por *Thematic Equative* y *Predicated Theme*. Las estructuras por las que se realizan estos elementos son sintagmas nominales en ambos idiomas, a pesar de que el inglés presenta una diferencia estadística significativa, el castellano, es su lugar equilibra esta situación con el uso de verbos.

En este tipo de textos, al contrario que los otros géneros analizados hay otro tipo de frases además de las declarativas, de ahí que se pueda analizar el uso de formas verbales en la posición de Head en las oraciones imperativas y preguntas en inglés. La elección de los elementos experienciales es mucho más variada, encontrándonos: *actors*, *carriers*, *senser*, *sayers* y *goal*, además de la atención que requiere en castellano el uso de *processes* en esta posición.

En cuanto a la *Generic Structure Potential* (GSP) y los patrones de progresión temática, las cartas al director están estructuradas en cuatro movimientos. Los títulos resumen el motivo de la carta. A pesar de que la extensión de texto es muy corta, se puede encontrar una introducción del tema, presentación de la argumentación y cierre. Todos ellos son de tipo derivados, en este caso particular la corta extensión del texto no permite desarrollar patrones ni más extensos ni más complejos, siendo en muchos casos solo una frase el único componente de cada *move*.

## La anotación automática del Tema

El último paso de este proceso está dedicado a la exploración de la automatización de la anotación automática. El objetivo de este esfuerzo es hacer disponible un corpus anotado fiable para futuras investigaciones. Y, lo que es más, intentar conseguir este objetivo a través de la anotación automática.

El corpus sobre el que se va a realizar la anotación automática es el mismo que el que se ha usado para la anotación manual. Esto permite comparar los resultados de ambos procesos y probar si existen diferencias. Las guías que se han validado en la primera parte del estudio, y se han puesto en práctica en la segunda parte a partir de la anotación manual, toman de nuevo un papel relevante, no sólo ayudando en la definición de las reglas sino también en la documentación del estudio para el beneficio de los futuros usuarios de este corpus (Dipper et al., 2004).

La herramienta que va a servir como soporte para realizar este proceso tiene que cubrir un número mínimo de requisitos: sencillez de uso, interfaz amigable, multitarea y multilenguaje (Orasan, 2003). Estos puntos son clave para conseguir que el estudio sea útil, reproducible, flexible e interactivo. Útil y accesible a todos los usuarios, no sólo a aquellos expertos en informática (Eckart de Castilho, 2014). Reproducible como máxima interpretación de la fiabilidad (Krippendorff, 2004). Flexible para permitir cambios en el uso de los componentes de las herramientas de análisis. E interactivas, esto es, que el estudio provea la suficiente documentación como para ser usado por otros. Con todos estos puntos en mente, la selección de la herramienta de anotación es GATE (General Architecture for Text Engineering) (Cunningham et al., 2011).

Los datos son preparados para ser tratados de forma automática (formato, metadatos...) (véase el capítulo 23) y cargados en GATE. Este preprocesamiento consiste en el tratamiento

que se le da al corpus antes de la anotación final. En este caso, la anotación final es la anotación temática, pero antes de que esta tarea tenga lugar, la información relevante tiene que extraerse del corpus, y esto se hace mediante este preprocesamiento. Para poder anotar automáticamente el tema es necesario identificar las unidades de análisis, los *tokens*, las anotaciones de *part of speech* y el *parsing*. En inglés hay herramientas disponibles para desarrollar estas tareas de forma eficaz, pero en castellano, en el momento en el que se hizo el estudio no se encontraron herramientas para una anotación de *part of speech* y *parsing*.

El preprocesamiento que se lleva a cabo en inglés anota los rasgos básicos en los que se van a basar las reglas automáticas. Aquí comienza la “traducción” de las reglas a JAPE. Tras este proceso un conjunto de reglas de JAVA contenidas en una API anotarán tras el pipeline del bootstrapping (sentence splitter, tagger, part-of-speech tagger and parser) el campo temático: Textual Theme (*Binder, Linkers*), Interpersonal Theme (*EngSta, ComMod*), PreHead (*Circ, CCL*) y Head (*Thematic Equative, Predicated Theme y Existential clauses*). La “traducción” de las guías a JAPE demostró ser un trabajo especialmente duro<sup>1</sup>. El alto nivel de la anotación discursiva necesita de un amplio conocimiento lingüístico que ha de ser transferido al ordenador, que es un lienzo en blanco.

Este proceso iterativo de traducción da como resultado la API que anotará en inglés el tema. Tras la comparación de los resultados obtenidos en la anotación automática con los de la anotación manual se demuestra que el número de elementos anotados en cada categoría no presenta una diferencia estadísticamente significativa entre las dos anotaciones. La única categoría en la que se encuentran dificultades es la de Interpersonal Theme, algo que podía ser esperado ya que el nivel de acuerdo entre los anotadores en esta categoría era el más bajo. Este éxito supera los antiguos escollos encontrados en la anotación automática del tema (véase Schwarz et al., 2008).

En el caso de castellano las herramientas disponibles en GATE, en el momento del estudio, no incluyen un part-of-speech tagger o un parser que realice anotaciones sintácticas lo suficientemente fiables como para usarlas de base para la creación de reglas automáticas. La solución dada a esta situación es llevar a cabo una anotación semi-automática de los elementos temáticos, los resultados obtenidos serán los que se comparen con los anotados manualmente en el capítulo 18 (manual annotation). Los resultados demuestran que se trata de una anotación fiable llevada a cabo en Head, PreHead, Interpersonal Theme, Textual Theme y Thematic Field. De este modo conseguimos nuestro objetivo de obtener un corpus anotados con datos fiables en castellano.

La anotación automática de ambos corpus permitirá estudios más profundos y pormenorizados del tema. La anotación llevada a cabo en formato electrónico permite el uso de ANNIC (ANNotations-In-Context) que es un “full-featured annotation indexing and retrieval system built on top of Apache Lucene - a high performance full-featured search engine implemented in JAVA, which supports indexing, searching and querying of large document collections” (Cun-

---

<sup>1</sup>El código JAPE lo ha desarrollado Sara Lana, en un esfuerzo común para compartir conocimiento lingüístico e informático) (véase el capítulo 24.2)

ningham et al., 2011) [completa herramienta de anotación indexación y recuperación construida sobre Apache Lucene una herramienta de búsqueda implementada en JAVA que permite la indexación, búsqueda y pregunta (querying) de largas cantidades de documentos]. Esto hace posible la explotación de las relaciones de las recientes anotaciones del tema con cualquier otro elemento anotado (por ejemplo, buscar las relaciones de realización de Head como NP etc...)

Tanto el uso de soportes informáticos como la anotación automática debería extenderse a los estudios lingüísticos. Es, bajo este prisma bajo el que se presentan los resultados de este trabajo, fiable y accesible, asegurando la reproducción de los experimentos y su disponibilidad para la comunidad científica. Este procedimiento resulta ser más rápido y eficiente que la anotación manual sin verse la eficacia comprometida.

El estudio consigue el objetivo validar empíricamente aspectos el modelo de tematización propuesto por Lavid et al. (2010d) implementando la metodología basada en el análisis contrastivo de corpus mediante la anotación en inglés y castellano. El corpus periodístico bilingüe comparable se ha caracterizado temáticamente en sus subgéneros: artículos, editoriales y cartas al director. El mismo corpus se ha anotado automáticamente en sus características temáticas a nivel de frase y de discurso, probando su eficacia, a través de la comparación del resultado obtenido manual y automáticamente.

Futuros trabajos podrán orientarse a superar las limitaciones encontradas en cuanto a la disponibilidad de recursos de procesamiento para la anotación automática del castellano y mejorar el desarrollo de las reglas. Puede extenderse además el estudio a otros géneros más allá de los periodísticos y explorar así la identificación automática de género a través de la anotación automática del tema.

# C

## Appendix

## C.1 Annotation Schemes and Guidelines

### **Definition of Thematic Field (English)**

Thematic Field: Initiating clause span of varying length up to and including the first nuclear constituent [FNC] in main clause (in bold in the examples), or one of the following:

- Predicated Theme construction [PT]
- “There” in Existential clauses.

Examples of Thematic field (in bold) ending in [FNC]:

- (1) [FNC:] **the cat** is on the mat
- (2) [FNC:] **Eating** is vital
- (3) [FNC:] **that he refused to do it** worried me
- (4) [FNC:] **Of unequal relevance** is ...
- (5) **On the table** [FNC:] stood a lamp
- (6) **But, surprisingly, before the meeting** [FNC:] **everybody** was glad to hear the news.
- (7) [FNC:] **What I want** is you
- (8) **In my opinion**, [FNC:] **Real Madrid**, their players have been holding up a banner

Examples of Predicated Theme Construction [PT] and “There” in Existential clauses:

- (9) **In fact** [PT:] **It is love** that makes the world go round
- (10) **When I arrived**, [THERE:] **there** were three people waiting for the bus

**Figure C.1** Annotation scheme of Thematic Field English



**Definition of Thematic Field (Spanish)**

Clause-initial material which goes from the beginning of the clause complex up to and including the first nuclear experiential constituent [FNC] in the main clause. The FNC can be realized by either lexical or morphological means.

Examples of Thematic field (in bold) ending in [FNC]:

- (1) [FNC:] **El gato** *está en la alfombra*
- (2) [FNC:] **Cansadísimo** *llegué ayer a casa*
- (3) [FNC:] **Se** *está muy bien aquí*
- (4) [FNC:] **se** *lo di ayer (le-allomorph)*
- (5) **Se** [FNC:] **me** *cayó* (*se* is here part of the Verbal Group *caerse* and is, therefore, pre-Head [it fulfils no participant role]).
- (6) **En realidad**, [FNC:] **corriendo** *no se consigue nada*
- (7) [FNC:] **que me digas eso** *significa que no me has entendido*
- (8) **Teng-**[FNC:] **o frío** (verbal base is pre-Head; both together, ITF)
- (9) [FNC:] **Ten** *cuidado!*
- (10) **Las vacaciones**, [FNC:] **todo el mundo** *sueña con ellas*
- (11) [FNC:] **Lo que necesitas** *es amor* (Thematic Equative)

**Figure C.2** Annotation scheme of Thematic Field Spanish

## C. Appendix

---

### TAGSET DEFINITIONS for THEME CATEGORIES IN ENGLISH

#### (Declarative clauses)

##### 1. THEMATIC HEAD (TH)

The **Thematic Head** is defined as the first nuclear constituent (not circumstantial) element in the clause. This can be a *Participant*, a *Process*, or the “*There*” element in existential clauses.

When the Thematic Head is a *Participant*, it can be realized as:

- a Noun Group. (e.g. The cat is on the mat; Peter is at home; She saw him yesterday)
- an Adverbial Group (e.g. Tomorrow is a holiday)
- a Non-finite Clause (e.g. Eating is vital ; To live is to die)
- a Nominal That-Clause (e.g. That he refused to do it worried me)

When the Thematic Head is a *Process*, it is realized as a verbal form, preceded by a Pre-Head element, such as for example a Circumstance (e.g. *On the table **stood** a lamp*) or an auxiliary.

When the clause is existential, the Thematic Head is realized by the “*There*” element (e.g. ***There** were three people waiting for the bus*)

##### 2. PREHEAD (PH)

The **Pre-Head** element is any circumstantial and/or finite element preceding the Thematic Head. This includes the following realisations (in bold face):

- Adverbial Groups (e.g. ***tomorrow** there will be another meeting*)
- Prepositional Phrases (e.g. ***On your right** you can see the Royal Palace*)
- Circumstantial clauses (e.g. ***After dropping her off**, he continued his trip*)
- Finite verbal forms, i.e. auxiliaries, preceding the lexical verb: (e.g. ***Should** you decide to leave the country, **Had** I known you were so near, I would have flown to meet you*)

##### 3. PREDICATED THEME (PT)

This construction is also known as “it-cleft” and consists of an ‘it+ be+ relative clause’ (e.g.: ***It is you** who are to blame*)

##### 4. TEXTUAL THEME (TT)

Elements which are instrumental in the creation of the logical connections in the text, such as linkers, binders or correlatives. These include:

- Linkers (paratactic nexus) (e.g. And don’t tell me you didn’t know; but let’s change the topic)
- Binders (hypotactic nexus) (e.g. However, the situation now is different; now we needed to promote the event, secondly, you should go to a doctor).
- Correlatives: (not only...but; either...or) (e.g. Not only didn’t he call but also forgot completely about us; either you’re with us or you’re against us.)

##### 5. THERE STRUCTURE (THERE)

This is a type of construction which occurs in existential clauses. It is realized by the “*There*” element. (e.g. ***There** are three boys waiting at the door*)

“**There** has been a dog barking all night”

##### 6. INTERPERSONAL THEME (IT)

These are elements which express the attitude and the evaluation of the speaker with respect to his/her message. These include:

- Vocatives, i.e., any item used to address (e.g. Tom! This is a nice surprise; Sir, could you follow me, please?)
- Comment Adjuncts (e.g. Surprisingly he didn’t mention anything; understandably, he kept a low profile)
- Modal Adjuncts (e.g. Probably that’s the only lesson we learned; Surely you didn’t do that!)

Figure C.3 English tag-set definition for Theme categories

## TAGSET DEFINITIONS for THEME CATEGORIES IN SPANISH

### (Declarative clauses)

#### 1. THEMATIC HEAD (TH)

The **Thematic Head** is defined as the first nuclear constituent (not circumstantial) element in the clause realised by either lexical or morphological means. When the Thematic Head is a *Participant*, it can be realized as:

- a Noun Group. (e.g. *El gato está en el sofá*)
- an Adverbial Group (e.g. *Mañana es festivo*)
- a Non-finite Clause (e.g. *Comer es vital ; Vivir es morir*)

When the Thematic Head is a *Process*, it is realized as a the grammatical part of the verb (e.g. *Saldre-mos a las tres*)

When the clause is existential, the Thematic Head is realized by the “Hay” element (e.g. *Hay tres personas esperando el autobús*)

#### 2. PREHEAD (PH)

The **PreHead** element is any circumstantial and/or finite element preceding the Thematic Head. This includes the following realisations (in bold face):

- Adverbial Groups (e.g. *Mañana habrá otra reunión*)
- Prepositional Phrases (e.g. *A la derecha está el Palacio Real*)
- Circumstantial clauses (e.g. *Después de descansar, Paco continuó su camino*)
- Clitic *se* when pronominal

#### 3. PREDICATED THEME (PT)

A construction that consists of two parts: (1) an initial thematic segment consisting of the copular verb followed by the element in Focus; and (2) a rhematic segment realized by a relative-like clause, as shown below:

“*Fue Fermín el que me dejó triste*”

#### 4. TEXTUAL THEME (TT)

Elements which are instrumental in the creation of the logical connections in the text, such as linkers, binders or correlatives. These include:

- Linkers (paratactic nexus) (e.g. *Y no me digas que no lo sabes*)
- Binders (hypotactic nexus) (e.g. *De todos modos, la situación ahora ha cambiado*)
- Correlatives: (not only...but; either...or) (e.g. *No solo no llamó sino que se marchó*)

#### 5. THERE STRUCTURE (THERE)

This is a type of construction which occurs in existential clauses. It is realized by the “Hay” element. (e.g. *Hay tres personas esperando el autobús*)

#### 6. INTERPERSONAL THEME (IT)

These are elements which express the attitude and the evaluation of the speaker with respect to his/her message. These include:

- Vocatives, i.e., any item used to address (e.g. *¡Tomás! Esto es una sorpresa muy agradable.*)
- Comment Adjuncts (e.g. *Sorprendentemente, él no mencionó nada*)
- Modal Adjuncts (e.g. *Probablemente sea lo único que ha aprendido*)

Figure C.4 Spanish tag-set definition for Theme categories

**ENGLISH Extended TAGSET DEFINITIONS for Textual Theme Types (definitions based on Halliday and Matthiessen (2004) IFG**

**Textual Theme (TT):**

Elements which are instrumental in the creation of the logical connections in the text, such as linkers, binders or correlatives:

- TT-Link
  - Linkers (paratactic nexus). Parataxis is the linking of elements of equal status (Halliday & Matthiessen, 2004, p. 384)
    - e.g.:
      - [TT-Link:] **And** don't tell me you didn't know; **but** let's change the topic
- TT-Bind
  - Binders (hypotactic nexus). Hypotaxis is the binding of elements of unequal status (Halliday & Matthiessen, 2004, p. 384)
    - e.g.:
      - [TT-Bind:] **However**, the situation now is different; **now** we needed to promote the event, **secondly**, you should go to a doctor.
- TT-Cor
  - Correlatives: (not only ... but; either ... or)
    - e.g.:
      - [TT-Cor:] **Not only** didn't he call but also forgot completely about us; **either** you're with us **or** you're against us.
      - He's [TT-Cor:] **either** holidaying **or** he's on another job

**Figure C.5** English tag-set definition for Textual Themes

**SPANISH Extended TAGSET DEFINITIONS for Textual Theme Types (definitions based on Halliday and Matthiessen (2004) IFG**

**Textual Theme (TT)**

- TT-Link
  - linkers (paratactic nexus) (e.g. [TT-Link:] **Y** no me digas que no lo sabes)
- TT-Bind
  - binders (hypotactic nexus) (e.g. [TT-Bind:] **De todos modos**, la situación ahora ha cambiado).
- TT-Cor
  - correlatives: (e.g. [TT-Cor:] **Tal** era el miedo **que** no salió de casa)

**Figure C.6** Spanish tag-set definition for Textual Themes

**ENGLISH Extended TAGSET DEFINITIONS for Interpersonal Theme Types (definitions based on Halliday and Matthiessen (2004) *IFG* and on an adaptation of Hyland's (2005) *Stance and engagement: a model of interaction in academic discourse*.)**

**Interpersonal Themes (IT)**

These are elements which express the attitude and the evaluation of the speaker with respect to his/her message:

- IT-Voc
  - Vocatives, i.e., any item used to address.
    - e.g.:
      - [IT- Voc:] **Tom!** This is a nice surprise.
      - [IT- Voc:] **Sir,** could you follow me, please?
- IT-Com
  - Comment adjuncts.
    - e.g.
      - [IT- Com:] **Surprisingly** he didn't mention anything.
      - [IT- Com:] **Understandably,** he kept a low profile.
- IT-Mod
  - Modal adjuncts.
    - e.g.:
      - [IT- Mod:] **Probably** that's the only lesson we learned
      - [IT- Mod:] **Surely** you didn't do that!
- IT-Eng
  - Engagement: The writer acknowledge and connect to others, recognizing the presence of the reader, focusing their attention, acknowledge their uncertainties and guiding them to interpretations.
    - e.g.:
      - [IT-Eng]: **It is reported that** 16% of American households use a cell phone exclusively for their telephone service.
      - [IT-Eng]: **It is clear that** our Government is not working for us.
- IT-Sta
  - Stance: The ways writers stamp their personal opinion onto their arguments:
    - e.g.:
      - [IT-Sta]: **It seems that** the Republicans have already blinked in the most feared immediate confrontation
      - [IT-Sta]: **It may be that** appropriate prosecutions would result in managers and frontline staff serving well-deserved prison sentences.

**Figure C.7** English tag-set definition for Interpersonal Themes

**SPANISH Extended TAGSET DEFINITIONS for Interpersonal Theme Types (definitions based on Halliday and Matthiessen (2004) *IFG* and on an adaptation of Hyland's (2005) *Stance and engagement: a model of interaction in academic discourse*.**

**Interpersonal Themes (IT):**

These are elements which express the attitude and the evaluation of the speaker with respect to his/her message:

- IT-Voc
  - Vocatives, i.e., any item used to address
    - e.g.:
      - [IT-Voc] **¡Juan!** Esta es una gran sorpresa.
      - [IT-Voc] **Señor,** ¿puede acompañarme, por favor?
- IT-Com
  - Comment adjuncts
    - e.g.:
      - [IT-Com] **Sorprendentemente** no mencionó nada al respecto.
      - [IT-Com] **Comprensiblemente,** mantuvo un perfil bajo.
- IT-Mod
  - Modal adjuncts
    - e.g.:
      - [IT-Mod] **Probablemente** esta ha sido la única lección que hemos aprendido
      - [IT-Mod] **¡Seguro** que no has hecho eso!
- IT-Eng
  - Engagement: The writer acknowledge and connect to others, recognizing the presence of the reader, focusing their attention, acknowledge their uncertainties and guiding them to interpretations.
    - e.g.:
      - [IT-Eng]: **Se dice que** algo tan insignificante como el aleteo de una mariposa puede ocasionar un huracán en el otro extremo del mundo
      - [IT-Eng]: **Está claro que** nuestro Gobierno ya no trabaja para nosotros.
      - [IT-Eng]: **Lo importante en realidad es** parar los desahucios
- IT-Sta
  - Stance: The ways writers stamp their personal opinion onto their arguments:
    - e.g.:
      - [IT-Sta]: **Parece que** hemos olvidado que su negocio siempre ha sido pagar poco por el dinero que les confiamos y cobrar mucho por el dinero que les pedimos prestado.
      - [IT-Sta]: **Opino que** si se quiere homenajear a alguien por su valía, no hace falta otorgarle un dinero que hubiera podido servir para cubrir la subsistencia de mucha gente que tiene que mendigar.

**Figure C.8** Spanish tag-set definition for Interpersonal Themes

**ENGLISH Extended TAGSET DEFINITIONS for Pre-Head Theme Types (definitions based on Lavid, Arús and Moratón (2013) *Investigating thematic choices in two newspaper genres: an SFL-based study analysis***

**Pre-Head Themes (PH):**

The Pre-Head element is any circumstantial and/or finite element preceding the Thematic Head.

- PH-Circ
  - Circumstantial elements that can be realized by Adverbial or Prepositional Groups:
    - e.g:
      - Adverbial groups: [Ph-Circ:] **Afterwards** there will be another meeting
      - Prepositional phrases: [Ph-Circ:] **On your right** you can see the Royal Palace
- PH-CCL
  - Circumstantial elements realized by a dependent clause:
    - e.g:
      - [Ph-CCL:] **After dropping her off**, he continued his trip
- PH-Finite
  - As auxiliaries preceding the lexical verb:
    - e.g:
      - [Ph-Finite:] **Should** you decide to leave the country, please let me know.
      - [Ph-Finite:] **Had** I known you were so near, I would have flown to meet you.

**Figure C.9** English tag-set definition for PreHeads

**SPANISH Extended TAGSET DEFINITIONS for Pre-Head Theme Types (definitions based on Lavid, Arús and Moratón (2013) *Investigating thematic choices in two newspaper genres: an SFL-based study analysis***

**Pre-Head Themes (PH):**

The Pre-Head element is any circumstantial and/or finite element preceding the Thematic Head.

- PH-Circ
  - Circumstantial elements that can be realized by Adverbial or Prepositional Groups:
    - e.g.:
      - Adverbial groups: [Ph-Circ:] **Después** habrá otra reunion
      - Prepositional phrases: [Ph-Circ:] **A su derecha**, pueden ver el Palacio Real.
- PH-CCL
  - Circumstantial elements realized by a dependent clause:
    - e.g.:
      - [Ph-CCL:] **Tras dejarle en casa**, Mario continuó su camino.
- PH-Finite
  - As auxiliaries preceding the lexical verb:
    - e.g.:
      - [Ph-Finite:] **Han** viajado también a Estambul y Jerusalén.
  - As auxiliaries preceding the lexical verb:
    - e.g.:
      - [Ph-Finite:] **Agrav**-a la situación el hecho de que llueva.
  - ‘Se’ markers are considered Pre-Head when pronominal or passive:
    - e.g.:
      - [Ph-Finite:] **Se hall**-a ahora en un proceso en un proceso de expansion internacional.

**Figure C.10** Spanish tag-set definition for PreHeads



**ENGLISH Extended TAGSET DEFINITIONS for Thematic Head Types (definitions based on Halliday and Matthiessen (2004) *IFG* and Martin, Matthiessen and Painter (1997) *Working with Functional Grammar*)**

**All examples include the defined participant in thematic position**

1. **Actor:** the participant doing the deed in a material processes, as in [Actor:] Peter went home, [Actor:] Mary received the letter or [Actor:] John gave Mary a kiss
2. **Goal:** the participant impacted by a doing in a material process, as in [Goal:] Mary was kissed by Peter, [Goal:] the letter was put in the mail or [Goal:] the bathrooms are cleaned hourly
3. **Beneficiary:** the participant benefiting (positively or negatively) from the doing in a material process, as in [Beneficiary:] Mary was given a letter, [Beneficiary:] he was granted a scholarship or [Beneficiary:] they were inflicted a crushing defeat
4. **Range** (or Scope): the participant that construes the domain over which the process takes places, as in [Range:] that mountain is climbed mostly on its northern side, or construes the process itself, either in general or in specific terms, as in [Range:] showers should be taken in the morning.
5. **Senser:** the participant sensing in a mental process, as in [Senser:] She likes ice-cream, [Senser:] I can't see the light, [Senser:] she knows a lot of stories, [Senser:] they prefer to stay
6. **Phenomenon:** the participant being sensed in a mental process, as in [Phenomenon:] he is hated everywhere, [Phenomenon:] deer can be seen crossing the fields, [Phenomenon:] that's well known by everybody or [Phenomenon:] that ring is very much coveted
7. **Sayer:** the participant saying, telling, stating, informing, asking, threatening, suggesting and so on in a verbal process, as in [Sayer:] She never tells the truth, [Sayer:] they ordered me to leave or [Sayer:] She threatened to kill herself
8. **Verbiage:** the content of saying in a verbal clause, when expressed as a nominal group, as in [Verbiage:] that story has been told many times, [Verbiage:] questions will be asked or [Verbiage:] that word was never uttered by me
9. **Receiver:** the addressee of a speech interaction in a verbal process, as in [Receiver:] I was told to leave at once, [Receiver:] the kids were told a story or [Receiver:] she was asked her name
10. **Token:** the participant representing the expression, symbol, form, name, function, position or actor in an identifying relational process. Identifying relational processes are reversible and the Token tends to appear in the first position with respect to the Value, as in: [Token:] Mary is the best, [Token:] green means "go" or [Token:] She played the leading role. The Token is also the participant that tends to go first in possessive and circumstantial identifying relational processes, as [Token:] they own the house or [Token:] tomorrow is January the 1st, respectively.
11. **Value:** the participant representing the content, symbolized thing, meaning, referent, filler, holder of position or role in an identifying relational process. Identifying relational processes are reversible and the Value appears in initial position when the process is reversed, as in [Value:] the best one is Mary, [Value:] "go" is symbolized by green or [Value:] the leading role was played by her. The Value is also the participant that goes first in possessive and circumstantial identifying relational processes when these are reversed, as in [Value:] the house is owned by them or [Value:] January the 1<sup>st</sup> is tomorrow, respectively.
12. **Carrier:** the participant to which an Attribute is assigned in an attributive relational process, whether intensive, possessive or circumstantial. These relational processes are not easily reversed. Examples: She is quite wise in general, [Carrier:] I have a guitar or [Carrier:] the movie is about a multimillionaire
13. **Attribute:** what is assigned to the Carrier in an attributive relational process, whether intensive, possessive or circumstantial. As attributive processes are not easily reversed, Attributes are not found in thematic position except in exclamations such as How [Attribute:] clever she is!
14. **"There":** The starting element in an existential process. It is not a participant. Examples: There is a hair in my soup or there are many people here
15. **Process:** a whole process, whether material, mental, verbal, relational or existential.
16. **None**

**Figure C.11** English tag-set definition for Head experiential roles

## C. Appendix

---

### SPANISH Extended TAGSET DEFINITIONS for Thematic Head Types (definitions based on Halliday and Matthiessen (2004) *IFG* and Martin, Matthiessen and Painter (1997) *Working with Functional Grammar*)

All examples include the defined participant in thematic position.

1. **Actor**: the participant doing the deed in a material processes, as in: [Actor:] El niño tiene anginas ('the kid has tonsillitis') or [Actor:] Los tres recibieron un premio ('the three of them were given [lit. received] an award')
2. **Goal**: the participant impacted by a doing in a material process, as in: [Goal:] El premio lo entregó el presidente de la academia ('the award was delivered by the president of the academy') or [Goal:] las paredes han de pintarse cada dos años ('the walls are to be painted every other year')
3. **Beneficiary**: the participant benefiting (positively or negatively) from the doing in a material process, as in: [Beneficiary:] A mí me dieron tres entradas ('I was given three tickets') or [Beneficiary:] me han concedido una beca ('I've been granted a scholarship')
4. **Range** (or Scope): the participant that construes the domain over which the process takes places, as in: [Range:] Esa pared fue escalada por primera vez en 1915 ('that wall was first climbed in 1915') or or construes the process itself, either in general or in specific terms, as in: [Range:] Esta música es tocada en funciones sociales ('this music is played in social events').
5. **Senser**: the participant sensing in a mental process, as in: [Senser:] A mi hermano le encanta este osito ('my brother loves this teddy bear') or [Senser:] los daltónicos no pueden percibir algunos colores (color blind people cannot make out some colors')
6. **Phenomenon**: the participant being sensed in a mental process, as in: [Phenomenon:] estas galletas no me gustan mucho ('I don't really like these cookies')
7. **Sayer**: the participant saying, telling, stating, informing, asking, threatening, suggesting and so on in a verbal process, as in: [Sayer:] Los borrachos siempre dicen la verdad ('drunken people always tell the truth') or [Sayer:] El juez ordenó evacuar la sala ('the judge ordered to create the room')
8. **Verbiage**: the content of saying in a verbal clause, when expressed as a nominal group, as in: [Verbiage:] Esa historia ya me la han contado ('I've already been told that story') or [Verbiage:] Eso no me lo dices otra vez (you're not telling me that again')
9. **Receiver**: the addressee of a speech interaction in a verbal process, as in: [Receiver:] Me dijeron que me olvidara de ello ('I was told to forget about that') or [Receiver:] Siempre nos cuentas historias (you're always telling us stories')
10. **Token**: the participant representing the expression, symbol, form, name, function, position or actor in an identifying relational process. Identifying relational processes are reversible and the Token tends to appear in the first position with respect to the Value, as in: [Token:] Juan es el responsable ('John is the responsible one [the person in charge]'). The Token is also the participant that tends to go first in possessive and circumstantial identifying relational processes, as: [Token:] alguien debe tener el dinero ('someone must have the money')
11. **Value**: the participant representing the content, symbolized thing, meaning, referent, filler, holder of position or role in an identifying relational process. Identifying relational processes are reversible and the Value appears in initial position when the process is reversed, as in: [Value:] el responsable es Juan ('the responsible one [the person in charge] is John'). The Value is also the participant that goes first in possessive and circumstantial identifying relational processes when these are reversed, as in: [Value:] el dinero lo debe tener alguien (lit. 'the money, someone must have it')
12. **Carrier**: the participant to which an Attribute is assigned in an attributive relational process, whether intensive, possessive or circumstantial. These relational processes are not easily reversed. Examples: [Carrier:] este niño es bastante bueno, en general ('this kid is quite good, in general') or [Carrier:] mi madre tiene dos guitarras ('my mother has two guitars')
13. **Attribute**: what is assigned to the Carrier in an attributive relational process, whether intensive, possessive or circumstantial. As attributive processes are not easily reversed, Attributes are not found in thematic position except in exclamations such as: ¡Qué [TH-Attribute:] inteligente es! ('How intelligent she is'). In other constructions where the Attribute appears in thematic position is it of the Absolute theme type, and in that case it does not fulfill a transitivity role: [TH-absolute:] Inteligente, no lo es mucho ('intelligent, she's not really')
14. **"Hay", "había", "habrá", "hubo", etc.**: The starting element in an existential process. It is variable with respect to tense but invariable with respect to person. The final inflection does therefore not serve to track participants, so the whole unit is Theme. Examples: [Existential:] Hay un pelo en mi sopa ('there is a hair in my soup')
15. **Process**: the inflected ending in a process; in perfect tenses, it is the inflected part of the Finite. This is a very usual kind of Theme in Spanish, the lexical part being Pre-Head: [Process:] Llegam-os enseguida ('[we] arrived quickly') or [Process:] H-an encontrado petróleo ('they've found oil')

Figure C.12 Spanish tag-set definition for Head experiential roles



## C.2 Annotation Experiments

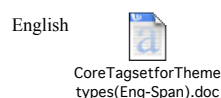
### AGREEMENT STUDY 1 TASK 1

Annotator's name:

The purpose of this exercise is to identify the clause span that functions as Thematic Field within the clause complex in English. The rest of the clause span will be the Rhematic Field.

This analysis is carried out on English Commentaries and News Reports.

The definition of Thematic field and its realisations can be found clicking on the icon below.



There are six texts, the titles are listed below. To access the corresponding template for your annotation, please click on the link in blue.

#### TEXT 1

##### The Bigger Issue in Sudan

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in the right column the span that functions as Thematic field:  
[Thematic Field \(click here to start with the annotation\)](#)

#### TEXT 2

##### The Vanishing Bomb

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in each clause the span that functions as:  
[Thematic Field \(click here to start with the annotation\)](#)

#### TEXT 3

##### The Limits of Energy Innovation

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in each clause the span that functions as:  
[Thematic Field \(click here to start with the annotation\)](#)

#### TEXT 4

##### Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair - The Times online October 26, 2008

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in each clause the span that functions as:  
[Thematic Field \(click here to start with the annotation\)](#)

#### TEXT 5

##### Opec hawks want to cut oil production to keep up price

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in each clause the span that functions as:  
[Thematic Field \(click here to start with the annotation\)](#)

#### TEXT 6

##### Microsoft results offer hope to tech sector

Please annotate as follows:

1. Read the clause in each cell.
2. Underline and copy in each clause the span that functions as:  
[Thematic Field \(click here to start with the annotation\)](#)

Figure C.13 Instructions of Agreement Study 1 Task 1

## TEXT 4

**Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair** - The Times online October 26, 2008  
(English News Report)

#	Text	
1	Dominique Strauss-Kam, the French head of the International Monetary Fund, escaped dismissal for a one-night stand with a subordinate today, but was denounced by board members for a "serious error of judgment."	Dominique Strauss-Kam, the French head of the International Monetary Fund,
2	The fund's managing director was allowed to remain in his post after an investigation cleared him of accusations of harassment, favouritism and abuse of power over his affair with Piroska Nagy, a Hungarian economist.	
3	But his reputation took a battering in the statement by the IMF board, which said the relationship "was regrettable and reflected a serious error of judgment on the part of the managing director."	
4	Shakour Shaalan, the fund's executive director, said IMF staff, and notably female staff, "are not at all happy" with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet <i>le grand séducteur</i> .	
5	"The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said Mr Shaalan. "We will continue to work with him."	
6	Mr Strauss-Kam, 59, said in a statement: "I very much regret the incident and I accept responsibility for it."	
7	The investigation was launched when Mario Blejer, a senior Argentine economist, said Mrs Nagy, his wife, had been seduced at the Davos international forum.	
8	A US law firm was asked to probe accusations that Mrs Nagy, who left the IMF in August, had been forced out by Mr Strauss-Kahn, possibly with an exorbitant pay-off.	
9	Morgan, Lewis & Bockius said there was no evidence to support the claims.	
10	The decision was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remodel and regulate capitalism in the wake of the financial crisis.	
11	French commentators hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is now behind us. We love each other as much as we did on the first day."	

[Back to index](#)

Figure C.14 Annotation sheet of Agreement Study 1 Task 1

## C. Appendix

### Text 5 Opec hawks want to cut oil production to keep up price

Instructions:

1. Read the clause in each cell.
2. Underline in the left cell and copy in the right cell the clause span that functions as **PreHead (PH)**

#	Text	
1	Opec's mettle will be tested next week when the cartel assembles in Vienna at a hastily convened meeting to discuss the economic crisis and whether the exporters' club should take action to arrest the fall in the price of crude oil.	
2	Mounting alarm over the dwindling rent from oil sales prompted Opec yesterday to bring forward an emergency meeting scheduled for mid-November to next Friday.	
3	Together the cartel members have lost about a billion dollars of revenue a day since crude price peaked in July.	
4	The benchmark US Light crude price fell \$5 per barrel on the New York Nymex exchange yesterday to \$69.15 (£40.05), less than half the value at oil's peak, while London's Brent crude futures contract fell to \$65.70.	
5	Evidence of rising oil stocks in America and weakening demand for energy in China's mills and workshops has speeded up oil's decline, but there is no agreement within the cartel over price strategy.	
6	Opec hawks, such as Iran and Venezuela, want to shore up the price by turning off the taps, but neither country can afford to cut output and forgo revenues.	
7	Both are failing even to produce at their agreed quotas because of weak investment in their oil industries, so the burden of any cut will fall on Kuwait and Saudi Arabia.	
8	The Saudis would not like to see the price falling below \$60 per barrel, a level at which most economists reckon the kingdom can support its profligate princes, the social security bills and invest in the oil industry.	
9	However, the Saudis will be listening to the grinding cogs of the slowing Asian industrial juggernaut with mounting concern.	
10	A deep recession might precipitate an oil price collapse	
11	It is Saudi Aramco that has been responsible for the weakening oil price, pushing more oil on the market since the energy summit in Jedda.	
12	According to Platts, the oil pricing publisher, even as the price tumbles Saudi Aramco has supplied every refiner's request to lift Saudi crude next month.	
13	The wider question is what impact falling crude will have on nonOpec oil output	
14	Russia's budget depends on a \$70 per barrel price, according to Aleksei Kudrin, the Finance Minister	
15	But Russian oil output is already falling because of weak investment, as is Mexican production	
16	The major Western oil companies can probably stomach prices dipping temporarily to \$60 for a few months and still pay their dividends, but a prolonged slump at \$50 per barrel would lead to the scrapping of investments followed by another cycle of shortages and soaring prices	

**Figure C.15** Annotation sheet of Agreement Study 1 Task 2

## AGREEMENT STUDY 2 TASK 1

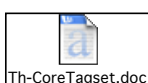
Annotator's name:

The purpose of this exercise is to classify the underlined span in each clause (simple or complex) as one of the following types: Thematic Head, PreHead, Textual Theme and Interpersonal Theme. If you think the underlined span belongs to none of these types, classify the span as "none".

This analysis is carried out in English commentaries and News Reports.

The definitions of each category (Thematic Head, PreHead...) can be found clicking on the icon below.

ENGLISH



These 5 categories are displayed in a drop down list on the right cell. As in the image below

Text	
<p><u>and Russian presidents to renew strategic arms reductions</u> has revived war arms.</p> <p>ated: nuclear weapons may come into the possession of states that might use s –creating new threats of unimaginable proportion.</p> <p>o, <u>the elimination of nuclear arms</u> is no longer the idea only of populists pted by professionals – politicians known for their realism and academics ty.</p> <p>i = <u>which served the goal of deterrence during the Cold War, when the</u> <u>ing blocs</u> – answered the needs and risks of the time. Security rested on a</p>	<div> Thematic Head PreHead Interpersonal Theme Textual Theme Thematic Equative Predicated Theme Absolute Theme </div>

### TEXT 1

#### The Bigger Issue in Sudan

Please annotate as follows:

1. Read the clause in each cell and inspect the underlined thematic span.
2. Classify the underlined thematic span by writing its type in the right column. Possible types are: Thematic Head (TH), PreHead (PH), Textual Theme (TT), Interpersonal Theme (IT). If you think the span is none of these types, classify the span as "none".

[Labelling \(click here to start with the annotation\)](#)

### TEXT 2

#### The Vanishing Bomb

Please annotate as follows:

1. Read the clause in each cell and inspect the underlined thematic span.
2. Classify the underlined thematic span by writing its type in the right column. Possible types are: Thematic Head (TH), PreHead (PH), Textual Theme (TT), Interpersonal Theme (IT). If you think the span is none of

**Figure C.16** Instruction sheet of Agreement Study 2 Task 1

## C. Appendix

### TEXT 4

**Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair** - The Times online October 26, 2008

(English News Report)

#	Text	
1	Dominique Strauss-Kam, the French head of the International Monetary Fund, escaped dismissal for a one-night stand with a subordinate today, but was denounced by board members for a "serious error of judgment."	
2	The fund's managing director was allowed to remain in his post after an investigation cleared him of accusations of harassment, favouritism and abuse of power over his affair with Piroska Nagy, a Hungarian economist.	
3	But his reputation took a battering in the statement by the IM F board, which said the relationship "was regrettable and reflected a serious error of judgment on the part of the managing director."	
4	Shakour Shaalan, the fund's executive director, said IM F staff, and notably female staff, "are not at all happy" with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet le grand seducteur.	
5	"The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said Mr Shaalan. "We will continue to work with him."	
6	Mr Strauss-Kam, 59, said in a statement: "I very much regret the incident and I accept responsibility for it."	
7	The investigation was launched when Mario Blejer, a senior Argentine economist, said Mrs Nagy, his wife, had been seduced at the Davos international forum.	
8	A US law firm was asked to probe accusations that Mrs Nagy, who left the IMF in August, had been forced out by Mr Strauss-Kahn, possibly with an exorbitant pay-off.	
9	Morgan, Lewis & Bockius said there was no evidence to support the claims.	
10	The decision was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remodel and regulate capitalism in the wake of the financial crisis.	
11	French commentators hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is now behind us. We love each other as much as we did on the first day."	

[Back to index](#)

**Figure C.17** Annotation sheet of Agreement Study 2 Task 1



Annotation Exercise 2: Labelling


Annotator's name:

The purpose of this exercise is to classify the Thematic Head (underlined in each clause complex) as one of the following types : Actor, Goal, Beneficiary, Senser, Phenomenon, Sayer, Verbiage, Receiver, Token, Value, Carrier, Attribute, Process, "There" . If you think the Thematic Head belongs to none of these types, classify it as "none".

This analysis is carried out in English commentaries and News Reports.

The definitions of each category can be found clicking on the icon below.

<u>International Monetary Fund</u> , escaped dismissal for a one-night ed by board members for a "serious error of judgment." nain in his post after an investigation cleared him of accusations ver his affair with Pirooska Nagy, a Hungarian economist. nt by the IMF board, which said the relationship "was ient on the part of the managing director." aid IMF staff, and notably female staff, "are not at all happy" es in France have earned him the epithet <u>le grand seducteur</u> . s. I don't think that we can ask him to do more at this time," said ." much regret the incident and I accept responsibility for it." jer, a senior Argentine economist, said Mrs Nagy, his wife, had at Mrs Nagy, who left the IMF in August, had been forced out pay-off. dence to support the claims. where Mr Strauss-Kahn is seen as a key figure in President m in the wake of the financial crisis. Anne Sinclair, Mr Strauss-Kahn's wife, a former television	<div>▼</div> <div>Actor</div> <div>Goal</div> <div>Beneficiary</div> <div>Range</div> <div>Senser</div> <div>Phenomenon</div> <div>Sayer</div> <div>Verbiage</div> <div>Receiver</div> <div>Token</div> <div>Value</div> <div>Carrier</div> <div>Attribute</div> <div>There</div> <div>None</div>
---	---

ENGLISH  Document

TEXT 1

The Bigger Issue in Sudan

Please annotate as follows:

1. Read the clause in each cell and inspect the underlined Thematic Head
2. Classify the Thematic Head by writing its type in the right column. Possible types are specified at the beginning of this exercise with definitions in the attached file.

[Labelling \(click here to start with the annotation\)](#)

TEXT 2

Figure C.18 Instruction sheet of Agreement Study 2 Task 2

## C. Appendix

### TEXT 4

**Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair** - The Times online October 26, 2008  
(English News Report)

#	Text	
1	Dominique Strauss-Kam, the French head of the International Monetary Fund, escaped dismissal for a one-night stand with a subordinate today, but was denounced by board members for a "serious error of judgment."	
2	The fund's managing director was allowed to remain in his post after an investigation cleared him of accusations of harassment, favouritism and abuse of power over his affair with Piroska Nagy, a Hungarian economist.	
3	But his reputation took a battering in the statement by the IM F board, which said the relationship "was regrettable and reflected a serious error of judgment on the part of the managing director."	
4	Shakour Shaalan, the fund's executive director, said IM F staff, and notably female staff, "are not at all happy" with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet le grand seducteur.	
5	"The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said Mr Shaalan. "We will continue to work with him."	
6	Mr Strauss-Kam, 59, said in a statement: "I very much regret the incident and I accept responsibility for it."	
7	The investigation was launched when Mario Blejer, a senior Argentine economist, said Mrs Nagy, his wife, had been seduced at the Davos international forum.	
8	A US law firm was asked to probe accusations that Mrs Nagy, who left the IMF in August, had been forced out by Mr Strauss-Kahn, possibly with an exorbitant pay-off.	
9	Morgan, Lewis & Bockius said there was no evidence to support the claims.	
10	The decision was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remodel and regulate capitalism in the wake of the financial crisis.	
11	French commentators hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is now behind us. We love each other as much as we did on the first day."	

[Back to index](#)

**Figure C.19** Annotation sheet of Agreement Study 2 Task 2

### Annotation Exercise 3: Labelling

Annotator's name:

The purpose of this exercise is to classify the Textual Themes (underlined in each clause complex) as one of the following types : Binder, Linker or Correlative . If you think the Textual Theme belongs to none of these types, classify it as "none".

This analysis is carried out in a number of clauses complexes in English selected from Editorials, News Reports and Letters to the Editor

The definitions of each category (Binder, Linker or Correlative) can be found clicking on the icon below.



The annotator selects the correct label from the drop down list on the right cell.

<p>;, sooner rather than later, as the only outside power sufficiently trusted by both sides to</p> <p>tatements of respective party positions.</p> <p>rievously damaged by their part in the sub-prime mortgage fiasco, he made the</p> <p>gy, and his credibility.</p> <p>a risk to the economy, he added that "most of the exchange rate movements that we have</p> <p>es meant to boost the economy.</p> <p>ie question of currency intervention, Shoichi Nakagawa, Japan's Finance Minister, said</p> <p>e rate and that he was he "watching with great interest."</p> <p>ned at wealthy individuals, it comes at a time when evidence of minimal UK corporation</p> <p>on has catered to the top of the political agenda</p>	<div> TT-Bind  TT-Link  TT-Cor  None </div>
--	---

Figure C.20 Instruction sheet of Agreement Study 2 Task 3

## C. Appendix

### Textual Theme Types

#	Text	
1	<b>But</b> his reputation took a battering in the statement by the IMF board, which said the relationship "was regrettable and reflected a serious error of judgment on the part of the managing director."	
2	<b>And</b> if there is, why did no one look at it?	
3	<b>And</b> this month Mr Dilnot warned politicians to avoid mixing up debt with the deficit after the Prime Minister claimed, erroneously, in a party political broadcast that the Coalition was "paying down Britain's debts".	
4	<b>As a result</b> it is not inconceivable that Moscow will emerge, sooner rather than later, as the only outside power sufficiently trusted by both sides to mediate between them.	
5	<b>As</b> regards America's fiscal problems, both were mere re-statements of respective party positions.	
6	<b>Even though</b> the reputations of the ratings agencies were grievously damaged by their part in the sub-prime mortgage fiasco, he made the maintenance of the triple-A rating a cornerstone of his strategy, and his credibility.	
7	<b>Although</b> Mr Draghi said that appreciation of the euro was a risk to the economy, he added that "most of the exchange rate movements that we have seen were [...] the result of domestic macro economic policies meant to boost the economy."	
8	<b>Although</b> the Bank of Japan has remained tight-lipped on the question of currency intervention, Shoichi Nakagawa, Japan's Finance Minister, said today that he saw "excessive volatility" in the yen's exchange rate and that he was he "watching with great interest."	
9	<b>Although</b> the MPs' report is largely focused on schemes aimed at wealthy individuals, it comes at a time when evidence of minimal UK corporation tax payments by multinationals such as Starbucks and Amazon has catapulted tax to the top of the political agenda.	
10	<b>And</b> forgivable though it might be for a prime minister to give the impression to his domestic audience that he has the power to do something and has everything under control, it is not only Britons who were caught up in the drama in the desert.	
11	<b>And</b> so it has proved.	
12	<b>And third</b> , former funders, sponsors and prize donors should use every means at their disposal, including the law, to recover the money that they clearly paid out under false pretences.	
13	<b>And</b> yesterday Mark Simmonds, Foreign Office minister for Africa, spoke of providing training and support for the Malian army.	
14	<b>As</b> Con Coughlin pointed out on these pages yesterday, Britain has only eight of these aircraft, and France none; America has roughly 200.	
15	<b>As far as</b> Ed Miliband, the Labour leader, was concerned yesterday, our relationship with the EU is fine and dandy.	
16	<b>As such</b> , the arrangement is best terminated.	
17	<b>As</b> we live longer, the pressure on the health service mounts.	
18	<b>But</b> analysts have said this scheme has made lenders less reliant on the need to attract savers' deposits, meaning that savings rates have plummeted.	
19	<b>But</b> bullish forecasts from the Bundesbank over an immediate return to growth for Germany, after a worse-than-expected 0.6 per cent slide between October and December, put sterling on the back foot again yesterday as it sank close to 15-month lows against the euro	
20	<b>But</b> Russian oil output is already falling because of weak investment, as is Mexican production	
21	<b>But</b> , he flagged up the Government's FirstBuy scheme to help people get on the housing ladder, which will provide further support to first time buyers.	
22	<b>Even if</b> £30 million was the correct figure, it would seem a small price to have paid to help the electorate exercise its democratic right on a substantially more informed basis than was the case.	
23	<b>Even if</b> Lance Armstrong had come across as the most contrite of penitents – which he did not – an interview with Oprah Winfrey on American TV is not sufficient to allow the cyclist to draw a line under his disgrace and "move on".	
24	<b>Finally</b> , AQIM and its allies in the region would like nothing more than to see a major deployment of Western troops, who would become vulnerable targets for attack on unfamiliar terrain.	
25	<b>Finally</b> , quite apart from the economic arguments, design and the creative arts are simply good for Britain.	
26	<b>First</b> , he must frame his policy as an attempt to save our relationship with Europe.	

Figure C.21 Annotation sheet of Agreement Study 2 Task 3

Annotation Exercise 04: Labelling types of Interpersonal Themes

Annotator's name:

The purpose of this exercise is to classify the Interpersonal Themes (underlined in each clause complex) as one of the following types: Vocatives, Comment Adjuncts, Modal Adjuncts, Engagement and Stance. If you think the Interpersonal Theme belongs to none of these types, classify it as "none".

This analysis is carried out in a number of clauses complexes in English selected form Editorials, News Reports and Letters to the Editor

The definitions of each category (Vocative, Comment adjunct, Modal adjunct, Engagement, Stance) can be found clicking on the icon below.



The annotator selects the correct label from the drop down list on the right cell.

<u>and Balls were happy to sit</u> for abolishing the 10p tax rate.	
literally, in that the direct responsibility for horse labelled as beef	
ency was set up in 2002, national governments have been deprived of	
gency action to provide their own assurances to their customers.	

- IT-Voc
- IT-Com
- IT-Mod
- IT-Eng
- IT-Sta
- None

Figure C.22 Instruction sheet of Agreement Study 2 Task 4

## C. Appendix

### Interpersonal Theme Types

#	Text	
1	In fact, <b>it seems that</b> the Republicans have already blinked in the most feared immediate confrontation, over a renewal of the US federal debt ceiling, without which the country might have defaulted.	
2	<b>It is therefore understandable that</b> the Obama administration is worried about the possibility of a referendum on the UK's continued membership of the 27-nation group.	
3	So <b>it is a shame</b> we cannot say the same about our transport network or our schools.	
4	Under Article 8, <b>it is clear that</b> the right to a family life is not absolute but qualified.	
5	<b>True</b> , he criticised Gordon Brown (in whose Cabinet he and Ed Balls were happy to sit) for abolishing the 10p tax rate.	
6	<b>It may be that</b> respondents answered the question somewhat literally, in that the direct responsibility for horse labelled as beef ending up on supermarket shelves lies with the companies.	
7	<b>It has been claimed that</b> , since the European Food Safety Agency was set up in 2002, national governments have been deprived of the power to regulate food testing.	
8	<b>Of course</b> , food suppliers should not make false claims.	
9	And <b>it is interesting that</b> supermarkets have had to take emergency action to provide their own assurances to their customers.	
10	<b>True</b> , the GOP has retained control of the House of Representatives, but largely thanks to its diligent gerrymandering of Congressional districts; in the overall House vote, the Democrats won half a million more.	
11	<b>Wisely</b> , Mr Obama made no attempt on Tuesday to lay down specific guidelines to Congress, where a bipartisan consensus is already taking shape (thanks, in large part, to the Republicans' need to boost the party's dismal standing among Hispanic voters).	
12	But <b>it is a pity that</b> fear of libel suits – or was it a last sliver of face-saving? – forced the redaction of the pithier remarks from Jeremy Paxman and others.	
13	<b>For goodness sake</b> let's have elections combined – if you go out to elect a councillor or MP you may as well pick a police commissioner too.	
14	<b>Perhaps</b> James Dyson was a little harsh towards poets; poetry, literature, music and paintings all feed our spirit.	
15	Hamas, <b>whether we like it or not</b> , shows no sign of weakening.	
16	<b>Let us hope that</b> , for once, lessons really have been learnt by officialdom.	
17	<b>Surely</b> canon law does not transcend the law of the land.	
18	<b>It is humiliating that</b> money-grubbing of this kind is allowed at a national level.	
19	<b>It is wrong, and unjustifiably costly</b> , to have benefits that everybody can claim.	
20	<b>Unfortunately</b> for the fox, the majority do not understand the countryside and are too easily influenced by television programmes out to prove a point.	

Figure C.23 Annotation sheet of Agreement Study 2 Task 4

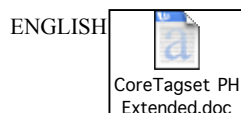
## Annotation Exercise 05: Labelling types of PreHead

Annotator's name:

The purpose of this exercise is to classify the Pre-Head Themes (underlined in each clause complex) as one of the following types: Circumstantial groups, Circumstantial clauses or Finite verbal forms . If you think the Pre-Head Theme belongs to none of these types, classify it as "none".

This analysis is carried out in a number of clauses complexes in English selected form Editorials, News Reports and Letters to the Editor

The definitions of each category (Circumstantial groups, Circumstantial clauses or Finite verbal forms) can be found clicking on the icon below.



The annotator selects the correct label from the drop down list on the right cell.

sale to residents for £1 in the Granby 'Four Streets' and Arnside Road in	
<u>nd petrol making life harder for savers</u> , rates on savings accounts have been	
struggle to make any real returns.	
<u>a long periods of stability interspersed by sudden and dramatic revaluations or</u>	
<u>than markets as most developed countries have already lost their AAA,"</u> he	
e and the pound falls." she said.	

PH-Circ  
PH-CCL  
PH-Finite  
None

Figure C.24 Instruction sheet of Agreement Study 2 Task 5

## C. Appendix

### Pre-Head Types

#	Text	
1	<u>For the full-year ending June 2009</u> , Microsoft forecasts earnings per share to range between \$2.00 to \$2.10 on revenue from \$64.9 billion to \$66.4 billion.	
2	<u>In a conference call with investors</u> , Mr Liddell said Microsoft will "tailor our business to whatever the economy brings."	
3	<u>"The market environment has changed over the last two weeks and the expectations for capital levels have changed following massive capital injections in financial institutions worldwide,"</u> he said.	
4	<u>In an interview published over the weekend</u> , Santander UK's chief executive Ana Botin said: "Right now, our thinking is that organic growth is what we are aiming for.	
5	<u>Earlier on Monday</u> , Ewald Nowotny, another ECB policymaker, said that recent currency movements were not reason to cause alarm.	
6	<u>Now</u> , the council proposes to sell off a cluster of houses for just £1 to residents - on the condition they will be brought up to a decent standard.	
7	<u>As part of the initial pilot scheme</u> , 20 houses will be offered for sale to residents for £1 in the Granby 'Four Streets' and Arnside Road in Kensington.	
8	As well as <u>rising living costs such as food, energy bills, rents and petrol making life harder for savers</u> , rates on savings accounts have been plummeting in recent months, meaning they face an even tougher struggle to make any real returns.	
9	<u>"[Sterling's] performance over the past twenty years has seen long periods of stability interspersed by sudden and dramatic revaluations or devaluations. Rarely have the trends been gentle,"</u> he said.	
10	<u>"I'd say the impact will be more on domestic political debate than markets as most developed countries have already lost their AAA,"</u> he said.	
11	<u>"Pension funds may benefit from loss of AAA if gilt yields rise and the pound falls,"</u> she said.	
12	<u>As David Cameron spelt out in his Commons statement today</u> , the complex is far out in the desert, 18 hours by road from Algiers, and it had been overrun in what appears to have been a large-scale and carefully planned operation by Mokhtar Belmokhtar, an Algerian who may have ties to al-Qa'ida.	
13	<u>By 2032</u> , the over-75s are expected to comprise 22 per cent of the population, up from 17 per cent now, so we need to act fast.	
14	<u>Between the two inaugurations of Barack Obama</u> , expectation and reality have been turned on their head.	
15	<u>With the secular Tuaregs then swiftly pushed aside by their former allies</u> , extremist militants took control of a swathe of territory as big as France.	
16	<u>Last week</u> , everything changed.	
17	<u>Given such dangers</u> , suggestions that Mr Hollande's first foray into foreign policy is simply to distract attention from his woeful poll ratings are hardly credible.	
18	<u>From September</u> , schools in England and Wales will abandon the existing staff salary structure that effectively guarantees an automatic annual pay rise for all teachers, regardless of merit.	
19	<u>In Libya</u> , the Islamists of Ansar al-Sharia played a key role in the attack on the American consulate in Benghazi, which killed the US Ambassador to Libya.	
20	<u>Due to our relatively robust economy</u> , the large communities of Romanians and Bulgarians already here, the far higher wages on offer and the ease with which new arrivals can access public services and benefits, Britain is a natural destination for such migrants.	
21	<u>Having started the week by offering to provide France with two C-17 transport planes to assist with its intervention in Mali</u> , Mr Cameron ended it anxiously trying to discover the fate of the British hostages taken captive after the attack by Islamist terrorists on BP's In Amenas gas field.	
22	<u>While Britain will be wary of further entanglements after the controversies of Afghanistan and Iraq</u> , there is no doubt that our Armed Forces will – as in Mali – be called upon to play a part in containing the Islamist threat, though the tangled alliances, shifting frontiers and guerrilla tactics that our forces will encounter could make it alarmingly hard to develop the necessary clarity about our aims, tactics and (perhaps most importantly) our exit strategy.	
23	<u>As Con Coughlin pointed out on these pages yesterday</u> , Britain has only eight of these aircraft, and France none; America has roughly 200.	
24	<u>If Mr Cameron is serious about confronting the threat posed by Islamist terrorists</u> , he should make sure our Armed Forces have the resources and manpower to deny them the safe havens they crave	

Figure C.25 Annotation sheet of Agreement Study 2 Task 5



# D

## Appendix

### D.1 Index of Texts

#### English News Reports

Text	Year	Title	Author	Source	Cl	W
Report 1	2008	Dominique Strauss-Kahn, head of the IMF, escapes dismissal over affair	Adam Sage	Times online	11	341
Report 2	2008	Opec hawks want to cut oil production to keep up price	Carl Mortished	Times online	16	422
Report 3	2008	Microsoft results offer hope to tech sector	Suzy Jagger and Mike Harvey	Times online	14	315
Report 4	2008	Dutch Government to inject 10bn of fresh capital into ING savings bank	Lilly Peel	Times online	15	347
Report 5	2008	G7 'preparing to drive down the yen'	Leo Lewis	Times online	14	566
Report 6	2008	Ukraine agrees terms of 16.5bn rescue by IMF	Lilly Peel	Times online	5	121

## D. Appendix

Report 7	2008	Barclays rejects government funding, secures 5.8bn from Qatar and Abu Dhabi	Grainne Gilmore	Times online	21	593
Report 8	2008	US economy officially on brink of recession	Gary Duncan	Times online	18	606
Report 9	2013	Narrow win for 'Merkel's Mac' in Lower Saxony	Tony Patterson	The independent	6	203
Report 10	2013	Georgian protesters want President Saakashvili out	Reuters	The independent	5	80
Report 11	2013	Top executive suppressed report damning Barclays	James Moore	The independent	16	429
Report 12	2013	Think-tank calls for new strategy to fire up the economy	Ben Chu	The independent	11	326
Report 13	2013	Santander eyes a double bank buy with 2bn tag	James Thompson	The independent	6	167
Report 14	2013	Tesco opens fifth dotcom store	James Thompson	The independent	6	153
Report 15	2013	UK firms to bid for Japan's nuclear clean-up	Mark Leftly	The independent	6	127
Report 16	2013	Hilco UK emerges among favourites to give HMV a spin	James Thompson	The independent	11	264
Report 17	2013	UK seems likely to lose EU battle over limits on bonuses	Nick Goodway	The independent	5	152
Report 18	2013	Cameron backs Agusta Westland despite Indian bribery scandal	Tom Bawden	The independent	8	209
Report 19	2013	Green hails LA Topshop while he eyes Dreams	BN	The independent	8	232
Report 20	2013	Pound takes another hit on German resurgence	Russell Lynch	The independent	14	419
Report 21	2013	Striking gold: Royal Mint moves back into Indian market	Russell Lynch	The independent	9	211
Report 22	2013	Mario Draghi says talk of currency war 'excessive'	Telegraph Staff and agencies	The telegraph	10	268
Report 23	2013	Lower income earners being squeezed hardest	Jessica Winch	The telegraph	17	424
Report 24	2013	Pound land: Derelict houses in Liverpool to be sold for just £1	Telegraph staff	The telegraph	14	415
Report 25	2013	MPs call for tax avoiders to be 'named and shamed'	Richard Blackden	The telegraph	11	335
Report 26	2013	A third of Britons fail to save	Agencies	The telegraph	19	555

Report 27	2013	Bovis and Persimmon post jump in profits	Rachel Coopers and agencies	The telegraph	14	408
Report 28	2013	City refuses to panic over the loss of UK's AAA credit rating	Russell Lynch	The independent	5	191
Report 29	2013	Stop ministers seeing official stats early – MPs	Ben Chu	The independent	14	373
Report 30	2013	Markets set for sterling sell-off after AAA setback for Britain	Ben Chu	The independent	20	515

## Spanish News Reports

Text	Year	Title	Author	Source	Cl	W
Report 1	2008	Una noche sin consecuencias para Strauss-Kahn	El País	El Pais on line	17	566
Report 2	2008	El recorte de producción de la OPEP hunde el petróleo	El País	El Pais on line	12	386
Report 3	2008	Los resultados de Microsoft brillan gracias al Windows Vista	EFE	El Pais on line	11	426
Report 4	2008	Holanda inyectará 10.000 millones en el grupo ING	Isabel Ferrer	El Pais on line	12	417
Report 5	2008	El G7 reafirma su interés en que el sistema financiero sea “fuerte y estable”	Agencias	El Pais on line	12	388
Report 6	2008	El FMI anuncia ayudas financieras para Ucrania y Hungría	Agencias	El Pais on line	12	305
Report 7	2008	Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown	EFE	El Pais on line	5	185
Report 8	2008	La crisis financiera lleva a Estados Unidos al borde de la recesión	Sandro Pozzi	El Pais on line	33	780
Report 9	2013	El Eurogrupo aprobará hoy 1.865 millones de euros para la banca española	Efe	El mundo	12	459
Report 10	2013	Inspectores sin fronteras	Raquel Villacéja	El mundo	21	523
Report 11	2013	La economía china crece un 7,8% en 2012, su menor subida en 13 años	Efe	El mundo	12	388

## D. Appendix

Report 12	2013	Banco Popular gana 133 millones con la venta de su negocio de recobro	Europa Press	El mundo	6	213
Report 13	2013	La Caixa anticipa una gradual recuperación' de la situación económica	Europa Press	El mundo	15	540
Report 14	2013	Lagarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída	María Ramirez	El mundo	24	607
Report 15	2013	La 'Troika' concederá a Grecia seis meses en los que no le reclamará más medidas	Efe	El mundo	9	293
Report 16	2013	BCE: la confianza de los mercados financieros ha aumentado notablemente	Efe	El mundo	9	267
Report 17	2013	Catalunya Banc y NCG sufrirán pérdidas de hasta 20.000 millones en 2012	Efe	La vanguardia	17	678
Report 18	2013	Las pérdidas de Unnim Banc casi se duplican en 2012	Efecom	La vanguardia	10	373
Report 19	2013	El banco francés Crédit Agricole multiplica por cuatro sus pérdidas	Efecom	La vanguardia	7	297
Report 20	2013	Hollande reconoce que Francia crecerá menos de lo previsto en 2013	Europapress	La vanguardia	11	292
Report 21	2013	Bolivia promete a España llegar a una "solución satisfactoria" para indemnizar a las empresas expropiadas	Europapress	La vanguardia	11	530
Report 22	2013	Alemania alcanza nuevo récord histórico de empleo en el último trimestre 2012	Efe	La vanguardia	8	317
Report 23	2013	Botín y Sáenz cobraron 3 y 8,2 millones en 2012, casi un 30% menos	Efecom	La vanguardia	10	373
Report 24	2013	Orizonia comunica a los trabajadores que se salvarán 800 empleos con la venta de parte de Orbest	Europa Press	La vanguardia	8	400
Report 25	2013	Moody's: la liquidez y financiación siguen siendo los retos de la banca española	Agencias	El pais	9	447
Report 26	2013	CaixaBank quiere suprimir 3.000 puestos de trabajo	Efe Europa Press	El mundo	12	364

Report 27	2013	Irlanda recorta los sueldos públicos a partir de 65.000 euros y congela el resto	Europa Press	El mundo	8	283
Report 28	2013	Villar-Mir analiza una posible operación con Colonial	Efe	La vanguardia	9	317
Report 29	2013	El número de viviendas hipotecadas bajó un 32,7% en 2012	Europa Press	La vanguardia	11	353
Report 30	2013	La prima de riesgo española se dispara a 400 puntos por los resultados en Italia	Efe	El pais	6	183

## English Editorials

Text	Year	Title	Author	Source	Cl	W
Editorial 1	2013	David Cameron's mixed signals on terrorism	Editorial	The independent	23	625
Editorial 2	2013	A dilemma that wont go away	Editorial	The independent	16	345
Editorial 3	2013	Obama must seize the chance to make history	Editorial	The independent	31	593
Editorial 4	2013	Stop all the demonishing of new EU migrants	Editorial	The independent	32	596
Editorial 5	2013	A confession is not enough	Editorial	The independent	9	217
Editorial 6	2013	Mali's Islamists are too dangerous to be ignored	Editorial	The independent	30	582
Editorial 7	2013	Still an island of ignorance about Europe	Editorial	The independent	5	154
Editorial 8	2013	Blame the tax loopholes	Editorial	The independent	16	321
Editorial 9	2013	Our interests come first, not those of America	Editorial	The Telegraph	17	447
Editorial 10	2013	Britain is ready to rock the boat in the EU	Editorial	The Telegraph	33	684
Editorial 11	2013	Fair pay for teachers	Editorial	The Telegraph	15	304
Editorial 12	2013	France moves in	Editorial	The Telegraph	14	285
Editorial 13	2013	As the terror threat shifts to Africa, we must meet it with great care	Editorial	The Telegraph	30	693
Editorial 14	2013	Heads in the sand over Eastern immigration	Editorial	The Telegraph	23	450
Editorial 15	2013	Britain is ill-prepared for this deadly new world	Editorial	The Telegraph	19	670
Editorial 16	2013	This winters tale	Editorial	The Telegraph	10	245

## D. Appendix

Editorial 17	2013	Our private schools deserve political praise	Editorial	The Telegraph	31	679
Editorial 18	2013	India, a useful friend	Editorial	The Telegraph	9	290
Editorial 19	2013	Courts and Parliament must work together	Editorial	The Telegraph	18	445
Editorial 20	2013	Ed Miliband hasn't learnt any economic lessons	Editorial	The Telegraph	20	456
Editorial 21	2013	Britain's hypocrisy over Sri Lanka	Editorial	The independent	11	309
Editorial 22	2013	The market needs watching	Editorial	The independent	27	602
Editorial 23	2013	Governments should not own banks	Editorial	The independent	19	329
Editorial 24	2013	A President at the very height of his power	Editorial	The independent	27	579
Editorial 25	2013	William Hague has been outflanked on Syria	Editorial	The independent	25	622
Editorial 26	2013	The Italian joker	Editorial	The Telegraph	9	305
Editorial 27	2013	Osborne needs to budget for economic growth	Editorial	The Telegraph	15	435
Editorial 28	2013	We all pay the cost of the BBC's failures	Editorial	The Telegraph	25	604
Editorial 29	2013	Time running out for talks with Iran	Editorial	The independent	20	345
Editorial 30	2013	Lessons from Pollard that the BBC still has to learn	Editorial	The independent	8	237

## Spanish Editorials

Text	Year	Title	Author	Source	Cl	W
Editorial 1	2009	Yak-42: justicia a medias	Editorial	El Pais on line	16	488
Editorial 2	2009	Sonrojo en Westminster	Editorial	El Pais on line	10	386
Editorial 3	2009	Sigue imputado	Editorial	El Pais on line	13	484
Editorial 4	2009	Proyectos divergentes	Editorial	El Pais on line	14	395
Editorial 5	2009	Medidas improvisadas	Editorial	El Pais on line	16	393
Editorial 6	2009	Equilibrio de poderes	Editorial	El Pais on line	14	372
Editorial 7	2009	El reto de Pyongyang	Editorial	El Pais on line	13	378
Editorial 8	2009	Camps y Torquemada	Editorial	El Pais on line	13	442
Editorial 9	2013	Terrorismo en el Sahel	Editorial	La Vanguardia	18	572
Editorial 10	2013	Grecia al límite	Editorial	La Vanguardia	10	412
Editorial 11	2013	Grave error en Nissan	Editorial	La Vanguardia	13	404
Editorial 12	2013	Basta ya de corrupción	Editorial	La Vanguardia	17	544
Editorial 13	2013	Italia y el cambio de tendencia	Editorial	La Vanguardia	28	534

Editorial 14	2012	Egipto, de la paradoja a la gobernabilidad	Editorial	La Vanguardia	23	552
Editorial 15	2012	El futuro de la sanidad pública	Editorial	La Vanguardia	14	428
Editorial 16	2012	Las expropiaciones de Evo Morales	Editorial	La Vanguardia	9	311
Editorial 17	2013	El listón de los imputados	Editorial	La Vanguardia	44	1067
Editorial 18	2013	España gana con el presupuesto de la UE	Editorial	La Vanguardia	20	566
Editorial 19	2013	Túnez, transición encallada	Editorial	La Vanguardia	11	421
Editorial 20	2013	El papado de Benedicto XVI	Editorial	La Vanguardia	44	1074
Editorial 21	2013	La cultura, apuesta estratégica	Editorial	La Vanguardia	25	566
Editorial 22	2013	Juicio militar en Rabat	Editorial	El Pais on line	18	382
Editorial 23	2013	Exportamos bien	Editorial	El Pais on line	19	485
Editorial 24	2013	Escuchas y fallos	Editorial	El Pais on line	20	471
Editorial 25	2013	Berlusconi todavía	Editorial	El Pais on line	13	391
Editorial 26	2013	El caso ‘Hellín’	Editorial	El Pais on line	12	375
Editorial 27	2013	Europa vigila	Editorial	El Pais on line	14	456
Editorial 28	2013	Medidas contundentes	Editorial	El Pais on line	18	596
Editorial 29	2013	Cita en Berlín	Editorial	La Vanguardia	18	548
Editorial 30	2013	El euro fuerte	Editorial	La Vanguardia	8	251

## English Letters to the Editor

Text	Year	Title	Author	Source	Cl	W
Letters to the editor 1	2012	A disappointing day for democracy	Steve Gilbert	the independent	7	214
Letters to the editor 2	2012	BBC inspires all classes	Michael J Marsden	the independent	5	144
Letters to the editor 3	2012	Dangerous ways to raise money	John Fleming	the independent	10	217
Letters to the editor 4	2012	Holland Park to Roxy Music	Andy Mackay	the independent	5	111
Letters to the editor 5	2012	Intellectual vacuum	Tony Slater	the independent	5	90
Letters to the editor 6	2012	Now can we see a politician take responsibility?	Malcolm Macintyre-Read	the independent	2	121
Letters to the editor 7	2012	Peer brings twitter to hell	Alice Renton	the independent	3	79
Letters to the editor 8	2012	Prosperity and peace for Gaza	David Mc-Dowall	the independent	6	118

## D. Appendix

Letters to the editor 9	to	2012	Put Starbucks out of its misery	Alan Stedall	the independent	4	137
Letters to the editor 10	to	2012	Schools in need of discipline	Kartar Uppal	the independent	4	113
Letters to the editor 11	to	2012	Baffling rail fares	Gordon Casely the telegraph	4	80	
Letters to the editor 12	to	2012	Correct etiquette for cyclists crossing the road	Andrew M Courtney	the telegraph	4	125
Letters to the editor 13	to	2012	How can it be right to send this SAS hero to jail for 18 months?	Jasper Archer	the telegraph	5	142
Letters to the editor 14	to	2012	How PCC electors were kept in the dark	Peter T D Williams	the telegraph	13	312
Letters to the editor 15	to	2012	Law on women bishops	John D Frew	the telegraph	5	130
Letters to the editor 16	to	2012	Native ash tree hope	Julian Freeman-Attwood	the telegraph	9	197
Letters to the editor 17	to	2012	No purple pride	Gregory Shenkman	the telegraph	8	109
Letters to the editor 18	to	2012	Overburdened forces	Sqn Ldr Dave Tisdale RAF (retd), Captain John Marshall RN, Surgeon Rear Admiral Ralph Curr, Christopher Samuel, Tim Dainton, Mike Parnell, Lt Col Malcolm Lindley	the telegraph	4	196
Letters to the editor 19	to	2012	Terrible cycle attacks in Israel, retaliation and recruitment to Hamas	Neville Teller	the telegraph	11	226
Letters to the editor 20	to	2012	The Arab spring has brought a new extremism to bear against Israel	Mike Davidson	the telegraph	6	145
Letters to the editor 21	to	2012	Deciding who benefits	Ben Crompton	The telegraph	8	124
Letters to the editor 22	to	2012	Can the police force be trusted to investigate his own failures?	Robin Jowitt	The telegraph	4	86
Letters to the editor 23	to	2012	Hospital consultants already carry our work commitments at weekends	Jonathan Frappell	The telegraph	10	199



Letters to the editor 24	2012	Dementia screening	Dr Martin Brunet, GP, and others	The telegraph	11	238
Letters to the editor 25	2012	Memorable Christmas song are things of the past	Mark Heane	The telegraph	2	40
Letters to the editor 26	2012	The ban on hunting is leading to an increase in weak and deased foxes	John Lewis	The telegraph	6	120
Letters to the editor 27	2012	The Government has no mandate to redefine the meaning of marriage	David Burrows MP (Conservative) and others	The telegraph	13	269
Letters to the editor 28	2012	Women's premiums	Vernon Stradling	The telegraph	9	136
Letters to the editor 29	2012	Traditional institution of marriage is central to the welfare of society	Cardinal Cormac Murphy-O'Connor	The telegraph	11	340
Letters to the editor 30	2012	Why Lord Heseltine was right in calling the Civil Service amateurish	Stanley Brodie QC	The telegraph	17	274
Letters to the editor 31	2012	Bullies and victims in the hoax call row	Justin Wollin	The independent	7	195
Letters to the editor 32	2012	Changing American's gun culture	David Ashton	The independent	11	223
Letters to the editor 33	2012	Clegg's welcome word on drugs	Rod Raso	The independent	16	242
Letters to the editor 34	2012	EBacc plans are unfair on the arts	William Gaver and others	The independent	13	249
Letters to the editor 35	2012	Gay wedding in church? Are you sure?	David P Negus	The independent	5	206
Letters to the editor 36	2012	How did prankst get through?	Paula Jones	The independent	8	185
Letters to the editor 37	2012	Limit access to guns but in wont end massacres	Andy Knott	The independent	7	173
Letters to the editor 38	2012	Poor penalised for phoning HMRC	John Healey MP	The independent	12	234
Letters to the editor 39	2012	Settle this Plebgate stand-off now	Mark Campbel	The independent	7	224
Letters to the editor 40	2012	The right to be a flintlock musket	Dudley Dean	The independent	7	170
Letters to the editor 41	2013	Free labour to stack shelve	Professor Tom Simpson	The independent	8	183
Letters to the editor 42	2013	Supermarkets must take blame for horse meat	Roland Courtney	The independent	9	199

## D. Appendix

Letters to the editor 43	to	2013	Water companies call the shots	Alan Mitcham	The independent	11	240
Letters to the editor 44	to	2013	Why Scotland should go it alone	W B McBride	The independent	11	200
Letters to the editor 45	to	2013	Richard III, legislator and killer	Valerie Crews	The independent	4	71
Letters to the editor 46	to	2013	The Government must have known about the damaging culture in the National Health Service	Angus McPherson	the telegraph	3	60
Letters to the editor 47	to	2013	After the EU speech and gay marriage, what next for Cameron?	David Saunders	the telegraph	3	73
Letters to the editor 48	to	2013	If doctors and nurses can be struck off, why are managers immune?	Don Cole	the telegraph	6	108
Letters to the editor 49	to	2013	We don't want mangey foxes anywhere near u	Peter Owen	the telegraph	2	38
Letters to the editor 50	to	2013	Tory U-turn on inheritance tax pledge puts pressure on George Osborne	Alasdair Ogilvy	the telegraph	5	123

## Spanish Letters to the Editor

Text	Year	Title	Author	Source	Cl	W
Letters to the editor 1	2012	¿Cómo están ustedeees?	Amadeu Branera Gubern	elpaionline	9	208
Letters to the editor 2	2012	Contra la amnesia	Enrique Linaza Iglesias	elpaionline	11	273
Letters to the editor 3	2012	Cooperación internacional	José María Martínez Cava	elpaionline	9	156
Letters to the editor 4	2012	Euro por receta	Rafael González	elpaionline	8	196
Letters to the editor 5	2012	La violencia en Gaza	Benjamín Telias Simunovic	elpaionline	9	212
Letters to the editor 6	2012	La preocupación del ciudadano catalán	Elisenda Boldú	elpaionline	8	192
Letters to the editor 7	2012	Las instituciones siguen copadas por varones	Pilar Aguilar Carrasco	elpaionline	12	240
Letters to the editor 8	2012	Otras políticas sí son posibles	Patricio Simó Gisbert	elpaionline	12	256

Letters to the editor 9	2012	¿Quiénes son los morosos?	Merche Castillejo Valdés	elpaisonline	7	166
Letters to the editor 10	2012	Un cierre incomprensible	Juan Antonio Fernández Cordon	elpaisonline	9	201
Letters to the editor 11	2012	El arte no entiende de razas ni sexos	Rodrigo Moreno	elmundoonline	7	245
Letters to the editor 12	2012	El fracaso del independentismo	Eva Garrido	elmundoonline	9	213
Letters to the editor 13	2012	El santuario de ETA y las FARC	Esperanza García	elmundoonline	4	83
Letters to the editor 14	2012	Explotar la desgracia ajena en televisión	Diana Pinto	elmundoonline	9	166
Letters to the editor 15	2012	La conciencia por encima del cargo	María Muñoz	elmundoonline	5	136
Letters to the editor 16	2012	La doble moral de la casta política	Ángel Jiménez	elmundoonline	6	184
Letters to the editor 17	2012	La rebelión global de la naturaleza	Laura González	elmundoonline	5	87
Letters to the editor 18	2012	Las televisiones, un paso por detrás	Manuel Cañedo	elmundoonline	5	150
Letters to the editor 19	2012	Protesta del embajador israelí	Raphael Schutz	elmundoonline	6	146
Letters to the editor 20	2012	Willy Toledo y la muerte de Zapata	Javier Álvarez	elmundoonline	6	178
Letters to the editor 21	2012	Con todo respeto ¿esto es ridículo!	Cristóbal Gil Cebrián	Elpais online	5	119
Letters to the editor 22	2012	Contra el cambio de programa de Mas	Domingo González Pulido	elpais online	4	265
Letters to the editor 23	2012	Es preciso bucear otros fondos	Luis Esquena Romaguera	Elpais online	4	135
Letters to the editor 24	2012	Invertir en ciencia	Juan Domenech Oliva	Elpais online	8	248
Letters to the editor 25	2012	La lotería y la salida de la crisis	José A Lorenz	Elpais online	4	130
Letters to the editor 26	2012	La nueva Consitutución Egipcia	Ali Salem Iselmu	Elpais online	6	181
Letters to the editor 27	2012	Las causas del fracas escolar	David Rabadá I Vives	Elpais online	13	293
Letters to the editor 28	2012	Por una educación para los alumnos	Álvaro Gil Rúi	Elpais online	9	241
Letters to the editor 29	2012	Recortes también en residencias	José Ripoll Conesa	Elpais online	8	186

## D. Appendix

Letters to the editor 30	2012	Un trato vejatorio	Jesús Sánchez-Ajofrín Reverte	Elpais online	9	125
Letters to the editor 31	2012	¿Ballenas y ballenos?	Jose Antontio Casanovas	La van-guardia online	4	138
Letters to the editor 32	2012	Guerra de impuestos	David Pérez Tallón	La van-guardia online	7	188
Letters to the editor 33	2012	Deberían sentir miedo	Emili Roca Miguel	La van-guardia online	12	211
Letters to the editor 34	2012	El frustrado puede matar	Paulino Castells	La van-guardia online	9	271
Letters to the editor 35	2012	Inidgnación social	Jose Manuel Pena	La van-guardia online	4	200
Letters to the editor 36	2012	Lista de defraudadores	Pedro García	La van-guardia online	4	153
Letters to the editor 37	2012	Rajoy dice la verdad	José M. Martínez Cava	La van-guardia online	6	118
Letters to the editor 38	2012	Matanza de inocentes	María Fané	La van-guardia online	5	121
Letters to the editor 39	2012	No todo vale	Arturo Piñana	La van-guardia online	5	144
Letters to the editor 40	2012	Premios con dinero	José M Martínez Orozco	La van-guardia online	3	129
Letters to the editor 41	2013	Voy a dejar de votar	Joan Francés Tutusaus	La van-guardia online	6	125
Letters to the editor 42	2013	Ciclistas arrollados	Luis M. <sup>a</sup> Peiz Soler	La van-guardia online	12	200
Letters to the editor 43	2013	Todo un Pontífice	Marta Benito	La van-guardia online	5	117
Letters to the editor 44	2013	Medidas preventivas	Victoria García Torres	La van-guardia online	8	177

---

Letters to the editor 45	2013	Chorizos británicos	J. Antonio Fernández	La van-guardia online	4	135
Letters to the editor 46	2013	Vigilar a los empresarios	Patricio Arribas y Atienza	Elpais online	2	136
Letters to the editor 47	2013	Habr� que regenerar la pol�tica	Jorge F. S�nchez Pe�a	Elpais online	8	137
Letters to the editor 48	2013	Tasa Robin Hood	Emilia S�nchez-Pantoja	Elpais online	9	322
Letters to the editor 49	2013	La gala de los Premios Goya	Paula Yuste Ruiz	Elpais online	10	251
Letters to the editor 50	2013	Cuesti�n de sentido com�n	Marisa Mart�nez L�pez	Elpais online	16	280

---

## **D.2 Annotated News Reports in English**

C1# Thematic Field		SYNTACTIC REALIZATION				SYNTACTIC REALIZATION				Experiential Role		Text	
TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	Grd.	Stance	Engagement	Grd.	Detail	Experiential Role	Text
1										NP+App	Proper Noun+Det Adj Noun+Prop Proper Noun	Actor	Dominique Strauss-Kahn, the French head of the International Monetary Fund, escaped dismissal for a one-night stand with a subordinate today, but was denounced by board members for a "Serious error of judgment".
2										NP	Det Noun Adj Noun	Goal	The fund's managing director was allowed to remain in his post after an investigation cleared him of accusations of harassment, favoritism and abuse of power over his affair with Priscilla Nagy, a Hungarian economist.
3	But									NP	Det Noun	Carrier	But his reputation took a battering in the statement by the IMF board, which said the relationship "was negligible and reflected a serious error of judgment on the part of the managing director."
4										NP+App	Proper Noun+Det Noun Adj Noun	Sayer	Shikour Shaulin, the fund's executive director, said IMF staff, and notably female staff, "are not at all happy" with Mr Strauss-Kahn, whose amorous adventures in France have earned him the epithet of grand seducer.
5										Clause		Verbiage	"The managing director has expressed his regrets. I don't think that we can ask him to do more at this time," said Mr Shaulin. "We will continue to work with him."
6										NP+App	Proper Noun+ Noun	Sayer	Mr Strauss-Kahn, 49, said in a statement "I very much regret the incident and I accept responsibility for it."
7										NP	Det Noun	Goal	The investigation was launched when Mario Blejer, a senior Argentine economist, said Mrs Nagy, his wife, had been seduced at the Davos international forum.
8										NP	Det Adj Noun Noun	Receiver	A US law firm was asked to probe accusations that Mrs Nagy, who left Mr Strauss-Kahn in 2005, had been seduced by Mr Strauss-Kahn, possibly with an exorbitant pay-off.
9										NP+NP+NP	Proper Noun+Proper Noun+Proper Noun	Sayer	Morgan, Lewis & Bockius said there was no evidence to support the claims.
10										NP	Det Noun	Goal	The decision was greeted with relief in France, where Mr Strauss-Kahn is seen as a key figure in President Sarkozy's drive to remodel and modernize the government.
11										NP	Adj Noun	Sayer	French commentators hailed the role played by Anne Sinclair, Mr Strauss-Kahn's wife, a former television presenter, who wrote in her blog: "Everyone knows that these are things which occur in the life of all couples. This one night adventure is now behind us. We love each other as much as we did on the first day."

Opce hawks want to cut oil production to keep up price

C1# Thematic Field		INTERP.				PREHEAD				SYNTACTIC REALIZATION				HEAD		SYNTACTIC REALIZATION				Experiential Role		Text
TEXTUAL		Stance	Engagement							Stance	Engagement											
1															Opce's mettle				NP	Proper Noun Noun	Goal	Opce's mettle will be tested next week, when the cartel assembles in Vienna for its biennial meeting to discuss whether the exporters' club should take action to arrest the fall in the price of crude oil.
2															Mounting alarm over the dwindling rent from oil sales				Nonfinite	Nonfinite	Actor	Mounting alarm over the dwindling rent from oil sales prompted Opce yesterday to bring forward an emergency meeting scheduled for late November to next Friday.
3								Together							Accompaniment- verb				NP	Det Adj Noun	Actor	Together the cartel members have been a billion dollars off the benchmark US light crude price oil since barges on the New York Nymex exchange yesterday to \$59.15 (£40.05), less than half the value at oil's peak, while London's Brent crude futures contract fell to \$55.70.
4															The benchmark US light crude price				NP	Det Noun Noun Adj Noun	Actor	The benchmark US light crude price oil \$55 per barrel on the New York Nymex exchange yesterday to \$59.15 (£40.05), less than half the value at oil's peak, while London's Brent crude futures contract fell to \$55.70.
5															Evidence of (...) and workshops				Nonfinite+Nonfinite	Nonfinite+Nonfinite	Actor	Members of Opec's oil cartel, in Kuwait and Venezuela, demand for a cut in production to shore up the price of oil.
6															Opce hawks, such as Iran and Venezuela				NP+App	NP+PP	Senser	Opce hawks, such as Iran and Venezuela, want to shore up the price by turning off the taps, but neither country can afford to cut output and risk angering its own producers.
7															Both				NP	Pronoun	Actor	Investment in their oil industries, so the burden of any cut will fall on Kuwait and Saudi Arabia.
8															The Saudis		hedge		NP	Det Proper Noun	Senser	The Saudis would not like to see the price falling below \$60 per barrel a level at which most economists reckon the kingdom can support its budget.
9	However,														The Saudis				NP	Det Proper Noun	Senser	However, the Saudis will be basing on the growing costs of the slowing Asian industrial juggernaut with mounting concern.
10															A deep recession		hedge		NP	Det Adj Noun	Actor	A deep recession might precipitate an oil price collapse.
11															It is Saudi Arabia				Clause	Predicated Theme	Token	It is Saudi Arabia that has been responsible for the weakening oil price, pushing more oil on the market since the energy summit in Jeddah.

#### D. Appendix

12	According to Platts, the oil pricing publisher, even as the price tumbles, Saudi Aramco is supplying every other's requests to fill barrels made next month.	Evidence	Evidence: Ashp	Saudi Aramco	NP	proper Noun	Actor	According to Platts, the oil pricing publisher, even as the price tumbles Saudi Aramco is supplying every other's requests to fill barrels made next month.
13				The wider question	NP	Der-Jedl Noun	To-ten	The wider question is what impact falling crude will have on non-opec oil output
14				Russia's budget	NP	proper Noun Noun	Carrier	Russia's budget depends on a \$70 per barrel price, according to Alken
15	But			Russian oil output	NP	proper Noun Noun	Actor	But Russian oil output is already falling because of weak investment in Mexican production
16				The major Western oil companies	NP	Der-Jedl Adj Noun	Actor	The major Western oil companies can probably stomach price dipping temporarily to \$60 for a few months and still pay their dividends, but a prolonged slump at \$50 per barrel would lead to the scrapping of investments followed by another cycle of shortages and soaring prices

## Microsoft results offer hope to tech sector

CTP Thematic Field			PREHEAD				SYNAPTIC REALIZATION				SYNAPTIC REALIZATION				Text	
TEXTUAL	INTERP		PREHEAD		SYNAPTIC REALIZATION		SYNAPTIC REALIZATION		SYNAPTIC REALIZATION		SYNAPTIC REALIZATION		SYNAPTIC REALIZATION		Text	
	Source	Engagement	Source	Engagement	Source	Engagement	Source	Engagement	Source	Engagement	Source	Engagement	Source	Engagement	Text	
1															Microsoft, the world's biggest computer company, offered a rare glimmer of hope to Wall Street after it unveiled better than expected quarterly profits.	
2															The software giant, based in Seattle, Washington, made net profits of \$429 million, or 16 cents a share, with its revenue rising to \$4.29 billion, a reported for the same period the year before.	
3															Revenue for the quarterly period rose 7 per cent to \$1.66 billion.	
4															Wall Street is seared about how much the US recession will erode Microsoft's corporate profits over the next few years.	
5															Microsoft executives said they will be hard hit by the economic downturn because their customers will cut back on IT spending.	
6	However,														However, Thursday's numbers show that Microsoft managed to find a way to make money even as its customers shelled out less for servers and other business programs.	
7															Ivan Trim, an analyst at Morgan Stanley, said: "If you are a Microsoft investor you should be happy that Microsoft thinks they will weather the storm relatively well."	
8		For the full year ending June 2009			Time										Microsoft's revenue rose 2009. Microsoft forecast earnings per share to rise to between \$2.40 to \$2.10 on revenue from \$64.9 billion to \$66.4 billion.	
9															Microsoft's earnings estimates called for earnings per share from \$2.12 to \$2.18 and revenue rising from \$67.3 billion to \$68.1 billion.	
10															Microsoft's guidance for the current quarter was marginally lower than Wall Street was expecting.	
11															Its shares rose then dipped in extended trading as investors digested the news.	
12		"We (...) in particular"			Quote										"We, the most companies, saw a slowdown at the end of the quarter in particular," Chief Financial Officer Chris Liddell said.	
13															Mr. Liddell said the worst hit among Microsoft's customers in the just-concluded quarter were small and medium-sized businesses that had been slow to recover from the economic gloom and who perhaps make decisions on a month-to-month basis."	
14															In a conference call with investors, Mr. Liddell said Microsoft will "batter our business to whatever the economy brings."	

**Dutch Government to inject €10bn of fresh capital into ING savings bank**

CTF	Thematic Field												Text
	TEXTUAL	INTERP.	PREHEAD		SYNACTIC REALIZATION		HEAD		SYNACTIC REALIZATION		EXPERIENTIAL ROLE		
			Source	Engagement	Source	Engagement	Detail		Source	Engagement	Detail		
1							The Dutch Government					Sayer	The Dutch Government announced last night that it would inject €10 billion (€7.8 billion) of fresh capital into ING, the Dutch savings bank with more than a million British customers.
2							The details of the deal, which was announced last night by ING chief executive Vitorino, would be revealed on the weekend.					Verbiage	The details of the deal, which was agreed over the weekend, were announced last night by Vitorino, ING chief executive. Vitorino said the bank had agreed to inject €10 billion of new capital into ING.
3							The Government					Actor	The Government will buy €10 billion of subordinated bonds in ING.
4							The deal					Actor	The deal could leave the state with as much as an 8.5 per cent stake in ING, which would mean ING would have two of the 12 seats on its supervisory board.
5							ING					Carrier	ING's rescued UK banks, ING will be free to pay dividends, while it owes money to the Dutch Government.



6	However,							the bank				NP	Det Noun		Sayer	However, the bank said it was skipping its final dividend for 2008 and executive bonuses because of the "exceptional circumstances".
7								The bonds				NP	Det Noun		Carrier	The bonds are redeemable at any time at 150 per cent of the issue price and carry no voting rights.
8								The capital injection				NP	Det Adj Noun		Actor	The capital injection will boost ING's core Tier 1 ratio, a measure of its financial strength.
9								Mr Bos				NP	Proper Noun		Sayer	Mr Bos said the deal was a pre-emptive measure to a rate cut as soon as "the financial hurricane" needed.
10							This is a healthy company,	Quote				NP	Pro noun		Sayer	"This is a healthy company," he said. "We wanted to be able to give it extra buffers to withstand market conditions in those extraordinary times."
11								Mr Timmer				NP	Proper Noun		Sayer	Mr Timmer said the bank had the trust of its customers and had not seen a large outflow of funds.
12						The market environment has changed over the last two weeks and the expectations for capital levels have changed following massive capital injections in financial institutions worldwide.	Quote					NP	Pro noun		Sayer	"The market environment has changed over the last two weeks and the expectations for capital levels have changed following massive capital injections in financial institutions worldwide," he said.
13								The move				NP	Det Noun		Actor	The move follows ING's revelation after trading on Friday that it had been forced to raise capital.
14								its shares				NP	Det Noun		Actor	Its shares had already plummeted 22.5 per cent in a 13-year low on rumours But it was short of capital.
15								ING				NP	Proper Noun		Carrier	ING has 85 million customers and 130,000 staff around the world.

## G7 'preparing to drive down the yen'

[illegible]

### D. Appendix

[illegible]

## Ukraine agrees terms of \$16.5bn rescue by IMF

C14	Thematic Field		PREHEAD				SYNTACTIC REALIZATION				Experiential Role		Text
	TEXTUAL	INTERP.	Source	Engagement			HEAD	Stance	Engagement	Grat.	Detail		
	1						The International Monetary fund			NP	Proper Noun	Actor	The International Monetary fund has reached agreement to provide up to \$100 billion for a \$ 6.6 billion (F10.2 billion) loan to help to ease the effects of the financial crisis.
	2						The 24-month deal, subject to parliament passing new bank laws			NP+NP+App	NP+NP+finite	Carrier	The 24-month deal, subject to parliament passing new bank laws, requires IMF board approval.
	3						Ukraine			NP	Proper Noun	Carrier	Ukraine is facing a hard predicament any decision in these years, its currency has been devalued by a fifth in a record low and the price of wheat for most export has fallen.
	4						Ukraine and Hungary, which due to be a substantial finance package" from the IMF within days.			NP+NP+App	Proper Noun + Proper Noun+Clause	Carrier	Ukraine and Hungary, which is due to add a substantial financing package" from the IMF within days, are set to become the first nations in the former Eastern bloc to gain IMF help in the credit crisis.
							Belarus, like Iceland			NP+App	Proper Noun+PP	Actor	Belarus, like Iceland and Pakistan, is also seeking an IMF loan.

Barclays rejects government funding, secures £5.8bn from Qatar and Abu Dhabi

CIP		Thematic Field		SYNTHETIC REALIZATION				SYNTHETIC REALIZATION				Expertise Role		Text
TEXTUAL	INTERP.	Source	Engagement	PREHEAD	Source	Engagement	Grd.	Detail	HEAD	Source	Engagement	Grd.	Detail	
1									Barclays			NP	Proper Noun	Barclays is to raise £7.5 billion from investors after securing a £5.8 billion cash injection from investors in Qatar and Abu Dhabi, in exchange for a stake of over 30 per cent in the British lender.
2									The bank			NP	Det Noun	The bank announced this morning that it will raise up to £2 billion from Qatar and Abu Dhabi, in exchange for a stake of over 30 per cent in the British lender.
3									It			NP	Pronoun	It will also raise £2.5 billion from Sheikh Mansour Bin Zayed Al Nahyan, a member of Abu Dhabi's royal family, which would give him a 2 per cent stake in the bank.
4									Barclays			NP	Proper Noun	Barclays is to raise £7.5 billion from investors after securing a £5.8 billion cash injection from investors in Qatar and Abu Dhabi, in exchange for a stake of over 30 per cent in the British lender.
5									The plan			NP	Det Noun	The plan will help John Velez, Barclays's chief executive, to regularise the bank without government help.
6									The bank			NP	Det Noun	The bank announced this morning that it will raise up to £2 billion from Qatar and Abu Dhabi, in exchange for a stake of over 30 per cent in the British lender.
7									Qatar Holdings, and HH Sheikh Mansour Bin Zayed Al Nahyan			NP+NP	Proper Noun+Proper Noun	Qatar Holdings, and HH Sheikh Mansour Bin Zayed Al Nahyan will raise £2.5 billion from investors after securing a £5.8 billion cash injection from investors in Qatar and Abu Dhabi, in exchange for a stake of over 30 per cent in the British lender.
In addition									Qatar Holdings			NP	Proper Noun	In addition, Qatar Holdings will receive £66 million for having helped

[illegible]

## US economy officially on brink of recession

CLIP Thematic Field				SYNAPTIC REALIZATION						Experiential Role		Text	
TEXTUAL	INTERP.		PREHEAD		SYNAPTIC REALIZATION		HEAD		SYNAPTIC REALIZATION		Detail		Experiential Role
	Sence	Engagement	Sence	Engagement	Sence	Engagement	Sence	Engagement	Sence	Engagement			
1							The US economy		NP		Det Adj Noun	Carrier	The US economy is on the brink of officially entering recession after a slump in Americans' consumer spending saw its gross domestic product (GDP) decline for the first time since the aftermath of the terrorist attacks of September 11, 2001.
2							Warnings (...) prove long and deep.		Infinite		Infinite	Goal	Warning that the American economy is threats in the wake of a new recession, and fears that this will prove long and deep, were reinforced after official figures showed that it contracted at an annual rate of 0.2 per cent in the past quarter.
3							The drop in US national income and output – by 0.1 per cent – in a quarter-on-quarter basis.		NP+App		NP+PP	Actor	The drop in US national income and output – by 0.1 per cent on a quarter-on-quarter basis – marked an abrupt reversal after growth at a brisk 0.7 per cent quarterly rate in the previous three months.
4							US economists		NP		Adj Noun	Sayer	US economists predicted that a far more brutal slump was now set to follow in the present quarter and into next year "The drop is a harbinger of much worse things to come. We now expect that GDP will contract at an annual rate of 0.5 per cent in the third quarter, and fall another 0.8 per cent in the fourth quarter of next year, generating the worst recession since 1981-82," Abiel Reimbert, of JPMorgan Chase in New York, said.
5							The main drag on America's performance		NP+App		NP+PP	Actor	The main drag on America's performance in the past three months, came as consumer spending slumped at an annual rate of 3.1 per cent.
6							R		NP		Pro noun	Carrier	It was the first drop since the end of 1991, and the biggest since 1989's second quarter.
7							Economic activity		NP		Adj Noun	Goal	Economic activity was also sapped as businesses cut back on their investment spending – the first fall since the end of 2006 – which fell at a 1.6 per cent annual rate, after rising by an annual 2.5 per cent in the second quarter. Investment in equipment and software fell by an annual 5.5 per cent.
8							The grim news		NP		Det Adj Noun	Actor	The grim news came after the US Federal Reserve stepped up its drive to shore up the rapidly faltering US economy on Wednesday with a surprise 50-cent interest rate cut.
9							American Express, the credit card group,		NP+App		proper Noun-NP	Actor	American Express cut interest rates and gave a \$1 billion bonus to its staff, announcing a call of 7,000 jobs or nearly 10 per cent of its workforce, in a move that could cost the company as much as \$1 billion. The move comes as it seeks to cut costs by \$1.8 billion (£1.1 billion) next year, while Motorola, the mobile phone company, is looking to shed 3,000 posts, or about 10 per cent of its workforce.
10							Markets		NP		Noun	Carrier	Markets found solace, however, in the third-quarter fall in US GDP being smaller than predicted, and in hopes that worldwide rate cuts led by the Fed could forestall a global recession.
11			On Wall Street		Panic-PP		US blue-chip shares		NP		Adj Adj Noun	Actor	On Wall Street, US blue-chip shares closed up more than 2 per cent, after a record 491-point slide last week, ending a losing streak of having risen 400 points in the previous two days.
12							Anxiety over the global outlook.		NP+App		NP+PP	Goal	Anxiety over the global outlook were stalked after the latest snapshot of economic sentiment in the euro-zone showed confidence among businesses and consumers at its lowest since 1993.

13								Remonides said that this was a clear signal that the 15-nation bloc was taking into recession.	Sayer	NP	Noun		
14								Germany and France joined Japan in signalling further moves to a recession.	Actor	NP+NP	Proper Noun+Proper Noun		
15								holder economic conditions to try to stave off a recession.	Carrier	NP	Proper Noun		
16								Berlin is set to introduce a stimulus package worth up to £15 billion next	Sayer	NP	Proper Noun		
17								In France, President Sarkozy said a fund for domestic industries would be set up in three weeks.	Actor	NP	Proper Noun		
18								In Japan, the government unveiled a ¥26.9 trillion (£69 billion) stimulus package to boost the economy, including fresh spending of ¥5 billion to encourage companies to buy government bonds.	Actor	NP	Det Noun		
								Mr Aso, the Japanese Prime Minister, said: "A harsh storm seems only a matter of time."	Sayer	NP+App	Proper Noun+NP		

## Narrow win for 'Merkel's Mac' in Lower Saxony

C14 Thematic Field		INTERP.			PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			EXPERIMENTAL ROLE			TEXT
TEXTUAL		Stance	Engagement																	
1																			Chancellor Angela Merkel's ruling coalition partners were on course to secure a wide-thin one-seat majority in Lower Saxony last night, in a ballot seen as a bellwether for the autumn general election.	
2				In a surprise result					Place		Place								In a surprise result, Mr Merkel's liberal Free Democrat partners in government re-entered parliament in Germany's second most populous state and secured a narrow victory. But they still need to secure the necessary 5 per cent of the vote to retain seats.	
3																			The outcome was expected to lead to the reinstatement of Ms Merkel's candidate in Lower Saxony – the conservative state prime minister Dayla McAllister, whose father was born in Glasgow and who ran his election campaign in the state capital of Hanover.	
4																			Last night's results suggested the Free Democrats would gain more than 10 per cent of the vote – a result which would enable them to form a new state government with Ms Merkel's conservative Christian Democrats, who did worse than expected by winning only 39 per cent of the vote.	
5																			It remains unclear whether the governing coalition will secure a place of honour in the state government after the September election, a Free Democrat party spokesman said after the polling closed.	
6																			Mr McAllister, 41, has been tipped as a successor to Ms Merkel.	

## Georgian protesters want President Saakashvili out

CPI		Thematic Field													
TEXTUAL	INTERP.	PREHEAD		SYN/CPIFC REALIZATION				HEAD		SYN/CPIFC REALIZATION				Experiential Role	Text
		Source	Engagement					Source	Engagement	Grnl.	Detail	Source	Engagement		
1															Crowds gathered outside the presidential offices in Tbilisi yesterday to demand the removal of President Mikheil Saakashvili, whose term in office they say should be ending.
2															Protesters said the President's second term should expire today, five years from his inauguration.
3															The constitution demands that the presidential term in five years, but does not call for elections until October this year.
4															Mt Saakashvili's National Movement party was defeated in parliamentary elections in October.
															Bidzina Ivanashvili became Prime Minister, forcing a difficult

## Top executive suppressed report damning Barclays

CTP		Thematic Field		SYNOPSIS										Text	
TEXTUAL	INTERP.	PREHEAD		SYNOPSIS REALIZATION			SYNOPSIS REALIZATION			EXPERIMENTAL ROLE		Text			
		Source	Engagement	Source	Engagement	Grat.	Source	Engagement	Grat.	Detail					
1										Bureaucracy	Proper Noun	linget	Bureaucracy was hit by further damaging revelations about the culture that was allowed to flourish in certain parts of the bank, yesterday after it emerged that a senior executive had suppressed a damaging report into Barclays Wealth in America		
2										The report, compiled by consultancy firm Genesis Ventures, Management	NP+App	Verbiage	The report, compiled by consultancy firm Genesis Ventures, was said by a Sunday newspaper to have sharply criticised a "revenue at all costs" strategy and a "culture that is high-risk and actively hostile to compliance"		
3										Genesis Ventures' report was initially only supplied to Andrew Timney, chief operating officer of the private investment division, and he suppressed it	NP	Verbiage	Management was said to have "ruled with an iron fist", acting to remove any potential threat to the bank's reputation		
4										Genesis Ventures' report was initially only supplied to Andrew Timney, chief operating officer of the private investment division, and he suppressed it	NP	Goal	Genesis Ventures' report was initially only supplied to Andrew Timney, chief operating officer of the private investment division, and he suppressed it		
5										His departure was announced to staff in an email a week ago, and it is the revelations come days after new chief executive Antony Jenkins wrote to all 140,000 staff members, telling them that if they didn't buy into his clean-up plan they would be kicked out	NP	Verbiage	His departure was announced to staff in an email a week ago, and it is the revelations come days after new chief executive Antony Jenkins wrote to all 140,000 staff members, telling them that if they didn't buy into his clean-up plan they would be kicked out		
6										The revelations	NP	Actor	The revelations come days after new chief executive Antony Jenkins wrote to all 140,000 staff members, telling them that if they didn't buy into his clean-up plan they would be kicked out		
										He	NP	Carrier	He is due to unveil a new strategy in February, which will see the bank		

8						Sarif at Canary Wharf	NP+App	Noun+PP	Verbage	Staff at Canary Wharf are said to be nervously awaiting a fresh round of lay-offs.
9						The revelations	NP	Det+Noun	Carrier	The revelations follow The Independent's exclusive interview with a whistleblower last July in which he talked of a culture of fear at the investment banking division
10						The Independent	NP	Proper+Noun	Sayer	He also accused some of his own staff of covering up other former members of staff in the division who sold similar stories.
11						One	NP	Pronoun	Sayer	One said yesterday: "This bubble up what I have been saying. The attitude and ethics of the big guys up top more than filtered in the bottom."
12						New rule breaches brought to light by such a report	Cause hedge	Cause	Carrier	NPQ rules breaches brought to light by such a report would be covered up by regulators who fined the bank £290m after its traders were found to have attempted to manipulate Libor interest rates last year.
13	Last night, however	Time				People close to the bank	NP+App	NP+PP	Sayer	Last night, however, people close to the bank stressed Mr Timony's departure was a reflection of the reforms being pushed through by Mr Barclays spokesman.
14						A bank's spokesman said Mr Timony had "walked out just clearly last week," that transformation of the bank's culture was "imminent".	NP	Det Proper+Noun+Noun	Sayer	A Barclays spokesman said Mr Timony had "walked out just clearly last week," that transformation of the bank's culture was "imminent".
15						Independent reports, like the one published in early 2012 for Barclays Wealth in America	NP+App	NP+Cause	Senser	Independent reports, like the one commissioned in early 2012 for Barclays Wealth in America, are expected to identify areas where change is required and to recommend remedial steps.
16						These types of exercise	NP+App	NP+PP	Carrier	These types of exercise never result in comfortable reading, but we have no choice but to accept them as a necessary step towards taking the necessary steps to address the issues raised, he said.

## Think-tank calls for new strategy to fire up the economy

C14 Thematic Field									
TEXTUAL	INTERP.		PREHEAD		SYNTHETIC REALIZATION		HEAD		Text
	Source	Engagement	Source	Engagement	Source	Engagement	Source	Engagement	
SYNTHETIC REALIZATION									
DETAIL									
1							The Government	directive	The Government should give the Bank of England and its incoming Governor Mark Carney a radical new mandate to get the UK economy moving again, a respected think-tank will argue today.
2							The Firm& Young ITEM Club	Proper Noun- Proper Noun	The Firm& Young ITEM Club says Threadneedle Street's existing 2 per cent inflation target is "not fit for purpose" and should be replaced with a "flexible GDP" target.
3							This	badge	It will require the Bank's Monetary Policy Committee (MPC) to announce that the government has a certain rate of inflation each year (in other words without taking into account the impact of rises in the cost of living).
4			In its winter 2012-13 forecast published today		Evidence- Noun/It is		the ITEM Club	Proper Noun	In its winter 2012-13 forecast published today, the ITEM Club argues that the existing inflation targeting regime has broken down.
5							This		This is because a global commodity price shock in recent years has compelled the MPC to keep monetary policy too tight to boost the ailing domestic economy while at the same time allowing monthly price rises to fall below the 2 per cent target.
6							It		It argues that the usual defence of the inflation target – that it is simple for the public to understand – no longer applies.
7							Recent inflation target	Adj Noun Noun	Recent inflation target overshoots and debates about the way that inflation is measured have undermined the credibility of the inflation target.
8							There	badge	It would seem to be little reason to suppose to target an indicator that is very hard to measure. Let alone control it against.
9							Pressure		Pressure is building for a new mandate for the bank, as the UK economy continues to stagnate five years after recession hit. Official figures released on Friday are expected to show that the economy contracted in the final quarter of 2012.
10							The Canadian Mr Carney		It is the idea of a money (GDP) ring, also known as a nominal GDP target, last month for economies that have already seen interest rates pushed close to zero.
However,							some current and former members of the MPC have given the prospects for a new mandate a mixed reception.	Np-pp	However, some current and former members of the MPC have given the prospects for a new mandate a mixed reception.

## Santander eyes a double bank buy with £2bn tag

CU #	Thematic Field			SYNTHETIC REALIZATION															Text
	TEXTUAL	INTERP.		PREHEAD			SYNTHETIC REALIZATION			HEAD	SYNTHETIC REALIZATION			EXPERIENTIAL ROLE					
		Stance	Engagement	Stance	Engagement	Detail	Stance	Engagement	Grat.		Stance	Engagement	Grat.	Stance	Engagement	Grat.			
1											Samander				NP	Propr Noun	Senior	Samander is considering a £2bn acquisition of Clydesdale and Yorkshire banks as part of plans to ramp up its UK presence.	
2											The UK division of the Spanish banking giant				NP-PP	NP-PP	Phenomenon	The UK division of the Spanish banking giant is understood to have been bought out by the UK's second largest bank, the Royal Bank of Scotland in October.	
3											The Australian bank				NP	Det Adj Noun	Verbiage	The Australian bank is reported to be under pressure from shareholders to offload the loss-making banks, although Clydesdale and Yorkshire banks have a healthy share of the market for small and medium-sized enterprises (SME) and corporate lending.	

5	In an interview published over the weekend, Samarander UK's chief executive Amit Boini said: "Right now, our thinking is that organic growth is what we are aiming for."	NP+App	Proper Noun Adj Proper Noun Noun	Sayer	In an interview published over the weekend, Samarander UK's chief executive Amit Boini said: "Right now, our thinking is that organic growth is what we are aiming for."
6					We are currently working on plans to grow our retail business, our SME business, we are building a team for large corporates in the UK, which we didn't have before."

## Tesco opens fifth dotcom store

Thematic Field	CITE					
	INTER.		PREHEAD		SYNACTIC REALIZATION	
TEXTUAL	Stance	Engagement	Stance	Engagement	Grat.	SYNACTIC REALIZATION
						Detail
						Experiential Role
1						Tesco
2						The opening of the highest-tech facility in Crawley
3						Tesco
4						Inside the M25
5						Ocado
6						Barney Burgess, Tesco's chief operating officer for grocery home shopping.

## UK firms to bid for Japan's nuclear clean-up

Thematic Field		INTERP.		SYNTHETIC REALIZATION				SYNTHETIC REALIZATION				Text		
		TEXTUAL		PREHEAD	Source	Engagement	Detail	HEAD	Source	Engagement	Detail			Experiential Role
1	1							British engineers Aunee, Babcock International, and Atkins						British engineers Aunee, Babcock International, and Atkins are believed to be critical in the start decommissioning work, estimated to be worth at least £2.2bn in Japan in a 10-year and 100,000-man effort.
	2							The new Japanese government						The new Japanese government is thought to be preparing to build a new nuclear power plant in the north-eastern Dutch plant, which was overwhelmed by a tsunami in 2011, and other reactors in seismically endangered areas.
	3							A nuclear source						A nuclear source said bids could be no later for the clean-up work before the end of the year, with British groups in a strong position due to all the decommissioning work that has been undertaken in the UK.
	4							US-owned Energy Solutions						US-owned Energy Solutions will also be interested.
	5							This is a huge opportunity						This is a huge opportunity, claimed the source.
								Rheine						Japan could start making some real progress on decommissioning

**Hilco UK emerges among favourites to give HMV a spin**

Thematic Field											
CT#	TEXTUAL	INTERP.		PREHEAD			SYNTACTIC REALIZATION			EXPERIENTIAL ROLE	Text
		Stance	Engagement	Stance	Engagement	Detail	Stance	Engagement	Detail		
1											Jacky Pottery-owner, Hilco UK, has emerged as the frontrunner to buy HMV's music chain, which will boost a high street already underpin by administrations in 2013.
2											The music and film chain's troubles were made worse yesterday when DVD rental chain Blockbuster confirmed plans to close 129 of its 528 stores over the coming weeks.
3											HMV collapsed on Monday when DeKretz was appointed as the 92-year-old company's new chairman.
4											The move put more than 4,000 jobs at risk, although HMV's 223 stores continue to trade.
5											On Friday, Hilco was locked in talks with DeLoraine which are expected to continue by trade.
6											The restructuring specialist has already managed a turnaround in sales at the music chain's parent company, the business from the London-based entertainment group in 2011.
7	However,										However, Hilco faces competition from other players for the UK business and the structure of any deal is likely to be complex, partly because HMV has many loss-making stores that a new owner is likely to jettison.

## UK seems likely to lose EU battle over limits on bonuses

## Cameron backs AugustaWestland despite Indian bribery scandal

## Green hails LA Topshop while he eyes Dreams

C1	Thematic Field		PREHEAD				SYNTACTIC REALIZATION				SYNTACTIC REALIZATION				Text	
	TEXTUAL	INTRPR.	Source	Engagement			Source	Engagement	Grad.	Source	Engagement	Grad.	Detail	Experiendial Role		
1										Sir Philip Green			NP	Proper Noun	Sayer	Sir Philip Green has bought of "Hustlers" trading at the combined value of \$10 million. The two-story store in New York, Chicago and Las Vegas emerged the billionaire is considering acquiring the beds retailer Dreams
2										The retail lycoson			NP	Det Adj Noun	Actor	The retail lycoson opened the 30,000 square foot store in L.A. in a fashion last Thursday, alongside the actress Kate Bosworth, in the latest expansion of this Atradi empire, which also includes Bhs, Dorothy Perkins, Matal, Schindler and Barton
3										Sir Philip			NP	Proper Noun	Sayer	Sir Philip Green has bought of "Hustlers" trading at the combined value of \$10 million. The two-story store in New York, Chicago and Las Vegas emerged the billionaire is considering acquiring the beds retailer Dreams
4										The two-story store			NP	Det Adj Noun	Extrior	The two-story store is Atradi's fourth in the United States, following openings in New York, Chicago and Las Vegas

#### D. Appendix

5					The LA store is the first since Sir Philip sold 25 per cent of Topshop and Topman to the US private-equity firm Leonard Green & Partners in a deal that valued the two brands at £70m in December	Taken	
6					The businessman declined to comment yesterday on speculation that he was interested in buying shares in the London-based fashion group, which has been awarded a royal warrant by its debut of £12m	Sayer	
7					Mike Clare, the founder and former chief executive of Dreams, Sun European Partners, the owner of sofa chain SCS, and the private-equity outfit Emerald Capital, said that the pair had bought the shares as a through to second-round bids around the end of this month	Carrier	
8					HIG Europe, the owner of mattress specialist Silentnight, is out of the running	Carrier	

## Pound takes another hit on German resurgence

CTP Thematic Field						SYNAPTIC REALIZATION							Text	
TEXTUAL	INTERP.		PREHEAD		SYNAPTIC REALIZATION		MEAD		SYNAPTIC REALIZATION		Experiential Role		Text	
	Source	Engagement	Source	Engagement	Grk.	Engagement	Grk.	Detail	Source	Engagement	Grk.	Detail		
1							The pound			NP		Ded Noun	Ackv	The pound took a fresh beating yesterday as the prediction of a German economic resurgence triggered a sterling sell-off and experts said the currency was at risk of a "large-scale devaluation"
2							Sterling			NP		Proper Noun	Ackv	Sterling falls only April's yen as the worst performer against a basket of international currencies this year as a 45 per cent decline fuels import price inflation and pushes up the cost of electricity for hundreds of thousands of households
3	Bat						Sallukh forecasts from the Bundesbank over an immediate return to growth for Germany, after a recession, after a period of 0.6 per cent slide between October and December,			NP+App		NP+PP	Ackv	Bat bullish forecasts from the Bundesbank over an immediate return to growth for Germany, after a worse-than-expected 0.6 per cent slide between October and December, put sterling on the back foot again yesterday as it sank close to 15-month lows against the euro
4			In another day of turbulence in international currency markets, the euro slipped further against the dollar, ending at 86½, within sight of its 15-month high of 87.2p hit earlier this month		Tine		the euro			NP		Ded Noun	Ackv	In another day of turbulence in international currency markets, the euro slipped further against the dollar, ending at 86½, within sight of its 15-month high of 87.2p hit earlier this month
5			Against the dollar		Manner		the pound			NP		Ded Noun	Ackv	Against the dollar, the pound took more pain, slipping to a seven-month low
6							The Bundesbank's primary report			NP		Ded Proper Noun Noun Noun	Sayer	The Bundesbank's primary report said Germany's economy would resume growth in the current quarter, with activity "gradually picking up pace in the remainder of the year"
7							Investment bank UBS forecast sterling is likely to be the next major currency that depreciates strongly	bodge		NP		Adj Noun Proper Noun	Sayer	An investment bank UBS forecast sterling is likely to be the next major currency that depreciates strongly
8			Head of exchange strategy		Role		Manoor Mohi-uddin			NP		Proper Noun	Sayer	Head of foreign exchange strategy Manoor Mohi-uddin said: "As central banks tolerate higher levels of inflation, the pound is set to weaken further across the board particularly against our favourite G4 currency, the US dollar. The pound seems clearly at risk of following the yen down."
9							The bank			NP		Ded Noun	Sayer	The bank has recommended clients buy options allowing them to sell the pound at \$148 in six months' time, making them profits if the pound falls even further below that level
10							Dealers			NP		Noun	Senior	Dealers also noted weekend comments from Bank of England rate-setter Martin Wolf, who warned the pound was still too high to help the UK economy rebalance effectively
11							The continued pressure on the currency			NP+App		NP+PP	Ackv	The continued pressure on the currency comes after its biggest weekly loss since June last year amid gloom over weak growth prospects
12							The Bank of England			NP		Proper Noun	Sayer	The Bank of England has signalled it is willing to tolerate higher inflation than previously, but it has also indicated that it will not do as the European Central Bank makes explicit commitments to prop up eurozone stragglers and preserve the single currency
13							The yen			NP		Ded Noun	Ackv	The yen also fell as the G20 finance ministers in Moscow shed away from criticism of Japan for weakening the currency
							Prime Minister Shinzo Abe			NP+App		NP+Proper Noun	Sayer	Prime minister Shinzo Abe has repeatedly called for a cheaper currency to help Japanese exporters before his bid for election

## Striking gold: Royal Mint moves back into Indian market

	CI # Thematic Field								Text					
	TEXTUAL		INTERP.		PREHEAD		SYNTHACTIC REALIZATION		HEAD		SYNTHACTIC REALIZATION		Experiential Role	
1														The Royal Mint is heading back into the world's biggest gold market after nearly a century's absence as it announced plans to produce



[illegible]

## Mario Draghi says talk of currency war 'excessive'

C17 Thematic Field		SYNTACTIC REALIZATION										Text	
TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION			Experiential Role	Text	
		Stance	Engagement	Stance	Engagement	Grat.		Stance	Engagement	Grat.			
1	Although			Mr Draghi said that appreciation of the euro was a risk to the economy.		Contingency		Contingency-Clause	he		NP	Pronoun	Although Mr Draghi said that appreciation of the euro was a risk to the economy, he added that "most of the exchange rate movements that we have seen were [...] the result of domestic macro-economic policies meant to boost the economy."
2				In this sense, I find really excessive any language referring to currency wars.		Quote		Quote	he		NP	Pronoun	"In this sense, I find really excessive any language referring to currency wars", he said.
3													"I urge all parties to exercise very, very strong verbal discipline. I think the less we talk about this the better."
4													"Looking at the nominal and real exchange rate of the euro, we see that by and large, it is around its long-term average."
5	Earlier on Monday,					Time		Time-AbsNP			NP	Proper Noun+NP	Earlier on Monday, Ewald Nowotny, another ECB policymaker, said that recent currency movements were not reason to cause alarm.
6				There is a euro appreciation against the yen but not a dramatic event.		Quote		Quote	he		NP	Pronoun	"There is a euro appreciation against the yen but not a dramatic event," he told reporters.
7													The euro hit a 15-month high against the dollar earlier this month complicating the ECB's policy making tasks by weighing on growth and feeding expectations that it may have to take fresh policy action although some ECB members are against that
8													Analysts have also played down the prospect of a currency war.
													Julian Jessop, chief global economist at Capital Economics, said that fears that the world is on the verge of currency wars were mostly based

## D. Appendix

10	The US and more recently Japan have both been accused of housing competitive advantage. But their main aim has been to boost domestic demand, which should ultimately benefit trading partners too. Indeed, Japan is arguably still not doing enough," he said.	Quote	Quote	NP	Der Noun	Sayer	
----	---	-------	-------	----	----------	-------	--

Lower income earners being squeezed hardest

C14 Thematic Field			SYNTACTIC REALIZATION				SYNTACTIC REALIZATION				Experiential Role		Text	
TEXTUAL	INTERP.		PREHEAD	Stance	Engagement	Grd.	HEAD	Stance	Engagement	Grd.	Detail			
1							Market's household finance index for February				NP+App	NP+PP	Actor	Market's household finance index for February found a growing gap between the pressure felt those on lower incomes and higher earners, who saw their financial situation easing.
2			Overall			Manner	the index				NP	Det Noun	Carrier	Overall, the index was unchanged at 37.7 in February, which is well below the 50 mark which shows that families' finances are improving rather than getting worse.
3							One third of households				NP+App	NP+Pp	Sayer	One third of households reported that their situation worsened during the month, while just one in 14 saw it get better.
4							Those with incomes of £15,000-£23,000				NP+App	NP+PP	Senser	Those with incomes of £15,000-£23,000 saw the strongest squeeze on their finances in the survey's four-year history, while those earning below £10,000 recorded the sharpest deterioration in their budgets for 14 months.
5	In contrast						households in higher income brackets				NP+App	NP+PP	Senser	In contrast, households in higher income brackets saw the deterioration in their finances slow down, while the highest earners, on £57,751 and over, saw the joint lowest squeeze on their budget in a year.
6						Time Manner carrier	Time Makers carrier and author of the report				NP+App	Proper Noun+NP	Sayer	Those who were most squeezed at Mark's research report said there was no let-up in the cost of living in the UK household finances during February, as higher living costs and muted wage trends combined to reduce cash availability at the fastest pace since mid-2012.
7							Rhyme					Quote		"Worsening consumer finances are likely to further rein in spending on retail, construction, health and social services, leading to be the most downbeat of the four quarters," says the report. It also says that the most pessimistic about their job security and workplace activity in February."
8							Rising rents and food costs				NP		Actor	Rising rents and food costs have placed added pressure on families, while a string of energy companies recently announced price hikes.
9			Meanwhile,			Time	wages				NP	Noun	Carrier	Meanwhile, wages have remained stagnant.
10							Separate research for VoucherCodes				NP+App	NP+PP	Actor	Separate research for VoucherCodes found consumers have so far only cut their spending on items such as new clothes and eating out has been reduced by 42pc and 31pc respectively, with respondents blaming high living costs and stagnant wages.
11							people working in retail, construction, health and social services				NP		Goal	According to Mark's research, people working in retail, construction, education, health and social services tended to be the most downbeat about their finances, while those working in IT and finance were the least pessimistic.
12			According to Mark's research,			Evidence	Evidence+AdvP				NP+App	NP+Noun+Pp	Carrier	Retail workers were the most pessimistic about job security.
13							Retail workers				NP	Adv Noun	Token	Four out of 10 households across the survey expect their finances to worsen over 2013, while 21pc predict an improvement.
14							Four out of 10 households across the survey				NP+App	NP+PP	Senser	The lowest earners were the most downbeat, with 57pc anticipating a deterioration, while those on the highest incomes had a neutral outlook, with similar numbers in this bracket predicting either a fall or an improvement.
15							The lowest earners				Np	Det Adj Noun	Token	The study also showed the sharpest deterioration in families' cash availability came from those with the type of people reporting a decline in their monthly survey and 1,500 people across Britain about their finances.
16							The study				NP	Det Noun	Actor	
							The monthly survey				NP	Det Adj Noun	Sayer	

Pound land: Derelict houses in Liverpool to be sold for just £1

C#		Thematic Field		INTERP.			PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			EXPERIENTIAL ROLE			Text
TEXTUAL		STANCE	ENGAGEMENT				STANCE	ENGAGEMENT	Grat.				STANCE	ENGAGEMENT	Grat.							
1																	NP-App	NP-Non-finite		Actor	The decision to let the houses go at the rock-bottom price comes after months of delays caused by the council's decision to break off talks with developer Leader1	
2																	NP	Det Noun		Carrier	The firm had been in line to redevelop hundreds of houses in the 'Granby Triangle' area of Liverpool since it was bought after the company failed to meet deadlines for signing the contract.	
3																	NP	Det Noun		Carrier	Since it was revealed the deal had fallen through in November last year, the boarded-up houses have remained empty.	
4																	NP	Det Noun			Now, the council proposes to sell off a cluster of houses for just £1 to residents - on the condition they will be brought up to a decent standard.	
5																	NP	Adv Noun		Carrier	Private landlords will also be able to bid for the tender to refurbish some of the vacant houses and then buy the freehold for £1.	
6																	NP	Proper Noun		Sayer	Last night, deputy mayor and council finance chief Clr Paul Brant said: "This allows people who may be excluded from mortgages but have construction skills to play a part in the regeneration of their communities."	
7																		Quote			It's been proven to work in other parts of the region. We've seen that the private sector model has not succeeded so far and, through this way of doing things, we can regenerate the area and create jobs and wealth."	
8																	NP	Noun		Actor	Residents are faced with a mix of anger and disappointment when the £1-a-house scheme fell through last year.	
9																	NP	Pronoun		Sayer	They said they had been uneasy that a private firm with no track record in social housing had been awarded the tender.	
10																	NP-App	NP- Proper Noun		Sayer	Local resident Theresa McDermott said: "This is a much better deal than we were offered. Obviously, the delays because of negotiation with Leader1, but although there's some uncertainty at the moment it's positive."	
11																	NP	Det Noun		Goal	As part of the initial pilot scheme, 20 houses will be offered for sale to residents for £1 in the Granby Four Streets' and Armistead Road in Kensington.	
12																	NP	Det Noun		Actor	In the 'Webster Triangle' in Picton the council will either partner up with housing associations or 'dispose of the properties to private landlords'.	
13																	NP-App	Proper Noun-clause		Sayer	Council Labour Democrat group leader Clr Richard Kemp, who has previously called into question the council's decision to get involved with private firms in the delivery of social housing, said he hoped the scheme would succeed, adding: "I think this is a good idea, provided there's a solid basis for it."	
14																		Quote			"Either housing associations or private individuals need to be doing this, as there's no profit to be made out of this kind of scheme, as we've seen	

MPs call for tax avoiders to be 'named and shamed'

C#	Thematic Field	INTRE:			PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			DETAIL			Experiential Role	Text
		TEXTUAL	Source	Engagement	Source	Engagement	Grat.	Source	Engagement	Grat.	Source	Engagement	Grat.	Source	Engagement	Grat.	Source	Engagement	Grat.		
1												Her Majesty's Revenue & Customs			NP	Proper Noun			Actor	Her Majesty's Revenue & Customs is also being a game of 'cat and mouse' with firms that promote tax avoidance schemes to individuals, a new report from the Public Accounts Committee (PAC) argues, in the latest salvo in the growing debate over tax laws.	
2	Although										Contingency-Clause			NP	Pronoun				Actor	Although the MPs' report is largely focused on schemes aimed at wealthy individuals, it comes at a time when evidence of minimal UK corporation tax payments by multinationals such as Starbucks and Amazon has catapulted tax to the top of the political agenda.	
3											Time-PP			NP-App	Proper Noun+NP				Sayer	At the start of his three-day tour of India boosting British trade, David Cameron, the Prime Minister, said 'some forms of tax avoidance have become so aggressive that there are moral questions that we have to answer about whether we want to encourage or allow that sort of behaviour'.	
4														NP-App	Np+Clause				Sayer	The report from the PAC - which took evidence from HMRC, the Treasury and the four largest accountancy firms - attacks tax authorities and calls for a new approach to harness public opinion to deter the use of tax avoidance schemes.	

### D. Appendix

[illegible]

### A third of Britons fail to save

#	Outer theme				Inner theme										Text	
	Tactical	Interpersonal			Pre-head	Synthetic Realisation	Quote	Element	Stance	Reader	NP	Det Noun	Set Types	Participants		
		Stance	Reader	Stance												
1					Almost one third of the UK population, equating to 15 million people, is not managing to put away anything in savings										Sayer	Almost one third of the UK population, equating to 15 million people, is not managing to put away anything in savings, a report warned today.
2								The tough economy and the need to help other family members struggling with high living costs							Token	The tough economy and the need to help other family members struggling with high living costs means that many people are living on a "financial precipse", with little or nothing to fall back on, Scottish widows said.
3								Some 31% of people			NP+App	NP+PP			Sayer	Some 31% of people said that they are not currently placing any cash aside, although this is a marginal improvement compared with a year ago when 32pc of consumers said this.
4					However,			only one in nine of those surveyed			AdvP				Semiser	However, only one in nine of those surveyed expects the economy to pick up this year and 17pc have no money to fall back on at all, the report added.
5					Of the two-thirds of people who are managing to add to their nest eggs,	Degree	Degree	32pc			NP	ProNoun			Currier	Of the two-thirds of people who are managing to add to their nest eggs, 32pc have a total pot of less than £1,000 - which would not even be enough to pay someone's average mortgage and council tax costs for a month - the report said.
6								One quarter of people surveyed who have families			NP+App	NP+PP			Sayer	One quarter of people surveyed who have families said they had handed out loans to their children, which were often just to help them cope with high living costs as wages stagnate.
7								Parents			NP	Noun			Sayer	Parents said that they had given out loans averaging £15,000 to help their children take major steps such as getting on the property ladder or going to university, showing an 11pc year-on-year increase in terms of lending.
8								Some 24pc of parents			NP+App	NP+PP			Sayer	Some 24pc of parents said this support had forced them to cut back on their own savings and one in 12 had stopped saving altogether.
9								Older generations			NP	Adv Noun			Semiser	Older generations are also feeling the strain, with grandparents saying they have lost £3,665 on average to their grandchildren.
10								Two-thirds of people across the survey			NP+App	NP+PP			Sayer	Two-thirds of people across the survey said that a general lack of any spare cash is holding back their ability to save.
11					As well as rising living costs such as food, energy bills, rents and petrol prices, the report said, many people are also struggling to make ends meet, plummeting in recent months, meaning they face an even tougher struggle to make any real returns.	Contingency	Contingency	Three on savings accounts			NP+App	NP+PP			Actor	As well as rising living costs such as food, energy bills, rents and petrol prices, the report said, many people are also struggling to make ends meet, plummeting in recent months, meaning they face an even tougher struggle to make any real returns.
12								A Government scheme launched in August has given lenders access to cheap finance to help borrowers.			NP+App	NP+Non finite			Actor	A Government scheme launched in August has given lenders access to cheap finance to help borrowers.

But analysts have said this scheme has made lenders less reliant on the credit ratings of savers' deposits, meaning that savings rates have plummeted.																			Sayer	Noun		
The choice of tax free has on the market has shone by a fifth compared with February 2012, recent research by financial information website Moneyfacts has found.																			Actor	NP+App		
The tax deals on offer this year are also seen likely to allow people to transfer savings in from older tax than those on the market in spring 2012. The website said.																			Carrier	NP+App		
High street banks have also reported a trend towards people using any extra money to pay down their debts as consumers continue to act cautiously in the difficult economy.																			Sayer	NP	Apd Noun Noun	
Iain McGowan, head of savings and investments at Scottish Widows, said: "When we face such immediate financial commitments, such as mortgage payments and day-to-day living expenses, then it is undoubtedly necessary to give these pressing needs priority."																			Sayer	NP+App	Proper Noun+NP	
"However, taking a wholly short-term view of our finances will mean we are ill-equipped for the financial needs and challenges that lie ahead in the long term."																					Quale	
This report is based on research from more than 5,000 people in December.																			Carrier	NP	Def Noun	

Bovis and Persimmon post jump in profits

#	Outer Theme				Inner Theme					Head				Participants		Text	
	Textual	Element	Stance	Reader	Pre-head	Stance	Reader	Synthetic Realization	Element	Stance	Reader	NP+NP	Proper Noun+Proper Noun	NP+PP	Se Types	Participants	
1									Both Persimmon and Bovis			NP+NP				Sayer	Both Persimmon and Bovis said their average selling price had increased by 10% in the course of last year, underlining the buildings' continued recovery.
2									Persimmon's average selling price in the second half of the year			NP+App		NP+PP		Carrier	Persimmon's average selling price in the second half of the year was £176,199, up 5pc on 2011, due to an increase in the number of larger family homes sold by Persimmon.
3									Bovis			NP		Proper Noun		Goal	Bovis had been boosted by a 5pc increase in its average sale price to £176,199, up 5pc on 2011, due to an increase in the number of larger family homes sold by Persimmon.
4									Householders			NP		Noun		Actor	Householders have sought to capitalise on the recovery of the property market thanks to snapping up cheap development land at the depths of the financial crisis, selling more expensive family homes rather than flats, and building homes in the more affluent south, where house prices have stayed strong.
5									Persimmon			NP		Proper Noun		Sayer	Persimmon said it continued to see healthy demand for traditional three and four bedroom houses across the country.
6									Builders			NP		Noun		Beneficiary	Builders have been keen to take advantage of the recovery in the market for lending and newbuild which will see a rise in mortgage lending, according to the industry.
7									We welcome the Government			Quote				Verbage	"We welcome the Government initiatives to support customer activity through FirstBuy and NewBuild. The Government's £80bn funding for lending scheme should also support an increase in mortgage lending at more attractive interest rates," said Mike Farley, Persimmon chief executive.
8									Revenue at Bovis								From a slow start in the fourth quarter of 2012, we anticipate this scheme will support a further gradual improvement in the mortgage market through 2013, mitigating to some extent the impact of the continuing uncertain economic conditions."
9									Revenue at Bovis			NP+App		Noun+PP		Actor	Revenues at Persimmon ticked up to £1.2bn from £1.15bn last time while Bovis revenues were £425.5m compared to £364.8m last time while
10									Bovis revenues			NP		Proper Noun+adj		Carrier	Bovis revenues were £425.5m compared to £364.8m last time while profits came in at £54m compared to £33m in 2011, largely thanks to a rise in completions and average selling prices.
11									Bovis Homes has made significant progress in 2012, delivering a strong backdrop of challenging, but stable, market conditions", said David Ritchie, Bovis chief executive.							Verbage	"Bovis Homes has made significant progress in 2012, delivering a strong improvement in revenue, profits and return on capital employed against a backdrop of challenging, but stable, market conditions", said David Ritchie, Bovis chief executive.
12									He			NP		Pronoun		Sayer	He added that a continuing lack of availability during the year of high loan to value mortgage products had constrained activity for new build homes.
13	But								He			NP		Pronoun		Sayer	But, he flagged up the Government's FirstBuy scheme to help people get on the housing ladder, which will provide further support to first

## D. Appendix

City refuses to panic over the loss of UK's AAA credit rating

#	Outer Theme				Inner Theme				Head				Text			
	Textual	Textual	Interpersonal	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	NP	Set types	Participants	Text
1										The City			NP-App	Det+Noun	Actor	The City refused to panic over the loss of the UK's AAA rating today as shares headed higher and sterling avoided a major sell-off.
2										The first chance for shares in London to react to the historic downgrade of the UK's credit rating agency Moody's late on Friday saw the FTSE 100 blue-chip benchmark gain 29.39 points to 6865.95.			NP+NP+Infinitive		Senser	The first chance for shares in London to react to the historic downgrade of the UK's credit rating agency Moody's late on Friday saw the FTSE 100 blue-chip benchmark gain 29.39 points to 6865.95.
3										The pound			NP	Det+Noun	Actor	The pound touched a 16-month low against the euro but largely held its ground against the dollar, staying above \$1.51 as analysts said the long-anticipated ratings cut had been priced in. The government's cost of borrowing for 10 years barely moved.
4										Glenn Urickie, senior dealer at foreign exchange specialist Moneycorp, said the market was "not reacting to the downgrade as much as it was reacting to the impact of this when it is a symptom not a cause. The key problem is that in the absence of bad news on the Eurozone people are re-examining their sterling holdings."			Proper+Noun+NP		Sayer	Glenn Urickie, senior dealer at foreign exchange specialist Moneycorp, said the market was "not reacting to the downgrade as much as it was reacting to the impact of this when it is a symptom not a cause. The key problem is that in the absence of bad news on the Eurozone people are re-examining their sterling holdings."
5										CMC Markets analyst Michael Hewson			NP+Proper+Noun		Sayer	CMC Markets analyst Michael Hewson the pound was suffering due to the Bank of England's willingness to tolerate higher inflation and call for more money printing by the Governor Sir Mervyn: "It's not the downgrade which is hurting sterling," he added.

Stop ministers seeing official stats early – MPs

Outer Theme		Inner Theme						Head		Text		
#	Interpersonal		Pre-Head			Head			Participants		Text	
	Element	Stance	Reader	Prehead	Stance	Reader	Element	Stance	Reader	Se types		
1							Ministers		directive	Noun	Actor	Ministers should no longer be allowed privileged early access to official statistics. A Commons Committee argues today.
2							A report by the House of Commons Public Administration Select Committee warns that the system whereby ministers are sent sensitive releases – including GDP figures – from the Office for National Statistics (ONS) 24 hours before they are made public risks undermining confidence in official statistics.			NP+PP	Sayer	A report by the House of Commons Public Administration Select Committee warns that the system whereby ministers are sent sensitive releases – including GDP figures – from the Office for National Statistics (ONS) 24 hours before they are made public risks undermining confidence in official statistics.
3	It is not appropriate that						ministers		directive	Noun	Carrier	It is not appropriate that ministers should have lengthy prior access to certain statistics but other interested parties do not, it says.
4							The Commons committees, which is chaired by the Conservative MP Bernard Jenkin, urges the Government to introduce legislation to scrap the pre-release system "at the earliest opportunity".			NP+Clause	Sayer	The Commons committees, which is chaired by the Conservative MP Bernard Jenkin, urges the Government to introduce legislation to scrap the pre-release system "at the earliest opportunity".
5							The recommendation			Noun	Verbiage	The recommendation will be welcomed by the chairman of the UK Statistics Authority, Andrew Dilnot, who is charged with overseeing the work of the ONS.
6							Mr Dilnot			Proper Noun	Sayer	Mr Dilnot has argued on several occasions that pre-release is damaging and should be dismantled.
7				Time-Clause			George Osborne and David Cameron			NP+NP Proper Noun+Proper Noun	Goal	The night before the ONS's first estimate of GDP growth for the final quarter of 2012 were released last month, George Osborne and David Cameron were photographed in a restaurant in Davos, Switzerland, apparently in a good mood, leading some to the assumption that the pre-released figures must have been decent.
8							they			Pronoun	Actor	In the event they showed that the UK economy declined by 0.3 per cent over the three months.
9							The committee			Det Noun	Sayer	The committee also argues that there is a potential clash between the Statistics Authority's two functions of being responsible for the production of official statistics and also regarding their quality.
10							It			Pronoun	Sayer	It recommends that there should be a greater physical separation of the workplaces of the regulators and the producers.
11							One of the responsibilities of the head of the Statistics Authority, which was established by the 2007 Statistics Act, is to correct ministers when they make misleading claims regarding official figures.			NP+PP	Carrier	One of the responsibilities of the head of the Statistics Authority, which was established by the 2007 Statistics Act, is to correct ministers when they make misleading claims regarding official figures.
							Mr Dilnot			Proper Noun	Sayer	In December Mr Dilnot revealed the Coalition for claiming that health spending had increased in real terms since 2010.

[illegible]

## Markets set for sterling sell-off after AAA setback for Britain

#	Chatter Theme			Inner Theme									Participants	Test	
	Textual	Interpersonal		Prehead	Stance	Pro-Mood	Reader	Syncretic Realization	Element	Stance	Reader	NP			Proper Noun
1									sterling			NP	Proper Noun		Goal
2									The pound			NP	Det Noun		Actor
3									The agency			Np	Det Noun		Sayer
4									Asia and European financial markets			Np+Np			Carrier
5									Simon Derrick, chief currency strategist of BNY Mellon			NP+App	Proper Noun+NP		Sayer
6								Quote	the			NP	Pronoun		Sayer
7									Andrew Sentance, a former member of the Monetary Policy Committee			NP+App	Proper Noun+NP		Sayer
8								Quote	the			Np	Pronoun		Sayer
9									Jim O'Neill, the head of Goldman Sachs' asset management arm			NP+App	Proper Noun+Np		Sayer
10								Quote	the			Np	Pronoun		Sayer

[illegible]



## **D.3 Annotated News Reports in Spanish**

D. Appendix

Una noche sin consecuencias para Straus-Kahn. (27/10/2008) Elpais online

Cl#	Thematic Field												Text			
	Outer Thematic Field				Inner Thematic Field				Engagement							
	TEXTUAL	Type	INTERP.	Type	PREHEAD	Stance	Engagement	Grat.	HEAD	Stance	Engagement	Synthetic Realization		Detailed	Semantic Realization	
1												theme			None	Un "serio error de juicio", algo "inminente", pero sin sombra de acoso, favoritismo o abuso de autoridad
2												NP+App	NP+NP		recipient	La aventura amorosa de Dominique Strauss-Kahn, director gerente del Fondo Monetario Internacional, con una empleada recién casada: el sábado (traducción del domingo en España) con la sentencia definitiva del comité ejecutivo de la institución financiera
3												NP+App	NP+NP		Actor	La relación de Strauss-Kahn con la economista húngara Preetia Nagu en enero pasado (cuando el director del FMI viajó a París para una reunión de la OPEP) fue una de las más polémicas de su vida
4												NP	Det Noun Adj		Goal	La investigación independiente encargada por la institución a un despacho de abogados sostiene que el ex ministro francés (afable, simpático y volarante para sus defensores; dilatante, aficionado a negocios poco transparentes y excesivamente hedonista, según sus detractores) se ajustó en su conducta a los estándares de la institución
5												NP+App	NP+NonFinite		payer	Para el caso de una "su promiscua" y la consecuencia para la reputación, de la institución que representa.
6	Paga	gender			Manner-PP							NP	Det Noun		payer	Los abogados de Morgan Lewis afirman que no hubo conductas reprochables más allá del "error personal" que el propio Strauss-Kahn, de 59 años, ha reconocido: no prestó a la mujer para obtener sexo ni intervino en su decisión de cambiar de trabajo ni favoreció una compensación mayor de la usual cuando abandonó el FMI en agosto.
7												NP+App	NP+Proper Noun		payer	Además, la recomendación de Strauss-Kahn para que se concediera una beca en el FMI a la hija de unos amigos, Emilie Byhet, es una "práctica común" en la institución.
8	Además	gender										Clause			Carrier	Byhet había colaborado en la reciente campaña del ex ministro para ser designado candidato socialista a la presidencia de Francia en 2007 frente a Ségolène Royal.
9												NP	Proper Noun		Actor	Poco después, Strauss-Kahn sustituyó a Rodolfo Ruiz en el FMI, con el apoyo explícito del presidente francés, Nicolas Sarkozy.
10					Poco después,							NP	Proper Noun		Actor	La clase política francesa, tan dividida en su moral sexual de la que abunda en la institución, se temió de inmediato la sombra del ex presidente del Banco Mundial Paul Wolfowitz, quien hubo de dejar el cargo en 2007 cuando se supo que había favorecido a su novia Shaha Ali Raza.
11												NP+App	NP+NP		payer	Y ello en un momento en que el convulso panorama financiero necesita, más que nunca, tranquilidad y buenos alimentos en forma de miles de millones de euros y dólares para devolver la confianza a la moneda.
12	Pero	gender										Verb	NP		Actor	En todo caso, el propio Strauss-Kahn anunció un principio de acuerdo con Ucrania para un préstamo de hasta 16.500 millones de dólares (algo más de 13.000 millones de euros) en 24 meses, condicionado a que se cumpliera una serie de requisitos sobre la regulación del sector bancario en ese país.
13	Y	gender										NP	Det Adj Noun		Carrier	No es el único que ha tenido que recurrir al FMI.
14												NP	NP		process	También Hungría, Islandia, Bielorrusia y Pakistán han pedido ayuda a Strauss-Kahn para tratar de salvar la moneda.
15												NP	Det Noun		payer	
16												Verb	NP		process	
17												NP+NP+NP+NP	NP+NP+NP+NP		payer	

El recorte de producción de la OPEP hunde el petróleo. El crudo cae a 62 dólares después del anuncio del cartel de que rebaja su oferta en 1,5 millones de barriles al día (24/10/2008) Elpais online

Cl#	Thematic Field										Text										
	TEXTUAL		INTERP.		PRE-HEAD		SYNTACTIC REALIZATION		HEAD			SYNTACTIC REALIZATION		EXPERIMENTAL RULE							
			Stance	Engagement			Stance	Engagement	Grat.	Detail		Stance	Engagement	Grat.	Detail	Experimental Rule					
1																El barrio de petróleo	NP+App	NP+App	NP+PP	Actor	El barrio de petróleo se ha desarrollado hoy en los mercados de materias primas después de que la Organización de Países Productores de petróleo haya confirmado en una cumbre extraordinaria convocada para hoy que rebaja en 1,5 millones de barriles al día su producción de crudo para adecuar su nivel de oferta a una caída de la demanda "muy profunda" de lo esperado por el colapso de la crisis que azota a las economías de los países industrializados. El precio del petróleo cayó a 62 dólares por barril en los mercados de futuros.
2																El recorte	NP	NP	Det Noun	Actor	El recorte se materializó como respuesta a los niveles de oferta de los países productores del mes de septiembre y en una efectiva a partir del 1 de noviembre.
3																A cada país	NP	NP	Pronoun	Beneficiary	A cada país se le ha asignado una cuota en proporción a su producción, de modo que los mayores recortes se les asignó a Arabia Saudí (466.000 barriles) e Irán (199.000 barriles). Venezuela dejó de producir petróleo por un tiempo.
4																los precios del petróleo	NP+App	NP+App	NP+PP	Pronoun	Para la OPEP, los precios del petróleo han sufrido un "terremoto" de nuevo tras el precedente en velozidad y magnitud "por lo que se encuentra "precipitado", lo que los niveles actuales podrían "poner muchos proyectos en peligro de suspensión o retraso", lo que generará "cortes en el suministro".
5																La actual crisis financiera	NP	NP	Det Adj Noun	Beneficiary	La actual crisis financiera, señala la organización, "está teniendo un impacto patente en la economía mundial".
6																Esta ralentización de la	NP	NP	NP+PP	Beneficiary	Esta ralentización de la demanda está "afectando" la situación del mercado, que ha obligado a estar "sobreabastecido" de petróleo, afirma.
7																la OPEP	NP+App	NP	NP+PP	Beneficiary	La OPEP explicó que el 17 de diciembre celebrará una reunión extraordinaria en la que se revisará la decisión adoptada hoy.
8																El anuncio	NP	NP	Det Noun	Actor	El anuncio ha provocado una reacción inmediata en los mercados internacionales.
9																El anuncio de la OPEP de rebaja su oferta de crudo en 1,5 millones de barriles al día	NP+App	NP+App	NP+PP	Actor	El anuncio de la OPEP de rebaja su oferta de crudo en 1,5 millones de barriles al día desde el 15 de febrero de 2007, cuando ofreció a 65,07 dólares.

“Los resultados de Microsoft brillan gracias al Windows Vista” El líder mundial de software se vuela en sus accionistas tras mejorar sus beneficios un 47% entre julio y diciembre de 2007 (25/01/2008) El país online

**“Holanda invertirá 10.000 millones en el grupo ING”** La entidad logra un acuerdo para una inversión de capital estatal (19/10/2008) El país online

Syntactic Field															
CI#	TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD	SYNTACTIC REALIZATION			Experiential Role	Text	
			Stance	Engagement		Stance	Engagement	Grail.	Detail		Stance	Engagement	Grail.	Detail	
1										El grupo holandés de banca y seguros ING		NP-PP			El grupo holandés de banca y seguros ING se consiguió un acuerdo con el Gobierno y el banco central de Holanda para fortalecer el capital del grupo financiero a las 8 crisis.
2	En concreto									las autoridades holandesas		NP		Det Noun Adj	En concreto las autoridades holandesas invierten 10.000 millones de euro en el capital según han anunciado en rueda de prensa el primer ministro, el director del banco central holandés y el director del grupo financiero ING.
3								Time-PP	Time-PP	el Estado		NP		Proper Noun	A partir de esta inversión de capital, el Estado portará los riesgos con ING que tendrán derecho a voto en decisiones "fundamentales" que se refieren a inversiones que afecten a más de 25% del capital propio de la empresa.
4										El Gobierno		NP		Proper Noun	El gobierno recibe "seguridades" por valor de 10.000 millones de euros, las cuales "tienen las mismas características" que las acciones de ING, las cuales son acciones de garantía de Fintanx, también los miembros del grupo financiero ING.
5										El presidente de ING, Michael Tannous		NP-App		NP-PP	El presidente de ING, Michael Tannous, ha asegurado en esta comparecencia que la ayuda a responder a "circunstancias excepcionales", y que la cantidad de 10.000 millones de euros es "suficientemente cómodo como para absorber las actividades del futuro", causadas por la crisis global.
6										El presidente del ING (banco holandés), Nout Wouda		NP-App		NP-PP	El presidente del ING (banco holandés), Nout Wouda, ha comentado que este apoyo financiero a ING es "una muestra de confianza" por parte de los accionistas de ING y permitirá al grupo tener "una excelente condición para continuar haciendo negocios".
7										Las acciones de la entidad		NP-App		NP-PP	Las acciones de la entidad cayeron el pasado viernes un 27%, prácticamente ante los rumores de que se estaba negociando la venta de la entidad.

8				El valor bursátil del banco	NP-App	NP-IP	Actor	El valor bursátil del banco se ha visto reducido en un 73% en lo que va año. ING tiene en España 1,7 millones de clientes, a través de su filial ING Direct, primer banco por Internet del país gracias a productos de ahorro populares como la Cuenta e el Depósito. <a href="#">Ver artículo</a>
9				se trata	NP	Proccoun	Process	El banco holandés ING Direct, primer banco por Internet del país gracias a productos de ahorro populares como la Cuenta e el Depósito. <a href="#">Ver artículo</a>
10				El pasado 4 de octubre	NP	Proper Noun	Source	El banco holandés ING Direct, primer banco por Internet del país gracias a productos de ahorro populares como la Cuenta e el Depósito. <a href="#">Ver artículo</a>
11				La inyección de capital	NP-App	NP-App	Currier	El pasado 4 de octubre, el banco holandés ING Direct, primer banco por Internet del país gracias a productos de ahorro populares como la Cuenta e el Depósito, anunció la adquisición de los activos de esta entidad en su totalidad.
12				Verb-tense	Verb	3 p sing	Process	La inyección de capital anunciada este domingo para ING supone la rpt del fondo que el gobierno de Amsterdam había anunciado para el rescate de entidades financieras en crisis.

El G7 reafirma su interés en que el sistema financiero sea "fuerte y estable". El Grupo de los Siete se muestra preocupado por la "excesiva volatilidad" del yen (27/10/2008) El país online

CI#		Discursive Field		SYNTACTIC REALIZATION ON										Text	
TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION ON		EXPERIMENTAL ROLE		Text			
		Stance	Engagement	Stance	Engagement	GrA	Detail	Stance	Engagement	GrA	Detail				
1							El G7		NP		Proper Noun	Sayer	El G7 ha informado hoy su composición, de que existía un sistema financiero internacional "fuerte y estable" que había permitido a los países ricos "ganar el dinero de la globalización". El euro, el dólar y el yen, que en los últimos meses se ha apreciado con fuerza frente al dólar y sobre todo frente al euro.		
2							El Grupo de los Siete, integrado por EEUU, Japón, Francia, Reino Unido, Italia, Canadá y Alemania, ha divulgado un comunicado de urgencia sobre los últimos movimientos en los mercados de divisas, de apenas un paréntesis.		NP+App		NP+Infinitive	Sayer	El Grupo de los Siete, integrado por EEUU, Japón, Francia, Reino Unido, Italia, Canadá y Alemania, ha divulgado un comunicado de urgencia sobre los últimos movimientos en los mercados de divisas, de apenas un paréntesis.		
3							Los países ricos, dice la nota, siguen "aguardando de cerca los movimientos y cooperando de manera prudente".		NP+App		NP+Clause	Actor	Los países ricos, dice la nota, siguen "aguardando de cerca los movimientos y cooperando de manera prudente".		
4							El G7 norte atlántico se reúne todos los años y los miembros tienen un "interés común" de que haya un sistema financiero internacional "fuerte y estable".		NP		Proper Noun	Sayer	El G7 norte atlántico se reúne todos los años y los miembros tienen un "interés común" de que haya un sistema financiero internacional "fuerte y estable".		
5							"Estamos preocupados por la reciente excesiva volatilidad en el tipo de cambio del yen y sus posibles implicaciones sobre la estabilidad económica y financiera", indican los ministros de finanzas y gobernadores de los bancos centrales del G7.		Verb		3 p plu presente indicativo	Verbalage	"Estamos preocupados por la reciente excesiva volatilidad en el tipo de cambio del yen y sus posibles implicaciones sobre la estabilidad económica y financiera", indican los ministros de finanzas y gobernadores de los bancos centrales del G7.		
6		En el último mes,			Time	Time-PP	El yen		NP		Proper Noun	Actor	En el último mes, el yen se ha apreciado el 13% frente al dólar y el 27% frente al euro, lo que ha propiciado fuertes descensos en la bolsa de Tokio y caídas en los beneficios de las empresas. El índice Nikkei de la bolsa japonesa ha caído un 2,3% y el índice S&P 500 de Wall Street a 1712,49 puntos.		
7							El dólar		NP		Proper Noun	Actor	El dólar ha bajado en Tokio y en Wall Street a 118,05 yentos, en fuerte descenso con respecto a los valores que se registraban a comienzos de año.		
8		En un año,			Time	Time-PP	El yen		NP		Proper Noun	Actor	En un año, el yen se ha apreciado el 18% frente al dólar y el 28% frente al euro, sobre todo debido a los movimientos registrados desde agosto pasado, hace apenas dos meses.		
9							Japón		NP		Proper Noun	Sayer	Japón anuncia mayor control en la venta de acciones.		
10							El primer ministro japonés, Taro Aso, ha anunciado hoy que su Gobierno buscará un mayor control de la venta de acciones a corto plazo, después de que la bolsa japonesa haya perdido alrededor de la mitad de su valor desde principios de año.		NP+App		NP+Proper Noun	Sayer	El primer ministro japonés, Taro Aso, ha anunciado hoy que su Gobierno buscará un mayor control de la venta de acciones a corto plazo, después de que la bolsa japonesa haya perdido alrededor de la mitad de su valor desde principios de año.		
11		Según la agencia nipona Kyodo, Aso también tratará de inspeccionar más fondos para capitalizar las instituciones financieras japonesas.			Evidence	Evidence-PP	Aso		NP		Proper Noun	Actor	Según la agencia nipona Kyodo, Aso también tratará de inspeccionar más fondos para capitalizar las instituciones financieras japonesas.		
12		El gobierno japonés					El mercado de valores y fortalecer el sistema financiero japonés, al proporcionar liquidez al		NP		Det Noun Adj	Verbal	El gobierno japonés, al proporcionar liquidez al mercado de valores y fortalecer el sistema financiero japonés, al proporcionar liquidez al		

**El FMI anuncia ayudas financieras para Ucrania y Hungría. El organismo llega a un principio de acuerdo con Kiev para el préstamo de 12.990 millones de euros (27/10/2008) El país online**

Thematic Field		SYNTACTIC REALIZATION										Text			
CI#	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION ON			Experiential Role	Text		
			Stance	Engagement	Stance	Engagement	Grat.	Detail	Stance	Engagement	Grat.	Detail			
1									El fondo Monetario Internacional (FMI)			NP	Proper Noun	Sayer	El Fondo Monetario Internacional (FMI) ha anunciado este domingo en un comunicado que ha llegado a un principio de acuerdo con Ucrania para el préstamo de 15.500 millones de dólares (12.590 millones de euros) para ayudar a Ucrania a superar la crisis.
2	Además								el director gerente del organismo, Dominique Strauss-Kahn,			NP+App	NP+proper Noun	Sayer	Además, el director gerente del organismo, Dominique Strauss-Kahn, ha prometido una "importante ayuda financiera" a partir de hoy para ayudar a Ucrania a superar la crisis y la economía de este país con urgencia.
3									El programa de la ayuda financiera y económica tratando la liquidez y el soporte financiero, poniendo el acento a la combinación del préstamo y reduciendo la inflación, proteger el crecimiento de Ucrania etc.			Verb	3 p sing presente indicativo	Verbage	"El programa de las autoridades busca la vuelta de Ucrania a la estabilidad financiera y económica tratando la liquidez y el soporte del sector financiero, poniendo el acento a la combinación interna y reduciendo la inflación", protege el crecimiento de Ucrania etc.
4									El programa (el acuerdo) la defendió ante el importante descenso de la producción (...) en la medida de lo posible				3 p sing presente indicativo	Verbage	"Al mismo tiempo, (el acuerdo) la defendió ante el importante descenso de la producción (...) en la medida de lo posible", indica.
5									El FMI en Kiev y las autoridades ucranianas han llegado un acuerdo para un paquete con las condiciones de devolución necesarias teniendo en cuenta los efectos combinados del derrumbe de los precios del acero y la inestabilidad del crédito internacional.			NP+NP	Actor	El FMI en Kiev y las autoridades ucranianas han llegado un acuerdo para un paquete con las condiciones de devolución necesarias teniendo en cuenta los efectos combinados del derrumbe de los precios del acero y la inestabilidad del crédito internacional.	
6									El préstamo equivale al 800% de la cuota que Ucrania aporta al FMI.				Det Noun	Carrier	El préstamo equivale al 800% de la cuota que Ucrania aporta al FMI.
7									Esta cuota es determinada para cada año en función de su peso en la economía mundial y de la cantidad de			NP	Det Noun	Carrier	Esta cuota es determinada para cada año en función de su peso en la economía mundial y de la cantidad de

[illegible]

"Barclays recibe 9.200 millones de fondos soberanos para evitar acudir al fondo de Brown" (31/02/2008) Elpais online

Cl#	Thematic Field												Text				
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION		DETAIL			Experiential Role			
			Stance	Engagement			Stance	Engagement	Gr-d.	Detail	Stance	Engagement	Gr-p	Detail	Experiential Role		
1														Proper Noun	Actor	El Barclays, recordó en su mayor parte a inversores del Golfo Pérsico para llevar a cabo una ampliación de capital de hasta 7.900 millones de libras (8.325 millones de euros). Ha comunicado hoy el banco británico el regulador de la Bolsa de Londres.	
2														NP	Det Noun	Actor	La entidad emitirá 3.000 millones de euros en opciones sobre acciones, destinadas a esos inversores árabes y otros 4.300 millones de libras (5.479 millones) en títulos convertibles de forma posterior.
3														NP+App	NP+pp	Actor	Se trata según el canal de 2.800 millones de libras (3.567 millones de euros) para el fondo del Golfo Pérsico y 1.500 millones de libras (1.919 millones) de acciones convertibles a convertirse en acciones.
4														NP+App	Proper	Actor	De este modo, el Barclays, que tras la ampliación estará controlado en más de un 30% por esos inversores del Golfo, se abre a nuevo la ayuda ofrecida por el Gobierno británico para incrementar la liquidez de los bancos del país.
5														NP+NP+NP	NP+NP+NP	Actor	Frente al rechazo de Barclays, Royal Bank of Scotland, Lloyds y HBOS, se han accedido a acudir a este

**La crisis financiera lleva a Estados Unidos al borde de la recesión (31/10/2008)** Elpais online

CI#		Thematic Field		SYNTACTIC REALIZATION										Text	
FENIXIAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION ON			Experiential Role		Text		
		Stance	Engagement	Stance	Engagement	Detail	Stance	Engagement	Detail	Stance	Engagement	Detail			
1.										NP			Estados Unidos dio ayer el primer paso hacia la recesión.		
2.										NP			El producto interior bruto (PIB) registró una contracción del 0,2% en el tercer trimestre del año. Equivalente a un ritmo anual de caída del 0,2%, el peor dato desde la última crisis económica en 2001.		
3.										NP-App			El motivo de este frenazo radica en la mayor caída del consumo -motor de la primera economía del mundo- y en la pérdida de confianza por parte de la ciudadanía ante la devaluación de la moneda, en medio de la peor crisis financiera desde la Gran Depresión.		
4.										NP			Los economistas anticipan ya que la contracción continuará hasta la primavera, es el que será el período más prolongado de recesión desde la II Guerra Mundial. El paro podría alcanzar entre el 8% y el 9%.		
5.										NP			Con todo, el frenazo es menor de lo que prevén los analistas, y quizá por eso las Bolsas estadounidenses no muestran demasiada preocupación.		
6.	Pero									NP			Este período se produjo sobre todo en septiembre, cuando el caso volvió a dominar los mercados financieros tras el rescate de las hipotecas de la Fannie Mae y Freddie Mac, el colapso del banco de inversión Lehman Brothers y la intervención pública en la aseguradora AIG.		
7.										NP			Este fenómeno se debe al hecho de que Estados Unidos creció a un ritmo anual del 2,2% en el segundo trimestre.		
8.		Con el pretexto de la libertad de expresión, las restricciones para acceder al crédito y el desplome de las inversiones,	Manner	Manner PP-NP						NP			Con el pretexto de la libertad de expresión, las familias no tienen otra salida en la presente coyuntura que mirar a sus bolsillos y recluirse allí donde pueden para hacer frente a un futuro incierto, sobre todo en el empleo.		
9.										NP-App			El gasto de los consumidores, del que dependen dos tercios partes del crecimiento, cayó un 3,1%.		
10.		Es	Verb	Web-ser						NP			Es el primer recorte trimestral en el consumo desde el último trimestre de 1991 y el mayor desde la primavera de 1980.		
11.										NP			El Departamento de Comercio advierte que esta caída en el gasto de los hogares restó 2,25 puntos al PIB.		
12.	Y	es	verb	Verb-ser						NP			Y es que las finanzas de las familias están bajo una tensión considerable, por la continua pérdida de empleo y de cada día de las inversiones.		
13.										NP-App			El dato del crecimiento económico revela en este sentido que los ingresos personales cayeron un 8,7% en el tercer trimestre, el peor dato desde que se empezó a utilizar este indicador en 1947 y que contrasta con el incremento del 19,1% del segundo trimestre, cuando experimentó a favor los primeros estímulos fiscales al crecimiento.		
14.										NP			Los hogares no son los únicos que miran hacia el último dólar en la coyuntura actual.		
15.										NP			La empresa está buscando como ahorrar por todos los frentes.		
16.										NP			La del tercer trimestre fue la primera reducción en el gasto desde final de 2006, del 1%.		

#### D. Appendix

[illegible]

**El Eurogrupo aprobará hoy 1.865 millones de euros para la banca española**

C #	Framete Field	INTERP.	PRI-HEAD	SYNACTIC REALIZATI ON	HEAD	SYNACTIC REALIZATI ON	Text
	TEXTUAL	Stance	Engagement	Stance	Engagement	Detail	Experiential Role
1							Los ministros de Economía y Finanzas de la zona del euro (Eurogrupo) darán este lunes su visto bueno al desembolso del segundo tramo de la ayuda financiera a España, un total de 1.865 millones de euros, para sanear la banca no nacionalizada.
2							Todo apunta a que el Eurogrupo dará su visto bueno a la operación, que será confirmada por el consejo de administración del Mecanismo Europeo de Estabilidad (MEDE) el 28 de enero y que beneficiará a Iberbank, Caja3, BMN y España-Duero (CEISS).
3							La operación conlleva unas estrictas condiciones impuestas por Bruselas, que prevén que Caja3 se integre en Ibercaja y que imponen una importante reducción del balance de las otras tres entidades para 2017.
4							Otro de los puntos destacados de la agenda del Eurogrupo es la recapitalización directa de la banca, sobre la que los diecisiete tendrán una discusión política en la que abordarán la delicada cuestión de los activos heredados, es decir, la posibilidad de aplicar la recapitalización directa de manera retroactiva.
5							Los ministros estudiarán si los países que accedan a la recapitalización directa de la banca desde los fondos de rescate europeos tendrán que aportar parte de la financiación, según fuentes comunitarias, una participación que podría situarse entre el 5 % y el 15 %.
6							Retroactividad.
7							La retroactividad mantiene enfrentados a los países favorables a que se permita de una manera completa -para que los rescates dejen de computar como deuda soberana- y a los que, liderados por Alemania, prefieren que tenga un alcance limitado y sujeto a condiciones.
8	Además,						Además, el Eurogrupo deberá elegir mañana a su nuevo presidente, en sustitución del luxemburgués Jean-Claude Juncker, un puesto para el que solo se ha presentado un candidato, el ministro de Finanzas holandés, Jeroen Dijsselboem, y que, según fuentes diplomáticas, tiene el camino bastante despejado.
9	Tampoco						Tampoco se descartará, no obstante, que las reticencias planteadas por Francia, que exige que se desarrolle un debate "en profundidad" sobre sus propuestas para el nuevo mandato, retrase finalmente ese nombramiento.
10							La petición de un rescate financiero de Chipre también será discutida por los ministros de la zona del euro, aunque la decisión se retrasará hasta después de las elecciones que se celebrarán en el país a finales de febrero, según las mismas fuentes comunitarias.
11							La decisión final no se tomará bajo el Gobierno actual, señalaron las fuentes, que añadieron que esperan que "en algún momento en torno a mediados de marzo haya posibilidad de llegar a un acuerdo sobre el programa".

Grecia	NP	Proper Noun	Carrier	Grecia también forma parte de la agenda del Eurogrupo, pero no se esperan decisiones al respecto, sino que los Diecisiete harán una puesta en común sobre la situación y el avance de la agenda reformista del Gobierno de Atenas para cumplir con los compromisos adquiridos a cambio del segundo rescate financiero.
--------	----	-------------	---------	--

## inspectores sin fronteras

[illegible]

La economía china crece un 7,8% en 2012, su menor subida en 13 años

Thematic Field				Text
CI #	INTERP.	PREHEAD	SYNTACTIC REALIZATION	
	TEXTUAL		HEAD	SYNTACTIC REALIZATION ON
				Experimental Role

Grecia	NP	Proper Noun	Carrier	Grecia también forma parte de la agenda del Eurogrupo, pero no se esperan decisiones al respecto, sino que los Diecisiete harán una puesta en común sobre la situación y el avance de la agenda reformista del Gobierno de Atenas para cumplir con los compromisos adquiridos a cambio del segundo rescate financiero.
--------	----	-------------	---------	--

## inspectores sin fronteras

[illegible]

La economía china crece un 7,8% en 2012, su menor subida en 13 años

Thematic field				Text	
CI #	INTPR.	PREHEAD	SYNTACTIC REALIZATION	HEAD	SYNTACTIC REALIZATION
	TEXTUAL				Experimental Rule



1										La economía china		NP				Actor	La economía china creció un 7,8% en 2012, su menor expansión en 13 años, con una subida del producto interior bruto (PIB) del 7,9% en el último trimestre del año.
2										La cifra		NP				Carrier	La cifra está muy por debajo del 9,3% de crecimiento que la segunda economía mundial registró en 2011, y aún más lejána al 10,4% que alcanzó en 2010.
3										se encuesta	Verb					Process	No obstante, se encuentra por encima del 7,5% de crecimiento fijado por el régimen comunista durante la Asamblea Popular de marzo del pasado año, por lo que los analistas creen que allana el camino para que se reponga durante este año.
4										Por encima del 10%, la economía se ha estado estabilizando	Verb					Verbage	"Por encima de todo, la economía se ha estado estabilizando", aseguró el Buró en el comunicado divulgado este viernes.
5										La ciudad institución	NP					Player	La ciudad institución señaló que la ralentización del crecimiento se produjo en un contexto de "miedo en el exterior y alicción doméstica", en referencia a la crisis de deuda que aún arrastran sus socios europeos y estadounidenses y a su efecto en el país asiático.
6										el Buró	NP					Player	Con forma simultánea, el Buró informó que el PIB creció un 7,9% en el último trimestre del año, impulsado por las medidas de estímulo del Gobierno, lo que terminó con la tendencia de declive de los siete trimestres previos.
7										El PIB	NP					Actor	En 2012, el PIB alcanzó 51.093 billones de yuanes (8.28 billones de dólares, 6,19 billones de euros).
8										el Buró	NP					Player	Por otra parte, el Buró también anunció, entre otras, las cifras de la producción industrial china, que ralentizó su crecimiento hasta un 10 % en 2012, 3,9 puntos porcentuales por debajo de lo alcanzado en 2011.
9										las ventas al por menor	NP					Actor	Asimismo, las ventas al por menor registraron una subida del 15,2% en diciembre, en contraste con el 14,9 % de noviembre, lo que supuso un repunte final del 14,3 % en 2012, aun así, 2,8 puntos porcentuales por debajo de 2011, rebajó la citada fuente.
10										la inversión china en el sector inmobiliario	NP+App					Carrier	En cambio, la inversión china en el sector inmobiliario aumentó un 16,2% en 2012 con respecto al año previo, pese a la campaña gubernamental para desinflar la burbuja inmobiliaria.
11										El sector inmobiliario	NP					Token	El sector inmobiliario es el más importante de la economía china, ya que supone más del 10% del cómputo total del PIB.
12										los principales mercados bursátiles chinos	NP					actor	Antes de que se publicaran estas cifras, los principales mercados bursátiles chinos abrieron optimistas, con la Bolsa de Shanghai subiendo un 0,48 % y la de Hong Kong aumentando un 0,73 en la apertura de sesión.

**Banco Popular gana 133 millones con la venta de su negocio de recobro**

CI#	Thematic Field		SYNTACTIC REALIZATION										Text				
	TEXTUAL	INTERP.	Stance	Engagement	PRE-HEAD	Stance	Engagement	Grat.	Detail	HEAD	Stance	Engagement		Grat.	Detail	SYNTACTIC REALIZATION	
1													NP	Proper Noun	Actor	Banko Popular ha llegado a un acuerdo con el Grupo EOS para la venta de la actividad en España del recibo de deuda incumplida por cuenta de terceros, de 0 a 90 días.	
2													NP	Det Noun	Carrier	La operación está valorada en 135 millones de euros y reportará a la entidad plusvalías cercanas a los 133 millones de euros.	
3					Según ha informado este lunes el banco que preside Ángel Ron a la Comisión Nacional del Mercado de Valores (CNMV), esta transacción se enmarca dentro de su plan de negocio 2012-2014.					Evidence Clause			NP	Det Noun	Carrier	Según ha informado este lunes el banco que preside Ángel Ron a la Comisión Nacional del Mercado de Valores (CNMV), esta transacción se enmarca dentro de su plan de negocio 2012-2014.	
4					De acuerdo con los términos de la operación, el Banco Popular, que para esta operación ha sido asesorado por Atlas DC Advisory, transmitirá a Lykeco Inversiones 2013, entidad del Grupo EOS, el negocio del Centro de Análisis y Recuperación de Incumplimientos, y ésta se encargará del recibo de las deudas, en exclusiva, durante un periodo inicial de diez años.					Evidence NP			NP+App	NP+Clause	Actor	De acuerdo con los términos de la operación, el Banco Popular, que para esta operación ha sido asesorado por Atlas DC Advisory, transmitirá a Lykeco Inversiones 2013, entidad del Grupo EOS, el negocio del Centro de Análisis y Recuperación de Incumplimientos, y ésta se encargará del recibo de las deudas, en exclusiva, durante un periodo inicial de diez años.	
5														NP+App	NP+App	Carrier	EOS, perteneciente al grupo empresarial de servicios financieros OTO, ya es proveedor del Banco Popular tras haber suscrito en 2009 un acuerdo para la adquisición de la filial de recibo del Banco Pastor, acuerdo que continuará vigente, según ha indicado la entidad.
6					Tras hacerse pública la transacción, las acciones de la entidad subían más de un 4% en Bolsa, después de ser negativas en la última sesión.					Time PP			NP+App	NP+App	Actor	Tras hacerse pública la transacción, las acciones de la entidad subían más de un 4% en Bolsa, después de ser negativas en la última sesión.	

La Caixa anticipa una 'gradual recuperación' de la situación económica

CI #	Phonetic Field			Text
	INTERP.	PRE-HEAD	HEAD	
	TEXTUAL			SYNTACTIC REALIZATION ON Experimental Rule

### D. Appendix

[illegible]

**aguarde pide al BCE y a los Gobiernos que no se relajen para evitar una recaída**

Thematic Field													Text	
CI #	TEXTUAL	INTERP.	PRE-HEAD		SYNTACTIC REALIZATION		HEAD	SYNTACTIC REALIZATION ON		Experiential Role				
		Stance	Engagement		Stance	Engagement	Grat.	Detail		Stance	Engagement	Grat.	Detail	
1												NP+App	Proper Noun+NP	Christine Lagarde, la directora gerente del Fondo Monetario Internacional, aseguró este jueves que las economías de la zona euro han "evitado el colapso", pero ni los gobiernos ni los bancos centrales deben "relajarse" a causa de "una recesión".
2												NP	Proper Noun	Lagarde se queja, en particular, de que faltan "planes a medio plazo" y de que las decisiones tardan demasiado en tomarse.
3							Time	Time PP				NP	Def+Noun Pro+Noun Adj	En su primera rueda de prensa del año, la ex ministra de Economía francesa alabó la acción de los bancos centrales como la clave que paró el hundimiento mundial y animó al Banco Central Europeo a que siga interviniendo en el mercado.
4	En cambio			fine			Verb					Verb		En cambio, fue menos positiva al hablar sobre los gobiernos, que han tomado algunas decisiones correctas, pero "a menudo en el último momento" y con una visión solo "a corto plazo" que no redujera el peso de la deuda pública en las

La 'Troika' concederá a Grecia seis meses en los que no le reclamará más medidas

C#	Thematic Field		SYNTACTIC REALIZATION										Text	
	TEXTUAL		INTERP.		PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION			EXPERIMENTAL ROLE
			Source	Engagement		Detail	Source	Engagement	On.	Detail				
1									NP-App	NP-MonCoinfin	Actor			La Troika de acreedores formada por la Comisión Europea, el Banco Central Europeo y el Fondo Monetario Internacional ha decidido otorgar a Grecia una moratoria de seis meses durante los que no le exigirá nuevas medidas de austeridad.
2			Según informaciones del diario griego "kathimerini"		Evidencia	Evidencia-PP			NP-App	NP-PP	Carrier			Según informaciones del diario griego "kathimerini", el objetivo de esta moratoria es dar a Grecia un respiro para poder llevar a la práctica las reformas fiscales y estructurales ya comprometidas.
3									NP	Det Houn	Carrier			El diario basa sus informaciones en un alto funcionario del Ministerio de Finanzas, quien asegura que la decisión se tomó la semana pasada en una reunión de la Troika.
4									NP-App	NP-PP	Carrier			Este hecho no implica un relajamiento de las condiciones pactadas, sino que persigue, por el contrario, dar a Grecia la posibilidad para que cumpla estrictamente con lo acordado, y sus problemas de deuda no se conviertan en tema de la campaña electoral en Alemania, donde en septiembre se celebran comicios legislativos.

### D. Appendix

5							Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.	NP+Classé	Journal	Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.
6	Según señaló el año pasado, el gobierno de "Kathimerini", la troika está disgustada porque el año pasado Hacienda solo ha podido llevar a cabo el 30% de las inspecciones previstas.	Evidence	Evidence+Classé	la troika	NP	Proper Noun	Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.	NP	Journal	Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.
7	Esta semana el Fondo Monetario Internacional desbarró 3.240 millones de euros dentro del plan de rescate a Grecia.	Time	Time+NP	el Fondo Monetario Internacional	NP	Proper Noun	Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.	NP	Journal	Entre las medidas en las que la troika quiere ver avances figura la lucha contra el fraude fiscal, el cumplimiento de los objetivos de recaudación y la privatización de empresas públicas.
8				La directora gerente del FMI, Christine Lagarde, destacó que "el programa está avanzando en la dirección adecuada, con notables progresos en el ajuste fiscal y mejora en la competitividad de sus costes laborales".	NP+App	NP+Proper Noun	La directora gerente del FMI, Christine Lagarde, destacó que "el programa está avanzando en la dirección adecuada, con notables progresos en el ajuste fiscal y mejora en la competitividad de sus costes laborales".	NP+App	Journal	La directora gerente del FMI, Christine Lagarde, destacó que "el programa está avanzando en la dirección adecuada, con notables progresos en el ajuste fiscal y mejora en la competitividad de sus costes laborales".
9	No obstante, Lagarde señaló que "queda mucho por hacer", sobre todo en la reducción de las "barreras a la competitividad" y en la puesta en práctica de los planes de privatización.			Lagarde	NP	Proper Noun	No obstante, Lagarde señaló que "queda mucho por hacer", sobre todo en la reducción de las "barreras a la competitividad" y en la puesta en práctica de los planes de privatización.	NP	Journal	No obstante, Lagarde señaló que "queda mucho por hacer", sobre todo en la reducción de las "barreras a la competitividad" y en la puesta en práctica de los planes de privatización.

8CE: la confianza de los mercados financieros ha aumentado notablemente

CI#	Thematic Field		SYNTACTIC REALIZATION										Text	
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION		EXPERIENTIAL ROLE			
			Source	Engagement		Source	Engagement		Source	Engagement	Goal	Detail		
1											NP App	NP App	Actor	La confianza de los mercados financieros ha aumentado "de forma significativa" y algunos indicadores coyunturales se han estabilizado aunque en niveles reducidos informa el Banco Central Europeo (BCE) en su boletín mensual de enero.
2											NP	Proper Noun	layer	El BCE prevé que "más avanzado 2013 debería comenzar una recuperación gradual, a medida que la orientación acomodaticia de la política monetaria, la gradual mejoría de la confianza de los mercados financieros y la menor fragmentación empiezan a mostrar su efecto sobre la demanda interna privada".
3	Audiencia										NP App	NP App	Carrier	Además, el fortalecimiento de la demanda externa debería sostener el crecimiento de las exportaciones.
4											NP App	NP App	Actor	Los riesgos en torno a las perspectivas económicas de la zona del euro siguen a la baja.
5											NP	Proper Noun	Senser	El BCE considera necesarias "nuevas contracciones en otras áreas de la política económica para garantizar una sólida estabilización de los mercados financieros y una mejora de las perspectivas de crecimiento".
6											Verb	Verb	Process	Deberán aplicarse con rapidez reformas estructurales adicionales destinadas a aumentar la flexibilidad, el dinamismo y la competitividad de la economía de la zona del euro, según la entidad monetaria.
7											NP	Proper Noun	layer	El BCE destaca "las reformas de los mercados de productos encaminadas a fomentar la competencia y la competitividad, así como medidas para mejorar el funcionamiento de los mercados de trabajo".
8											NP	Det Noun	Actor	Estas reformas elevarán el potencial de crecimiento y el empleo en la zona del euro y mejorarán la capacidad de ajuste de los países de la zona.
9											NP	Proper Noun	layer	El BCE señala que "la reciente disminución sustancial de los rendimientos de los bonos soberanos debería verse regalada por nuevos progresos en materia de saneamiento presupuestario conforme a los compromisos adquiridos en el marco del Pacto de Estabilidad y Crecimiento".

**Catalunya Banc y NCG sufrirán pérdidas de hasta 20.000 millones en 2012**

CI#	Dynamic Index		SYNTACTIC REALIZATION				IRL AD	Engagement				Text
	TEXTUAL	INTERP.	IRL HEAD	Source	Engagement	Grat.		Source	Engagement	Source	Engagement	
1												Catalunya Banc, NCG Banco, dos de las cuatro entidades españolas nacionalizadas, convalidó el plan de 2012 con unas pérdidas de hasta 20.000 millones de euros, según los planes de reestructuración publicados hoy por la Comisión Europea.
2												La oficina del comisario europeo de Competencia, Joaquín Almunia, hizo pública hoy la versión no confidencial de los planes de reestructuración de estas dos entidades, aprobados el pasado 28 de noviembre por Bruselas, pero aún no ha divulgado los informes de los otros dos bancos nacionalizados, Bankia y Banco de Valencia.
3												Según ese documento, Catalunya Banc registrará unas pérdidas antes de impuestos de entre 11.000 y 12.000 millones de euros en 2012, pero reducirá esas cifras, ya que en 2013 hasta 2015 entre 700 y 1.200 millones de euros, como se puede ver en el informe de 2014. El terreno a terreno positivo en 2015 con un beneficio antes de impuestos de 200 a 300 millones de euros.
												Según el informe de Catalunya Banc, el plan de reestructuración de la entidad

### D.3 Annotated News Reports in Spanish

[illegible]

Las pérdidas de Unnim Banc casi se duplican en 2012

CIB	Dynamic Business		SYNCRITIC REALIZATION				REALD		Engagement			Experiential Risk		Text
	EXTERNAL	INTERNAL	Source	Engagement	Source	Engagement	Goal	Source	Engagement	Source	Engagement	Source	Engagement	
1								Unim Banc, entidad propietaria del BBVA.		NP-App	Proper Noun-PP		Actor	Unim Banc entidad propiedad del BBVA, registro en 2012 unas pérdidas del 864 millones de euros, un 84 % más respecto al año anterior. En 2013, los resultados de la entidad mejoraron, con unos beneficios de 273,3 millones de euros, un 53 % más respecto al de 2012. En 2014, el resultado del grupo ha caído un 38,5 %, hasta los 198 millones, ha informado la entidad a la Comisión Nacional del Mercado de Valores (CNMV).
2								Unim		NP	Proper Noun		Actor	Por su parte, el resultado de explotación de Unim Banc ha arrojado al cierre del último ejercicio unas pérdidas de 909,5 millones de euros, casi el triple que el año anterior, cuando el resultado también fue negativo, aunque en 368,2 millones.
3								el resultado de explotación de Unim Banc		NP-App	NP-PP		Actor	En los últimos doce meses, Unim ha realizado dotaciones por importe de 474,2 millones, cinco veces más que en 2011, para sanear su cartera inmobiliaria, cubrir el deterioro de activos financieros y por los costes de reestructuración laborales, entre otros apartados.
4							Time	Unim	Time-PP	NP	Proper Noun		Actor	En este período, la entidad ha logrado reducir casi en un 7 % los gastos de personal, que se han situado en 382,1 millones de euros, frente a los 196,4 millones de 2011, mientras que los gastos generales de administración han bajado un 14,4 %, hasta los 63,4 millones.
5							Time	la entidad	Time-PP	NP	Det Noun		Actor	El BBVA completó en julio del pasado año la compra del 100 % de Unim Banc, la entidad que fue nacionalizada y subastada por el Estado, tras haber obtenido las autorizaciones pertinentes de las autoridades comunitarias.
6								El BBVA		NP	Proper Noun		Actor	El BBVA, que ha mantenido la marca Unim en las oficinas, se adjudicó la entidad catalana a principios de 2012 a un precio de un euro.
7								Unim en las oficinas, el grupo que preside Francisco González ganó 11,6 millones en 2012, un 44,2 %		NP-App	Proper Noun		Actor	
								El grupo que preside		NP-App	NP-PP		Actor	
								El grupo que preside		NP-App	NP-PP		Actor	

### D. Appendix

9	A finales del pasado año				BBVA y los representantes de sus trabajadores				BBVA y los representantes de sus trabajadores pactaron la salida de 1.169 empleados de Unim, con el fin de dar cumplimiento a los ajustes exigidos por Bruselas tras la adquisición de la entidad catalana por parte del banco.
10			Time	Time-pp					El plan, que incluía el cierre de más de 330 oficinas, contemplaba además prejubilaciones, relocalaciones y bajas incentivadas que se consumirán progresivamente hasta mediados de 2014.

**El banco francés Crédit Agricole multiplica por cuatro sus pérdidas**

C1#	Thematic Field												
	TEXTUAL	INTERP.	PRE-HEAD		SYNTACTIC Realization		HEAD	Engagement			Experiential Role	Text	
			Source	Engagement	Source	Engagement	Grat.		Source	NP+App	NP+Prope P Noun		
1									El banco francés Crédit Agricole			Actor	El banco francés Crédit Agricole registró una pérdida neta de 6.471 millones de euros en 2012, una cifra 4,4 veces superior al negativo de 1.470 millones de euros del año anterior, lo que le ha llevado a anunciar un nuevo plan de inversión para 2016.
2													
3							Time	Time Advpp	el banco	NP	Det Noun	Actor	Solo en el tercer trimestre, según un comunicado difundido hoy por la entidad, el banco perdió 3.982 millones de euros, un 29,8 por ciento más que en el mismo periodo del año anterior.
4									El descenso en su resultado anual	NP+App	NP+PP	Carrier	El descenso en su resultado anual, según el grupo, deriva tanto de los resultados de los últimos tres meses como del impacto vinculado a las decisiones tomadas en los trimestres anteriores, como la venta de su filial griega Emporiki al banco Alpha Bank.
5									El ejercicio 2012	NP	Det Noun	Goal	El ejercicio 2012 estuvo marcado por las medidas tomadas para adaptarse "al nuevo entorno financiero y regulatorio", detalló el banco, que cita entre otras el recorte de su actividad operativa o la cesión de algunas de sus actividades.
6									El producto neto bancario de Crédit Agricole	NP+App	NP+PP	Actor	El producto neto bancario de Crédit Agricole cayó en 2012 un 15,8 por ciento, hasta 16.315 millones de euros, mientras que el resultado de explotación se desplomó un 80,2 por ciento, hasta 542 millones.
7									El banco	NP	Det Noun	Sayer	El banco precisó no obstante que a lo largo de 2012 reforzó su solidez financiera, y que la proporción entre el capital de máxima calidad y los activos ponderados por riesgo (Core tier 1 ratio) se situó en el 9,2 por ciento, frente al 8,6 por ciento registrado a finales de diciembre de 2011.
8	Asimismo								añadió				Asimismo, añadió que aquí a 2016 pondrá en marcha un programa de ahorro de gastos estimado en 650 millones de euros que afecta al sector informático, inmobiliario y de inversiones, y sobre el que no precisó el impacto eventual que esos planes podrían tener sobre el empleo.

**Hollande reconoce que Francia crecerá menos de lo previsto en 2013**

Thematic Field		C14									
TEXTUAL	INTERP.	PREHEAD	STANCE	ENGAGEMENT	SYNTACTIC REALIZATION	HR AD	STANCE	ENGAGEMENT	EXPERIMENTAL ROLE	Text	
1						El presidente de Francia, François Hollande				El presidente de Francia, François Hollande, ha anunciado que la eurozona, según se calcula, cumpliría su predicción de crecimiento para 2013, que recoge su crecimiento del producto interior bruto (PIB) del 0,5%.	
2						Para el año 2013 todo el mundo sabe que no alcanzaremos el 0,8% que estaba previsto", afirmó Hollande en una comparecencia conjunta en Atenas con el primer ministro griego, Antonis Samaras.				"Para el año 2013 todo el mundo sabe que no alcanzaremos el 0,8% que estaba previsto", afirmó Hollande en una comparecencia conjunta en Atenas con el primer ministro griego, Antonis Samaras.	
3		A este respecto			Matter: PP	informó				A este respecto, informó de que la nueva predicción de crecimiento se establecerá a finales de mes de marzo por parte del Alto Consejo de las Finanzas Públicas y determinará lo que tendrá que hacer Francia en materia presupuestaria en 2014, según las recortes de la Comisión Europea.	
4	Aun así	recalcó			Verb	recalcó				Aun así, recalcó que Francia y Grecia no se encuentran en la misma situación que países de la eurozona como España, Portugal o Italia, que están en recesión.	
5						"Francia es uno de esos países en que, en términos de crecimiento, la actividad está en una situación menos mala"					
6						El presidente francés				"Francia es uno de esos países en que, en términos de crecimiento, la actividad está en una situación menos mala". El presidente francés subrayó que en Europa es necesario más crecimiento, una condición para que se puedan cumplir los objetivos de déficit público.	
7		Así,			Manner	recalcó				Así, recalcó que es la confianza, tanto de inversores como de consumidores, la que permite ese crecimiento.	
8	Por otra parte					Hollande				Por otro parte, Hollande mostró su apoyo a Grecia, que considera debe ser respaldada por toda Europa, especialmente a través de la rápida movilización de fondos estructurales y del pacto de crecimiento.	

## D.3 Annotated News Reports in Spanish

9	Además								afirmó				Verb	3		Process	Además, afirma que desea que el país permanezca en la zona euro, y que nunca ha dejado de convencer de ello a sus socios.
10													Verb	3		Process	En este sentido, indicó en que los esfuerzos realizados por Grecia "son considerable" y resaltó que ningún pueblo de Europa ha soportado una prueba de estas características.
11													Quote			Verbiage	"Debemos estar a su lado", agregó.

Bolivia promete a España llegar a una "solución satisfactoria" para indemnizar a las empresas apropiadas

CT#	Thematic Field	PREHEAD	INTERP.	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Experiential Role	Text
	TEXTUAL					Gram.					
1											El Gobierno boliviano se ha comprometido ante el ministro de Asuntos Exteriores y de Cooperación, José Manuel García-Margallo, a llegar a una "solución satisfactoria" para las empresas apropiadas en el país europeo de la zona euro.
2											El jefe de la diplomacia española ha conversado esta tarde con el ministro boliviano de la Presidencia, Juan Ramón Quintana, quien ha reconocido "retrasos" en los procesos iniciados para pagar un justiprecio a Red Eléctrica e Iberdrola, afectadas por sendas nacionalizaciones en mayo y diciembre pasados.
3											Quintana ha prometido al ministro, siempre según las mismas fuentes, la voluntad del Ejecutivo de Evo Morales de "acelerar" estos dos procesos y de llegar en los tres casos, también en lo que se refiere a la reciente expropiación de Sabsa, filial de Abertis, a una "solución satisfactoria".
4											La conversación entre ambos se produce después de que el Gobierno español haya amenazado en un comunicado oficial con "replantearse el conjunto de las relaciones bilaterales" con Bolivia tras esta tercera nacionalización de una empresa española realizada sin previo aviso y de una forma que ha irritado sobremanera al Gobierno español, que ve innecesario haber ordenado a las fuerzas de seguridad custodiar las instalaciones de las empresas para garantizar la ejecución de la medida.
5	De hecho										De hecho, la diplomacia española llegó a pedir públicamente a las autoridades bolivianas "más cuidado en las formas" a la hora de proceder a estas nacionalizaciones, pero el Gobierno de Morales hizo caso omiso y tras anunciar la nacionalización de Sabsa desplegó tropas del Ejército en los aeropuertos gestionados por la filial de Abertis.
6											El Gobierno ha reiterado en diversas ocasiones que no cuestiona el derecho soberano de un país de nacionalizar empresas en sectores estratégicos, pero lo que sí exige es que se pague una indemnización justa a la compañía afectada por la medida, lo que no ha ocurrido hasta la fecha con ninguna de las empresas españolas afectadas.
7											La expropiación de Sabsa ha sido la gota que ha colmado el vaso y lo que ha llevado al Gobierno a amenazar con revisar el conjunto de las relaciones bilaterales con Bolivia, sin descartar medidas en materia de cooperación al desarrollo.
8	De hecho										De hecho, la delegación del PP en el Parlamento Europeo ha registrado sendas preguntas en la Eurocámara dirigidas a la Comisión y al Consejo de la UE, con el fin de que estas dos instituciones valoren si "se pueden mantener" los programas de cooperación al desarrollo con Bolivia así como el acuerdo preferencial en materia de aduanas si La Paz prosigue con sus "ataques" a empresas europeas.
9											Mañana a las 17.00 horas la embajadora de Bolivia en España ha sido convocada a una reunión en el Ministerio de Exteriores para que dé "explicaciones" al secretario de Estado de Cooperación Internacional y para Iberoamérica, Jesús García, sobre esta última nacionalización.
10											En función de esas explicaciones, el Gobierno español "actuará en consecuencia", en palabras del propio García-Margallo.
11											Desde que llegó al poder en 2006, el presidente boliviano, Evo Morales, ha nacionalizado una veintena de empresas de distintos sectores, principalmente el energético, que estaban en manos de compañías extranjeras, incluidas españolas.

Alemania alcanza nuevo récord histórico de empleo en el último trimestre 2012

CT#	Thematic Field	PREHEAD	INTERP.	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Experiential Role	Text
	TEXTUAL					Gram.					
1											Alemania alcanzó un nuevo récord histórico de empleo en el último trimestre de 2012 con un total de 41,9 millones de personas cotizando en el sistema de seguridad social alemán, anunció este martes la Oficina Federal de Estadística (Destatis).
2											Se trata de la cifra más alta de empleo en Alemania en comparación con todos los trimestres desde la reunificación del país en octubre de 1990, según subrayaron los técnicos de Destatis desde su sede en Wiesbaden, en el centro del país.

### D. Appendix

[illegible]

otín y Sáenz cobraron 3 y 8,2 millones en 2012, casi un 30% menos

Lexemic Field		SYNTACTIC REALIZATION										Experiential Role		Text
CI #	TEXTUAL	INTERP.	Source	Engagement	PREHEAD	Source	Engagement	Grnd.	HEAD	Source	Engagement	NP+Prop or Noun	Beneficiary	
1										El presidente del Banco Santander, Emilio Botín, recibió en 2012 una atribución total de 9 millones de euros, lo que supone una reducción del 52% respecto a un año antes, en tanto que el consejero delegado, Alfredo Sáenz, cobró 8,2 millones, el 25% menos.		NP+Prop or Noun	Beneficiary	
2			Así		Manner				Manner-Advb	se		NP	Sayer	Así se indica en el Informe de la Comisión de Nombramientos y Retribuciones publicado hoy por la entidad, en el que se pone de manifiesto que en 2012, la remuneración total del consejo de administración del Banco Santander alcanzó 27,52 millones de euros, el 34,6 % menos que los 41,29 millones que percibieron en 2011.
3										Los descuentos en la retribución de Botín y Sáenz		NP+App	Carrier	Los descuentos en la retribución de Botín y Sáenz se deben sobre todo, a que en ambos casos, la retribución variable se ha reducido en un 50 %.
4			Según los mismos datos publicados hoy por la entidad del resto de consejeros ejecutivos del Banco,		Evidence-NonFinite	Evidence				el vicepresidente del grupo, Matías Rodríguez Inciarte		NP+Prop or Noun	Beneficiary	Según los mismos datos publicados hoy por la entidad, del resto de consejeros ejecutivos del Banco, el vicepresidente del grupo, Matías Rodríguez Inciarte, cobró 5,1 millones de euros, el 14,31 % menos.
5			Por su parte,		Matter-pp	Matter				Ana Patricia Botín, consejera delegada del banco Santander UK, la filial británica Santander,		NP+App	Beneficiary	Por su parte, Ana Patricia Botín, consejera delegada del banco Santander UK, la filial británica Santander, percibió 5,1 millones de euros, un 2,39 % más que los 5,010 de hace un año.
6			Otro de los consejeros							ejecutivos del Banco Santander Juan Rodríguez Inciarte		NP+Prop or Noun	Beneficiary	Otro de los consejeros ejecutivos del Banco Santander Juan Rodríguez Inciarte cobró 2,9 millones de euros en 2012, lo que representa una reducción del 13,9%.
7										La entidad		NP	Sayer	La entidad explica en su informe que la remuneración fija de los consejeros ejecutivos son las mismas que las percibidas desde 2009, con la sola excepción de Ana Patricia Botín-Sáenz de Sautuola y O'Shea por "sus actuales responsabilidades como consejera delegada de Santander UK".
8										Los consejeros ejecutivos		NP	Beneficiary	Los consejeros ejecutivos reciben el 40 % de la remuneración variable en el año, la mitad en efectivo y la mitad en acciones, y el 60 % restante durante tres años, también a partes iguales en efectivo y acciones.
9										Esto		NP	Token	Esto significa que el importe en efectivo recibido en 2012 por Emilio Botín por sus funciones ejecutivas fue de 1,62 millones de euros y de 4,4 millones de euros en el caso de Alfredo Sáenz.
10			En 2012		Time-pp	Time				el Banco Santander		NP	Token	En 2012, el Banco Santander logró un beneficio neto de 2.205 millones de euros, el 59 % menos que un año antes, debido a la dotación de 18.800 millones a provisiones para cubrir su riesgo inmobiliario en España.

Orizónia comunica a los trabajadores que se salvarán 800 empleos con la venta de parte de Orbest



## D.3 Annotated News Reports in Spanish

C1#	Thematic Field		SENTENCE REALIZATION										Text				
	TEXTUAL	INTERP.	STANCE	ENGAGEMENT	PREHEAD	STANCE	ENGAGEMENT	SYNTHACTIC REALIZATION	HEAD	STANCE	ENGAGEMENT	EXPERIMENTAL ROLE					
1													La dirección de Orizonia ha comunicado a los trabajadores del grupo la noticia que se van a saldar 600 empleados en el Expediente de Regulación de Empleo (ERE), que afectará a unos 4.000 empleados, gracias a la venta de parte de la portafolio Orizon.	Sayer	Según informó uno de los trabajadores de Iberojel Trascontinental quien esta tarde ha acudido a la reunión que se ha celebrado en la sede de Palma y a la que también han asistido los representantes de los tres grupos de inversión (Carlyle Group, ICG Equity Fund y Vista Capital), que son los que tienen la propiedad de la empresa.		
2					Según informó uno de los trabajadores de Iberojel Trascontinental									NP+App		NP+PP	Process
3								Evidence	Evidence-clause	ha acudido				Verb		Tomás Toth ha señalado que la empresa mantiene la confidencialidad en cuanto a qué empresa será la que adquiere parte de Orizon, si bien algunos trabajadores ya han apuntado que podría tratarse del Grupo Barceló y que, en concreto, adquiriría la parte de Portugal, lo que permitiría mantener los citados empleos, dado que muchos servicios de este país se gestionan desde Palma de Mallorca.	
4	Además				durante el encuentro,			Time	Time-pp	los directivos de Orizonia				NP+App	NP+PP	Sayer	Además, durante el encuentro, los directivos de Orizonia han revelado que hay otras vías de negociación con otras empresas para adquirir parte del grupo, aunque no han concretado de qué compañías se trata, razón por la cual los trabajadores han criticado que existe "mucho secretismo".
5					A lo largo de la reunión, a la que han asistido diez trabajadores y cinco representantes de Orizonia, incluidos los propietarios,					la dirección							A lo largo de la reunión, a la que han asistido diez trabajadores y cinco representantes de Orizonia, incluidos los propietarios, la dirección ha asegurado que la nómina de este mes de febrero está más o menos asegurada, al tiempo que les han anunciado que el ERE, que afectará a unos 4.000 trabajadores, se presentará la semana que viene, junto a la suspensión de pagos.
6								Time	Time-NP+App	la dirección				NP	Det Noun	Sayer	Los trabajadores han indicado que los directivos de Orizonia les han señalado que el juez que se encargue del ERE intentará arreglar las cosas lo antes posible, para que antes del mes de marzo esté todo solucionado, de manera que los empleados que formen parte de una de las empresas del grupo y no tengan ninguna actividad, se les va a dar vacaciones para que no tengan que asumir el gasto de acudir a su puesto de trabajo.
										Los trabajadores				NP	Det Noun	Sayer	
7					Tras las vacaciones, según ha informado												Tras las vacaciones, según ha informado Toth, la empresa les otorgará un permiso retributivo para evitar estos gastos de acudir a trabajar.
					Toth			Time	Time-pp+App	la empresa				NP	Det Noun	Actor	
8										La dirección de Orizonia y los trabajadores				NP+NP		Actor	La dirección de Orizonia y los trabajadores se volverán a reunir este miércoles en Palma de Mallorca para darles más detalles de la operación.

Moody's: la liquidez y financiación siguen siendo los retos de la banca española

Thematic Field		SENTENCE REALIZATION										Text		
C1#	TEXTUAL	INTERP.	STANCE	ENGAGEMENT	PREHEAD	STANCE	ENGAGEMENT	SYNTACTIC REALIZATION	HEAD	STANCE	ENGAGEMENT	EXPERIMENTAL ROLE		
								Grd.						
1										La agencia Moody's			Det Noun Proper Noun NP Sayer	La agencia Moody's ha advertido este miércoles de que la banca española seguirá enfrentándose "a retos de liquidez y de financiación" pese a los recientes signos positivos, como la aceleración de la actividad en los mercados a principios de 2013.
2	Y es que									"habrá un entorno operativo negativo"			Quote Verbiage	Y es que según las previsiones de Moody's, en los próximos 12 meses, "habrá un entorno operativo negativo" tanto para la banca como para la deuda soberana española, algo "que continuará pesando sobre la confianza del mercado y por tanto, impedirá lograr condiciones de financiación normalizadas".
3										Moody's			NP Proper Noun Sayer	Moody's ha hecho esta consideración en el informe titulado "Bancos españoles: persisten los retos de liquidez y financiación a pesar de los recientes signos positivos", publicado este miércoles en el que pese a estas advertencias, ha resaltado que desde el inicio del año, varios de los principales bancos españoles han accedido al mercado mediante emisiones de deuda senior y créditos hipotecarios, y que incluso algunas entidades han anticipado el préstamo concedido por el Banco Central Europeo (BCE).

### D. Appendix

4																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**CaixaBank quiere suprimir 3.000 puestos de trabajo**

CIT	Descriptive Field	SYNERGIC REALIZATION						Experiencial Role	Text						
		INTERP.		PROJEAD		REALIZATION									
	TEXTUAL	Suave	Engagement	Suave	Engagement	Gral.	INTAD	Suave	Engagement						
1															La dirección de CaixaBank ha planteado a los representantes de su plantilla la supresión de unos 3.000 puestos de trabajo para "reestructurar" la entidad tras la absorción de Banca Cívica y del banco de Valencia.
2															La entidad recuadra en un comunicado la reciente integración del grupo de cajas de ahorros y la próxima de la antigua filial de Banca.
3	Por ello,														Por ello el objetivo es "adaptarse al entorno actual y mejorar la eficiencia de sus recursos".
4															El banco ganó 230 millones de euros en 2012, un 78,2% menos que el año anterior, tras destinar 10.299 millones a provisiones para sanear el riesgo inmobiliario, cuatro veces más que las dotaciones del año anterior.
5															Beneficiary
6						Time	Time-pp								A finales de 2012, tras sus dos últimas adquisiciones, el grupo de La Caixa contaba con 32.625 empleados, 5.632 más que un año antes, y con 6.342 sucursales, la red de oficinas más extensa del país tras crecer en 1.146 locales.
7															La entidad catalana se adjudicó el intervenido Banco de Valencia en diciembre de 2012, cuando fue subastado por el Fondo de Reestructuración Ordenada Bancaria (FROB).
8						Manner	Manner-Adop								Precisamente el banco valenciano ha sufrido dos expedientes de Regulación de Empleo (ERE) desde noviembre: uno de 360 personas bajo la gestión del FROB y otro de 890 afectados, más del 50% de su plantilla, anunciado este pasado mes de enero.
9						Time	Time-NP								Meses antes, a principios de 2012, CaixaBank se hizo con la ayudada Banca Cívica.
															Esta entidad, que tenía 8.000 trabajadores hasta febrero del pasado año,

10								CasaBank				Proper Noun	Sayer	Gusbank precisa que ha abierto un periodo de negociaci3n con los sindicatos "que se podr3a prolongar hasta finales de abril", y remarca que se propone "analizar todas la f3rmulas posibles y acordar las mejores medidas para ajustar su plantilla a las necesidades del negocio".
11								Comisiones Obreras (CCOO)				Proper Noun	Sayer	Comisiones Obreras (CCOO) ha advertido de que no aceptar "salidas forzadas" y considera "excesiva" la cifra planteada de despidos, pues en su opini3n el banco tiene capacidad de generar beneficios y no ha recibido ayudas p3blicas.
12													Sayer	En un comunicado, el sindicato ha afirmado que "la plantilla est3 haciendo un esfuerzo de contenci3n salarial importante" desde el convenio de 2011, junto con prolongaciones de jornada, adem3s de graves tensiones en los centros de trabajo.

landia recorta los sueldos públicos a partir de 65.000 euros y congela el resto

CT#	Thematic Field	INTERP.				PREHEAD				SYNTACTIC UTILIZATION				HEAD				Engagement				Experiential Role				Text
		TEXTUAL																								
1																								El Gobierno irlandés ha anunciado una reforma del marco salarial de los funcionarios del país para los tres próximos años que incluye un recorte de entre el 5,5% y el 10% para los sueldos más altos.		
2																								El paquete de medidas, comunicado por el ministro irlandés de Gasto Público y Reforma, el laborista Brendan Howlin, prevé también una ampliación de la jornada laboral en el sector público.		
3																								Los funcionarios que ganan entre 65.000 y 80.000 euros anuales verán rebajados sus emolumentos un 5,5%, mientras que aquellos que ingresen entre 80.000 y 150.000 euros anuales, el recorte será del 8%.		
4																								En el caso de los funcionarios con salarios anuales de entre 150.000 y 185.000 euros, el recorte será del 9% y para a los que ganen más de 185.000 euros el recorte será del 10%.		
5																								El Gobierno irlandés suspenderá también cualquier incremento para todos los salarios que superen los 65.000 euros anuales.		
6	Por su parte																							Por su parte, los funcionarios que ganen menos de 35.000 euros anuales verán congelados sus sueldos por tres meses, mientras que aquellos con ingresos de entre 35.000 y 65.000 euros al año verán congelados dos veces sus salarios por un período total de seis meses.		
7																								Con este plan, que no ha sido respaldado por todos los grandes sindicatos irlandeses, Dublín pretende ahorrar 1.000 millones de euros en los tres próximos años.		
8	Por otro lado																							Por otro lado, el Gobierno pretende que los funcionarios que trabajen 35 horas semanales o menos deberán a partir de ahora trabajar 37 horas a la semana, mientras que aquellos que trabajen más de 35 horas deberán trabajar 39 horas y los que trabajen más de 39 horas deberán trabajar 41 horas a la semana. Aunque se evaluará el tratamiento de los sindicatos.		

**Villar-Mir analiza una posible operación con Colonial**

CI#	Thematic Field		Text									
	TEXTUAL	INTERP.	PREHEAD	SYNTACTIC REALIZATION	HEAD	Engagement			Experiential Role	Text		
		Stance	Engagement		Engagement	Goal		Stance				
1												El Grupo Villar Mir ha reconocido hoy que ha estado analizando una posible operación con inmobiliaria Colonial, sin que a la fecha actual existan acuerdos ni negociaciones de ningún tipo con los accionistas de dicha empresa.
2												Colonial ha subido alrededor de un 10% a media sesión de la jornada de Bolsa de este lunes tras publicar se informaciones sobre una eventual compra de la inmobiliaria patrimonialista.
3			En un comunicado									En un comunicado, el grupo que dirige Juan Miguel Mir ha salido al paso de informaciones de prensa que apuntaban a una posible compra de la inmobiliaria Colonial y a las negociaciones mantenidas por la empresa catalana para buscar inversores interesados en adquirir la empresa.
4			En este sentido,									En este sentido, el Grupo Villar Mir ha desmentido que esté negociando algún acuerdo sobre la compra de activos de Colonial.
5												Sin embargo, el grupo ha indicado en un comunicado que cualquier posible operación no tendría relación alguna, ni directa ni indirecta, con Grupo OHL, en el

### D. Appendix

[illegible]

**El número de viviendas hipotecadas bajó un 32,7% en 2012**

CI#	The main Field	INTERVIEW			SYNTHETIC SITUATION			HEAD			Experimental Role			Text
		TEXTUAL	Source	Engagement	PREHEAD	Source	Engagement	Grat.	Source	Engagement	Source	Engagement	Experimental Role	
1									El número de hipotecas constituidas sobre viviendas	NP-App	NP-PP	Actor	El número de hipotecas constituidas sobre viviendas descendió un 32,7% en 2012 respecto al año anterior, con lo que prácticamente iguala el retroceso registrado en 2011, que fue del 32,6%.	
2									la caída experimentada en 2012 en el número de viviendas hipotecadas	NP-App	NP-PP	Actor	La caída experimentada en 2012 en el número de viviendas hipotecadas se suma a los descensos registrados en los cinco años anteriores y es el más pronunciado de este que estalló la crisis.	
3	En concreto					Time	Time-PP		la disminución	NP	Det Noun	Actor	En concreto, en 2011 la disminución fue del 32,6%, en 2010 la caída fue del 7%, en 2009 fue del 22%, en 2008 fue del 32% y en 2007, del 6,7%.	
4														
5						Evidence	Evidence-PP		el importe medio de las hipotecas constituidas sobre viviendas	NP-App	NP-NonFinite	Actor	Según los datos difundidos este martes por el Instituto Nacional de Estadística (INE), el importe medio de las hipotecas constituidas sobre viviendas bajó un 7,8% en 2012, mientras que el capital prestado se redujo un 38% en el conjunto del año.	
6						Evidence	Evidence-PP		El número de hipotecas sobre fincas rústicas y urbanas (dentro de éstas últimas se incluyen las viviendas)	NP-App	NP-PP	Actor	De acuerdo con los datos del organismo estadístico, durante 2012 el número de hipotecas sobre fincas rústicas y urbanas (dentro de éstas últimas se incluyen las viviendas) descendió un 29,5%, respecto a 2011.	
7									El capital de los créditos hipotecarios concedidos	NP-App	NP-PP	Actor	El capital de los créditos hipotecarios concedidos bajó un 33,4% en el conjunto del año, en tanto que el importe medio de las hipotecas constituidas sobre el total de fincas cayó un 5,6%.	
8	Por su parte								El INE	NP	Proper Noun	Sayer	El INE ha precisado que el próximo 6 de marzo ofrecerá una información más detallada de los resultados anuales de esta estadística.	
9									El número de hipotecas constituidas sobre viviendas	NP-App	NP-PP	Actor	Por su parte, el número de hipotecas constituidas sobre viviendas se situó en 121.377 en diciembre de 2012, con una caída del 27,6% respecto a mismo mes de 2011.	
10									El retroceso interanual de diciembre, con el que las viviendas hipotecadas	NP-App	NP-Claus	Carrier	El retroceso interanual de diciembre, con el que las viviendas hipotecadas encadenan ya 32 meses de caídas consecutivas, es menos pronunciado que el de noviembre, cuando las viviendas hipotecadas se redujeron más de un 31%.	
11						Manner	Manner-Adv		la cifra total de hipotecas sobre viviendas	NP-App	NP-PP	Token	Aun así, la cifra total de hipotecas sobre viviendas es la más baja en cualquier mes dentro de la serie histórica comparable, iniciada en 2003.	
									El importe medio de las hipotecas constituidas sobre viviendas alcanzó el pasado mes de diciembre los 101.460 euros, un 3% menos que en igual mes de 2011, mientras que el capital prestado se redujo un 29,5%, hasta los 1.783 millones de	NP-App	NP-PP	Actor	El importe medio de las hipotecas constituidas sobre viviendas alcanzó el pasado mes de diciembre los 101.460 euros, un 3% menos que en igual mes de 2011, mientras que el capital prestado se redujo un 29,5%, hasta los 1.783 millones de	

a prima de riesgo española se dispara a 400 puntos por los resultados en Italia

CI#	Thematic Field	SYNTACTIC REALIZATION										Text					
		INTERP.		PREDICAD		ENGAGEMENT		SYNTACTIC REALIZATION		HEAD			ENGAGEMENT		EXPERIENTIAL ROLE		
	TEXTUAL		Stance	Engagement						Stance	Engagement						
1																La situación de ingobernabilidad en la que ha quedado Italia tras las elecciones ha sacudido los mercados de deuda.	
															NP+clause	Actor	
2																La prima de riesgo de la tercera economía de la zona euro se ha disparado más de 40 puntos en la apertura de este martes, hasta los 335 puntos básicos, mientras la española ha llegado a tocar los 412 puntos, un vertiginoso ascenso desde los 361 puntos del cierre del lunes.	
																NP+App	Actor
3																Verb	Process
																	En ambos casos, se trata del mayor nivel desde diciembre.
4																	La prima, el índice que marca la diferencia de rentabilidad de los bonos españoles a 10 años con los equivalentes alemanes, se ha disparado hasta los 415 puntos pasadas las 8:30 horas, mientras que el interés del bono escalaba hasta el 5,614%, frente al 5,11% del cierre de ayer.
																NP+App	Actor
5																	El riesgo país no alcanzaba esos niveles desde que, a finales de diciembre, comenzó a retroceder tras alcanzarse el acuerdo en Estados Unidos para evitar el 'abismo fiscal'.
																NP	Actor
6																	Por su parte, la prima de riesgo italiana se situaba en 344,80 puntos básicos, con una rentabilidad del 4,907%, tras haber abierto en 302,20 puntos básicos.
																NP+App	Carrier

## **D.4 Annotated Editorials in English**

## Editorial: David Cameron's mixed signals on terrorism

Cl #	Thematic Field										Text				
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance		Engagement	SYNTACTIC REALIZATION	Detail	Experiential Role
1		It must be granted that	Eng						the situation in and around the huge In Amenas gas plant in southern Algeria			Np+App	Np+PP	Senser	It must be granted that the situation in and around the huge In Amenas gas plant in southern Algeria was confused, and remained so last night even as we went to press.
2					As David Cameron spelt out in his Commons statement today,		shared knowledge	Evidence	Evidence-Clause			Np	Det Noun	Carrier	As David Cameron spelt out in his Commons statement today, the complex is far out in the desert, 18 hours by road from Algiers, and it had been over an in what appears to have been a large-scale and carefully planned operation by Mokhtar Belmokhtar, an Algerian who may have ties to al-Qaida.
3									Communications appear, at best,	attitude		Np+App	Np+PP	Carrier	Communications appear, at best, to have been intermittent.
4					Even allowing for all this, however, and for the difficulty that the leader of any country faces when there are fears for the safety of its nationals abroad, it is not only from Algeria that the signals have been less than distinct in recent days.		shared knowledge	Contingency	Contingency Non-finite			clause	clause	Carrier	Even allowing for all this, however, and for the difficulty that the leader of any country faces when there are fears for the safety of its nationals abroad, it is not only from Algeria that the signals have been less than distinct in recent days.
5									Those emitted from London, at least in so far as they have reached the public domain, Mr Cameron	hedge		Np+App	NP+Clause	Carrier	Those emitted from London, at least in so far as they have reached the public domain, have not been exactly crystal-clear either.
6		True,	Com									Np	Proper Noun	Sayer	True, Mr Cameron sought to quell some of the more fevered speculation when he stressed today that the blame for what had happened rested firmly with the terrorists and no one else.
7									That			NP	Pronoun	Carrier	That sounded like an attempt to dispel the impression that the UK Government held the Algerian authorities responsible for any hostage deaths, amid a chorus of British headlines condemning a 'botched' operation.
8									The Prime Minister			Np	Proper Noun	Senser	The Prime Minister was patently unhappy that Algeria had not informed him before its forces embarked on their assault, but understood that this was neither the time nor the place to pin blame.
9	Elsewhere, however,								his statement			Np	Pronoun Noun	Actor	Elsewhere, however, his statement left some disturbing questions about the Government's precise attitude to terrorism and hostage-taking.
10									The official view, shared by successive governments			Np+app	Np+Clause	Carrier	The official view, shared by successive governments, has long been that the UK does not negotiate and does not give in to terrorists.
11									Rheme			Rheme		None	Nor does it pay ransoms, on the grounds that this only encourages further kidnapping.
12	Yet								this message of no compromise			Np+App	Np+PP	Carrier	Yet this message of no compromise was not the one that seemed to hover around the UK position as articulated by Mr Cameron.
13					Instead			Manner	Manner-Advb			NP	Pronoun	Sayer	Instead, he cited a British offer of technical and intelligence support, including hostage negotiation and rescue teams, which – he said with transparent regret – had been declined.
14									Sensible	attitude		Np+app	Np+PP	Carrier	Sensible restraint in response to a hostage crisis is one thing, but what role might British negotiators have played?.
15									There		Question	Verb	question	Existent	Is there one rule in principle for how Britain behaves on its own territory, or when the hostages can be counted on the fingers of one hand, and another in practice when there may be dozens of British nationals in danger and there are doubts about the ability of a sovereign foreign government to rescue them?.
16									There			There	There	Existent	There was a hint of arrogance in Mr Cameron's offer that could well have smacked, to those on the receiving end, of colonial-style condescension.
17	And				for a prime minister to give the impression to his domestic audience that he has the power to do something and has everything under control,			Contingency	Contingency Clause		shared knowledge	NP	Pronoun	Goal	And forgivable though it might be for a prime minister to give the impression to his domestic audience that he has the power to do something and has everything under control, it is not only Britons who were caught up in the drama in the desert.
18									The captives			NP	Det Noun	Carrier	The captives also included Americans, Japanese and Belgians.

### D. Appendix

19										all their governments		question	verb	question	Acute	Did all of their governments offer help, and what might have resulted if any or all had moved to intervene unilaterally?
20										There	hedge		There	there	Existent	There may be some very preliminary lessons here.
21										one			NP	Pronoun	Carrier	One is that UK governments still have to learn to be less insular and must appreciate the effect of their words on others.
22										Another			NP	Pronoun	Carrier	Another is that ministers must guard against saying one thing and doing another in responding to terrorism.
23	And, lastly								given that the spread of multinationals makes it likely that the citizens of several countries will be caught up in similar incidents in future, some standard mechanism is needed to ensure that their governments do not act precipitately and at cross purposes.	Contingency Clause			NP	Pronoun Adj Noun	Attribute	And lastly, given that the spread of multinationals makes it likely that the citizens of several countries will be caught up in similar incidents in future, some standard mechanism is needed to ensure that their governments do not act precipitately and at cross purposes.

**A dilemma that went g away**

Cl	Thematic Field		SYNTACTIC REALIZATION												Text
	TEXTUAL	INTERP.	PREHEAD	STANCE	ENGAGEMENT	GRAT.	DETAIL	HEAD	STANCE	ENGAGEMENT	GRAT.	DETAIL	EXPERIENTIAL	ROLE	
1	As		we live longer,			shared knowledge	Manner	Manner-Clause				Np-App	Np-PP	Actor	As we live longer, the pressure on the health service mounts.
2									The number of hospital stays by elderly patients, by over-75s in particular				NP+App	Actor	The number of hospital stays by elderly patients, by over-75s in particular, has soared by about a third since 2000.
3									It			Np	Pronoun	Actor	It makes for a lopsided use of resources, not just because other people are prevented from using these beds but also because, as Sir David Nicholson, the new head of the NHS Commissioning Board, says in this newspaper, prolonged hospital stays are bad for old people, too.
4									Sir David			NP	Proper Noun	Sayer	Sir David points out that some 40 per cent of all patients in the average general hospital now have dementia.
5	Yet								this growing body of often confused people	attitude		Np-App	Nonfinite-PP	Senser	Yet, this growing body of often confused people find themselves all but abandoned in the alien environment of modern hospitals that "are fast moving and organised round getting a diagnosis".
6			These are "bad places for old, frail people,			shared knowledge	Quote		he			NP	Pronoun	Sayer	These are "bad places for old, frail people," he says.
7									A wealth of evidence			Np-App	Np-PP	Actor	A wealth of evidence backs up his belief that protracted hospital stays lead to poor outcomes, as well as being very expensive.
8	While		all patients are exposed to infection in hospital			shared knowledge	Contingency	Contingency Clause	the elderly			NP	Det Noun	Actor	While all patients are exposed to infection in hospital, the elderly especially risk losing their independence and eating and drinking poorly.
9									A 2011 report by the care sector watchdog, the Quality Care Commission, said more than half of hospitals in England were failing key standards in care for the elderly in terms of nutrition and dignity.			Np-App	Np-Propor Noun	Sayer	A 2011 report by the care sector watchdog, the Quality Care Commission, said more than half of hospitals in England were failing key standards in care for the elderly in terms of nutrition and dignity.
10									Some			Np	Pronoun	Carrier	Some did not just have problem with resources, they also had a bad attitude.
11									what	question		Question	Interrogative Pronoun	Attribute	What to do?
12			Last year				Time	Time-NP	the King's fund			NP	Proper Noun	Sayer	Last year, the King's fund said that if all primary care trusts followed the practice of the best, and reduced the admission levels for chronically ill elderly patients, the NHS could free up more than 7,000 beds.
13									That			Np	Pronoun	Carrier	That is a laudable goal, but if beds for the elderly are cut without the development of more integrated allie-hospital care services, they will be even worse off.
14									Sir David			Np	Proper Noun	Sayer	Sir David promises more community-based treatment centres for elderly patients, which sounds right.
15									The problem			Np	Det Noun	Carrier	The problem is not going to go away.
16			By 2032				Time	Time-pp	the over-75s			Np	Det Noun	Phenomenon	By 2032, the over-75s are expected to comprise 22 per cent of the population, up from 17 per cent



Obama must seize the chance to make history

Cl	Thematic Field	SYNTACTIC REALIZATION				SYNTACTIC REALIZATION				HEAD	SYNTACTIC REALIZATION				Experimental Role	Text
		TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	Grat.	Detail	Grat.	Engagement	Stance	Grat.	Detail	
1						Between the two inaugurations of Barack Obama				Time-pp				Np+Np	Noun+Noun	Actor
2	When					America's first black president took the oath of office on that frigid, unforgettable 20 January 2009 in Washington	attitude		Time	Time-clause				Np	Det Noun	Carrier
3														Question	Interrogative Pronoun	Attribute
4														NP	Pronoun Det Noun	Actor
5	Yet					today			Time	Time-Advb				NP	Det Noun	Carrier
6	But													NP	Det Noun	Carrier
7	In fact													NP	Pronoun	Carrier
8										The President				NP	Det Noun	Senser
9										He				NP	Pronoun	Actor
10						Four years of unrelenting, unbending	attitude		Time	Time-NP				NP	Noun-Adj	Sayer
11										The window of opportunity				Np+App	Np+PP	Carrier
12	But					between now and mid-2014			Time	Time-PP				NP	Proper Noun	Carrier
13						For one thing			Contingency	Contingency-pp				NP	Pronoun	Carrier
14										A strengthening economy, and America's under-reported but vastly significant progress towards energy self-sufficiency				Np+Nonfinite	Np+Nonfinite	Actor
15						As for the Republicans		shared knowledge	Contingency	Contingency-Advpp				NP	Det Noun	Actor
16			It is worth remembering too that							Americans as a whole				Np+App	Np+PP	Senser
17										Democrats				NP	Noun	Actor
18	In fact		it seems that							the Republicans				NP	Det Noun	Behaver
19			it shows							Their proposal				NP	Det Noun	Carrier
20	But									the party				NP	Pronoun	Senser
21										Mr Obama				NP	Proper Noun	Actor
22										A "grand bargain"				NP	Det Adj Noun	Carrier
23							hedge			Blame		question		Question		None
24										fracture				NP	Noun	Actor
25										He				NP	Pronoun	Carrier
26										Many Republicans				NP	Pronoun Noun	senser

## D. Appendix

27											Measures to tackle climate change					NP+App	NP+PP	Carrier	Measures to tackle climate change are a more distant, but not wholly unattainable, goal.
28	But										Mr. Obama					NP	Proper Noun	Actor	But Mr. Obama must seize the moment.
29	If										Contingency- Clause					NP	Pronoun	Carrier	If this first term revealed a flaw, it was a passivity based on assuming his opponents would be susceptible to reason.
30	But										Contingency- Clause					NP	Det Noun	Actor	But here too – if the urgency with which he is pressing for gun control is any indicator – that lesson has sunk home.
31	And										Contingency- Advb					NP	Pronoun	Actor	And if so, he might deliver a second term that makes history.

### Stop all the demonising of new EU migrants

Thematic Field	Textual	Interp.	Stance	Engagement	Prehead	Stance	Engagement	Syntactic Realization	Head	Stance	Engagement	Syntactic Realization	Detail	Experiential Role	Text
1									The UK Independence Party			NP	Det Proper Noun	Actor	The UK Independence Party has added a click to its website to count down the days until 1 January 2014, when controls expire on immigration from Romania and Bulgaria, the two newest EU member states.
2		It is a shameless piece of scaremongering designed to support Ukip's alarmist claim that the change will result in an influx of 550,000 Romanian or Bulgarian migrants to the UK.							the change			NP	Pronoun	Carrier	It is a shameless piece of scaremongering designed to support Ukip's alarmist claim that the change will result in an influx of 550,000 Romanian or Bulgarian migrants to the UK.
3					Today			Time	the lobby group MigrationWatch			NP+App	NP+Proper Noun	Actor	Today the lobby group MigrationWatch plucked from the air another dubious figure, saying there will be around 50,000 new arrivals every year once restrictions are lifted.
4									It			NP	Pronoun	Sayer	It predicted "significant consequences" for housing, jobs, schools and hospitals.
5									The justification for this moral panic			NP+App	NP+PP	Carrier	The justification for this moral panic is the fact that the Government got the figures so wrong last time around.
6	When				Poland joined the EU,			Time	the Home Office infamously	attitude		NP	Det Proper Noun	Sayer	When Poland joined the EU, the Home Office infamously predicted that no more than 13,000 Poles would come to Britain annually.
7					Within two years			Time	264,560			NP	Pronoun	Actor	Within two years 264,560 had arrived.
8									Things			NP	Noun	Carrier	Things are bound to be even worse this time, the prophets of doom suggest, because the minimum wage in Romania and Bulgaria is half that in Poland.
9									Low-skilled immigrants			NP	Adj Noun Noun	Actor	Low-skilled immigrants will flood in to undercut British workers.
10									Such			NP	Pronoun	Carrier	Such is the myth.
11									The real impact, according to the Migration Observatory at Oxford University			NP+App	NP+NonFinite	Carrier	The real impact, according to the Migration Observatory at Oxford University, will be far smaller.
12									There		shared knowledge	There	There	Existent	There are many reasons to suppose they are right.
13	When				Poland and others joined the EU			Time	the UK			NP	Det Proper Noun	Token	When Poland and others joined the EU, the UK was the only large member-state to open its labour market in 2014, by contrast, all major European economies are lifting controls.
14									Most Romanians			NP	Adj Noun	Goal	Most Romanians, for instance, are expected to go to Italy, Spain and France, with which they have linguistic and cultural links, and where many already have relatives.
15									Those deterred by the high rate of youth unemployment in those countries			NonFinite	NonFinite	Carrier	Those deterred by the high rate of youth unemployment in those countries are more likely to head for Germany and the Netherlands than Britain.
16	Plus								Bulgarians and Romanians			NP+NP	Noun+Noun	Carrier	Plus Bulgarians and Romanians generally have a weaker command of English than Poles.

17	In any case	Eng	contrary to popular scare stories,	booster		Manner	Manner-AdvP	immigration			NP	Noun	Carrier	In any case, contrary to popular scare stories, immigration is, on balance, good for the British economy.
18								Many immigrants			NP	Adj Noun	Actor	Many immigrants have made huge contributions to our society.
19								Migrants			NP	Noun	Actor	Migrants create jobs as well as competing for them.
20								Immigrants			NP	Noun	Carrier	Immigrants often have a particular dynamism.
21								They			Np	Pronoun	Actor	They often do jobs that local people on benefits will not do.
22								They			Np	Pronoun	Actor	They perform vital tasks in the public services.
23								They			Np	Pronoun	Carrier	They are generally young, and have more children, which is good news, given the ageing population.
24								Businesses			NP	Noun	Sayer	Businesses report they have a strong work ethic.
25								The economy			NP	Det Noun	Actor	The economy grows faster because of their contribution.
26								All of that			Np+App	Pronoun+PP	Carrier	All of that is lost on the myth-peddlers who suggest immigrants are all on benefits – when 97 per cent of registered migrant workers have full-time jobs – or more prone to crime, when the Association of Chief Police Officers says offending rates among Polish, Slovak, Lithuanian, Romanian and Bulgarian incomers are pretty much in line with the rest of the population.
27	So							eries			Np	Noun	actor	So critics resort to sneats, confusing Romanians with Roma, or planting other dog-whistle racist innuendos – of which, it must be feared, we will hear more when the Australian spin-doctor, Lynton Crosby, masterminds the next Conservative election campaign.
28	If							here	shared knowledge		There	There	Existent	If the net effect of migration is positive, however, there will always be some who lose out.
29								Politicians, national and local	directive		NP+App	Pronoun Adj+Adj	Actor	Politicians, national and local, need to focus on criteria for social housing to quash any suspicion that migrants can jump the queue.
30								Extra support	directive		Np	Adj Noun	Goal	Extra support needs to be given to schools and hospitals in areas with large numbers of newcomers, and eligibility for benefits should require a minimum period of residency, with offspring who live abroad not eligible for child benefit.
31	But							all that			NP	Pronoun Pronoun	Goal	But all that can be done within a context of celebrating the virtues of immigration.
32								Demonising new arrivals			NonFinite	NonFinite	Carrier	Demonising new arrivals is not a useful response.

## A confession is not enough

CI #	Thematic Field		SYNTACTIC REALIZATION										Text		
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Explanatory Role	
			Stance	Engagement		Stance	Engagement	Grat.	Detail		Stance	Engagement	Grat.	Detail	
1	Even if				Lance Armstrong had come across as the most contrite of penitents – which he did not –	hedge			Contingency Clause	an interview with Oprah Winfrey on American TV			NP+App	NP+PP	Actor
2					–								Rheme		None
3										The only advance made by this much-anticipated encounter			NP+App	NP+Nonfinite	Carrier
4		Yes,	Com							his reputation			NP	Pronoun Noun	Goal
5	But									the damage that Armstrong has done to his sport			NP+App	NP+clause	Actor
6	First									he	hedge		Np	Pronoun	Goal
7	Second									he	hedge		NP	Pronoun	Goal
8	And third,									former funders, sponsors and prize donors	hedge		NP	Adj Noun Noun Noun Noun	Actor
9	Until				he feels, and the public feels, that Armstrong has paid a proper price for his cheating, this canker will not have been excised	booster			Time Clause	this canker			Np	Det Noun	Goal

## D. Appendix

Mali's Islamists are too dangerous to be ignored

Thematic Field		SYNTACTIC REALIZATION															Text
TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	Grd.	Detail	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Grd.	Detail	Experiential Role	
1										Military involvement in foreign conflicts	hedge		Np+App		Np+pp	Carrier	
2	But			With Islamist extremists in Mali threatening to overrun the country, destabilise the region, and hand a vast area of the Sahel to al-Qa'ida and their like		shared knowledge	Manner		Manner-NonFinite	the dangers of ignoring the West/African country's plea for help			Np+App		Np+pp	Carrier	
3										The French President's swift action			Np+App		det Adj Noun Noun noun	Carrier	
4										Mali's troubles			Np		Noun Noun	Actor	
5										A string of defeats by Tuareg separatists – aided by Islamist fighters returning home from post-Gaddafi Libya –			Np+App		Np+pp	Actor	
6				With the secular Tuaregs then swiftly pushed aside by their former allies,		shared knowledge	Manner		Manner-NonFinite	extremist militants			Np		Adj Noun	Actor	
7										The result			Np		Det Noun	Carrier	
8										Three Islamist groups			Np		Det Adj Noun	Carrier	
9										The International Criminal Court			Np		Det Proper Noun	Actor	
10	But although			the UN finally approved a plan to deploy African troops to the country in December,	hedge		Contingency		Contingency-clause	assistance for the beleaguered government in Bamako			Np+App		Np+pp	Goal	
11				Last week,			Time		Time-AdvP	everything			Np		Pronoun	Actor	
12				A rebel victory at the town of Konna						A rebel victory at the town of Konna			Np+App		Np+pp	Actor	
13										France's decision to send war planes to help re-take Konna and to send troops to protect Bamako			Np+App		Np+NonFinite	Token	
14	And although		booster	François Hollande is open to charges of precipitousness		shared knowledge	Contingency		Contingency-clause	he		shared knowledge	Np		Pronoun	Carrier	
15				With Mopti lost,						it	hedge		Np		Pronoun	Carrier	
16										Britain's offer of assistance in the form of two C-17 transport aircraft			Np+App		Np+pp	Carrier	
17				In part,			Manner		Manner+pp	such support		shared	Np		Adj Noun	Carrier	
18	But									there is also an element of moral obligation here.			Np	There	Existent		

19							With Mali's Islamists suggesting an influx of mercenary fighters from Libya – in which our involvement was pivotal – it is incumbent upon us, as well as France, to help deal with the consequences.	shared knowledge	Contingency	Contingency NonFinite				NP	Pronoun	Carrier	With Mali's Islamists suggesting an influx of mercenary fighters from Libya – in which our involvement was pivotal – it is incumbent upon us, as well as France, to help deal with the consequences.
20														Np+App	Np+PP	Actor	The Prime Minister's promise of "no boots on the ground" must be adhered to, however, regardless of how the conflict plays out.
21							Here		Place	Place+Advb	the matter			Np	Det Noun	Carrier	Here the matter becomes less straightforward.
22											The French Foreign Minister			Np	Det Proper Noun	Sayer	The French Foreign Minister yesterday stressed that the intervention would last a "couple of weeks" only.
23											It			Np	Pronoun	Carrier	It is difficult to be so sure, however.
24							So far		Time	Time+Advb	progress			Np	Noun	Carrier	So far, progress has been anything but smooth, with militants gaining ground in at least one part of the country yesterday.
25											The risk of military stalemate, of civilian casualties, and of retaliatory terrorist attacks in France (or elsewhere in Europe)			NP+NP	NP+NP	Carrier	The risk of military stalemate, of civilian casualties, and of retaliatory terrorist attacks in France (or elsewhere in Europe) are all too real.
26							Given such dangers	shared knowledge	Contingency	Contingency NonFinite				Np+App	NP+Clause	Carrier	Given such dangers, suggestions that Mr Hollande's first foray into foreign policy is simply to distract attention from his woeful poll ratings are hardly credible.
27											The near-unilateral rush in Bamako's aid			NP+App	NP+Proper Noun Noun	Carrier	The near-unilateral rush to Bamako's aid also sits uncomfortably with his promise to cut paternalistic ties with former French colonies.
28							The sooner... the better				France			NP		Actor	The sooner, therefore, that France is backed – politically – by the UN and – militarily – by the Economic Community of West African States, the better.
29							For all the difficulties of intervention in Mali,		Contingency	Contingency PP	the alternatives			NP	Det Noun	Carrier	For all the difficulties of intervention in Mali, the alternatives are worse.
30											There			There	There	Existent	There are many lessons to be learned from Afghanistan: the perils of a failed state providing a haven for al-Qaeda's fighters; the dangers of a failed state providing a haven for al-Qaeda's fighters.

## Still an island of ignorance about Europe

Cl.	Thematic Field		SYNTACTIC REALIZATION										Text
	TEXTUAL	INTERP.	PREHEAD		HEAD		SYNTACTIC REALIZATION		SYNTACTIC REALIZATION		Experiential Role		
			Stance	Engagement	Stance	Engagement	Detail	Grak.	Stance	Engagement	Grak.	Detail	
1		it would be reasonable to expect that	Sua			Shared knowledge	Contingency Clause	Contingency Clause			NP+App	NP+PP	Phenomenon
2	Yet	it is suddenly announced that	Eng				Time	Time-pp			Np	Pronoun	Goal
3											NP+App	NP+PP	Sayer
4							Verb				there	There	Existent
5	And if				Judge		Contingency Clause	Contingency Clause		Question	Question	Interrogative Pronoun	attribute

## D. Appendix

Blame the tax loopholes

Thematic Field		SYNTACTIC REALIZATION										Experimental Role	Text			
CI #	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Verb	Process	
1					Chalk up				Chalk up			Directive		NP+App	Process	
2									Some – including Goldman Sachs –					NP+App	Senser	
3									The prospect of too-big-to-fail banks using such a sly bit of avoidance to save their already over-remunerated staff a tax bill – and at the expense of the struggling British Exchequer –				NP+App	NP+NonFinite Conj Nonfinite	Actor	
4									Labour's Chris Leslie					Np	Adj Proper Noun	Sayer
5		Sadly,	Com						one			inclusive pronouns		Np	Pronoun	Senser
6		In any case	Com						there			shared knowledge		There	There	Existent
7									the same					NP	Det Noun	Carrier
8									The coffee giant					Np	Det Noun Adj	Carrier
9									What Britain needs					Clause	Token	
10		It was particularly foolish of	Eng		On income tax				the Chancellor					Np	Det Noun	Actor
11									He					Np	Pronoun	Target
12	And				so				it					Np	Pronoun	Actor
13					On corporation tax				what is required					Clause		Carrier
14									Ed Miliband's proposal to require corporations to publish a single figure showing how much corporation tax they pay					NP+App	NP+PP	Carrier
15									The objective of all politicians					NP+App	NP+pp	Carrier
16									A tax regime that relies on corporations choosing to do the right thing					NP+App	NP+Clause	Carrier

Our interests come first, not those of America

Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experimental Role	Text
CI	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experimental Role	Text
1									The former US secretary of state Henry Kissinger			NP+App	NP+Proper Noun	Goal	The former US secretary of state Henry Kissinger was once quoted as saying (though he cannot remember doing so): "If I want to talk to Europe, whom do I call?"
2									The answer,			NP	Det Noun	Carrier	The answer, in fact, has always been fairly straightforward.
3	If				an American president wanted the interests of his country represented in the chancelleries of the EU, then the first number on his list was always that of the British prime minister	hedge		Contingency	on his list			NP+App	NP+PP	Token	If an American president wanted the interests of his country represented in the chancelleries of the EU, then the first number on his list was always that of the British prime minister.
4		It is for that reason that	Eng						Washington			NP	Pronoun	Carrier	It is for that reason that Washington is so alarmed by talk of a British exit from Europe.
5					Without Britain			Manner	the EU			NP	Det Proper Noun	Actor	Without Britain, the EU would constitute a powerful political and trading bloc often hostile to America, but with Britain at the table, Washington is guaranteed, at the very least, a sympathetic hearing.
6		It is therefore understandable	Eng						the Obama administration			NP	Det Proper Noun	Senser	It is therefore understandable that the Obama administration is worried about the possibility of a referendum on the UK's continued membership of the 27-nation group.
7									What is harder to accept, however, is			Clause		Sayer	What is harder to accept, however, is the way Philip Gordon, the US assistant secretary of state for European affairs, voiced those concerns ahead of David Cameron's keynote speech on Europe, which is expected to pave the way for a referendum in 2018.
8	Since	it is hard not to conclude that	Sta		briefings by officials of his rank are normally off the record		shared knowledge	Time	his intervention			NP	Pronoun Noun	Goal	Since briefings by officials of his rank are normally off the record, it is hard not to conclude that his intervention was deliberately timed to influence this increasingly fractious debate.
9	When	Indeed	com		this newspaper reported before Christmas that the Americans were unhappy with Mr Cameron's approach.	self mention		Time	Washington			NP	Pronoun Noun	Actor	Indeed, when this newspaper reported before Christmas that the Americans were unhappy with Mr Cameron's approach, Washington dismissed a story that it is now happy to see broadcast widely.
10	However								the Americans			NP	Det Noun	Carrier	However, the Americans are not keen for us to stay in the EU for our interests, but their own. Mr Gordon said America wanted an outward-looking EU with Britain in it.
11									This			NP	Pronoun	Carrier	This may sound an attractive prospect, but it is not the direction of travel in Europe.
12					Over the next year or two,			Time	the eurozone countries			NP	Det Noun Noun	Actor	Over the next year or two, the eurozone countries will almost certainly forge a closer relationship that will make a renegotiation of the UK's position inevitable.
13									The status quo			NP	Det Noun Noun	Carrier	The status quo is simply not going to be an option, a point that the public understands and with which Mr Cameron is trying to come to terms.
14					Instead of trying to caricature this predicament as a Eurosceptic fantasy			Contingency	those who wish Britain to stay in Europe			NP+App	NP+Clause	Actor	Instead of trying to caricature this predicament as a Eurosceptic fantasy, those who wish Britain to stay in Europe should be exploring positive ways to make it happen.
15									Washington			NP	Pronoun Noun	Actor	Washington could better use its influence to persuade EU leaders to recognise and respond to the difficulties that Britain will have dealing with a new federation on the Continent.
16		we may well conclude that	Sta		As a nation			Role	saying in the EU			NP	Pronoun	Carrier	As a nation, we may well conclude that saying in the EU is in this country's long-term interests.
17	However								we		inclusive pronouns	NP	Pronoun	Actor	However, we must not conduct our foreign policy on the basis of what is right for America, but what is best for us.

Britain is ready to rock the boat in the EU

Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experimental Role	Text
CI	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experimental Role	Text
1									David Cameron			NP	Proper Noun	Actor	David Cameron gave a succinct explanation in the House of Commons yesterday for his determination to strike a new deal with the European Union.
2									He			NP	Pronoun	Sayer	He warned that the most dangerous position this country could take in the light of changes being forced by the eurozone sovereign debt crisis would be to stand back and say we are going to do nothing about it.
3									The Prime Minister			NP	Det Proper Noun	Sayer	The Prime Minister said that Britain had to fight for the new deal it wants "so then we can ask for the consent of the British people to settle this issue once and for all".
4									This			NP	Pronoun	Token	This will be the broad thrust of the much-anticipated speech on Europe that he will deliver in the Netherlands tomorrow.

### D. Appendix

[illegible]



CI #	Thematic Field		SYNTACTIC REALIZATION										Text				
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	Grat.	Detail	HEAD	Stance		Engagement	SYNTACTIC REALIZATION	Grat.	Detail
1											Hostility from the teaching unions to any proposal for school reform is normally an indication that the plan has merit.			NP+App	NP+NonFinite	Carrier	
2					so				Manner-Advb	it				NP	Pronoun	Carrier	So it is with Michael Gove's decision to press ahead with performance-related pay for teachers.
3					From September				Time	Time-PP	Schools in England and Wales			NP+APP	NP+PP	Actor	From September, schools in England and Wales will abandon the existing staff salary structure that effectively guarantees an automatic annual pay rise for all teachers, regardless of merit.
4											Head teachers			NP	Adj Noun	Beneficiary	Head teachers will instead be given the power to set salary levels that not only recognise the abilities of their staff through annual appraisals but also reflect conditions in the local teacher recruitment market.
5											Teachers			NP	Noun	Sayer	Teachers claim convincingly that such a merit-based system, which is commonplace in the private sector, will lead to "discrimination and unfairness".
6											They			NP	Pronoun	Senser	They are also angry at being singled out for such treatment – there are no such proposals for other parts of the public sector, such as hospitals or prisons.
7	But										there	shared knowledge	There	There		Existent	But there is good reason for that.
8											The system of automatic advancement			NP+App	NP+pp	Carrier	The system of automatic advancement has a deleterious effect on teaching.
9										It				NP	Pronoun	Actor	It allows poor teachers to prosper while the best teachers too often see their talent go unrewarded.
10											The Education Secretary			NP	Det Proper Noun	sayer	The Education Secretary says he wants the best teachers to be paid more without having to leave the classroom by seeking headships.
11											The new set-up			NP	Det adj Noun	Actor	The new set-up will also make it easier for good teachers to move between schools, creating a real market in teacher talent.
12											This			NP	Pronoun	Carrier	This is all to the good.
13											The current system			NP	Det adj Noun	Actor	The current system allows too many under-performing teachers to coast through their careers with head teachers almost powerless to do anything about it.
14											The new structure			NP	Det adj Noun	Actor	The new structure will give heads far greater powers over their own schools and that can only be to the benefit of the pupils.
15											Performance pay			NP	Noun Noun	Carrier	Performance pay is not just about the setting of salary levels; it should also be the catalyst for a wide-ranging cultural change in the teaching profession.

France moves in

CI #	Thematic Field															Text
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION Grat.	Detail	HEAD	Stance	Engagement	SYNTACTIC REALIZATION Grat.	Detail	Experiential Role	
1					To most people		shared knowledge	Contingency-pp	Contingency-pp	Mali			NP	Proper Noun	Carrier	To most people, Mali is a remote West African country known chiefly for the fabled city of Timbuktu.
2	Yet				on Saturday			Time	Time-pp	David Cameron			NP	Proper Noun	Actor	Yet on Saturday David Cameron offered François Hollande two C-17 transport planes to help French forces counter an Islamist rebellion.
3	And				yesterday			Time	Time-Advb	Mark Simmonds, Foreign Office minister for Africa			NP+App	Proper Noun+Proper Noun	sayer	And yesterday Mark Simmonds, Foreign Office minister for Africa, spoke of providing training and support for the Malian army.
4										Why		question	Question	Interrogative Pronoun	Attribute	Why has Britain become involved?
5										Mali			NP	Proper Noun	Carrier	Mali forms part of one of the most unstable regions in the world, the Sahel, running across Africa from Senegal to Eritrea.
6					Last March,			Time	Time-pp	Tuareg tribesmen and members of al-Qaeda in the Islamic Maghreb			NP+NP+App	NP+NP+PP	Actor	Last March, Tuareg tribesmen and members of al-Qaeda in the Islamic Maghreb seized control of the northern part of the country.
7					Last week			Time	Time-pp	they			NP	Pronoun	Actor	Last week they moved south, taking Koma and threatening Mopti, home to the only army garrison between them and the capital, Bamako.
8										The French president			NP	Det Adj Noun	sayer	The French president has responded with commendable speed, deploying helicopter gunships which have helped recapture Koma, thereby slowing up the dilatory response of the United Nations and the regional powers.

### D. Appendix

9									With a civilian administration controlled by a small, weak army		Manner-NonFinite	Mali						NP	Proper Noun	Carrier	With a civilian administration controlled by a small, weak army, Mali has the makings of a failed state which Islamic radicals could use as a springboard for attacks on other African countries, especially Nigeria, as well as on Europe.
10												Their military defeat and a negotiated settlement with more moderate Tuareg elements is, therefore, a strategic priority.						Np>Np>App	Np>Np>PP	Carrier	Their military defeat and a negotiated settlement with more moderate Tuareg elements is, therefore, a strategic priority.
11												Mr Hollande and his European partners						NP>NP	Proper Noun-PProun Adj Noun	Carrier	Mr Hollande and his European partners face a formidable challenge.
12												Seven French hostages						Np	Det Adj Noun	Goal	Seven French hostages are held by the rebels; the army needs building up almost from scratch – its leader, Capt Amadou Sanogo, fears that, having intervened, foreign powers will demand a return to civilian government.
13												Restoring representative rule and training and equipping soldiers to recover the north						Nonfinite-Nonfinite	Nonfinite-Nonfinite	Carrier	Restoring representative rule and training and equipping soldiers to recover the north will take time.
14												France, backed by Britain						NP>App	NP>NonFinite	Beneficiary	But in moving in, France, backed by Britain, has done the right thing.

As the terror threat shifts to Africa, we must meet it with great care

Cl.	Thematic Field		SYNTACTIC REALIZATION										Text				
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION		EXPERIMENTAL						
			Stance	Engagement			Stance	Engagement	Goal	Detail		Stance	Engagement	Goal	Detail	Experimental Role	
1														NP+App	NP+Clause	Actor	The tragic events in Algeria, where Islamists succeeded in seizing 700 hostages, of whom at least a dozen are dead, leaving a grisly trail of charred bodies, reinforce an important lesson for Britain and its allies: the centre of gravity of international jihad is shifting from Asia to North Africa.
2														NP	Pronoun	Actor	That brings the threat closer to what Malcolm Rifkind, the former foreign secretary, last week called the "soft underbelly of Europe", recalling a phrase of Churchill's.
3														Clause		Actor	It is this westward shift that makes France's military campaign in Mali against an array of Islamist insurgents so important.
4							For the same reason,			Contingency	Contingency			NP	Proper Noun	Senser	For the same reason, David Cameron is right to have extended logistical and other assistance to Paris.
5														NP	Noun Noun	Actor	Mali's troubles have been brewing for a long time.
6						After Gaddafi's regime was deposed in nearby Libya			Time	Time-Clause				NP+App	NP+pp	Actor	After Gaddafi's regime was deposed in nearby Libya, pro-regime fighters from the Tuareg ethnic group flooded into Mali, taking looted arms with them.
7														NP	Pronoun	Actor	They joined forces with Islamists to seize large sections of the country's north, only to find themselves sidelined by the better-armed fundamentalists.
8														NP+App	NP+pp	Actor	A coup by Malian officers last spring worsened the situation, producing further chaos.
9														NP+App	NP+pp	Sayer	The region's Economic Community of West African States pledged to send military forces, trained and helped along by the West, to restore order.
10	But													NP	Det Noun	Goal	But that plan was torn up last week, when Islamists threatened Mali's capital Bamako.
11														NP	Proper Noun	actor	France intervened at the request of the Malian president, and with the overwhelming support of the local people.

12										The Algerian attack was almost certainly planned before the French intervened, so it should not be interpreted as "blowback".	Det Adj Noun	Goal
13	But									But it highlights the connection between events in the region, and the type of threat that Mali's Islamist coalition might have posed had it not been stopped (for the time being, at least).	Pronoun	Sayer
14										Algeria was only the latest manifestation of an arc of instability from Yemen in the east to Mauritania in the west.	Proper Noun	Token
15										Even as al-Qaeda's core leadership has been devastated by drone strikes in recent years, its Yemeni affiliate, al-Qaeda in the Arabian Peninsula, has become the group most feared by Western counter-terrorism officials.	NP+App	Carrier
16										In Libya, the Islamists of Ansar al-Sharia played a key role in the attack on the American consulate in Benghazi, which killed the US Ambassador to Libya.	NP+App	Actor
17										There is also some evidence that those Libyan terrorists had been in contact with al-Qaeda in the Islamic Maghreb (AQIM), a group that forms part of the Islamist coalition in Mali, and that was singled out in the Prime Minister's address to the House of Commons on Friday.	There	Existent
18										The spillover from Mali could affect not just these countries, but also Mauritania, Morocco, Nigeria and Tunisia – the last a young, fragile democracy that could easily be shaken.	NP+App	Senser
19										Britain has thus far struck a careful and prudent balance in addressing these daunting challenges.	Proper Noun	Actor
20										North Africa's melange of jihadists must be weakened before it can enrich itself in Mali, and Britain should help France in those areas where it has a particular advantage: transport, intelligence, and possibly special forces.	NP+App	Goal
21										The crisis is also yet another reminder that Britain's influence with allies, and its own national security, require strong military capabilities of the sort that have been drastically shunk over the past two years.	NP	Carrier
22	But									But we in Britain should tread carefully.	NP	Pronoun
23										France, the former colonial power in Mali, possesses a depth of understanding of the region that we cannot hope to match.	NP+App	Carrier
24										Washington is said to have no more than an "impressionistic understanding" of the groups in northern Mali.	NP	Target
25										We will know even less.	NP	Pronoun
26	Moreover									Moreover, many of North Africa's jihadists are small, divided groups with only tenuous links to al-Qaeda and no reach into Europe.	NP	Carrier
27	Finally									Finally, AQIM and its allies in the region would like nothing more than to see a major deployment of Western troops, who would become vulnerable targets for attack on unfamiliar terrain.	NP+NP+App	Senser
28										Britain should therefore emphasise the role of African ground forces in liberating Mali, even as it plays an active role at arm's length.	NP	Actor
29										Algeria's apparently bungled rescue attempt does not inspire confidence.	Noun Adj Noun Noun	Actor
30	But									But only by deepening our security relationship with regional powers will we be able to ensure that jihadists, having been squeezed in Afghanistan and Pakistan, are given no respite closer to Europe.	Pronoun	Actor

## Heads in the sand over Eastern immigration

Cl.	Thematic Field		SYNTACTIC REALIZATION												Text			
	TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION				Experiential Role		
			Stance	Engagement		Stance	Engagement	Grat.	Detail	Stance	Engagement	Grat.	Detail					
1					In 2004,					Time-PP	Britain				NP	Proper Noun	sayer	In 2004, Britain opened its borders to the new Eastern European members of the EU.
2											The Civil Service				NP	Det Proper Noun	sayer	The Civil Service predicted that only 13,000 or so people would arrive: in fact, the eventual total topped a million.
3											This				NP	Pronoun	Token	This was not only one of the most dismal failures of foresight in Whitehall's history, but one of the most damaging.
4					While many of the new arrivals contributed to the economy,					Time	they				NP	Pronoun	Actor	While many of the new arrivals contributed to the economy, they also put intolerable stresses on unprepared public services and stoked such resentment that immigration became one of the most pressing (and most toxic) issues in politics.
5					ing					Verb	history				question		Verb	Is history about to repeat itself?
6					At the end of the year,					Time	controls on migration from Bulgaria and Romania				NP+App	NP+PP	Goal	At the end of the year, controls on migration from Bulgaria and Romania will be relaxed.
7					Again,					Time	the Civil Service				NP	Det Proper Noun	Carrier	Again, the Civil Service is entirely nonchalant, predicting a ripple of arrivals rather than a flood.

#### D. Appendix

8									Ministers.							Ministers, understandably, are not persuaded.
9	True,	Com							many Romanians and Bulgarians							True, many Romanians and Bulgarians already have the right to work here, the boom that drew so many Poles has ended, and this time Britain will not be alone in opening its borders.
10	Indeed	Com							the Romanian ambassador to London, Ion Jinga							Indeed, the Romanian ambassador to London, Ion Jinga, reassured this newspaper this week that the vast majority of those who want to come already have.
11	Yet								there	shared knowledge						Yet there are equally good reasons to believe there will again be a significant influx.
12									Contingency clause	shared knowledge						Due to our relatively robust economy, the large communities of Romanians and Bulgarians already here, the far higher wages on offer and the ease with which new arrivals can access public services and benefits, Britain is a natural destination for such migrants.
13									There	shared knowledge						There are also precedents from elsewhere.
14									Italy							Italy opened its doors to Romania in 2002; by 2011, 968,000 had arrived.
15									Similar numbers							Similar numbers poured into Spain.
16									there	shared knowledge						Then there are the 2.5 million Roma, who suffer persistent discrimination in their home countries and have already decamped in significant numbers.
17									the pressure group MigrationWatch UK							All told, the pressure group MigrationWatch UK expects the number arriving to be between 30,000 and 70,000 per year.
18									It							It is calling for ministers to extend controls on immigration for another five years, but there is little chance of the EU granting permission.
19	Like it or not, Eng								the borders							Like it or not, the borders will open – yet alarmingly, there is little sign that central or local government are taking the issue seriously.
20	Yes,	Com							Eric Pickles							Yes, Eric Pickles has warned of the impact on the property market and social housing, and Theresa May has promised to make it harder for new arrivals to use the NHS and the benefits system, but that hardly compensates for the more general complacency.
21									We	inclusive pronouns						We hope the optimists are right, and there is no repeat of the chaos of 2004.
22	But if								we	inclusive pronouns						But if they are not, we must be properly prepared.
23	Sadly,	Com							the Government							Sadly, at the moment the Government appears to have its head buried firmly in the sand.

Britain is ill-prepared for this deadly new world

[illegible]

8										Irrespective of how this particular incident is resolved,	Contingency Clause	Mr Cameron								Carrier		Irrespective of how this particular incident is resolved, Mr Cameron is right to highlight the fact that we can expect more such attacks in the years to come.
9										In particular,	Manner-PP	he								senser		In particular, he believes that al-Qaeda in the Islamic Maghreb and other north African terrorist groups, which thrive in the ungoverned space of failed states, now have the capacity to pose a "large and existential threat" to our well-being – particularly in Mali, where large tracts of the country have already become a safe haven for militants.
10										By seizing on long-standing local conflicts for their own ends,	Manner-PP	al-Qaeda and its allies								actor		By seizing on long-standing local conflicts for their own ends, al-Qaeda and its allies have added a fresh and disturbing dimension to the turmoil in the area, posing a threat far beyond the region itself.
11										The situation										Carrier		The situation is complicated by the recent seismic changes to the political landscape of the Middle East and north Africa in the wake of the Arab Spring. Islamists have attempted – largely successfully – to assert their influence in Tunisia, Libya and Egypt, and are playing an increasingly prominent role in Syria (undermining William Hague's argument that the West should arm the opposition there).
12										The overthrow of Colonel Gaddafi										carrier		The overthrow of Colonel Gaddafi has had particularly problematic consequences, leading to a collapse of government in Libya's restorer regions, and flooding the rest of the region (not least Mali) with looted arms and ammunition.
13	While									Britain will be wary of further entanglements after the controversies of Afghanistan and Iraq,	Time	there		shared knowledge						Existent		While Britain will be wary of further entanglements after the controversies of Afghanistan and Iraq, there is no doubt that our Armed Forces will – as in Mali – be called upon to play a part in containing the Islamist threat, though the tangled alliances, shifting frontiers and guerrilla tactics that our forces will encounter could make it alarmingly hard to develop the necessary clarity about our aims, tactics and (perhaps most importantly) our exit strategy.
14	Yet if									these are the frontiers of the future	Contingency Clause	the sight of one of our precious C-17s marooned on a French runway								Token		Yet if these are the frontiers of the future, the sight of one of our precious C-17s marooned on a French runway was hardly the best omen.
15	As									Con Coughlin pointed out on these pages yesterday.	Evidence	Britain								Carrier		As Con Coughlin pointed out on these pages yesterday, Britain has only eight of these aircraft, and France none; America has roughly 200.
16												It								Carrier		It is a handy shorthand for our respective military approaches – Britain running down its forces, banking the dividend from leaving Afghanistan and relying on reservists to fill the gaps in its ranks; the US turning to the Pacific and increasingly perturbed (as General Stanley McChrystal, the former Nato commander in Afghanistan, and Leon Panetta, the outgoing defence secretary, have made clear) by the way we have chosen to limit our capabilities, even as our responsibilities grow.
17										At the time of the Government's Strategic Defence and Security Review,	Time-PP	there								Existent		At the time of the Government's Strategic Defence and Security Review, there were warnings that it left Britain perilously ill-prepared to face the uncertainties of an increasingly chaotic world.
18												That analysis								Sayer		That analysis holds.
19	If									Mr Cameron is serious about confronting the threat posed by Islamist terrorists.	Contingency Clause	he								Actor		If Mr Cameron is serious about confronting the threat posed by Islamist terrorists, he should make sure our Armed Forces have the resources and manpower to deny them the safe havens they crave.

This writers tale

CI #	Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	DETAIL	EXPERIMENTAL	Text
1			It was good to see	Eng						so many top sporting events			NP	Adj Adj Noun	Actor	It was good to see so many top sporting events going ahead yesterday despite the snow.
2										Several Heineken Cup rugby matches			NP	Adj Proper Noun Adj Noun	Scope	Several Heineken Cup rugby matches were played on white pitches as teams tried their best to give spectators the entertainment they had paid to see.
3	So		it is a shame	Eng						we		inclusive pronoun	NP	Pronoun	Sayer	So it is a shame we cannot say the same about our transport network or our schools.
4						Not for the first time				the automatic response to bad winter weather			NP+App	NP+PP	Carrier	Not for the first time, the automatic response to bad winter weather has been to shut down, rather than carry on.
5										Heathrow Airport, which spent £36 million to avoid a repeat of the chaos in December 2010, has already cancelled hundreds of flights for today in anticipation of unsuitable conditions.			NP+App	NP+clause	actor	Heathrow Airport, which spent £36 million to avoid a repeat of the chaos in December 2010, has already cancelled hundreds of flights for today in anticipation of unsuitable conditions.
6										Many train operators			NP	Adj Noun Noun	sayer	Many train operators have announced in advance that they are running amended – ie, reduced – timetables.

## D. Appendix

7	And											some schools				NP	Adj Noun	Sayer	And some schools are warning that they might be shut for an entire week unless the cold snap relents.
8		Doubtless,	Com									movement				NP	Noun	Carrier	Doubtless, movement is difficult in rural areas after heavy snowfall, but is it really necessary for transport and schools in towns and cities to be so badly disrupted by what is a natural seasonal occurrence, albeit one experienced less regularly today than in the past?
9												the country				NP	Det Noun	Carrier	Fifty years ago this month, the country was in the grip of the Big Freeze, the coldest winter of the 20th century yet by a long way. The schools stayed open and the trains ran on time.
10	On the other hand											there				There	There	Existent	On the other hand, there was hardly any sport for two months, so maybe we are not such winnips today after all.

Our private schools deserve political praise

CI #	Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	DETAIL	Experimental Role	Text
1										Britain's independent schools			NP	Proper Noun Adj Noun	Carrier	Britain's independent schools are something that we ought to celebrate.
2										Their reputation for excellence			NP+App	NP+PP	Actor	Their reputation for excellence spans the globe.
3	Yet as					we report today	self-mention		Evidence-clause	headmasters			NP	Noun	Carrier	Yet, as we report today, headmasters are sick and tired of their schools being demonised both by elements of the political class and in society at large.
4										Tim Hands, the incoming chairman of the Headmasters' and Headmistresses' Conference, tells The Sunday Telegraph: "There is a tendency for Joe Public to think about independent schools with a 'them and us' mentality in which independent schools represent toffs and are therefore to be tilted at."			NP+App	Proper Noun+NP	Sayer	Tim Hands, the incoming chairman of the Headmasters' and Headmistresses' Conference, tells The Sunday Telegraph: "There is a tendency for Joe Public to think about independent schools with a 'them and us' mentality in which independent schools represent toffs and are therefore to be tilted at."
5										His comments			NP	Det Noun	Verbiage	His comments echo those of Frances King, headmistress of Roedean, who revealed in January that she was stepping down from her job, saying she was frustrated with always being "on the negative side of public opinion".
6										Anthony Seldon, the master of Wellington College in Berkshire			NP+App	Proper Noun+NP	Sayer	Anthony Seldon, the master of Wellington College in Berkshire, has spoken of a "suppressed hatred" in parts of society towards former public-school pupils.
7										This so-called "posh prejudice"			NP	Det Adj Noun	Goal	This so-called "posh prejudice" has not been helped by politicians.
8	If					anything	hedge		Contingency	they			NP	Pronoun	Senser	If anything, they have encouraged it.
9										we			NP	Pronoun	Carrier	We are used to seeing the Left attack private schools, despite many Labour MPs having benefited from them.
10										Ed Balls, Harriet Harman and Tessa Jowell			NP+App	Proper Noun+NP	Carrier	Ed Balls, Harriet Harman and Tessa Jowell were all privately educated, and the socialist firebrand Diane Abbott once said that the private school she sent her son to was "the making of him".
11	But									what is staggering, and disappointing, is			Clause	Det Noun Adj	Actor	But what is staggering, and disappointing, is the Coalition's failure to show moral support.
12										Mr Hands			NP	Proper Noun	Sayer	Mr Hands says that, despite being theoretically committed to choice in education, the Government has sent mixed messages to the independent sector and failed to let it "know where it stands".
13	When					he spoke at the Conservative Party conference in 2012			Time-clause	the Eton-educated David Cameron			NP	Det Noun Adj Proper Noun	Sayer	When he spoke at the Conservative Party conference in 2012, the Eton-educated David Cameron mentioned that he went to a "great school" but declined to name it.

14	At least									Mr Cameron's overly cautious choice of language is preferable to the opportunism of the Westminster-educated Nick Clegg.	Carrier	Proper Noun AAbb Adj noun prep noun								
15										In a speech last year, Mr Clegg suggested that the superior achievement of private schools was somehow 'corrosive' to society – only to indicate recently in a radio interview that he would happily send his own child to one if his wife decided on it.	Sayer	Proper Noun								
16										"I want my child to have the best possible education," he told callers.	Sayer	Pronoun								
17										"It's the most fundamental instinct anyone has."	Verbage									
18										Mr Clegg is absolutely right in his assessment of human nature.	Carrier	Proper Noun								
19	But									But how can wanting the best for your child be a "fundamental instinct" one minute and yet "corrosive" the next?	Attribute	Interrogative Pronoun								
20										Not for the first time can the Deputy Prime Minister be accused of a convenient U-turn.	Target									
21										Politicians on both sides of the divide seem almost paranoid about speaking the truth, which is that the financial autonomy of the independent sector has made it into a beacon of excellence.	Carrier	NP+pp								
22										Its academic results are superb – last summer, almost a third of privately educated teenagers gained straight A grades at A-level – and its contribution to Britain's culture profound (more than a third of British medal winners in the 2012 London Olympics were from private schools).	Carrier	Det Adj Noun								
23										Many independent schools engage with their local communities and next year Eton will sponsor a non-fee-paying school.	Actor	Adj adj Noun								
24										Far from being the preserve of "toffs", many also provide financial support for poorer pupils; in 2012 one third of Oxford University's bursaries to undergraduates from low-income backgrounds were distributed to students from independent schools.	Actor	Pronoun								
25										Some parents make huge financial sacrifices to send their children to a private school.	Actor	Det Noun								
26										There is nothing elitist about parents following their "fundamental instincts" and trying to get the best education for their child.	Existent	There								
27										A survey conducted last year found that 57 per cent of the public would send their child to a private school if they could afford it.	Actor	Det Noun								
28										This is not necessarily a judgment on the state sector, which continues to see exam results improve.	Carrier	Pronoun								
29										Rather, it is an endorsement of the high quality of private schools, their small class sizes and their emphasis on building confidence in young people.	Carrier	Pronoun								
30										Rather than ignoring or disparaging these great British institutions, politicians should hold them up as an example to learn from.	Actor	Noun								
31										Ensensible if the politicians in question happened to have attended ones.	None									

India, a useful friend

[illegible]

#### D. Appendix

[illegible]

**Courts and Parliament must work together**

Thematic Field		CLA														Text	
TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Experiential Role			
		Stance	Engagement		Stance	Engagement	Grat.	Stance	Engagement	Grat.	Stance	Engagement	Grat.			Detail	
																Few things perplex people more than our apparent inability to remove from the country foreign nationalists who have committed serious crimes here.	
1															Senser	Too many have been allowed to stay because they have put down roots and claim a right to family life under Article 8 of the European convention.	
2															Beneficiary	Last year, the Commons voted to amend the immigration rules to make it clear that anyone convicted of a crime attracting a prison term of a year or more should only exceptionally use the convention as a means to avoid deportation.	
3															Actor	Yet according to Theresa May, the Home Secretary, this expression of Parliament's will is routinely being ignored by the courts.	
4	Yet														Carrier		
5															sayer	"Some judges evidently do not regard a debate in Parliament on new immigration rules as evidence that Parliament wants to see those new rules implemented," she said.	
6															Carrier	This is an extremely serious accusation for which there is considerable anecdotal evidence, although Mrs May will be under pressure to produce chapter and verse in order to justify such an onslaught.	
7	Indeed	Con													Carrier	Indeed, the vehemence of the language she has used to criticise the judiciary is unprecedented from such a high-ranking member of the government.	
8															Senser	The British system of governance relies upon maintaining mutual trust and respect among its constituent parts – the executive, the legislature and the judiciary.	
9	But														Carrier	But this is another sign that it is beginning to break down, with judges critical of the avalanche of new laws coming out of Westminster, ministers openly derisive of civil servants, and MPs more	



10	In addition									the authority wielded by supranational judicial bodies such as the European Court of Justice and the European Court of Human Rights means that the traditional balance of powers is no longer easy to identify.	NP+App	NP+NonFinite	Token	In addition, the authority wielded by supranational judicial bodies such as the European Court of Justice and the European Court of Human Rights means that the traditional balance of powers is no longer easy to identify.
11	It is clear that	Eng								Under Article 8, it is clear that the right to a family life is not absolute but qualified.	NP+App	Carrier	Carrier	Under Article 8, it is clear that the right to a family life is not absolute but qualified.
12										The Commons has additionally stated that other than in very exceptional circumstances, foreigners who have committed serious crimes in this country, or who have attempted to cheat the immigration system, should be deported.	NP	Det Noun	Sayer	The Commons has additionally stated that other than in very exceptional circumstances, foreigners who have committed serious crimes in this country, or who have attempted to cheat the immigration system, should be deported.
13	If									If this is not happening then the judges who are wilfully ignoring the clear wishes of Parliament should be removed.	NP+App	NP+clause	Goal	If this is not happening then the judges who are wilfully ignoring the clear wishes of Parliament should be removed.
14										Mrs May is now preparing new legislation to leave no room for doubt.	NP	Proper Noun	Actor	Mrs May is now preparing new legislation to leave no room for doubt.
15										She is right to say judges do not make the laws but are there to uphold them.	NP	Pronoun	Carrier	She is right to say judges do not make the laws but are there to uphold them.
16	However	Eng								However, in the past when MPs have denounced judicial decisions, it has turned out that the courts have only been fulfilling what they believed Parliament wanted.	NP	Det Noun	Actor	However, in the past when MPs have denounced judicial decisions, it has turned out that the courts have only been fulfilling what they believed Parliament wanted.
17	After all	Mod								After all, it was our legislators who passed the Human Rights Act.	Clause		Actor	After all, it was our legislators who passed the Human Rights Act.
18										It could always be repeated if ministers want to send a real message of intent.	NP	Pronoun	Goal	It could always be repeated if ministers want to send a real message of intent.

Ed Miliband hasn't learnt any economic lessons

Thematic Field	TEXTUAL INTERP.	Stance	Engagement	PREHEAD	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experiential Role	Text
1									NP	Proper Noun	Actor	Ed Miliband travelled to Bedford yesterday to make a much-rumoured "major speech on the economy."
2									NP	Pronoun	Actor	He chose Bedford because it was there in 1957 that Harold Macmillan told the British people that they had "never had it so good".
3									Quote		Verbiage	Now, said Mr Miliband, they would never have it good again unless they voted Labour.
4									NP	Pronoun	Actor	He also borrowed a phrase from another Conservative prime minister, Benjamin Disraeli, for the refrain that ran throughout his address: "One Nation."
5									NP	Pronoun	Senser	You might have thought that here was a party leader, after many years in opposition, seeking to repair the failures of a government that had spent too long in office.
6	Yet								NP+App	NP+clause	Carrier	Yet many of the economic difficulties he railed against are of Labour's creation – not least the country's indebtedness and the red tape strangling the small businesses he now holds in such high regard.
7									NP+App	NP+pp	Carrier	Nowhere in Mr Miliband's speech was there a hint of contrition for Labour's handling of the economy or any apology for inflating the public sector to such a gargantuan size that it sucked the life out of private enterprise.
8									NonFinite		Carrier	To listen to Mr Miliband was to believe that the last government bequeathed a benign legacy that the Coalition has squandered, a world in which welfare spending was not out of control, where schools provided wonderful education for all, and our hospitals were paragons of care and compassion.
9	True	Com							NP	Pronoun	sayer	True, he criticised Gordon Brown (in whose Cabinet he and Ed Balls were happy to sit) for abolishing the 10p tax rate.
10				Isn't					NonFinite	Question	Carrier	Isn't trashing past leaders a somewhat distasteful Soviet-era trait?
11	But			not for a moment	Time-PP				NP	Pronoun	Senser	But not for a moment did he regret the overall direction taken by the last government.
12		Com							NP	Proper Noun	Carrier	Indeed, Mr Miliband appears to have learnt nothing: the first policy that he has earmarked for possible inclusion in the next election manifesto is classically Old Labour.

## D. Appendix

13										Labour	hedge	NP	Proper Noun	Actor	Labour would reintroduce the 10p income tax rate but fund it from a levy on £2 million houses.
14										This,		NP	Pronoun	Verbiage	This, he claims, would demonstrate that those who make a contribution to the economy will be rewarded.
15	So									Mr Miliband		NP	Proper Noun	Carrier	So, far from being a One Nation standard-bearer, Mr Miliband is in reality the champion of
16										people		NP	Noun	Actor	divisive old-style "us versus them" politics.
17										Recent figures	question	NP	Adj Noun	Actor	After all, don't people in large houses make a contribution to the economy?
18										Living standards		NP	Adj Noun	Actor	Recent figures showed that more than 60 per cent of the UK's income tax is paid by 14 per cent of taxpayers, and the numbers paying any tax at all have fallen by 1.4 million under the Coalition.
19										Whose		NP	Adj Noun	Actor	Living standards have declined, real incomes are back to where they were 10 years ago and the country is struggling to return to growth.
20										our memories	Question	Question	Interrogative Pronoun	Attribute	Whose fault is that?
										question		NP	Det Noun	Process	Are our memories so short that we will let Labour get away with such a cynical rewriting of recent history?

Britain's hypocrisy over Sri Lanka

CI #	Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experiential Role	Text
1										Few Western countries			NP	Adj adj Noun	Sayer	Few Western countries were louder than Britain in condemning the Sri Lankan army's shocking violations of human rights as it moved to suppress a bloody rebellion by ethnic Tamils demanding a separate state.
2										those words of reproach			NP+App	NP+pp	Carrier	Clearly, those words of reproach were either not intended to be taken that seriously, or we just move on fast from such squabbles because, as this newspaper reveals, the UK sold several million pounds worth of arms last year to a regime we castigated only a few years ago.
3										"arms" in this instance			NP+App	NP+pp	Carrier	Let anyone imagine that "arms" in this instance refers only to equipment of a purely defensive character, it should be pointed out that the sales included pistols, assault rifles, shotguns and ammunition.
4										The government			NP	Det Noun	Sayer	The government will no doubt respond that we have to engage with governments we don't necessarily like, that defence forms a vital part of our export trade and that if we hadn't sold arms to Sri Lanka, another less conscience-stricken country would have done so.
5										The government	hedge		NP	Det Noun	Actor	The government can also count on the lack of an outcry over the sales.
6										The Tamil cause in Sri Lanka			NP+App	NP+App	Actor	The Tamil cause in Sri Lanka never engaged the British public that much, not least because the so-called Tamil Tigers committed grisly crimes of their own.
7										Rheme			Rheme		None	Not are we alone in moderating our former criticism of the Sri Lanka authorities.
8	As								Time	nervousness in the West			NP+App	NP+App	Actor	As Beijing draws Colombo into its own economic and military orbit, nervousness in the West is growing about the implications of losing influence in Sri Lanka.
9										These			NP	Pronoun	Carrier	These are real considerations, but they beg the question of why we then claim to exercise stringent controls over arms sales and that there is an ethical dimension to them.
10										France, another developed democracy			NP+App	Proper Noun+NP	Actor	France, another developed democracy, makes no such claims.
11										Britain, on the other hand			NP+App	Proper Noun+NP	Actor	Britain, on the other hand, simply courts ridicule and bolsters a reputation for hypocrisy by denouncing a government for egregious violations of human rights while at the same time selling it guns.

The market needs watching

CI #	Thematic Field	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experiential Role	Text
1										The opinion we report on today	self mention		NP+App	NP+clause	Actor	The opinion poll we report on today finds that British people blame food companies rather than the Government for the horsemeat scandal.

[illegible]

### D. Appendix

27							It is Mr Paterson,		Clause	Predicated theme	Carrier	It is Mr Paterson, above all, who has questions to answer.
----	--	--	--	--	--	--	--------------------	--	--------	------------------	---------	--

Governments should not own banks

Thematic Field		SYNTACTIC REALIZATION										Text		
TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail	Experiential Role	Text
1								Rheme			rheme		None	Another year, another highly politicised row about bonuses at state-owned Royal Bank of Scotland.
2								Rheme			rheme		None	Nor is 2013 looking any less controversial than 2012.
3								The institution			NP	Det Noun	Actor	The institution bailed out to the tune of £45bn four years ago, and expected to report yet more losses next week, is all set to hand out an extra £250m to its staff.
4								That the Treasury is drawing up plans either to sell off the government stake or to give it away to taxpayers.			clause		Actor	That the Treasury is drawing up plans either to sell off the government stake, or to give it away to taxpayers, is therefore to be welcomed.
5	Indeed,	Com		with this year's jumbleaxe set to be even more jarring for taxpayers – in the light of the £30m fine for Libor-rigging levied this month –			Manner	Manner-Clause			NP+App	NP+PP	Carrier	Indeed, with this year's jumbleaxe set to be even more jarring for taxpayers – in the light of the £30m fine for Libor-rigging levied this month – the case for a swift exit is stronger than ever.
6								The Chancellor	hedge		NP		Actor	The Chancellor may have dodged the most obvious flat by specifying that the Libor penalty must be met by bank staff.
7	But							the sink over excess pay			NP+App	NP+PP	Actor	But the sink over excess pay continues nonetheless, and the £800,000 bonus due to chief executive Stephen Innes is the subject of particularly impassioned debate.
8	Yet							George Osborne's options			NP	Proper Noun Noun	Carrier	Yet George Osborne's options are limited.
9				Simply	booster		Manner	Manner-Adv			NP	Adj Noun	Actor	Simply slashing pay risks the recovery we, as shareholders, so desperately need.
10								Here			Adv/PP	Adv/b	Carrier	Here, in fact, is the nub: government ownership of a bank is fraught with tensions which cannot be resolved.
11	As			such			Manner	Manner-NP			NP	Det Noun	Carrier	As such, the arrangement is best terminated.
12								It			NP	Pronoun	Senser	It will hurt the Exchequer, of course.
13								Initial hopes of a profit from the rescue	attitude		NP+App	NP+PP	Carrier	Initial hopes of a profit from the rescue were swiftly punctured.
14	And although			the share price has shot up by more than a third in the past 12 months		shared knowledge	Contingency	Contingency-ly-Clause	There		There		Existent	And although the share price has shot up by more than a third in the past 12 months, there is another 40 per cent-plus to go before the state breaks even.
15								Rheme			rheme		None	No matter.
16								Rheme			Rheme		None	Better to cut RBS loose at a cost than to continue in attempts to reconcile the irreconcilable.
17								The Chancellor			NP	Proper Noun	Senser	The Chancellor will want to wait until the run-up to the 2015 election to put his plans into action.
18								So blatant a piece of political bribery	attitude		Adv/PP		Senser	So blatant a piece of political bribery only emphasises the need to get the Government out of banking as soon as may be.
19								The state			NP	Det Noun	Actor	The state should be gone from RBS before the next bonus season.

A President at the very height of his power

Thematic Field	Text
23	

TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Experimental Role		
		Stance	Engagement		Stance	Engagement	Grd.		Stance	Engagement	Grd.	Detail	Detail	Experimental Role		
1								what				Question	Interrogative Pronoun	attribute		What a difference an election victory makes.
2								President Obama's State of the Union – with the exception of his emotional appeal for action on gun control – may have lacked the soaring cadences of his second inaugural speech three weeks earlier.				Question	NP+App	Actor		President Obama's State of the Union address on Tuesday evening – with the exception of his emotional appeal for action on gun control – may have lacked the soaring cadences of his second inaugural speech three weeks earlier.
3	Instead		Mod					it				NP	Pronoun	Carrier		Instead, it was a long and eminently sensible list of goals, enumerated by a man confident in his power and visibly liberated by the knowledge that he will never face the voters again.
4								His central theme – the need to improve the lot of the middle class				NP+App	NP+NonFinite	Carrier		His central theme – the need to improve the lot of the middle class – may have been familiar, but it is no less true for that.
5								Nothing				NP	Pronoun	Actor		Nothing has more perniciously eroded national morale than the ever-widening gap between rich and poor and the decades-long stagnation, or worse, of the living standards of ordinary Americans.
6	Rightly		Com					Mr. Obama				NP	Proper Noun	Actor		Rightly, Mr. Obama pressed for improved education and job training, the closure of tax loopholes that favour the wealthy, and a substantial increase in the minimum wage.
7								His call for action on climate change, like his demand for tighter gun legislation				NP+App	NP+pp	Carrier		His call for action on climate change, like his demand for tighter gun legislation, was common sense.
8								How				Question	Interrogative Pronoun	Attribute		How much he will achieve is another matter, however.
9								Any deal to avoid impending automatic spending cuts of \$85bn and pave the way for a long-term deficit-reduction plan – the overriding domestic issue – the overriding domestic issue				Question	NP+App	Actor		Any deal to avoid impending automatic spending cuts of \$85bn and pave the way for a long-term deficit-reduction plan – the overriding domestic issue – will require Republican support.
10	Sadly,		Com					neither the President's speech, nor the Republican response from Senator Marco Rubio				AdjP		Actor		Sadly, neither the President's speech, nor the Republican response from Senator Marco Rubio, offered much sign of common ground.
11	As							both				NP	Pronoun	Carrier		As regards America's fiscal problems, both were mere re-statements of respective party positions.
12								the sense of a new assuredness in Mr. Obama				NP+App	NP+pp	Carrier		Even so, the sense of a new assuredness in Mr. Obama was unmissable.
13								the President's poise				NP	Det Noun Noun	Actor		In part, the President's poise stems from his clear victory in November, after a campaign fought on many of the topics of Tuesday's address.
14								The second factor				NP	Det Det Noun	Carrier		The second factor is the corollary of the above: the soundly defeated Republicans know something must now change.

### D. Appendix

[illegible]

William Hague has been outflanked on Syria

Thematic Field		SYNTACTIC REALIZATION								Text	
Cl #		PREHEAD	Stance	Engagement	Grat.	HEAD	Stance	Engagement	Grat.	Detail	Experiential Role
TEXTUAL	INTERP.										
1							Britain's inability to persuade its fellow EU countries to have weapons supplied to non-Islamist groups in the Syrian opposition was a diplomatic failure waiting to happen.				Carrier
2	The result is that		Eng				Britain				Goal
3		To invert Douglas Hurd's famous formulation, we are punching below our weight, pressing for aggressive solutions which we lack both the muscle to impose ourselves, and the powers of argument to convince our allies to support.				Contingency we y-Nonf finite					Actor
4		Rewind					Rewind				Process
5		Then				Time-Advb that	It was Russia that				Carrier
6		Fresh from our relatively painless involvement in the uprising that brought down Mummur Gaddafi, London, along with Paris and Washington, used all its diplomatic influence to secure a UN resolution that would hasten the Syrian dictator's fall.				Manner-Adjp	London, along with Paris and Washington				Actor

7	But									Russia (along with China)								NP+App	Proper Noun+pp	Actor	But Russia (along with China) blocked all those initiatives.
8		With hindsight	Eng							there								There	Exist	Exist	With hindsight there are several reasons why the UK can be grateful for its failure to secure a UN resolution.
9																		NP	Carrier	Carrier	One is that Syria is not Libya.
10																		NP	Proper Noun	Goal	Libya may be riven by tribal loyalties, but it is largely homogeneous in ethnic and religious terms.
11																		NP	Adj Adj Noun	Carrier	Solid democratic governance may take many years to put down roots there, but the sort of basic political stability that is a prerequisite for such a development does not seem completely out of the question.
12																		NP	Proper Noun	Carrier	Syria, by contrast, is a religious and ethnic patchwork.
13																		NP	Pronoun	Senser	No one can doubt that President Assad's Alawite clique has only clung to power through the ruthless use of force.
14	But																	There	Exist	Exist	But there can be no confidence that the fall of Mr Assad would be a prelude to peace.
15		Far more likely	Sta															NP+App	NP+clause	Carrier	Far more likely, it would be only the first act in a civil war that could make the massacres of the past two years look modest.
16																					In this context, Britain's urge to arm the secular rebel groups while keeping Jhadi ones at arm's length looks naive for a once big power with long experience in the Middle East.
17	Once																	NP	Pronoun	Actor	Once the weapons are in the theatre, they will remain in play.
18	And if																	There	Exist	Exist	And if they are in the hands of people we think of as the good guys today, there is no guarantee that they will remain so tomorrow.
19																		There	Exist	Exist	Then there is Russia.
20	While																	NP	Proper Noun	Actor	While Britain, tentatively supported by France, continues to try to hasten a military resolution to the conflict, Moscow has further undermined those efforts by forging ties of its own with some of the Syrian rebel groups, without sacrificing its support for Mr Assad.
21	As a result	It is not inconceivable that	Eng															NP	Proper Noun	Actor	As a result it is not inconceivable that Moscow will emerge, sooner rather than later, as the only outside power sufficiently trusted by both sides to mediate between them.
22																		NP	Proper Noun	Actor	Libya may have whetted the appetite of Britain and France for again exercising disproportionate influence in North Africa, but President Hollande has learned quickly from his Mali adventure how treacherous those sands can be.
23																		NP+NP	Proper Noun+Proper Noun	Actor	Former Secretary of State Hillary Clinton and Defence Secretary Leon Panetta wanted to supply Syria's rebels with arms, but President Obama, ever more reluctant to commit the US to conflicts like these, vetoed the idea.
24																		NP	Proper Noun	Senser	William Hague may well be motivated by the sort of humanitarian compulsions that were conspicuously absent from Conservative policy towards Bosnia 20 years ago.
25	But																	NP	Noun Noun	Carrier	But however righteous his motives, yesterday's debacle is a timely reminder that there are severe limits both to the role of force in resolving the conflicts that emerged from the Arab Spring, and to the

### The Italian joker

Thematic Field													Text		
CI	TEXTUAL	INTERP.	PREHEAD				SYNTACTIC REALIZATION				HEAD	EXPERIENTIAL			
			Stance	Engagement				Stance	Engagement	Grd.			Stance	Engagement	Role
												European			Actor
European leaders have been patting themselves on the back with increasing vigour in recent															

## D. Appendix

2																		The promise by Mario Draghi, the head of the European Central Bank, to drain the coffers dry before letting the single currency fail appeared to have stabilised the situation, or at least to have bought some time.
3																		Events in Italy, however, show the extent to which such hopes are built on sand rather than rock. In the run-up to the country's election, which concludes today, it was widely expected that the centre-left party, led by Pier Luigi Bersani, would form a coalition with the centrist grouping run by Mario Monti, the technocrat parachuted in to replace Silvio Berlusconi when Europe's leaders tired of his political and fiscal indiscipline.
4																		Despite some grumbling, and some populist gestures, enough of Mr Monti's programme of austerity and structural reform would be retained to satisfy the markets, and the country's paymasters.
5																		How shocking, and unnerving, for those in Brussels to see the Five Star Movement, a band of novices led by the outspoken comedian Beppe Grillo, seize the initiative on an anti-politics and anti-euro ticket.
6																		Thanks to a moratorium on opinion polling, nobody knows whether this guerrilla army will maintain its momentum, or even whether it will hold together after the election.
7	But																	But its success shows that a generation failed by the system – in the shape of sky-high unemployment and debt – is not willing to play along with politics as usual. It is not just Italy: look at events in Spain, or Greece.
8																		The national crises that threw the eurozone into turmoil have not gone away, the setting has been turned down to simmer rather than boil.
9																		All it will take is one election, or one government carving in to protest, and the heat will be back on.

Osborne needs to budget for economic growth

Thematic Field	SYNTACTIC REALIZATION													Text	
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	SYNTACTIC REALIZATION	Detail		Experiential Role
1					By his own standards,			Evidence-pp	Evidence-pp			NP	Proper Noun	Goal	By his own standards, Ed Balls has been relatively restrained in heaping calumny on George Osborne, following the news that Moody's, the ratings agency, is to downgrade Britain from AAA – its highest tier of creditworthiness – to AA1. Perhaps this was because the statement in which the agency justified its action will have made even less comfortable reading for the shadow chancellor than his opposite number. It blamed, first, the slow growth of the global economy – something beyond either man's power – and second, "the drag on the UK economy from the ongoing domestic public- and private-sector deleveraging process". In layman's terms, this means that the economy is still struggling to cope with the hideous legacy left by Labour in particular the ruined state in which it left the public finances. The report also makes it clear that the only reason Britain's stable outlook is stable rather than negative is because Moody's believes that the Coalition retains the political will to see through its deficit reduction programme – something few would expect Mr Balls to prioritise, given his long history of spending more on borrowing and taxing more. All the same, the damage to Mr Osborne is more than merely symbolic. Even though the reputations of the ratings agencies were grievously damaged by their part in the sub-prime mortgage fiasco, he made the maintenance of the triple-A rating a cornerstone of his strategy, and his credibility.
2		Perhaps	Mod						this			NP	Pronoun	Carrier	
3									It			NP	Pronoun	Sayer	
4					In layman's terms			Evidence-pp	this			NP	Pronoun	Token	
5								Evidence	The report			NP	Det Noun	Actor	
6					All the same,			Manner-AdvP	the damage to Mr Osborne	attitude		NP+App	NP+PP	Carrier	
7	Even though				the reputations of the ratings agencies were grievously damaged by their part in the sub-prime mortgage fiasco,	Shared knowledge	Contingency	Contingency	he			NP	Pronoun	Actor	



## We all pay the cost of the BBC's failures

501

### D. Appendix

[illegible]

## Time running out for talks with Iran

Thematic field															Text
CI	TEXTUAL			PREHEAD			SYNTACTIC REALIZATION			HEAD	SYNTACTIC REALIZATION			Experiential Role	
		INTERP.	Stance	Engagement		Stance	Engagement	Grat.		Stance	Engagement	Grat.	Detail		
1															

2	And								Despite tentatively positive noises from both Tehran and Washington in recent months, decades of stalemate can hardly inspire optimism.	Actor	NP+PP								
3									And the latest developments leave the likelihood of a deal slimmer than ever	Actor	Det Adj Noun								
4									Last week, the UN nuclear watchdog confirmed that Iran was going ahead with plans to install next generation centrifuges at its plant at Natanz.	Sayer	Det Proper noun Noun Adj								
5									The new equipment can enrich uranium at two or three times the rate of that it replaces, raising Western fears of an escalation of the weapons programme that Tehran consistently denies.	Actor	Det Adj Noun								
6	Even if							Stu	Even if – and it is a big if – Iran is not trying to build a weapon now, faster centrifuges could give it the ability to make a sudden "dash" for a bomb.	Actor	Proper Noun								
7									Then last weekend, the Iran Atomic Energy Organisation announced the discovery of sufficient raw uranium deposits to reble the country's stock.	Sayer	Proper Noun								
8									It also claimed to have identified 16 more sites for the civil nuclear power stations that Tehran claims to be the ultimate destination for its nuclear fuel.	Sayer	Pronoun								
9									Neither move inspires confidence ahead of this week's talks.	Senser	Noun								
10	And when								And when the so-called P5+1 group – Russia, China, Britain, France, Germany and the US – meet in Almaty both issues will be high on the agenda.	Carrier	Adj Noun								
11									There is no option but to press ahead as planned, however.	Existent									
12	And								And that means offering Iran more tempting incentives to scale back its enrichment programme.	Token	Pronoun								
13								Eng	It is not entirely unrealistic to hope Tehran might show some willingness to negotiate in return.	Actor	Proper Noun								
14									For all the recent posturing, sanctions are biting too hard to be tolerated indefinitely.	Actor	Noun								
15									It is as well to be realistic, though.	Carrier	Pronoun								
16									Ayatollah Ali Khamenei has little room to manoeuvre ahead of the presidential elections in June.	Carrier	Proper Noun								
17									There is still time for diplomatic bridge-building.	Existent									
18									Even with the new kit, Iran's bomb-making capability is not yet beyond control.	Carrier	Proper Noun Adj Noun								
19									Israel has backed off, at least in the immediate term.	Actor	Proper Noun								
20	But								But the game of cat-and-mouse that has gone on for so long cannot go on forever.	Carrier	Npr+PP								

## Lessons from Pollard that the BBC still has to learn

Thematic Field																	Text	
CI#	TEXTUAL		INTERP.		PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION				Experimental Role
			Stance	Engagement				Stance	Engagement	Grat.				Stance	Engagement	Grat.	Detail	
1												There	attitude	shared knowledge	There		Existent	There is probably no organisation on earth that would come up smelling only of roses if its underbelly were picked over as exhaustively as the BBC's editorial processes have been in the wake of the Jimmy Savile affair.
2												The decision to publish more than 1,000 pages of documents supporting Nick Pollard's report			NP+App	NP+NonFinite	Phenomenon	
																		The decision to publish more than 1,000 pages of documents supporting Nick Pollard's report can be regarded as brave, or foolhardy, or simply – as we would see it – the duty of a publicly funded organisation.

## D. Appendix

3	But	it is a pity that	Eng							But it is a pity that fear of libel suits – or was it a last sliver of face-saving? – forced the redaction of the pithier remarks from Jeremy Paxman and others.										Actor	
4										In the event, for all the muteness, back-passing and back-biting this mass of material exposes, it only reinforces Mr Pollard's conclusions: that the BBC's lines of accountability were deeply flawed and that members of its lavishly paid and top-heavy management saw it as their priority, in a crisis, to flout responsibility rather than take it.										Actor	
5	And									And here it might be pertinent to mention the recent appointment of James Purnell, former TV										None	
6										and Labour minister, as the BBC's director of strategy.										Beneficiary	
7										Mr Purnell, it appears, will not only receive a salary just short of £300,000, but will also qualify for up to two years' salary if he leaves.										Actor	
8										Less favourable terms are to be introduced for senior recruits in a few weeks.										None	
										Did old-style Corporation cronyism bring Mr Purnell in just under the wire?										None	

## **D.5    Annotated Editorials in Spanish**

## D. Appendix

### SPANISH EDITORIALS

#### Yak-42: justicia a medias

El tribunal admite que la falsa identificación fue intencionada pero no dice los motivos

CI #	Thematic Field	SYNTACTIC REALIZATION												Text		
TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Experiential Role		
		Stance	Engagement		Stance	Engagement	Gral.	Detail		Stance	Engagement	Gral.	Detail			
1									La falsa identificación de 30 de los 62 militares españoles muertos en el accidente del avión Yak-42, el 26 de mayo de 2003, cuando se apostaba a aterrizaj en la base turca de Trabzon, procedente de Afganistán			NP+App	NP+Clause	Carrier	La falsa identificación de 30 de los 62 militares españoles muertos en el accidente del avión Yak-42, el 26 de mayo de 2003, cuando se apostaba a aterrizaj en la base turca de Trabzon, procedente de Afganistán, no fue error, sino un acto intencionado, perpetrado a sabiendas.	
2									La Audiencia Nacional			NP	Proper Noun	Stayer		La Audiencia Nacional ni siquiera ha entrado a deliberar si se trató de una mera imprudencia: todas las pruebas le han llevado a considerar el hecho como un delito de falsedad en documento oficial realizado de modo consistente y con conocimiento de causa, es decir, con dolo.
3									Las consecuencias			NP	Det Noun	Carrier		Las consecuencias son importantes.
4				Agrava			Verb	Verb-agrar	3 p sing Presente Indicativo			Verb	3 p sing Presente Indicativo	process		Agrava laconducta de los acusados haciéndola mercedora de penas de prisión: tres años para el principal acusado, el general de Sanidad Vicente Navarro, y año y medio para el comandante y el capitán médicos José Ramón Ramírez y Miguel Ángel Sanz.
5				En el caso del general Navarro, ya expensas de posibles recursos,			Contingency	Contingency-PP+Conj+pp	la pena de tres años			NP+App	NP+PP	Carrier		En el caso del general Navarro, y a expensas de posibles recursos, la pena de tres años supone su ingreso en prisión.
6	No obstante,								el fallo de la Audiencia			NP+App	Det Noun+PP	Carrier		No obstante, el fallo de la Audiencia se queda en la mitad de las penas de cinco y seis años, respectivamente, solicitadas para este acusado por el fiscal. Y las acusaciones en representación de las familias de las víctimas.
7									El tribunal			NP	Det Noun	Stayer		El tribunal argumenta por qué: la falsa identificación de los 30 cadáveres responde "a una sola decisión delictiva, a una sola voluntad criminal".
8				cabe			Verb	Verb	3 p sing Presente Indicativo			Verb	3 p sing Presente Indicativo	process		No cabe, pues, hablar de un delito continuado.
9	Y			en lo que respecta a los otros dos acusados,			Manner	Manner+PP	la omisión de determinados datos en los informes de necropsia que realizaron			NP+App	NP+Clause	Stayer		Y en lo que respecta a los otros dos acusados, la omisión de determinados datos en los informes de necropsia que realizaron coadyuvó a "la acción falsaria de Navarro", pero no constituyó delito por sí misma.
10									El tribunal			NP	Det Noun	Stayer		El tribunal describe con detalle la conducta dolosa de Navarro, su voluntad de fallar a la verdad.
11	Pero								Nada			NP	Pronoun	Verbage		Pero nada dice sobre por qué actuó así, salvo la explicación del propio Navarro de que su misión era repatriar los cadáveres.
12									Desconocer todo sobre esa misión, si existió o no, y si respondió, de existir, a órdenes superiores, directas o indirectas			Nonfinite+finite clause	Carrier		Desconocer todo sobre esa misión, si existió o no, y si respondió, de existir, a órdenes superiores, directas o indirectas, constituye una grave laguna procesal que, de no corregirse, seguiría proyectando zonas de sombra sobre la tragedia del Yak-42.	
13	Quizás	Mod		para curarse en salud,			Contingency	Contingency-PP	el tribunal			NP	Det Noun	Actor		Quizás para curarse en salud, el tribunal se cuida de delimitar el objeto de su veredicto, negando que pueda extenderse "a valoraciones ajenas al objeto procesal ni a la conducta de personas no acusadas y sobre las que, en consecuencia, no se ha practicado prueba ni se ha formado convicción alguna".
14	Pero	es difícil admitir que	Eng						indagar sobre la existencia y naturaleza de la misión que parece haber condicionado la conducta del principal acusado	hedje		Nonfinite	Carrier			Pero es difícil admitir que indagar sobre la existencia y naturaleza de la misión que parece haber condicionado la conducta del principal acusado sea tarea ajena al tribunal, bien directamente, bien por deducción de testimonio.

## SPANISH EDITORIALS

## Yak-42: justicia a medias

El tribunal admite que la falsa identificación fue intencionada pero no dice los motivos

Thematic Field		PREHEAD										HEAD				SYNTACTIC REALIZATION				EXPERIENTIAL		Text	
Cl #	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD		Stance	Engagement	Grat.	Detail	SYNTACTIC REALIZATION		Stance	Engagement	Grat.	Detail	Stance	Engagement	Grat.	Detail	Role		
1																						Carrier	La falsa identificación de 30 de los 62 militares españoles muertos en el accidente del avión Yak-42, el 26 de mayo de 2003, cuando se apostaba a aterrizar en la base turca de Trabzon, procedente de Afganistán, no fue error, sino un acto intencionado, perpetrado a sabiendas.
2																						Sayer	La Audiencia Nacional ni siquiera ha entrado a deliberar si se trató de una mera imprudencia: todas las pruebas le han llevado a considerar el hecho como un delito de falsedad en documento oficial realizado de modo consciente y con conocimiento de causa, es decir, con dolo.
3																						Carrier	Las consecuencias son importantes.
4																						process	Agrava la conducta de los acusados haciéndola merecedora de penas de prisión: tres años para el principal acusado, el general de Sanidad Vicente Navarro, y año y medio para el comandante y el capitán médicos José Ramón Ramírez y Miguel Ángel Sanz.
5																						Carrier	En el caso del general Navarro, y a expensas de posibles recursos, la pena de tres años supone su ingreso en prisión.
6	No obstante,																					Carrier	No obstante, el fallo de la Audiencia se queda en la mitad de las penas de cinco y seis años, respectivamente, solicitadas para este acusado por el fiscal y las acusaciones en representación de las familias de las víctimas.
7																						Sayer	El tribunal argumenta por qué: la falsa identificación de los 30 cadáveres responde "a una sola decisión delictiva, a una sola voluntad criminal".
8																						process	No cabe, pues, hablar de un delito continuado.
9	Y																					Senser	Y en lo que respecta a los otros dos acusados, la omisión de determinados datos en los informes de necropsia que realizaron coadyuvó a "la acción falsaria de Navarro", pero no constituyó delito por sí misma.
10																						Sayer	El tribunal describe con detalle la conducta dolosa de Navarro, su voluntad de fallar a la verdad.
11	Pero																					Verbiage	Pero nada dice sobre por qué actuó así, salvo la explicación del propio Navarro de que su misión era reparar los cadáveres.
12																						Carrier	Desconocer todo sobre esa misión, si existió o no, y si respondió, de existir, a órdenes superiores, directas o indirectas, constituye una grave laguna procesal que, de no corregirse, seguirá proyectando zonas de sombra sobre la tragedia del Yak-42.
13		Quizás	Mod																			Actor	Quizás para curarse en salud, el tribunal se cuida de delimitar el objeto de su veredicto, negando que pueda extenderse "a valoraciones ajenas al objeto procesal ni a la conducta de personas no acusadas y sobre las que, en consecuencia, no se ha practicado prueba ni se ha formado convicción alguna".
14	Pero	es difícil admitir que	Eng																			Carrier	Pero es difícil admitir que indagar sobre la existencia y naturaleza de la misión que parece haber condicionado la conducta del principal acusado sea tarea ajena al tribunal, bien directamente, bien por deducción de testimonio.

### D. Appendix

15							Las familias de las víctimas	NP+App	Det Noun+PP	Carrier	Las familias de las víctimas tienen motivos para estar satisfechas. Sobre todo, tras los iniciales obstáculos de la propia justicia, reacia a actuar, y la escasa atención inicial de la opinión pública a su lucha por la exigencia de responsabilidades
16							La impunidad	NP	Det Noun	Carrier	La impunidad no ha prevalecido, pero queda pendiente saber por qué los acusados actuaron como lo hicieron.

## Sonrojo en Westminster

Los escandalosos gastos de los diputados británicos comienzan a pasar factura política (20/05/2009) Elpais online

Thematic Field		SENTENCE										Text			
Cl #	TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Experiential Role
			Stance	Engagement		Stance	Engagement		Stance	Engagement	Grad.	Detail			
1	Es poco probable que	Sita									NP+App	NP+PP	Actor	Es poco probable que el escándalo de los gastos de los diputados británicos a costa del contribuyente afecte decisivamente a la intención de voto de los ciudadanos, pero es seguro que ha asertado un golpe formidable a la reputación de un Parlamento considerado modélico en el exterior y honorable más allá de toda duda en el interior.	
2									attitude		Np+App	Np+AdvP+PP	Carrier	La insólita dimisión de su presidente, primera en más de tres siglos, con ser imprescindible por la incompetencia del laborista Michael Martin en el manejo del asunto y sus intentos para ocultar los abusos, no va a ser el fin de la historia.	
3											NP+App	Det Noun+Proper Noun	Stayer	El primer ministro Gordon Brown intentó salir a ver al paso del desastre anunciando un acuerdo interpartidista para la reforma del sistema, pero Westminster y los partidos en su conjunto, y eso es lo grave, han perdido la confianza pública.	
4											NP+App	Np+Clause	Carrier	Las noticias que el diario conservador Daily Telegraph viene publicando son motivo no sólo para desgastar los actuales mecanismos de control -que ni laboristas ni conservadores han tenido el menor interés en cambiar-, sino probablemente para iniciar por fin una reforma que acabe con los profundos anacronismos del sistema democrático más aliñado del mundo.	
5	Si		hedge		Contingency			Contingency- Clause			Verb	3 p sing Presente Indicativo	process	Si en ese sistema donde los Lorens todavía no son plenamente elegidos o sólo dos partidos tienen monopolícamente el control del poder, los Comunes pueden diseñar un tinglado de inaceptables gabelas para que los ciudadanos corran con los gastos particulares de sus representantes, vulgares o extravagantes, es que algo será no funciona.	
5											NP	Det Noun	Actor	Las corruptelas afectan a todos los partidos y a diputados de toda edad y condición.	
7	Peró								attitude		Np+App	NP+PP	Actor	Peró la losa del desprestigio de un pilar básico de la democracia cae más pesadamente sobre el laborismo germanote, a cuyo frente el desacreditado premier recoge la ira creciente de unos ciudadanos que, en medio del despliarse, hacen equilibrios para llegar a final de mes.	
8											NP	Proper Noun	Stayer	Brown ha rechazado adelantarse a unas elecciones, como le pide la oposición conservadora, que presumiblemente perderá por goleada. Los diputados británicos no están bien pagados, ni puede considerarse corruptos para los estándares de otros países europeos incluidos.	
9	Peró										NP	Pronoun	Actor	Peró muchos han cometido el error de considerar sus injustificables gastos consentidos como un complemento salarial de libre disposición.	
10	Y										NP	Det Noun	Actor	Y pocas cosas se toleran peor en una democracia que la sensación de que hay unas reglas para los que mandan y otras para la leonés.	

**Sigue imputado**

**Figure Imputatio**



CI #	Thematic Field	TEXTUAL	INTERP.	PREHEAD	SYNTACTIC REALIZATION				HEAD	SYNTACTIC REALIZATION				Experiential Role	Text
		Stance	Engagement		Stance	Engagement	Grat.	Detail		Stance	Engagement	Grat.	Detail		
1									El presidente de la Generalitat valenciana, Francisco Camps,			Np-App	Np-Propri Noun	Sayer	El presidente de la Generalitat valenciana, Francisco Camps, compareció ayer como imputado ante el Tribunal Superior de Justicia de la Comunidad Valenciana, del que salió, 45 minutos después, con la confirmación de que sigue imputado.
2									El tribunal			NP	Det Noun	Actor	El tribunal facilitó una escueta nota en la que afirma que "el proceso sigue su curso".
3									Esto			NP	Pronoun	Carrier	Esto significa que los indicios de delito de cohecho apreciados por el Juez Garzón antes de inhibirse en favor del Tribunal Superior (por la condición de aforados de tres de los investigadores) no han sido disueltos por la declaración del presidente; pese a lo que el mismo había dado a entender que ocurriría al manifestar que nada desahaba tanto como poder comparecer ante el juez para aclarar el asunto.
4									Los dirigentes populares			NP	Det Noun Adj	Senser	Los dirigentes populares habían pretendido confundir a la opinión pública presentando la condición de imputado como un mero detalle procesal derivado de la forma en que había instruido el caso el Juez Garzón.
5	Pero								el juez valenciano			NP	Det Noun Adj	Actor	Pero el juez valenciano no ha encontrado motivos para modificar esa condición, lo que en principio significa que aprecia los mismos indicios.
6	Si							Contingency	el juez			NP	Det Noun	Sayer	Si de la declaración se hubiera deducido una posición más favorable para el reo, el juez tendría que haberlo traducido en alguna medida.
7									3 p sing Presente Indicativo			Verb	3 p sing Presente Indicativo	process	No ha habido tal, aunque tampoco ha solicitado medidas cautelares, lo que habría indicado reforzamiento de los indicios contra el reo.
8									El empeño del PP en quitar importancia a la imputación			Np-App	NP-NonFinite	Carrier	El empeño del PP en quitar importancia a la imputación estaba motivado por su negativa a adoptar con alguien tan relevante en el partido las mismas medidas (suspensión de millonaria) aplicadas a otros imputados de menor nombrada.
9									El PP			NP	Proper Noun	Carrier	El PP tiene cada día más difícil mantener esa actitud, contraria a lo establecido en su propio código de conducta y que supone aplicar dos criterios diferentes en Madrid y en Valencia.
10	porque								Camps			NP	Proper Noun	Actor	Porque, entretanto, Camps sigue sin presentar facturas que, desmintiendo la documentación que figura en el sumario, acreditasen su afirmación de que él se pagó los trajes que presuntamente le regaló el jefe de la trama corrupta en Valencia.
11	y								3 p sing Presente Indicativo			Verb	3 p sing Presente Indicativo	process	y sigue también sin ofrecer una explicación que reste verosimilitud a la de que trajes y regalos pagaban un trato de favor en adjudicaciones.
12									la dirección del PP			NP-App	NP-PP	Actor	Desde que la Audiencia Nacional adoptó las primeras decisiones sobre el asunto, la dirección del PP ha ido cambiando de estrategia para, en último extremo, encontrarse en la actual situación: todas las medidas políticas que se ha resistido a adoptar contra los implicados se vuelven contra la totalidad del partido una vez que las actuaciones judiciales se han puesto en marcha.
13									La instrucción			NP	Det Noun	Actor	La instrucción, entretanto, sigue su curso, y las excusas que los populares han ido desgranando hasta ahora van perdiendo verosimilitud a medida que se van conociendo nuevos datos. Al contrario de lo que el comité de recepción a Camps ante las puertas del tribunal ha querido dar a entender, no se trata de un proceso por motivos políticos, sino de un proceso por corrupción en el que, eso sí, están implicados varios políticos.

## D. Appendix

CI #	Thematic Field	PREHEAD					SYNTACTIC REALIZATION					HEAD					SYNTACTIC REALIZATION					Text				
		TEXTUAL	INTERP.	Stance	Engagement		Stance	Engagement	Gral.	Detail		Stance	Engagement	Gral.	Detail		Stance	Engagement	Gral.	Detail		Stance	Engagement	Gral.	Detail	
1																										
2																										
3																										
4																										
5																										
6																										
7																										
8																										
9																										
10																										
11																										
12																										
13																										

En poco tiempo se verá	Time	Time-pp	se	NP	Pronoun	Senser
14						En poco tiempo se verá si la política de Netanyahu lleva a Israel a una decisión trascendental para la paz y la seguridad mundiales; la de si actuar o no por su cuenta contra Irán.

## Medidas improvisadas

Falta elaboración política en las propuestas para salvar los sectores del automóvil e inmobiliario (23/05 /2009) Elpais online

Thematic Field		Text															
Ci #	INTERP.	PREHEAD				HEAD				SYNTACTIC REALIZATION				Experiential Role			
		Stance	Engagement			Stance	Engagement	Grail.	Detail		Stance	Engagement	Grail.	Detail			
1										Una de las virtudes			NP+App	NP+pp	Goal	Una de las virtudes que con razón se le atribuyen al debate sobre el estado de la nación (también en este periódico) fue el grado de concreción de las propuestas económicas del presidente del Gobierno.	
2					attitude			Time	Time-NonFinite	Rodriguez Zapatero			NP	Proper Noun	Sayer	Tras un diagnóstico bastante acertado y un reconocimiento razonable de los errores de previsión, Rodríguez Zapatero anunció la rebaja paulatina de la designación fiscal por vivienda y la implantación de ayudas directas a la compra de automóviles para contrarrestar el brutal descenso de la demanda.	
3	Pero									ambas propuestas					Det Noun	Pero ambas propuestas perdieron fuerza en la tramitación parlamentaria, y en el caso de la medida estrella para reanimar el mercado automovilístico, se está desperdiciando por una gestión desautorizada con las comunidades autónomas.	
4	De forma que							Time	Time-Advb	los potenciales compradores, los concesionarios, los Gobiernos autónomos y el propio Gobierno central			NP-NP+NP+NP	NP+NP+NP+NP	Senser	De forma que hoy los potenciales compradores, los concesionarios, los Gobiernos autónomos y el propio Gobierno central desconocen exactamente quién pondrá los 2.000 euros de subvención prometidos, y los modelos y plazos incluidos.	
5										Rheme				Rheme		None	Un patron económico más ordenado, en resumen.
6								Verb	Verb-ser	3 p sing Presente indicativo			Verb	3 p sing Presente Indicativo	process	Es un propósito que identifí bien el programa electoral del PSOE en 2004, pero que ha sido olvidado en estos años.	
7							A ese propósito se	Contingency	Contingency-pp	se			NP	Pronoun	Actor	A ese propósito se dirigen algunas decisiones acertadas, como la supresión de la desgravación por compra de vivienda o la simplificación administrativa para la creación de empresas, incluida la extensión de las posibilidades de interllocución digital de los ciudadanos con el conjunto de las administraciones.	
8										La decepción			NP	Det Noun	Actor	La decepción ha venido luego, al saberse que gran parte de las medidas propuestas carecían de la mínima elaboración para aplicarse inmediatamente.	
9							A la improvisación del Ejecutivo se han abalido	Accompaniment	Accompaniment-ent-PP+Verb	3 p pl Presente indicativo			Verb	3 p plu Presente Indicativo	process	A la improvisación del Ejecutivo se han abalido las dificultades para conseguir apoyos parlamentarios para alcanzar la aprobación de las decisiones.	
10										Las prisas por hacer anuncios, por buscar sólo el efecto de los enunciados, sin dotarlos de contenido			NP+App	NP+NonFinite	Carrier	Las prisas por hacer anuncios, por buscar sólo el efecto de los enunciados, sin dotarlos de contenido, no es una buena forma de gobernar.	
11										Todo enunciado		directive	NP	Det Noun	Carrier	Todo enunciado ha de tener definido con precisión su itinerario de concreción y aplicación.	
12	y									Rheme				Rheme	None	Y en mayor medida si ésta depende del concurso de otras administraciones y de sus presupuestos.	
13										El momento				Det Noun	Sayer	El momento no aconseja ni excesos retóricos ni grandes representaciones.	
14							En mitad de una recesión	Place	Place-pp	ciertas medidas en ámbitos como la fiscalidad, los estímulos de la demanda o la reforma del sistema bancario			NP+App	NP+pp	Actor	En mitad de una recesión, ciertas medidas en ámbitos como la fiscalidad, los estímulos de la demanda o la reforma del sistema bancario exigen amplios acuerdos.	
15	y						. deben presentarse	directive	Verb	3 p plu	directive	Verb		3 p plu Presente Indicativo	process	Y, desde luego, deben presentarse una vez bien estudiadas y con la viabilidad de su aplicación asegurada.	
e										Gobierno y PP		directive		Np+np	Carrier	Gobierno y PP necesitan fortalecer la seriedad de sus compromisos con los electores antes de la celebración nacional de elecciones generales.	

## D. Appendix

### Equilibrio de poderes

Al contradecir al Gobierno, el Constitucional refuerza la legitimidad de la Ley de Partidos (23/05 /2009) Elpals online

Thematic Field	CI #	TEXTUAL	INTERP.	PREHEAD	Stance	Engagement	Gral.	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Gral.	SYNTACTIC REALIZATION	Detail	Experiential Role	Text
	1								El Tribunal Constitucional				NP	Proper Noun	Actor	El Tribunal Constitucional ha estimado el recurso de la coalición Iniciativa Interacionalista La Solidaridad con los Pueblos (IISP) contra su anulación por parte del Supremo, que la consideró un instrumento de ETA-Batasuna para burlar su prohibición y concurrir a las elecciones europeas.
	2			Con esta resolución se refuerza			Contingency-PP+Verb		se				NP	Pronoun	Actor	Con esta resolución se refuerza el equilibrio de poderes del sistema político, que incluye un Tribunal Constitucional que no necesariamente avala lo que defiende el Ejecutivo.
	3								La Ley de Partidos				NP	Proper Noun	Carrier	La Ley de Partidos es una norma excepcional motivada por una situación que también lo es.
	4								Esa excepcionalidad				NP	Det Noun	Token	Esa excepcionalidad implica dos riesgos principales: que se relativicen, en nombre de las razones de Estado, las exigencias propias del Estado de derecho y del procedimiento judicial, y que la excepcionalidad en principio prevista para el condeñado que dirige ETA se extienda a situaciones, grupos y personas cada vez más alejados de ese núcleo.
	5	Por una parte,							La sentencia				NP	Det Noun	Sayer	La sentencia incide en ambos aspectos.
	6						Mode-Adxb+Verb		se				NP	Pronoun	Senser	Por una parte, no se conforma con los indicios obvios de voluntad defraudatoria de Batasuna, y constata que no existe en el auto confirmación de que esa voluntad se haya materializado en la candidatura señalada.
	7	Por otra,		encuentra			Verb		3 p sing Presente Indicativo				Verb	3 p sing Presente Indicativo	process	Por otra, no encuentra indicios suficientes para acreditar la integración en una trama al servicio de ETA o de formaciones ligadas a ella.
	8								No cualquier actitud favorable al mundo de la izquierda abertzale				NP+App	NP+PP	Carrier	No cualquier actitud favorable al mundo de la izquierda abertzale implica formar parte del entramado terrorista.
	9								La legitimación que para la política antiterrorista supone que la legalización de partidos dependa de los tribunales, y no del Gobierno				NP+App	NP+Clause	Verbiage	La legitimación que para la política antiterrorista supone que la ilegalización de partidos dependa de los tribunales, y no del Gobierno, se ve confirmada por una sentencia que contradice la posición del Ejecutivo, cosa que no habla ocurrido hasta ahora.
	10								Ello				NP	Pronoun	Actor	Ello debilita las críticas nacionalistas de que la ley se aplicaba en función de conveniencias políticas del Gobierno de turno.
	11								El Constitucional				NP	Proper Noun	Sayer	El Constitucional viene a decir que, sin una base probatoria sólida de dependencia de ETA, la aplicación de la ley no puede limitar derechos fundamentales, como el de participación política.
	12			En el caso de ANV			Manner-PP		las pruebas decisivas de esa dependencia				NP+App	NP+PP	Actor	En el caso de ANV, las pruebas decisivas de esa dependencia aparecieron bastante después de las elecciones municipales.
	13	Por lo mismo,							la sentencia				NP	Det Noun	Actor	Por lo mismo, la sentencia protege lo que podía haber sido un flanco débil de la ley con vistas al recurso planteado ante el Tribunal de Estrasburgo, que se expanda, sin las garantías del Estado de derecho, a partidos legales preexistentes.
	14			Ya no puede			Time-PP+Verb		3 p sing Presente Indicativo				Verb	3 p sing Presente Indicativo	process	Ya no puede defenderse esa teoría.

### El reto de Pyongyang

El régimen norcoreano prosigue, contra todos, su alarmante camino hacia un arsenal nuclear (26/05 /2009) Elpals online

Thematic Field	CI #	TEXTUAL	INTERP.	PREHEAD	SYNTACTIC REALIZATION	HEAD	Experiential Role	Text

[illegible]

## Camps y Torquemada

**Camps y Torquemada**  
Elarajoy descalifica a la justicia mientras el líder valenciano busca la absolución en las urnas (26/05/2009) Elpais online

CI #	Thematic Field	Text
	TEXTUAL INTERP	
	PRE-HEAD	
	HEAD	
	SYNTACTIC	
	EXPOSITIONAL	

### D. Appendix

		Stance	Engagement	Grat.	Detail	Stance	Engagement	Grat.	Detail	Proper Noun	Actor	
1									El Partido Popular	NP		<b>El Partido Popular</b> ha cedido a la tentación de confundir unos eventuales resultados favorables en las próximas elecciones europeas con una absolución en los casos de corrupción que afectan a destacados dirigentes y cargos institucionales, como Francisco Camps.
2									Esta reacción	NP	Carrier	<b>Esta reacción</b> no constituye una sorpresa, puesto que se debía trasladar en las declaraciones de los líderes populares, incluido el propio Camps, desde que trascendieron las primeras decisiones de la Audiencia Nacional relacionadas con el caso <i>Gürtel</i> .
3	Pero				Degree-PP+Verb				3 p sing Presente Indicativo	Verb	process	<b>Pero más grave aún es</b> que esta forma de eludir responsabilidades tampoco representa una novedad en las estrategias del PP para enfrentar los casos de corrupción que le rodean. En este sentido, el propio Camps ha admitido que ofreció en los procesos judiciales contra el presidente de la Diputación de Castellón, Carlos Fabra, para quien el fiscal acaba de solicitar dos años de cárcel.
4									La dirección nacional de los populares, con Mariano Rajoy a la cabeza	NP+app	Actor	<b>La dirección nacional de los populares, con Mariano Rajoy a la cabeza</b> , ha decidido cerrar filas con el presidente de la Generalitat valenciana.
5								La incondicionalidad de este apoyo	NP+app	NP+app	Carrier	<b>La incondicionalidad de este apoyo</b> no es sólo asunto del PP y de sus complejos equilibrios internos, en los que Camps llegó a afirmarse como un sólido apoyo de Rajoy durante el aceso de Esperanza Aguirre y de los sectores más radicales del partido tras la derrota en las últimas elecciones generales; es también, y sobre todo, una cuestión que afecta al obligado respeto a los tribunales desde las partidas políticas.
6									Rajoy	NP	Sayer	<b>Rajoy</b> ha apuntado a los "torquemada del siglo XXI" como causantes de los problemas de Camps con la justicia.
7									Las palabras del líder de la oposición	NP+App	Sayer	<b>Las palabras del líder de la oposición</b> desafían gratuitamente la actuación de los jueces y, en concreto, del instructor del caso en el Tribunal Superior de Justicia de Valencia
8					Contingency- Clause+Verb				3 p sing Presente Indicativo	Verb	process	<b>Sea cual sea la hipótesis que se escoja, parece</b> inapropiado dar una fuerza política con vocación de Gobierno propagar insinuaciones sobre la justicia en lugar de exigir responsabilidades a sus militantes.
9									Eso	NP	Carrier	<b>Eso</b> es lo que establece el código de conducta del PP y esa es la decisión que ha adoptado en relación con los implicados en la Comunidad de Madrid.
10									Los problemas judiciales de Camps	NP+App	Carrier	<b>Los problemas judiciales de Camps</b> se deben exclusivamente a comportamientos sobre los que no ha ofrecido ninguna explicación.
11	y								el cerrado apoyo de Rajoy	NP+App	Toten	<b>Y el cerrado apoyo de Rajoy</b> sólo significa, por su parte, que el Partido Popular ha renunciado a exigírsela.
12					Manner-NonFinite				el PP y su dirección nacional	NP+NP	Carrier	<b>Al seguir este camino, el PP y su dirección nacional</b> no quedan a merced de los resultados electorales, como pretenden sus dirigentes en un intento de repetir la estrategia ensayada con Fabra; quedan a merced de lo que decidan los tribunales
13	y				Time-pp				lo que han decidido los tribunales	Clause	Actor	<b>Y de momento, lo que han decidido los tribunales</b> ya es ha llevado a proteger a destacados imputados por cohecho.

## Terrorismo en el Sahel

CI #	Thematic field										Text				
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION		Experiential Role		
			Stance	Engagement	Stance	Engagement	Detail	Stance	Engagement	Detail	Stance	Engagement	Detail		
1												NP+Alp	NP+pp	Actor	El asalto de un grupo de terroristas islámicos a la central ganista de An A menas, 1.600 kilómetros al sur de Argel, ha puesto sobre la mesa dos cuestiones vitales para Europa: la seguridad frente a Al Qaeda y la dependencia energética del norte de África, una región sacudida desde hace dos años

2		Durante años,			Time	Time-PP	el Gobierno argelino			NP	Det Noun	Sayer	Durante años, el Gobierno argelino se ha quejado de la poca ayuda que recibía de Europa para combatir a los grupos armados de inspiración islámica.
3		Para desespero de las autoridades argelinas	attitude		Manner	Manner-PP	los países europeos, entre ellos España			NP+App	NP+PP	Actor	Para desespero de las autoridades argelinas, los países europeos, entre ellos España, han pagado decenas de millones de euros para rescatar a sus nacionales secuestrados por diversas bandas criminales en el Sahel.
4							los terroristas			NP	Det Noun	Actor	Los terroristas han utilizado este dinero para afianzar sus posiciones en la región y cometer más atentados y secuestros.
5		Cargados con millones de euros,			Manner	Manner-NonFinite	los líderes de Al Qaeda en el Magreb islámico (AQMI)			NP+App	NP+PP	Carrier	Cargados con millones de euros, los líderes de Al Qaeda en el Magreb islámico (AQMI) no tuvieron ningún problema para acudir en el mercado negro las armas que volaron de los arsenales libios tan pronto como cayó el régimen de Gadaffi.
6							Los asaltantes de Ain Amenas			NP+App	NP+PP	Actor	Los asaltantes de Ain Amenas no han equipados sólo con armas ligeras, sino que, al parecer, llevaban lanzacohetes y explosivos muy sofisticados con los que pretendían hacer salir por los aires la cuarta planta productora de gas de Argelia.
7		De haber podido alcanzar este objetivo			Contingency	Contingency-NonFinite	el suministro de energía a Europa, especialmente a España,			NP+App	NP+AdvP	Goal	De haber podido alcanzar este objetivo, el suministro de energía a Europa, especialmente a España, se hubiera visto muy afectado.
8							La coalición internacional que ayudó a los rebeldes libios a derrocar a Gadaffi			NP+App	NP+Clause	Actor	La coalición internacional que ayudó a los rebeldes libios a derrocar a Gadaffi ha contribuido bien poco a que el nuevo Estado sea capaz de controlar su territorio.
9	Tan...que	Tan crítica es la situación en el sur del país, que	attitude		Degree	Degree-Clause	el Gobierno de Trípoli			NP+App	NP+PP	Senser	Tan crítica es la situación en el sur del país, que el Gobierno de Trípoli ha decidido cerrar las fronteras con Argelia, Níger, Chad y Sudán.
10							Ain Amenas			NP	Proper Noun	Carrier	Ain Amenas está a sólo un centenar de kilómetros de Libia y fue de allí, precisamente, de donde salieron los terroristas.
11	Ahora,	como demuestra la guerra en Mali			Manner	Manner-Clause	las bandas criminales que conforman AQMI	attitude		NP+App	NP+Clause	Carrier	Ahora, como demuestra la guerra en Mali, parece obvio que las bandas criminales que conforman AQMI tienen la fuerza necesaria para convertir el corazón del Sahel en un territorio hostil a Europa, dominado por la concepción más atrasada del islam.
12	Aun así a pesar de todo	lo que se juega,		Shared knowledge	Contingency	Contingency-Clause	la Unión Europea			NP	Proper Noun	Actor	Aun así, a pesar de todo lo que se juega, la Unión Europea ha dejado a Francia prácticamente sola en su lucha contra los islamistas y a favor de un país, Mali, que hasta hace muy poco pasaba por ser una de las democracias más estables del noroeste de África.
13		Es cierto que	Eng				Francia			NP	Proper Noun	Carrier	Es cierto que Francia tiene muchos más intereses en la región que el resto de sus socios europeos.
14							Las minas de uranio del norte de Níger,			NP+App	NP+PP	Actor	Las minas de uranio del norte de Níger, por ejemplo, alimentan las centrales nucleares sin las que la economía del país no podría funcionar.
15	Sin embargo						la Unión Europea	directive		NP	Proper Noun	Carrier	Sin embargo, la Unión Europea debería haber tenido un papel mucho más activo, y no sólo ahora, sino desde el estallido de las primaveras árabes, para garantizar que las transiciones políticas en el norte de África no facilitarían el auge del terrorismo.
16							Argelia, fiel a sus principios,			NP+App	Proper Noun+Np	Actor	Argelia, fiel a sus principios, ha resuelto por la fuerza el ataque a Ain Amenas. Consideran que no hay mejor estrategia para combatir el terrorismo.
17	Ahora, sin embargo,						el Gobierno argelino	directive		NP	Det Noun Adj	Actor	Ahora, sin embargo, el Gobierno argelino debería reforzar la seguridad en los yacimientos de gas y petróleo, y Europa debería ofrecer su ayuda siempre que sea requerida.
18							El terrorismo islámico, bien nutrido de armas y dinero, sólo será derrotado en el Sahel con una acción conjunta de los gobiernos de la zona, reforzada por Estados Unidos y la Unión Europea.			NP+App	NP+NonFinite	Goal	El terrorismo islámico, bien nutrido de armas y dinero, sólo será derrotado en el Sahel con una acción conjunta de los gobiernos de la zona, reforzada por Estados Unidos y la Unión Europea.

### Grecia al límite

## D. Appendix

	TEXTUAL INTERP.			PREHEAD			SYNTACTIC REALIZATION			HEAD			SYNTACTIC REALIZATION			Experiential Role		
	Stance	Engagement		Stance	Engagement		Gral.	Detail		Stance	Engagement	Gral.	Detail			actor		
1																		UNA serie de ataques con fusiles de asalto, bombas incendiarias y cocteles molotov se han sucedido en los últimos días en Grecia, un país que ha entrado en su sexto año de recesión y al que los ajustes y subidas de impuestos han llevado a una situación límite.
2																		La preocupación del Gobierno heleno ante esta escalada de violencia que ha afectado a políticos -como el primer ministro conservador, Andonis Samaras, cuya oficina fue tiroteada la madrugada del lunes-, a sedes de su partido, Nueva Democracia, a partidos de la oposición, como el Pasok, y a periodistas, ha llegado a un punto tal que incluso se habla de clima de guerra civil.
3	Aunque																	Aunque no ha habido de momento daños personales y todavía no se conoce la autoría de estos ataques -solamente muchas miradas se dirigen hacia el partido izquierdista Syriza, al que acusan de no condenar claramente el uso de la violencia o incluso, como hizo ayer el portavoz del Ejecutivo, de alimentarla con sus ataques verbales.
4																		Mañana, el Parlamento griego debe votar la apertura de una investigación contra diversos políticos, entre los que se hallan los ex primeros ministros Yorgos Papandreu y Lukas Papadimos, por sus responsabilidades al haberse negado a investigar un escándalo de evasión fiscal cuyas listas, publicadas, afectan a más de 2.000 ciudadanos griegos, entre ellos varios políticos y empresarios.
5																		Un escándalo que ha conmocionado a la opinión pública y que puede desestabilizar al Gobierno.
6																		Todo ello hace que Grecia viva unos momentos particularmente tensos.
Mientras																		Mientras las colas del paro no hacen más que crecer, los precios energéticos convierten la calefacción en un lujo y la economía sigue hundándose en un paisaje sin horizontes, la aparición de la violencia armada no ha hecho más que proyectar negros presagios sobre el futuro.
7																		
Mientras																		Mientras la mitad de la población griega rehúsa entre los desperdicios, la otra mitad, rechaza las subidas de impuestos, en una cruda imagen publicada por un conocido periodista.
8																		
																		La impresión en la calle es que el Gobierno de Samaras, una coalición de Nueva Democracia, el socialista Pasok e Izquierda Democrática, como los anteriores, apenas hace nada para resolver los problemas diarios de los griegos y sólo se preocupa por mantenerse en el poder sin que las numerosas huelgas y manifestaciones de protesta hayan logrado modificar un ápice la situación.
9																		
																		De ahí la enorme preocupación ante esta nueva escalada de violencia en el país, que ha superado todos los límites.
10																		



Grave error en Nissan

Cl #	Thematic Field										Text						
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION			HEAD		SYNTACTIC REALIZATION		Experiential Role					
			Stance	Engagement		Stance	Engagement	Gral.	Detail		Stance	Engagement	Gral.	Detail			
1										El fracaso inicial en las negociaciones entre la dirección de Nissan en Barcelona y los sindicatos			NP+App	NP+PP	Carrier	El fracaso inicial en las negociaciones entre la dirección de Nissan en Barcelona y los sindicatos para fabricar un nuevo modelo de automóvil fue un error muy grave.	
2										Los sindicatos			NP		Det Noun	Sayer	Los sindicatos rechazaron inicialmente las propuestas de la empresa exigidas en el último momento para aumentar el tiempo efectivo de trabajo por el mismo sueldo.
3		Realmente	Mod							este hecho			NP		Det Noun	Carrier	Realmente, este hecho supone un sacrificio que no le gusta a nadie, pero que han de aceptar buena parte de los trabajadores del país para poder mantener el empleo, en una Catalunya en la que hay más de 700.000 parados.
4	Sin embargo									la primera actitud de rechazo sindical			NP+App	NP+PP	Carrier		Sin embargo, la primera actitud de rechazo sindical fue matizada rápidamente por la asamblea de los trabajadores afiliados a Sigra-USOC, que el sábado se mostraron dispuestos a retomar el diálogo y a aceptar la última propuesta que la empresa puso sobre la mesa.
5					Sobre los empleados pesa	Place			Place-PP	pesa			NP		Det Noun	process	Sobre los empleados pesa la dura crisis por la que atraviesa la industria del automóvil en toda Europa, y especialmente en Francia, en donde la propia multinacional franco-japonesa, propietaria también de Renault, se plantea la supresión de 7.500 empleos en las plantas de dicho país.
6										La propuesta de fabricar un nuevo turismo en Barcelona, por más sacrificios que ello exija ahora			NP+App	NP+PP	Carrier		La propuesta de fabricar un nuevo turismo en Barcelona, por más sacrificios que ello exija ahora, supone la creación de cuatro mil puestos de trabajo (mil directos y tres mil indirectos).
7										La fabricación del nuevo modelo			NP+App	NP+PP	Token		La fabricación del nuevo modelo puede significar, además, inversiones para modernizar la planta de Barcelona, que se hará más competitiva, y reforzará la confianza de la multinacional japonesa para futuros proyectos.
8										La más que probable rectificación de la posición de los sindicatos	hedge		NP+App	NP+PP	Carrier		La más que probable rectificación de la posición de los sindicatos será una señal clara de la voluntad de los trabajadores de ponerse al frente de la crisis, aspecto que también deberá valorar la compañía a la hora de seguir contando con Barcelona como ciudad apta para la industria automovilística, que actualmente pasa por una delicada situación.
9										El futuro de este sector			NP+App	NP+PP	Carrier		El futuro de este sector depende de la mejora de la competitividad a través de mejoras tecnológicas y de reducción de costes, incluidos los salariales.
10		Es duro de aceptar,	Eng		hay	Shared knowledge			Verb	hay			Verb			Existent	Es duro de aceptar, pero no hay más alternativa en el actual mundo globalizado si se quiere mantener el empleo.
11										Las cúpulas sindicales y los trabajadores de Nissan			NP+NP		Actor		Las cúpulas sindicales y los trabajadores de Nissan han hecho bien al reflexionar en las últimas horas para modificar sus planteamientos iniciales e intentar que la multinacional acepte fabricar el nuevo modelo de turismo en Barcelona.
12					es	booster			Verb	es			Verb			process	Es mucho más que un automóvil lo que está en juego.
13					se trata				Verb	se trata			Verb			process	Se trata de seguir manteniendo las bases de la productividad del país.

It

Basta ya de corrupción

## D. Appendix

CI #	Thematic Field		Text											
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION		Experiential Role	Text		
			Stance	Engagement	Stance	Engagement	Gral.	Detail	Stance	Engagement	Gral.	Detail		
1					attitude			Time			NP	Det Noun	Actor	DESDE hace ya demasiado tiempo, los ciudadanos se desayunan cada día con un nuevo caso de corrupción política en España que primero produce perplejidad y después indignación.
2		Cunde la sospecha que	Eng	shared knowledge							NP+App	NP+pp	Actor	Cunde la sospecha que ciertos comportamientos irregulares en el entorno de los partidos se han desarrollado en un clima de total impunidad, inadmisibles en democracia.
3	Y										NP	Pronoun	Actor	Y se extiende el desánimo general ante el exceso de desparramo y irresponsabilidad con que, pese a las graves sospechas que se ciernen sobre ellos, muchos de esos partidos hacen frente a la aparición de esos casos.
4											NP	Det Noun	Carrier	El resultado es que se generaliza la impresión que es todo el sistema político el que está afectado, que la corrupción forma parte de la cultura política del país y que, por tanto, la desafección ciudadana -la desconfianza hacia los poderes públicos- está más que justificada.
5											NP	Det Adj Noun	Actor	El malestar ciudadano se ha acrecentado ante los datos que incriminan al ex tesorero del PP, Luis Bárcenas, que se acogió a la amnistía fiscal decretada por el Gobierno para, presuntamente, blanquear 10 millones de euros a través de una empresa.
6											NP+App	NP+pp	Actor	Una cantidad ingresada en cuentas en Suiza cuyo montante total podría superar los 22 millones.
7	Y							Verb			Verb		process	Y no habían transcurrido veinticuatro horas desde que los responsables del Gobierno y del PP negasen ninguna relación con las actividades de Bárcenas cuando, ayer, hubieron de salir de nuevo a la palestra para volver a negar que los miembros de la dirección del partido hubieran percibido sobresueldos procedentes de constructoras y empresas de seguridad.
8								Accompaniment					process	Al caso Bárcenas se añade la situación comprometida del presidente de la Comunidad de Madrid, Ignacio González, por el asunto del ático de lujo en Marbella.
9	Y							Place			NP+App	NP+NP	Actor	Y, fuera del ámbito del PP, el llamado caso Campoon -sobre presuntas comisiones a políticos de Galicia a cambio de concesiones públicas-, que implica a políticos del PSOE y cuyas ramificaciones en Catalunya han alcanzado también a Convergencia, o la imputación del alcalde de Sabadell, del PSC, por malversación de fondos públicos.
10											NP+App	NP+clause	Actor	Estos y muchos otros casos abiertos que abrazan toda la geografía española han contribuido a encender todas las alarmas y a aumentar el descrédito de la democracia.



#### D. Appendix

[illegible]

[illegible]

## Egipto, de la paradoja a la gobernabilidad

Thematic Field			SYNTACTIC REALIZATION										Text
CI #	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		SYNTACTIC REALIZATION		Experiential Role	Text	
			Stance	Engagement	Stance	Engagement	Grat.	Detail	Stance	Engagement			Grat.
1					attitude		Contingency	Egipto		NP	Actor	I GUAL que otros países que saborearon las mieles de la primavera árabe, Egipto se ha visto abocado a una situación paradójica: las revoluciones sociales emprendidas para acabar con regímenes autoritarios han acabado aupando al poder a mandatarios islamistas de imprecisa convicción democrática.	
2	Dicho de otro modo,							el levantamiento popular		NP	Actor	Dicho de otro modo, el levantamiento popular acabó con dictaduras y sentó las bases de la democracia, pero a la postre las urnas se inclinaron por gobiernos de perfil autoritario, cuyo compromiso con la pluralidad puede llegar a ser insuficiente	
3								Egipto es el paradigma de este proceso.		NP	Taken	Egipto es el paradigma de este proceso.	
4								Las revueltas populares, impulsadas por sectores jóvenes y laicos de la población		NP+App	Actor	Las revueltas populares, impulsadas por sectores jóvenes y laicos de la población, acabaron con el régimen de Hosni Mubarak, que llevaba 30 años en el poder.	
5								La plaza Tahrir, escenario de aquella revuelta		NP+App	Carrier	La plaza Tahrir, escenario de aquella revuelta, se ha convertido en un símbolo reconocido en todo el mundo árabe.	
5							Time	se		NP	Actor	Una vez ganada esta batalla, una vez que se hubo plantado cara al poder militar, se convocaron elecciones.	
7	y							estos comicios		NP	Goal	Y estos comicios fueron ganados por los Hermanos Musulmanes, que designaron a Mohamed Morsi como figura más adecuada para asumir la presidencia egipcia.	
8								Fue gracias al ejercicio democrático que		Cause	Actor	Fue gracias al ejercicio democrático que Morsi llegó al poder.	
9	Pero						Time	el presidente egipcio		NP	Actor	Pero, al poco de hacerlo, el presidente egipcio se arrogó poderes excepcionales -que lo situaban por encima del poder judicial-, y a los que a la postre debió renunciar, forzado por la presión popular- y auspició el redactado de una Constitución de corte islamista.	
10								Esta Carta Magna, sometida a referéndum en dos vueltas		NP+App	Beneficiary	Esta Carta Magna, sometida a referéndum en dos vueltas, ha recibido la sanción positiva -un 63,8%- del electorado egipcio.	
11	Bien es cierto que	Eng						la tasa de participación en tal referéndum		NP+App	Carrier	Bien es cierto que la tasa de participación en tal referéndum fue baja, inferior al 33%.	

#### D. Appendix

Pero			aun así, fue suficiente	Manner	Manner-PP+Verb	fue suficiente		Verb		process	
12											Pero, aun así, fue suficiente para que el Gobierno que encabezó Morsi se viera legitimado, tanto en lo referente a los cinco meses de gestión que lleva al frente del país, como en lo relativo a la mencionada Constitución, que en adelante se convertirá en el documento rector de la vida egipcia.
13						Mohamed Morsi		NP	Proper Noun	Sayer	Mohamed Morsi pronunció ayer un discurso televisado en el que hizo un llamamiento a la estabilidad nacional, pidió unidad a las diversas fuerzas políticas y aseguró que todas sus decisiones no tienen otro objetivo más que el servicio al país.
14						La oposición		NP	Det Noun	Senser	La oposición ve las cosas de otro modo.
15		A su entender,		Evidence	Evidence-PP	la Constitución impulsada por Morsi, además de verse referendada mediante "irregularidades generalizadas"		NP+App	Proper Noun+Nonfinite	sayer	A su entender, la Constitución impulsada por Morsi, además de verse referendada mediante "irregularidades generalizadas", anuncia un inaceptable futuro de ribetes teocráticos.
16	Tanto...que		Tanto es así, que	Degree	Degree-Clause	la atomizada oposición egipcia	attitude	NP	Det Adj Noun Adj	Actor	Tanto es así, que la atomizada oposición egipcia ha conseguido unirse en un Frente Nacional de Salvación, que reúne a liberales, socialdemócratas o comunistas, y cuyo objetivo es defender la pluralidad y la división entre política y religión, haciendo frente a la pulsión teocrática.
17			Una vez sancionada la Constitución,	Time	Time-NonFinite	la próxima batalla política que va a librarse en Egipto		NP+App	NP+clause	Carrier	Una vez sancionada la Constitución, la próxima batalla política que va a librarse en Egipto son las elecciones al Parlamento, previstas para dentro de dos meses.
18			Una vez más,	Time	Time-pp	las perspectivas de los candidatos islamistas		NP+app	NP+pp	Carrier	Una vez más, las perspectivas de los candidatos islamistas son halagüeñas.
19						Rheme		Rheme		None	Más que las de la oposición, por cierto.
20						Dichos comicios	hedge	NP	Det Noun	Actor	Dichos comicios podrían, por tanto, salirse con una nueva victoria para las huestes de los Hermanos Musulmanes.
21		Aun así,		Contingency	Contingency-Adbpp	Mohamed Morsi	directive	NP	Proper Noun	Carrier	Aun así, Mohamed Morsi debe ser consciente de que gobernar un país significa atender a todos sus ciudadanos, respetando la pluralidad y las libertades.
22						Este	directive	NP	Pronoun	Carrier	Este es el principio que debe guiar su acción.
23	y					esto		NP	Pronoun	Carrier	Y esto es así tanto por respeto al sistema democrático que lo encontró como para garantizar la gobernabilidad de su país, que es clave en Oriente Medio.

La sanidad pública

CI #	Thematic Field										Text			
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC/ REALIZATION			HEAD		EXPERIENTIAL Role				
			Stance	Engagement	Stance	Engagement	Grdl.	Detail	Stance	Engagement	Grdl.	Detail		
1	SE la conoce como	Eng	Shared knowledge						ley de Acompañamiento para los Presupuestos del 2013 de la Comunidad de Madrid,		NP	Pronoun	Goal	SE la conoce como ley de Acompañamiento para los Presupuestos del 2013 de la Comunidad de Madrid, fue aprobada el jueves por el pleno de la Asamblea e incluye elementos dispares, como la externalización de servicios sanitarios, la modificación de la ley del Juego para adecuarla a las exigencias de Eurovegas o la implantación del euro por receta.
2	Pero								Esta ley ómnibus		NP	Det Noun Adj	Carrier	Esta ley ómnibus presenta diversos aspectos polémicos.
3									entre todos ellos		PP	PP+ProNoun+PrtoNoun	Actor	Pero, entre todos ellos, quizás se lleven la palma los relativos a lo que una parte importante del colectivo médico madrileño ha calificado como el mayor plan de privatización de la sanidad pública puesto en marcha en España y como el paso inicial para una ulterior política de privatizaciones.
4	En efecto,	Mod	Shared knowledge						dicha norma		NP	Det Noun	Sayer	En efecto, dicha norma contempla la cesión a operadores privados de la gestión de seis hospitales y 27 centros de salud, que atienden a un total de 1,5 millones de personas.
5									Los políticos populares que rigen la Comunidad de Madrid		NP+App	NP+Clause	sayer	Los políticos populares que rigen la Comunidad de Madrid, cifran en 200 millones de euros el ahorro que permitirá la privatización

[illegible]

## Las expropiaciones de Evo Morales

CI #	Thematic Field											Text				
	TEXTUAL	INTERP.	PREHEAD			SYNTACTIC REALIZATION			HEAD				SYNTACTIC REALIZATION			Experiential Role
			Stance	Engagement		Stance	Engagement	Grat.	Detail		Stance	Engagement	Grat.	Detail		
1										Iberdrola			NP	proper Noun	Goal	Iberdrola ha sido la última víctima de la fiebre expropiadora y nacionalizadora del presidente de Bolivia, Evo Morales, siguiendo el ejemplo de algunos países bolivarianos del continente, como Venezuela y Ecuador, además de Argentina.
2										El líder boliviano			NP	Det Noun Adj	sayer	El líder boliviano anunció ayer la expropiación de las acciones de Iberdrola en dos distribuidoras de electricidad en las regiones de La Paz y Oruro, de una empresa de servicios y de una gestora de inversiones.
3					Como motivo,			Manner	Manner-pp	Morales			NP	Proper Noun	sayer	Como motivo, Morales ha esgrimido la necesidad "de que las tarifas sean equitativas en los dos departamentos y que la calidad del servicio sea uniforme en el área rural y urbana".
										Morales			NP	proper Noun	Actor	Morales ya expropió en mayo las acciones de la filial de Red Eléctrica Española en Bolivia, sin que hasta el

#### D. Appendix

5							Además de las eléctricas,		Morales				NP	Proper Noun	Actor	Además de las eléctricas, Morales ha nacionalizado una quinena de empresas de hidrocarburos, cementos y minas, entre otras, desde que llegó al poder en el 2006.
6									El Gobierno boliviano				NP	Det Noun Adj	ca rier	El Gobierno boliviano tiene un grave problema de déficit de generación de energía visto el alto problema de apagones y cortes de suministro eléctrico que se producen en el país, y decisiones como esta son una vía de escape para ocultar problemas internos.
7									El Ejecutivo de Morales				NP+App	NP+PP	sayer	El Ejecutivo de Morales anunció ayer una "justa remuneración" a Iberdrola, pero al no tratarse de un contencioso entre naciones sino de decisiones de un Gobierno contra empresas extranjeras, legalmente es muy difícil llegar a una solución eficaz.
8									La economía y la credibilidad de España				NP+NP		Actor	La economía y la credibilidad de España no viven su mejor momento, pero no es motivo para que dirigentes latinoamericanos en arranques de populismo quiten a las empresas españolas de forma arbitraria el control de filiales compradas en su día y que, como la compañía de José Ignacio Sánchez Galán, están ejerciendo su actividad con todas las garantías.
9							Así, sólo		se				NP	Pronoun	Actor	Así, sólo se crea inseguridad jurídica entre las compañías extranjeras y se alejan futuras inversiones.

### El listón de los imputados

CI #	Thematic Field				HEAD				Text			
	TEXTUAL	INTERP.	STANCE	ENGAGEMENT	PREHEAD	STANCE	ENGAGEMENT	EXPERIENTIAL	ROLE			
								NP+App	NP+PP	Actor	LOS escándalos por corrupción dominan el panorama político y ocupan diariamente la narración periodística de nuestra vida pública.	
1		No es extraño que	Eng				attitude	NP+App	NP+Nonfinite	Carrier	No es extraño que la ciudadanía, abrumada por tantas noticias desoladoras, esté perdiendo la confianza en el sistema político.	
2								NP+app	NP+PP	Actor	Los comentarios de ascensor han dejado de referirse al tiempo o a la crisis; con sorna o con indignación, con desoladora acidez o con irrepresible indignación, nuestro sistema político está siendo percibido por los ciudadanos de a pie, especialmente por los jóvenes, los más castigados por la crisis, como la reedición ibérica de la cueva de Ali Babá.	
3		Cuando la sospecha de que	Eng	Shared knowledge			los comportamientos irregulares en el entorno de los partidos	NP+App	NP+PP	Actor	Cuando la sospecha de que los comportamientos irregulares en el entorno de los partidos se han desarrollado en un clima de total impunidad.	
4							Se	NP	Pronoun	Actor	Se extiende entre los ciudadanos el desánimo general ante el exceso de desparramo e irresponsabilidad con que, pese a las graves sospechas sobre ellos, los partidos hacen frente a la aparición de casos de corrupción.	
5							Se	NP	Pronoun	Carrier	Se está generalizando la impresión de que es todo el sistema político el que está infectado; que la corrupción forma parte de la cultura política del país y que, por lo tanto, la desafección ciudadana -la desconfianza hacia los poderes públicos- está más que justificada.	
6							La primera víctima de la corrupción	NP+App	NP+PP	Carrier	La primera víctima de la corrupción es, por lo tanto, la democracia.	
7	Y	es que	Eng	Shared knowledge			los casos de corrupción	NP+App	NP+PP	Actor	Y es que los casos de corrupción no afectan sólo a los partidos políticos.	
8					En mayor o menor grado		alcanzan	Verb		process	En mayor o menor grado alcanzan a todas las instituciones,	
9							La presidencia del poder judicial	NP+App	NP+PP	Actor	La presidencia del poder judicial quedó meses atrás mancada por un lamentable caso de aprovechamiento privado del dinero público.	
10												



[illegible]

### D. Appendix

25	En este sentido, es preciso tomar medidas claras y muy contundentes para evitar que el ruido, los rumores y las falsas acusaciones se fundan con los casos reales y se mezclen con las investigaciones en curso afectando un clima irrespirable que podría colapsar nuestra vida democrática.	Eng	directive						tomar medidas claras y muy contundentes para evitar que el ruido, los rumores y las falsas acusaciones					NonFinite		Det Noun	Verbage	Actor	En este sentido, es preciso tomar medidas claras y muy contundentes para evitar que el ruido, los rumores y las falsas acusaciones se fundan con los casos reales y se mezclen con las investigaciones en curso afectando un clima irrespirable que podría colapsar nuestra vida democrática.
26	SON bienvenidas, por supuesto, las medidas que sugería esta semana la mesa institucional que, presidida por el presidente de la Generalitat en compañía de la máxima representación de la judicatura y de directores de oficinas de control, debatió diversos proyectos legislativos y judiciales que faciliten la limpieza del sistema.	Eng	shared knowledge						las medidas					NP				Verbage	SON bienvenidas, por supuesto, las medidas que sugería esta semana la mesa institucional que, presidida por el presidente de la Generalitat en compañía de la máxima representación de la judicatura y de directores de oficinas de control, debatió diversos proyectos legislativos y judiciales que faciliten la limpieza del sistema.
27									Apostar por la transparencia e introducir cambios en la legislación electoral servirá sin duda en el próximo futuro para renovar nuestros viecidos mecanismos de representación y control.					NonFinite			Carrier	Carrier	Apostar por la transparencia e introducir cambios en la legislación electoral servirá sin duda en el próximo futuro para renovar nuestros viecidos mecanismos de representación y control.
28	Pero								la urgencia y la gravedad del momento					NP+NP				Actor	Pero la urgencia y la gravedad del momento exigen medidas excepcionales y rápidas.
29									se trata					Verb				process	Se trata de preservar el prestigio de nuestro sistema democrático.
30									se trata					Verb				process	Se trata de afrontar con responsabilidad y realismo el escepticismo, la desconfianza y el resqueño con que la opinión ciudadana asiste al tremendo espectáculo de la corrupción en todas las instituciones.
31									El nudo de la cuestión					NP+App		NP+pp	Carrier	Carrier	El nudo de la cuestión, en estos momentos, es qué tipo de filo cortante debe exigirse a las instituciones y a los partidos.
32									la condena judicial					NP		Det Noun Adj	Token	Token	Hace unos años, la condena judicial significaba dimisión.
33									Rheme					Rheme				none	Cuando no inhabilitación.
34									Verb					Verb				Actor	Después subió el listón y el corte se afinó: la apertura de un proceso penal, esto es, el banquillo de los acusados, condujo a la dimisión.
35									El grave malestar social que hemos descrito anteriormente unido al encanamiento de la competición política	attitude				NP+App		NP+clause	Actor	Actor	El grave malestar social que hemos descrito anteriormente, unido al encanamiento de la competición política, parece desplazar ahora el listón más arriba: corte en el momento de la imputación, esto es, en la primera presunción de delito por parte del juez instructor.
36									Imputación					NP		Noun	Token	Token	Imputación: figura garantista del procedimiento penal que concede al afectado el derecho a declarar con la asistencia de un abogado.
37									esa					NP		Pronoun	Carrier	Carrier	¿Debe ser esa la hoja cortante?
38									lo fácil					NP		Pronoun Adj	Carrier	Carrier	Hoy, lo fácil sería decir que sí.
39									La plaza pública					NP		Det Noun Adj	Sensor	Sensor	La plaza pública quiere escarmientos.
40									La sociedad					NP		Det Noun	Carrier	Carrier	La sociedad esta necesitada de gestos de austeris.
41									rheme					Rheme			None	None	Imputación-dimisión, sí, en los casos de mayor relevancia.
42									El paradigma					NP			Carrier	Carrier	El paradigma no debiera ser inquisitorial.
43									Hay					Verb		Det Noun	Existent	Existent	Hay que preservar la libertad de los partidos.
									está					Verb			Carrier	Carrier	Cada palo deberá aguntar su vela, porque, en última instancia, está la sanción de las urnas.

España gana con el presupuesto de la UE

[illegible]

[illegible]

## D. Appendix

CI #	Thematic Field	PREHEAD	INTERP.	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Experiential Role	Text
	TEXTUAL	Stance	Engagement	Time	Time-PP				
1		DOS años después de la caída de la dictadura de Ben Ali.				Túnez		Actor	DOS años después de la caída de la dictadura de Ben Ali, Túnez vive horas de graves enfrentamientos e incertidumbre.
2						La huelga general de ayer -convocada por el principal sindicato y los partidos de la oposición- fue ampliamente seguida por la población.		Carrier	La huelga general de ayer -convocada por el principal sindicato y los partidos de la oposición- fue ampliamente seguida por la población.
3						El asesinato del abogado y líder de un partido de izquierda integrado en el Frente Popular, Chokri Belaid, cometido el miércoles, ha arivado las protestas de los sectores descontentos con el Gobierno, formado por los islamistas moderados del partido Enahda y dos formaciones laicas y progresistas.		Actor	El asesinato del abogado y líder de un partido de izquierda integrado en el Frente Popular, Chokri Belaid, cometido el miércoles, ha arivado las protestas de los sectores descontentos con el Gobierno, formado por los islamistas moderados del partido Enahda y dos formaciones laicas y progresistas.
4						El funeral de este dirigente congregó ayer a miles de personas en la capital tunecina.		Actor	El funeral de este dirigente congregó ayer a miles de personas en la capital tunecina.
5		Para hacer frente a la situación		Contingency	Contingency NonFinite	El primer ministro, Hamadi Yebeli		Sayer	Para hacer frente a la situación, el primer ministro, Hamadi Yebeli, anunció un profundo cambio de Gobierno que supondría relevar a los actuales ministros por tecnocratas independientes y, dentro de un tiempo, convocar elecciones.
6						Esta salida		Goal	Esta salida fue rápidamente rechazada por la dirección del propio partido de Yebeli, extremo que pone de manifiesto la creciente fractura interna del islamismo político, que en los comicios de octubre del 2011 recibió un gran apoyo popular para dirigir la transición a la democracia.
7						La Asamblea Constituyente que salió de las primeras elecciones tras la revolución del jazmin no ha logrado los acuerdos mínimos para pactar la nueva Constitución y se ha convertido en una institución inoperante, atenazada por las divisiones.		Actor	La Asamblea Constituyente que salió de las primeras elecciones tras la revolución del jazmin no ha logrado los acuerdos mínimos para pactar la nueva Constitución y se ha convertido en una institución inoperante, atenazada por las divisiones.
8						La pobreza y la corrupción castigan amplias zonas del país, circunstancia que explota el islamismo radical mientras los sectores urbanos y laicos, los jóvenes y las mujeres denuncian al Gobierno por no frenar la violencia islamista y por ceder a las demandas de las fracciones más reaccionarias de Enahda.		Actor	La pobreza y la corrupción castigan amplias zonas del país, circunstancia que explota el islamismo radical mientras los sectores urbanos y laicos, los jóvenes y las mujeres denuncian al Gobierno por no frenar la violencia islamista y por ceder a las demandas de las fracciones más reaccionarias de Enahda.
9						La pugna en el seno del islamismo moderado y el choque entre este y la mayoría de las formaciones laicas, así como la amenaza del extremismo salafista, conforman un cuadro de alto riesgo en la nación que fue punto de arranque de las primaveras árabes y que parecía estar más preparada para culminar con éxito el camino hacia un sistema de libertades.		Carrier	La pugna en el seno del islamismo moderado y el choque entre este y la mayoría de las formaciones laicas, así como la amenaza del extremismo salafista, conforman un cuadro de alto riesgo en la nación que fue punto de arranque de las primaveras árabes y que parecía estar más preparada para culminar con éxito el camino hacia un sistema de libertades.
10						Las clases medias que articulan los sectores más dinámicos de Túnez se debaten entre el malestar, la desilusión y la necesidad de frenar un islamismo violento que choca con una de las sociedades más abiertas y occidentalizadas de la región.		Carrier	Las clases medias que articulan los sectores más dinámicos de Túnez se debaten entre el malestar, la desilusión y la necesidad de frenar un islamismo violento que choca con una de las sociedades más abiertas y occidentalizadas de la región.
11						Lo que ocurre en este país norteafricano está lejos de la dinámica que se vive en Egipto, pero pone en evidencia que la transición a la democracia en la zona requerirá más tiempo y más esfuerzos de lo que muchos imaginaron cuando la gente perdió el miedo y ocupó las plazas.		Carrier	Lo que ocurre en este país norteafricano está lejos de la dinámica que se vive en Egipto, pero pone en evidencia que la transición a la democracia en la zona requerirá más tiempo y más esfuerzos de lo que muchos imaginaron cuando la gente perdió el miedo y ocupó las plazas.

El papado de Benedicto XVI

CI #	Thematic Field	PREHEAD	INTERP.	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Experiential Role	Text
	TEXTUAL	Stance	Engagement	Time	Time-PP				

2									El Papa			NP	Proper Noun	Sayer	El Papa anunció su dimisión ayer por la mañana, lo que causó enorme sorpresa en el orbe católico.
									Evidence-PP	Evidence		NP	Det Noun	Phenomenon	Según Federico Lombardi, portavoz del Vaticano, la noticia surgió de improviso incluso a los más allegados al Pontífice.
3												NP+App	NP+clause	Phenomenon	El hecho de que la anterior renuncia papal ocurriera hace nada menos que seis siglos -Gregorio XII, en 1415- no hizo sino acentuar la impresión en el seno de la Iglesia.
4												NP	Proper Noun	Carrier	En total, Benedicto XVI habrá permanecido en su puesto poco menos de ocho años.
5									Degree-PP	Degree		NP+App	Proper Noun+Clause	Sayer	Su Santidad, que cuenta 85 años, ha aducido motivos de salud para justificar la renuncia.
6									Place-PP	Place		NP	Pronoun	Sayer	En su comunicado se refirió al vigor físico y espiritual necesario para desempeñar el cargo, y precisó "que en los últimos meses ha disminuido en mí de tal forma que he de reconocer mi incapacidad para ejercer bien el ministerio que me fue encomendado".
7												NP	Det Noun	Sayer	A algunos analistas apuntan que este declive físico se ha podido ver agravado por las intrigas sucesorias vaticanas, que afloraron con el denominado caso Vatileaks, que resultan habituales cuando el Pontífice alcanza una edad avanzada, y son de difícil gestión, incluso en plenas facultades.
8												NP	Adj Noun	Carrier	Ningún papado es sencillo, y el de Benedicto XVI no ha sido excepción.
9												NP	Proper noun	Actor	Cuando fue elegido sucesor de Juan Pablo II y 265.º pontífice católico, el 19 de abril del 2005, la Iglesia capeaba una tremenda tormenta mediática.
10															
11												NP	Pronoun	Beneficiary	La alimentaban la difusión de numerosos casos de abusos sexuales cometidos por eclesiásticos durante decenios, y un creciente clamor popular que exigía el fin de estas conductas y de su encubrimiento.
												NP	Det adj ProperNoun	Actor	El nuevo Papa mostró entonces gran determinación y desbarbó a quienes acusaban a numerosas diócesis e incluso al Vaticano de proteger a los pederastas, al admitir que había "pecado y suciedad" en la Iglesia, al pedir perdón a las víctimas de los abusos y al introducir mecanismos de control y denuncia.
12												NP+App	NP+NonFinite	Actor	El revuelo causado por esta cuestión se prolongó hasta el inicio del presente decenio.
13												NP	Proper Noun	Actor	Pero Benedicto XVI mantuvo su firmeza y demostró, en este capítulo, una muy plausible voluntad regeneradora.
14												NP+App	NP+Proper Noun	Carrier	Cualquier balance del papado de Benedicto XVI debe tener muy en cuenta su formación, su perfil y, también, su pertenencia al ámbito de la cultura germánica.
15												NP+App	NP+Proper Noun	Carrier	Cualquier balance del papado de Benedicto XVI debe tener muy en cuenta su formación, su perfil y, también, su pertenencia al ámbito de la cultura germánica.

#### D. Appendix

					A diferencia de su antecesor, Juan Pablo I, que fue un hombre de acción, un viajero incansable y un gran comunicador, y es recordado, entre otros motivos, por su contribución a la caída de los regímenes comunistas. Benedicto XVI es, ante todo, un intelectual, un autor, un hombre de universidad y estudio, un académico riguroso y de acusado conservadurismo.					Benedicto XVI						NP	Proper Noun	Carrier	A diferencia de su antecesor, Juan Pablo I, que fue un hombre de acción, un viajero incansable y un gran comunicador, y es recordado, entre otros motivos, por su contribución a la caída de los regímenes comunistas. Benedicto XVI es, ante todo, un intelectual, un autor, un hombre de universidad y estudio, un académico riguroso y de acusado conservadurismo.	
16					No siempre	time				fue							Verb		process	No siempre fue así.
17					En tiempos del concilio Vaticano II, que abrió la Iglesia a la modernidad.	Time				Joseph Ratzinger							NP	Proper Noun	Carrier	En tiempos del concilio Vaticano II, que abrió la Iglesia a la modernidad. Joseph Ratzinger era un influyente teólogo progresista.
18	Pero				posteriormente .siendo catedrático de Dogmática en la Universidad de Tubinga	Time				vivió							Verb		process	Pero, posteriormente, siendo catedrático de Dogmática en la Universidad de Tubinga, vivió muy de cerca las revueltas estudiantiles, de componente marxista, e inició una progresiva evolución hacia posiciones más contenidas.
19					Tras afincarse en Roma en 1981, y ejercer veinte años como director de la Congregación para la Doctrina de la Fe, ese giro dio sus frutos	Time				Ratzinger							NP	Proper Noun	Actor	Tras afincarse en Roma en 1981, y ejercer veinte años como director de la Congregación para la Doctrina de la Fe, ese giro dio sus frutos: Ratzinger organizó y pilotó la atenuación de algunos de los ecos del Vaticano II.
20																				
21										Rheme							Rheme		None	De ahí su fama de conservador.
22		Es preciso recordar	Eng	directive						esta trayectoria							NP	Det Noun	Carrier	Es preciso recordar esta trayectoria para entender las coordenadas que definen el papado de Benedicto XVI.
23					En este sentido	Manner				Hay							Verb		Existente	En este sentido, hay que subrayar que su argumentario ideológico fue decisivo en la condena de la Teología de la Liberación, corriente cristiana izquierdista, no exenta de notas marxistas, que definía destacados movimientos sociales y eclesiales en América Latina.
24					Pasando de lo político a lo social	Manner				añadiremos							Verb		process	Pasando de lo político a lo social, añadiremos que Benedicto XVI ha preferido siempre la tradición.
25					ha rechazado	Verb				ha rechazado							Verb		process	Ha rechazado los derechos de los homosexuales .les prohibió explícitamente el ingreso en los seminarios-, el uso del preservativo y las demandas de igualdad de las mujeres para poder ordenarse como sacerdotes de la Iglesia; incluso amenazó de excomunión a los políticos defensores del derecho al aborto.
26		Cierto es que	Eng	shared knowledge						en su actualización de los pecados capitales, emitida en el 2008							NP+App	NP+Nonfinite	Actor	Cierto es que en su actualización de los pecados capitales, emitida en el 2008, sintonizaba con causas de significación progresista, e incluía entre ellos los datos al medio ambiente o las decisiones que ocasionan pobreza, injusticia o dificultad, en línea con la doctrina social de la Iglesia.
27	Pero									en ese mismo texto									Verbage	Pero en ese mismo texto criticaba experimentos científicos como la manipulación genética.

CI #	Thematic Field	TEXTUAL	INTERP.	PREHEAD	SYNTACTIC REALIZATION	HEAD	Engagement	Stance	Engagement	Stance	Engagement	Text
28	Aunque					Contingency						Aunque no ha superado el kilometraje recorrido por su antecesor, Benedicto XVI ha llevado a cabo también una considerable labor pastoral.
29						Contingency						Mención especial merecen sus contactos con la juventud, que han dado pie a reuniones multitudinarias, como la de Colombia en el 2005, con un millón y medio de participantes.
30						Contingency						O la de Madrid en el 2011, con similar afluencia.
31						Contingency						Esta pulsión viajera le llevó a Barcelona para la dedicación del gaudiniano templo de la Sagrada Familia: un acto relevante en el que el Papa de Roma entró en el templo de Gaudí rezando en catalán.
32						Contingency						Barcelona lo recordará siempre.
33						Contingency						Viajó también a Estambul y a Jerusalén.
34	Si					Contingency						Si el pontificado de Benedicto XVI empezó con los problemas ya apuntados, su tramo final se ha visto acompañado por el llamado caso Vatileaks.
35	Es decir,					Contingency						Es decir, por la filtración de documentos privados vaticanos o papales, referidos algunos de ellos a episodios de dudosa gestión en la Santa Sede.
36						Contingency						El mayor domo del Papa fue juzgado y encarcelado por ello.
37	Pero					Contingency						Pero prevaleció la sensación de que el caso no terminaba ahí, de que era un reflejo de las luchas de poder ante la sucesión papal, que ahora, tras la renuncia anunciada ayer, se ha precipitado y debe resolverse en meses y medio.
38						Contingency						Benedicto XVI ha dicho que su futuro está en un convenio de clausura.
39						Contingency						Quirá un apartamento en la propia Ciudad del Vaticano.
40						Contingency						Muy pronto desaparecerá, pues, de la escena terrenal, para entrar en la historia.
41	Pero					Contingency						Pero el Vaticano deberá seguir enfrentándose a los problemas de un mundo en transformación, amenazado por los fundamentalismos, en el que su futuro, y el de la Iglesia católica, pasa por elevar el número de vocaciones, pese a la secularización de la sociedad occidental.
42						Contingency						El Papa de la tradición y de la ortodoxia se va con un gesto valiente y moderno que conecta culturalmente con pulsiones que hoy están muy presentes en la sociedad occidental: saber marchar a tiempo, modular los poderes vitálicos.
43						Contingency						El gesto de Benedicto XVI es un gesto alocinado cuya trascendencia va más allá de los vastos confines materiales y espirituales de la Iglesia católica.
44						Contingency						Con su decisión, Joseph Ratzinger puede que abra paso a un nuevo agnoscimiento católico.

La cultura, apuesta estratégica

CI #	Thematic Field	TEXTUAL	INTERP.	PREHEAD	SYNTACTIC REALIZATION	HEAD	Engagement	Stance	Engagement	Stance	Engagement	Text
1						Contingency						DESTACADAS figuras de la escena cultural catalana se referían ayer en este diario a la reacción de este sector ante los recortes y las penalizaciones que está sufriendo.
2						Contingency						Una reacción que cabría considerar más moderada que la de los profesionales de la educación y la sanidad públicas, sectores que, junto al de la cultura, integran el trípede que sustenta el Estado de bienestar.
						Contingency						El tono queda de dicha reacción no debería, en ningún

#### D. Appendix

4																		Sus frutos son muy apreciados, tanto los intangibles como los tangibles.	Carrier		NP													Entre los primeros, su condición de fundamento social, de escuela de ciudadanía y de herramienta para la emancipación y la libertad personales.	
5																		Rheme			rheme												Entre los segundos, una serie de datos ya bien conocidos, pero cuyo recuerdo no es sobrero: el sector cultural español emplea a medio millón de personas (alrededor del 3% de la población activa) y motiva la actividad de más de 100.000 empresas.		
6																		el sector cultural	Det Noun Adj		NP												O, por decirlo en términos comparativos, el sector cultural está próximo, por su volumen, al energético o al agrícola.		
7																		otro/s			Verb												Otros a menudo que el futuro de nuestra sociedad depende del desarrollo de las industrias del conocimiento, y que la cultura es un elemento clave, tanto para el progreso cívico como por su capacidad para crear riqueza.		
8																		ningún partido	Adj Noun		NP												En épocas de bonanza, ningún partido discute estos principios, aunque las inversiones públicas culturales queden en España por detrás de las de otros países occidentales.		
9																		el sector cultural	Det Noun adj		NP												Pero, en lo que llevamos de crisis, además de padecer una sucesión de recortes en los presupuestos generales y un descenso del consumo cultural motivado por el empobrecimiento colectivo, el sector cultural ha sufrido iniciativas legales lesivas.		
10																																			
11																		Rheme			Rheme												Unas, por activa, como el aumento del IVA cultural en 13 puntos, situándolo en el 21%.		
12																		que	Pronoun		NP												Y otras, por pasiva, como la siempre demorada reforma de la ley de Mecenazgo que debería propiciar, de modo efectivo, la afluencia de recursos privados a la actividad cultural de consumo colectivo.		
13																		su contribución al PIB	NP+App		NP+App												Todo ello, a pesar de que su contribución al PIB sigue siendo superior a las subvenciones que recibe, incluso ahora, cuando el consumo cultural cae, arrastrado por la crisis.		
14																		Esta política restrictiva	Det Noun Adj		NP												Esta política restrictiva ha llevado el sector, al decir de los profesionales consultados por La Vanguardia, a una situación límite.		
15																		El margen para los recortes	NP+PP		NP+App												El margen para los recortes se ha agotado.		
16																		es			Verb												No es posible hacer más sin desmantelar el sistema cultural, desbichar el tejido creativo y, en suma, descuidar el humus cultural hasta producir en él un daño de difícil recuperación.		
17																		Algunas instituciones, como la Consejería de Cultura de la Generalitat	NP+PP		NP+App												Algunas instituciones, como la Consejería de Cultura de la Generalitat, han anunciado un plan de choque inmediato para levantar el sector cultural de su postración.		
18																		dicho plan	Det Noun		NP												Aunque difícilmente trazará aparejados mayores recursos, dicho plan merece la bienvenida.		
19																		Rheme			Rheme												Al menos, por lo que tiene de reconocimiento oficial de una situación crítica.		
20																		lo que ocurre en otras instancias políticas, donde la cultura no se conceptúa todavía como un bien común y deseable y sigue	Goal		Clause												Porque lo que no es de recibo es lo que ocurre en otras instancias políticas, donde la cultura no se conceptúa todavía como un bien común y deseable y sigue		



22									La cultura, como la ciencia, es una apuesta estratégica inenunciable para España.	Carrier	NP+PP	NP+App	
23								es	No es ni siquiera una opción más.	process		Verb	
24								se trata	Se trata, entrando el siglo XXI y asentada la sociedad postindustrial, de la opción.	process		Verb	
25									No darse cuenta de ello y seguir penalizándola es uno de los peores errores que pueden cometer los gobernantes.	Carrier		NonFinite	

## Juicio militar en Rabat

CI #	Thematic Field		HEAD										Text			
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	Stance	Engagement	Stance	Engagement		Experiential Role		
1												NP+App	NP+PP	Carrier	El campamento de protesta de Gdim Izik, en las alheras de El Aaiún, fue la mayor demostración saharaui desde que España entregó a Marruecos su última colonia en 1975.	
2					A partir de mediados de octubre de 2010.			Time					NP	Det Noun Noun	Actor	A partir de mediados de octubre de 2010, unos 20.000 saharauis se concentraron allí pacíficamente para plantear reivindicaciones sociales.
3					En la madrugada del 8 de noviembre,			Time-PP					NP+App	NP+PP	Actor	En la madrugada del 8 de noviembre, las fuerzas de seguridad marroquíes asaltaron el campamento.
4													PP+PP		Scope	En la operación y en los posteriores enfrentamientos en El Aaiún murieron 11 antidisturbios y dos civiles saharauis.
5					Previamente,			Time-advb					NP	Det Noun	Actor	Previamente, un joven murió en un control por el disparo de un gendarme.
6					Ventisiete meses después de aquellos hechos,			Time-NP				NP+App	NP+PP	Goal	Ventisiete meses después de aquellos hechos, 24 activistas saharauis, entre ellos un intelectual como Naama Asfari, han sido juzgados por un tribunal militar en Rabat.	
7					Excepto en dos casos	Manner		Manner-AdjP				NP+App	NP+PP	Carrier	Excepto en dos casos, las condenas por “violencia sobre agentes del orden con resultado de muerte” han resultado severas: 8 a cadena perpetua, 4 a 30 años y 10 a penas de entre 20 y 25 años.	
8											hedge	NP		Carrier	Las apurencias pueden ser engañosas.	
9												NP+NP		Sayer	Los abogados de la defensa y sus clientes pudieron explayarse ante el tribunal.	
10												NP	Det Noun	Sayer	Los activistas defendieron incluso con argumentos, y también a gritos, el derecho a la autodeterminación.	
11					había	shared knowledge		Verb				Verb		Existent	Había observadores internacionales presentes en la sala.	
12	Pero				más allá de las formas			Contingency AdvP				NP	Det Noun	Goal	Pero, más allá de las formas, los activistas fueron juzgados por una jurisdicción de excepción respecto a la que, en la práctica, no cabe recurso.	
13												NP	Det Noun	Actor	El tribunal se negó a investigar las denuncias de tortura que formularon los acusados.	
14												NP	Det Noun	Actor	Las condenas se fundamentan en las confesiones ante la policía judicial, que los activistas dicen inventadas o arrojadas bajo tortura.	
15					En las armas con las que fueron abatidos los antidisturbios	Place						Verb		process	En las armas con las que fueron abatidos los antidisturbios no aparecen las huellas dactilares de los acusados, ni tampoco se les reconoce en un vídeo del asalto proyectado en la sala de audiencia a instancia de la fiscalía.	
16												NP+App	NP+Clause	Actor	El desarrollo del mayor juicio de saharauis desde que Rabat administra el territorio muestra que si en Marruecos los derechos humanos siguen siendo una asignatura pendiente, en el Sáhara lo son mucho más.	
17												NP+App	NP+PP	Actor	El conflicto del Sáhara volverá en abrir a la mesa del Consejo de Seguridad de la ONU.	

## D. Appendix

	Es hora de que,	Eng	directive	para lograr un mayor respeto de los derechos humanos en ese territorio		Contingency	Contingency Non Finite	el contingente de la ONU (Minurso) allí desplegado		directive	NP+App	NP+PP	Actor	Es hora de que, para lograr un mayor respeto de los derechos humanos en ese territorio, el contingente de la ONU (Minurso) allí desplegado se encargue de supervisar su aplicación, como lo hacen otras muchas fuerzas de paz.
18														

Exportamos bien

Thematic Field												Text			
CI #	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD		Stance		Engagement	Exponential Role	Text		
			Stance	Engagement	Grat.										
Si					Contingency	Contingency	Contingency	los datos de 2012			NP+App	NP+PP	Sayer	Si España era el problema y Europa, la solución, los datos de 2012 confirman que también en términos económicos, la solución está en Europa.	
1								Rheme							Y en otros mercados.
2					Time		Time-PP	el sector exterior				Det Noun Adj	Token		En los últimos ejercicios el sector exterior es el único que regala alegrías a la economía.
3					Contingency	Contingency	Contingency	la recesión			NP	Det Noun	Carrier		Si no fuera por el dinamismo de sus exportaciones, la recesión sería mucho más profunda.
4								La balanza por cuenta corriente —esto es, el saldo de las operaciones de bienes, servicios y transferencias entre España y el resto del mundo—				NP+App		Actor	La balanza por cuenta corriente —esto es, el saldo de las operaciones de bienes, servicios y transferencias entre España y el resto del mundo— viene arrojando un signo positivo que convierte al sector exterior en el principal activo de la economía.
5					verb	verb		es			verb			Process	Es algo sustantivo, puesto que una de sus principales debilidades en la época de la burbuja estribaba en el extraordinario déficit del 10% del PIB en la balanza de pagos.
6								La balanza comercial			NP	Det Noun Adj		Actor	La balanza comercial sigue mejorando, pese a que la recesiva coyuntura de la eurozona aminore su capacidad de absorción de las exportaciones.
7					Time	Time	Time-PP	las importaciones			NP	Det Noun	Actor		Al tiempo, las importaciones siguen con fuerza a la baja —y ese es uno de los secretos del saldo— por el escaso pulso de la demanda.
8								La balanza de servicios			NP+App	NP+PP	Carrier		La balanza de servicios también resulta positiva, gracias a la industria turística.
9								la cuenta financiera			NP	Det Noun Adj	Actor		Y la cuenta financiera ha compensado en los últimos meses buena parte de la desastrosa fuga de capitales de la primera mitad del año pasado.
10					Contingency	Contingency		modular la política económica en función de los parámetros			Non Finite		Actor		Si el sector exterior es el único que tira de la economía, lo lógico sería modular la política económica en función de los parámetros que facilitan su buena conducta, a favor de las apuestas implícitas sectoriales que están en la base de los resultados comerciales: el sector agroalimentario, el del automóvil.
11	lo lógico sería	Eng			Contingency	Contingency		el presidente del Gobierno			NP+App	NP+Proper Noun	Sayer		Esperemos que el presidente del Gobierno ofrezca hoy alguna pista en este sentido durante el debate sobre el estado de la nación.
12					Time		Time-PP	los últimos meses			NP	Det Adj noun	Sayer		Por vez primera en la historia, los últimos meses arrojan saldos comerciales positivos respecto a los socios europeos.
13								La tasa de cobertura en el comercio de mercancías (valor de las exportaciones menos valor de las importaciones)			NP+App	NP+PP	Sayer		La tasa de cobertura en el comercio de mercancías (valor de las exportaciones menos valor de las importaciones) arroja índices históricos.
14					Contingency	Contingency		eso			NP	Pronoun	Carrier		Aunque en algunos meses eso se deba más a la debilidad de las importaciones que a la fortaleza de las exportaciones, y aunque la acelerada caída de la actividad europea en los últimos meses de 2012 haya acabado lastrando la potencialidad exportadora: un indicio más para cuestionar el

## Escuchas y fallos

535

### D. Appendix

[illegible]

## Berlusconi todavía

Cl #	Thematic Field										Text			
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION	HEAD	Stance	Engagement	Carrier	Experiential Role	
1					Hasta hace pocas semanas,			Time	Time-PP			NP+App	Carrier	Hasta hace pocas semanas, la posibilidad de que Silvio Berlusconi volviera a decidir la política italiana parecía negligible.
2												NP+App	Carrier	Su distancia con la alianza centroizquierdista de Pier Luigi Bersani, primera en intención de voto, era en enero de 11 puntos.
3												NP	Actor	Esa diferencia se ha reducido a cinco en los últimos sondeos publicados antes de las cruciales elecciones del 24 y 25 de febrero, gracias, entre otros factores, a las imposibles promesas fiscales y laborales de Berlusconi y a la repercusión del escándalo del Monte dei Paschi, el banco de Siena vinculado históricamente al centro-izquierda.
4					Según estas encuestas			Evidence	Evidence-PP			NP+App	Carrier	Según estas encuestas, los centristas agrupados en torno a Mario Monti figuran relegados a un cuarto puesto, tras el cómico antistema Berpe Grillo.
					Quince meses después de verse obligado a dimitir en medio de un total descrédito y la peor crisis financiera italiana contemporánea			Time	Time-clause	attitude		NP+App	Actor	Quince meses después de verse obligado a dimitir en medio de un total descrédito y la peor crisis financiera italiana contemporánea, la sombra del magnate calabga de nuevo.

		Es cierto que	Eng							Berlusconi		NP	Proper noun	Senser	Es cierto que Berlusconi ya no aspira a dirigir el Gobierno —condición impuesta por su aliada Liga Norte para evitar el desplome de su voto—, pero resulta obvio que una eventual victoria de la coalición derechista haría a Il Cavaliere su figura dominante.
6										Rheme		Rheme		None	Una perspectiva dramática para Italia y la Unión Europea.
7						De cumplirse los pronósticos		Contingency	Contingency-pp	el centro-izquierda de Bersani		NP+App	NP+Proper Noun	Actor	De cumplirse los pronósticos, el centro-izquierda de Bersani controlaría la Cámara de Diputados, gracias al premio en escabos que la embrollada ley electoral italiana otorga al partido más votado.
8	Pero									la mayoría del Senado, que como Cámara territorial exige la victoria en las regiones más importantes del país,		NP+App	NP+clause	Carrier	Pero la mayoría del Senado, que como Cámara territorial exige la victoria en las regiones más importantes del país, está presumiblemente fuera de su alcance.
9						Para manejar el Senado (el Gobierno necesita el voto de confianza de ambas Cámaras)		Contingency	Contingency Non finite	Bersani		NP	Proper Noun	Carrier	Para manejar el Senado (el Gobierno necesita el voto de confianza de ambas Cámaras), Bersani necesitaría el apoyo de Monti, pese a las reticencias que su talante reformista y su condición tecnocrática suscita entre los sindicatos y los aliados izquierdistas del jefe socialdemócrata.
10										Eso		NP	Pronoun	Actor	Eso permitiría a Monti exigir un papel dominante en la economía transalpina, a la que en buena medida ha devuelto la credibilidad.
11										Ningún desenlace electoral		NP	Adj Noun Adj	Carrier	Ningún desenlace electoral es peor para Italia y Europa que la posibilidad de que vuelva a las andadas Silvio Berlusconi, que encamagado en sus propios y turbios asuntos ha sido incapaz en ocho años de emprender una sola de las cruciales reformas que su país precisa.
12		En el mejor de los casos,	Eng							semejante escenario		NP	Adj Noun	Token	En el mejor de los casos, semejante escenario significaría el regreso de Italia a la inestabilidad política crónica y de nuevo la zozobra en los mercados y el euro.

El caso 'Helin'

CI #	Thematic Field		HEAD										Text
	TEXTUAL	INTERP.	PREHEAD		SYNTACTIC REALIZATION		HEAD				Experiential Role		
			Stance	Engagement	Stance	Engagement	Stance	Engagement	NP+App	NP+clause	Carrier		
1												El descubrimiento periodístico de que un asesino y exconvicto trabaja para la policía y la Guardia Civil es un hecho que produce un profundo desasosiego.	
2									NP+App	Proper Noun+clause	Goal	Emilio Hellin Moro, militante ultraderechista de fuerza Nueva que secuestró y mató en 1980 a Yolanda González por su supuesta —y finalmente falsa— vinculación con ETA, fue condenado a 43 años de cárcel por ese asesinato, que produjo una enorme alarma social en aquellos años de transición política, con terrorismo clara y grupos parapoliciales infiltrados en las fuerzas de seguridad del Estado.	
3						Time	Time+pp		NP	Proper Noun	Actor	Tras 14 años de cárcel, Hellin se cambió de nombre para ejercer libremente su profesión como experto informático, lo que es razonable y legítimo en un sistema democrático que apuesta por la rehabilitación y la inserción.	
4									clause		Carrier	Lo que requiere una investigación y la correspondiente depuración de responsabilidades es la colaboración que la empresa del expistolero presta a la lucha contra el crimen con acceso a datos confidenciales.	
5									NP	Proper Noun	Sayer	Hellin siempre se jactó de sus buenas relaciones con los agentes de la ley.	
									NP+App	Proper Noun+clause	Carrier	David Martínez Loza, el que desde Fuerza Nueva ordenó el asesinato de Yolanda González, era un guardia civil en excedencia y el propio Hellin se alojaba en casa de un	

#### D. Appendix

7							Inevitables	NP	Noun	Attribute	Inevitables son, por tanto, las sospechas que retornan a una de las etapas más oscuras de la historia política reciente; maxime cuando llueve sobre mojado.
8	porque						fue también la prensa	Clause		Actor	Porque fue también la prensa la que hace 20 años abrió los ojos a la policía dando con el paradero de Helin en Paraguay, donde se había refugiado tras huir de la cárcel.
9							Las primeras reacciones oficiales dan	NP	Der Adj Noun Adj	Carrier	Las primeras reacciones oficiales son, sin embargo, decepcionantes.
10						dan		Verb		process	Dan a entender que se desconocía la identidad del hombre que trabajaba para Interior, lo que evidenciara una grave negligencia, una ausencia de filtros que pone en riesgo la propia seguridad del Estado.
11							Patéticos	AdjP	Adj	Attribute	Patéticos resultan, por lo demás, las explicaciones del secretario de Estado de Seguridad, Francisco Martínez, quitando hierro al asunto alegando que Helin es uno más en su empresa y que fue contratado durante la etapa socialista; una buena coartada para alimentar la antipolítica, pero inútil para infundir confianza en torno a los vigilantes de la ley.
12							Esta inquietante situación			Carrier	Esta inquietante situación requiere sin duda una explicación seria desde el más alto nivel.

Europa vigila

CI #	Thematic Field	CI						Text		
		TEXTUAL	INTERP.	PREHEAD	SYNFACTIC REALIZATION	HEAD	Experiential Role			
		Stance	Engagement	Stance	Engagement	Stance	Engagement			
1						La Comisión Europea (CE)	NP	Proper Noun	Senser	La Comisión Europea (CE) no entiende el proceso de fusión de los reguladores independientes sectoriales que, sin excesivas prisas, pretendía poner en pie el Gobierno español prácticamente desde enero de 2012.
2						Neelie Kroes, vicepresidenta de la CE,	NP+App	Proper Noun+NP	Sayer	Neelie Kroes, vicepresidenta de la CE, ha hecho saber al ministro Soría que el proyecto de fusión causará, tal y como está redactado, una pérdida importante de independencia en las funciones de regulación y control de los mercados.
3						La amenaza de abrir un procedimiento de infracción si no se cambian los términos del proyecto de fusión de la Comisión Nacional de Competencia, la Comisión de Energía, los tres reguladores sectoriales más importantes, para formar un megaregulador llamado Comisión Nacional de Mercados y Competencia (CNMC), indican que la CE ha captado el excesivo sesgo político de la fusión y la irrelevancia práctica de la nueva CNMC.	NP+App	NP+clause	Sayer	
4						Las objeciones de fondo de Kroes	NP+App	NP+PP	Actor	Las objeciones de fondo de Kroes se centran principalmente en el hecho de que muchas de las competencias de investigación y sanción de los reguladores sectoriales, en particular las que atañen a las telecomunicaciones, se trasladan al Ministerio de Industria, y en que el cuerpo técnico de los reguladores, supuestamente independiente y que según los modelos vigentes de Europa es elegido por cada uno de los organismos sectoriales, sería decidido por el Consejo de Ministros.
5		En otras palabras,		Manner		el proyecto del Gobierno	NP+App	NP+PP	Carrier	En otras palabras, el proyecto del Gobierno pretende convertir unas instituciones orientadas a garantizar la competencia de los mercados, en otra dirección general más del Ministerio de Industria.
						El pretexto	NP	Det Noun	Carrier	El pretexto es el ahorro de gasto público, pero resulta indefendible desde el momento que la CNMC se pretende financiar con cargo a los Presupuestos mientras que los

[illegible]

## Medidas contundentes

[illegible]

### D. Appendix

[illegible]



				Demasiado riesgo de dinero sospechoso en quien fue gerente del PP durante casi dos decenios, senador por ese partido y ascendido a tesorero en 2008, como para que la sociedad pueda entender cualquier dilación indebida en la investigación y enjuiciamiento de los hechos.	Senser	
					Det Noun	
					NP	
					hedge	
					Contingency AdjP	
					Contingency	
				attitude		
				Demasiado riesgo de dinero sospechoso en quien fue gerente del pp durante casi dos decenios, senador por este partido y ascendido a tesorero en 2008, como para que		

Cita en Berlín

Thematic Field	Thematic Field										Text		
	TEXTUAL	INTERP.	Stance	Engagement	PREHEAD	Stance	Engagement	SYNTACTIC REALIZATION Grad.	HEAD	Stance		Engagement	Experiential Role
													El presidente del Gobierno español visita hoy Berlín para entrevistarse con la canciller alemana, Angela Merkel, antes del próximo Consejo Europeo, el día 7 de febrero en Bruselas.
1	Aunque				la visita no tiene un carácter extraordinario			Contingency	Contingency clause				Aunque la visita no tiene un carácter extraordinario, el actual contexto europeo y español le otorga una especial relevancia.
2									Los países			Det Noun	Los países celebran cada año cumbres bilaterales y la canciller Merkel ha tenido un especial interés en verse con Mario Monti, Mariano Rajoy y François Hollande, los principales gobernantes de la Europa latina, antes de una cumbre europea con dos asuntos de gran calado: el presupuesto de la Unión Europea para el 2013 y el desideratum del Reino Unido por la celebración de un referéndum, antes del 2018, sobre su pertenencia a la Unión.
3									Rajoy			Proper Noun	Rajoy no viaja a Berlín en las mejores condiciones posibles, pese a las previsiones de una ligera mejora de la economía española durante el segundo semestre del 2013.
4									El expediente español			Det Noun Adj	El expediente español sigue preocupando en Europa, pese a las notas positivas en la previsión macroeconómica, y la situación política interna se ha visto subitamente agravada por una sucesión de revelaciones periodísticas sobre una supuesta contabilidad paralela del partido gobernante en España.
5								Time	Time-PP			Noun	Hace tres meses, España era noticia en los principales medios de comunicación internacionales por el encuadre soberanista de las elecciones catalanas, y estos días vuelve a serlo por la presunta circulación de dinero negro por las ciberías del Partido Popular.
5					Hace tres meses								La prensa alemana de referencia ha concedido estos últimos días singular relevancia a las informaciones referidas a la tesorería del PP y al caso Gürtel, nombre que no deja de llamar la atención en el mundo germanico, puesto que significa corrección en su lengua.
7									La prensa alemana de referencia				No son buenas compañías para viajar a Berlín y pedir a la Cancillería una suavización de la política de austeridad en el sur de Europa.
8					son	booster		Verb	son	booster			La carpeta española pesa y a ella se sumó, ayer, la petición de dimisión de Rajoy formulada por el jefe de la oposición, Alfredo Pérez Rubalcaba.
9									La carpeta española			Det Noun Adj	Con esta iniciativa dominical, el líder del PSOE pone al presidente en la incómoda posición de tener que pronunciarse en Alemania -cuando comparezca ante la prensa junto a la canciller- sobre un asunto poco agradable de abordar en el extranjero.
					Con esta iniciativa dominical			Manner	el líder del PSOE				Con esta iniciativa dominical, el líder del PSOE pone al presidente en la incómoda posición de tener que pronunciarse en Alemania -cuando comparezca ante la prensa junto a la canciller- sobre un asunto poco agradable de abordar en el extranjero.

#### D. Appendix

[illegible]

## El euro fuerte

CI #	Thematic Field	SYNTACTIC REALIZATION												Text								
		TEXTUAL		INTERP.		PRIHEAD		STANCE		ENGAGEMENT		SYNTACTIC REALIZATION			HEAD		STANCE		ENGAGEMENT		EXPERIENTIAL ROLE	
1																						LA fortaleza de la moneda única está creando tensiones en la zona euro, donde países como Francia están alzando la voz sobre la repercusión que tiene sobre la competitividad de sus exportaciones, juicio que Alemania no comparte.
2																						Las declaraciones ayer del presidente del Banco Central Europeo (BCE), Mario Draghi, indican la necesidad de limitar diferencias internas y de encontrar un equilibrio con vistas a facilitar un acuerdo sobre el presupuesto de la UE 2014-2020, que se debate ahora.
3																						La reciente iniciativa de Japón de buscar un yen débil para mejorar sus ventas al exterior ha provocado una tormenta en los ámbitos exportadores aun cuando las políticas de flexibilidad monetaria seguidas en EE.UU. y Gran Bretaña hayan llevado a resultados similares.
4																						En Europa la preocupación se ha extendido tras ver el euro revalorizarse este año un 10% con el yen y un 6,5% con el dólar en seis meses.
5	Pero																					Pero frente a la posición de París, Berlín ha indicado que no encuentra problemas en el valor actual del euro.
6																						En esta polémica, el mercado no prevé reacciones hasta que el euro supere los 1,45 dólares y juzga la polémica entre los dos grandes países de la zona como un reflejo de la diferencia en competitividad y en la evolución de sus economías.
7																						España, con una exportación fuerte, no ha hecho mención sobre el valor de la divisa.
																						Draghi buscó el punto medio señalando que el tipo de cambio no es objetivo de la política monetaria, pero vigilará su impacto sobre el crecimiento.

## **D.6    Annotated Letters to the Editor in English**

## D. Appendix

### ENGLISH LETTERS TO THE EDITOR

A disappointing day for democracy. The Independent

#	Outer	Inner theme					Head					Text				
	Textual	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant				
1								I	self mention		NP	Pronoun I	Carrier			
2								Further signatories			NP	Adj Noun	Goal			
3								A figure in excess of £100m			NP+App	NP+PP	Verbiage			
4								The dedicated few	attitude		NP	Det Adj Noun	Actor			
5								How many		question	Interrogative	Interrogative	Attribute			
6	For goodness sake let's	Stance	directive	have			Verb	have			Verb	Verb	Process			
7				Fund			Verb	2p imp		directive	Verb	2p sing Imperative	Process			

BBC inspires all classes. The Independent

#	Outer	Inner theme					Head					Text				
	Textual	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant				
1								I	self mention		NP	Pronoun I	Senser			
2				Tell			Verb	2p imp		directive	Verb	2p sing Imperative	Process			
3								We	self mention		NP	Pronoun we	Senser			
4							Time-Advb	Sunday lunchtimes			NP	Noun	Scope			
5								I	self mention		NP	Pronoun I	Actor			

Dangerous ways to raise money. The Independent

#	Outer				Inner theme				Head				Text
	Textual	Stance	Element	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	
1									The Government			NP	The Government is giving serious consideration to charging to cross the Thames via the Blackwall Tunnel in east London. In order to create some spurious "ring-fencing" of funds towards the cost of a new Thames crossing further downstream.
2									Drivers			NP	Drivers would then be forced to either pay up or drive further into London seeking a free crossing route.
3									The Dartford crossing			NP	The Dartford crossing was fully paid up years ago but that toll is to be increased to £2.50.
4 Thus									the milch-cow motorist			NP	Thus, the milch-cow motorist would be asked to pay twice yet again to use longstanding infrastructure – once through normal fiscal channels and now with these tolls.
5									This government			NP	This government spouts about "fairness".
6		Eng	It is politically and ethically fair that						new projects			NP	It is politically and ethically fair that new projects be paid for either through taxation or through the marketplace with the issuing of gilts.
7									The electorate			NP	The electorate can then make judgements on such decisions via the ballot box.
8									Deviation from these traditional methods			NP+App	Deviation from these traditional methods also occurred with the building of new hospitals some years ago using Private Finance Initiative (PFI), originally under the Tories but also continued under Labour.
9									Those hospitals			NP	Those hospitals are now crippled with the paying back of the PFI debt for decades to come.
10 Once					governments begin to go "off-piste" by identifying certain projects as being the financial liability of specific sections of the population then		directive	Time	trouble			NP	Once governments begin to go "off-piste" by identifying certain projects as being the financial liability of specific sections of the population then trouble will surely ensue.

Holland Park to Row Music. The Independent

#	Outer				Inner theme				Head				Text
	Textual	Stance	Element	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	
1					In your feature on Holland Park School (15 November) you write that Bryan Ferry taught at the school in the 70s.			Place	you			NP	In your feature on Holland Park School (15 November) you write that Bryan Ferry taught at the school in the 70s.

### D. Appendix

2								Bryan		NP	Proper Noun	Actor	Bryan did teach pottery at a girls' school in south London in 1970, but the Roxy Music member who taught at Holland Park was me.
3								I	self mention	NP	Pronoun I	Actor	I spent a chaotic but very happy time as a supply teacher in the music department in 1971-72, where among other things I persuaded the school to buy an EMS suitcase synthesiser.
4								I	self mention	NP	Pronoun I	Actor	I still encounter ex-pupils, a few of whom remember my time there with enjoyment.
5								I	self mention	NP	Pronoun I	Actor	I left suddenly in February 1972 to go straight into the studio to record the first Roxy Music album.

intellectual vacuum

#	Outer		Inner theme								Head				Text
	Textual		PreHead		Syntactic Realization		Element	Stance	Reader	Syntactic Realization	Stance	Reader	Syntactic Realization	Participant	
1		Perhaps						James Dyson				NP	Proper Noun	Carrier	Perhaps James Dyson was a little harsh towards poets: poetry, literature, music and paintings all feed our spirit.
2	But				despite this	attitude		Contingency-pp	the human mind			NP	Det Adj Noun	Carrier	But despite this the human mind is little changed from 5,000 years ago.
3									What distinguishes mankind now from what we were then, for good or ill		shared knowledge	Clause	Wh cleft	Carrier	What distinguishes mankind now from what we were then, for good or ill, is science and technology.
4									advances			NP	Adj Noun	Actor	No advances could have been made without the work of inventors, be it paper, the printing press, television, modern medicine and agriculture, let alone mobile telephones and the internet.
5									Those studying the arts			NonFinite	NonFinite	Carrier	Those studying the arts have inventors to thank for the means to study.

Now can we see a politician take responsibility?

#	Outer	Inner theme					Head			Participant	Text
	Textual	Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader		
1	Textual										The politicians who lined up to demand the resignation of the BBC's recently appointed Director General, George Entwistle, because he was not aware of the content of every programme across BBC output, will be dancing in the aisles to celebrate the damage done to probably the only British public institution that still has a global reputation for basic honesty.
2											I wonder if they will now also demand the immediate resignation of Patrick McLoughlin, Secretary of State for Transport, who denied any responsibility for more than £100m compensation payments for the West Coast rail franchise fiasco, as he – the politician – could not be expected to know about every action

## D.6 Annotated Letters to the Editor in English

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant		
1							I	self mention		NP	Pronoun I Senser		
2	But						we		inclusive pronoun	NP	Actor		
3							The libel and slander laws	attitude		NP+App NP	Carrier		

I understand Neil March's discomfort about the BBC payout to Lord McAlpine (letter, 17 November), though what he does with the money is surely his business.  
But we should applaud Lord McAlpine for taking action to discourage thoughtless and irresponsible tweeting.  
The libel and slander laws exist to prevent damaging lies being told or repeated about citizens, rich or poor, whoever they may be: Twitter should be shown to be as subject to those laws as newspapers and television

Prosperity and peace for Gaza

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant		
1						Time	Time-Clause As long as the Gaza Strip remains under indefinite siege,	attitude violent clashes between Israel and the incarcerated population		NP+App NP+pp	Carrier		
2							This indefinite siege	attitude		NP	Carrier		
3							It			NP	Carrier		
4		whether we like it or not,					Hamas,		shared knowledge	NP+App NP Conj Clause	Senser		
5							The West			NP	Det Noun Actor		
6							It is only when Gaza prosper that	attitude		clause	Carrier		

As long as the Gaza Strip remains under indefinite siege, violent clashes between Israel and the incarcerated population are inevitable.  
This indefinite siege is immoral and criminal, since it is collective punishment.  
It is also stupid, since it resolves nothing, adds to mutual bitterness and renders the conflict more intractable.  
Hamas, whether we like it or not, shows no sign of weakening.  
The West should negotiate a deal with Hamas for international monitors, and then tell Israel it must lift the siege and allow Gaza free access by sea and air, as well as open land access with Egypt.  
It is only when Gaza prospers that those Palestinians wanting to hit Israel targets will no longer be tolerated.

Put Starbucks out of its misery

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant		
1							We		inclusive pronoun	NP	Actor		
2						Evidence- clause	It is reported that			NP	Det Adj Adj Noun Sayer		

We read that Starbucks has reported a taxable profit only once in its 15 years of operation in the UK; in consequence over this time it has paid just £8.6m in corporation tax.  
It is reported that its chief financial officer has said: "We're not at all pleased about our financial performance here."

## D. Appendix

3											These multinational companies	NP	Det Adj Noun	Actor	These multinational companies come to Britain to enjoy the benefits of trading in a secure, politically stable environment, yet do their very best to weasel out of making any realistic contribution to the upkeep of their host nation or the education and welfare of its people.
4											Contingency-NonFinite	Contingency		Carrier	Given the terrible trading performance of Starbucks, surely it would be an act of mercy to the shareholders of this multinational for our Government to introduce "coup de grâce" legislation to force the closure of its UK operations.

Schools in need of discipline

#	Outer Textual	Inner theme PreHead					Head					Text				
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant	Stance	Reader	Text
1									Richard Garner (News, 16 November)			NP	Sayer			Richard Garner (News, 16 November) reports that Buckingham University has recently shown that nearly 30 per cent of trainee teachers last year had not taken up teaching posts six months later, and that 52 per cent of teachers under the age of 60 are not within the classroom.
2								Evidence-PP	to argue that the Government's Teaching Agency should withdraw funding for places from teacher training providers who perform poorly in supplying recruits to the profession	hedge		Clause	Carrier			For Alan Smithers, one of the authors, to argue that the Government's Teaching Agency should withdraw funding for places from teacher training providers who perform poorly in supplying recruits to the profession is wholly inappropriate.
3									The biggest problem facing trainee teachers, and most teachers, is the poor behaviour that teachers experience day in, day out.	booster		NP+App nite	Token			The biggest problem facing trainee teachers, and most teachers, is the poor behaviour that teachers experience day in, day out.
4								Role	As a teacher, self mention	attitude		NP	Phenomenon			As a teacher, the depressing statistics do not surprise me.

Baffling rail fares

#	Outer Textual	Inner theme PreHead					Head					Text				
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant	Stance	Reader	Text
1									Keith Eaton's example of saving £38 on a Bristol to Paddington fare, by buying separate tickets for each half of the journey, is replicated all over Scotland.			NP+App Noun Clause	Carrier			SIR – Keith Eaton's example of saving £38 on a Bristol to Paddington fare, by buying separate tickets for each half of the journey, is replicated all over Scotland.



## D.6 Annotated Letters to the Editor in English

3	But				buy					Verb	Verb	2p imp			directive	Verb	2p sing Imperative	Process	But buy two tickets — Aberdeen-Stonehaven and Stonehaven-Dundee, and the fare on the same train falls to £18.
4												This practice				NP	Det Noun	Token	This practice is termed “split-ticketing”, and it adds a further hurdle for unwary passengers within an already labyrinthine fares system.

Correct etiquette for cyclists crossing the road

#	Outer Textual	Inner theme PreHead					Head					Text					
		Reader	Stance	Element	Reader	Stance	Syntactic Realization	Reader	Stance	Element	Reader		Stance	Syntactic Realizati	Participant		
1						May		question	verb	Interrogative I		self mention			NP	Pronoun I Sayer	May I add to Christopher Pratt's list of dangerous cycling habits (Letters, November 13) those cyclists who ride along the pavement and then veer across zebra crossings as if they had right of way?.
2						As a driver		self mention		Role-AdvbP I		self mention			NP	Pronoun I Actor	As a driver, I have often had to brake sharply to give way to a mounted cyclist appearing in front of me on a zebra crossing.
3										Cyclists					NP	Noun Actor	Cyclists only have precedence on the various forms of pedestrian crossings if they are carrying their machines – even dismounting and pushing their bicycle is not sufficient.
4	However					most			Degree	Degree-Adj					NP	Adj Noun Senser	However, most road users accept that these crossings are the safest way for a cyclist to cross a busy road, and if dismounted cyclists cross with the care exercised by most pedestrians, they are unlikely to bother anyone.

How can it be right to send this SAS hero to jail for 18 months?

#	Outer		Inner theme										Head			Text
	Textual		PreHead													
	Textual		Element	Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant			
1									The case of Sgt Danny Nightingale (report, November 11), the Special Air Service soldier jailed for possessing a “war trophy” pistol presented to him by the Iraqi army for outstanding service,			NP+App NP+Claus e	Phenomenon	The case of Sgt Danny Nightingale (report, November 11), the Special Air Service soldier jailed for possessing a “war trophy” pistol presented to him by the Iraqi army for outstanding service, made deeply distressing reading.		
2									Sgt Nightingale			NP	Proper Noun	Actor	Sgt Nightingale kept the pistol in a locked container at home, and has harmed nobody, but was jailed, kicked out of the Army, and his young, wholly innocent, family was made homeless.	
3						Meanwhile		Time	Time-Advb			NP+NP+N P	NP+NP Conj NP	Actor	Meanwhile, convicted terrorists, paedophiles, crooked politicians and rapists can roam our streets, incredulous at the generous hand they have been dealt by British justice.	
4			Of course,	Com	shared knowledge				our Army		inclusive pronoun	NP	Det Proper Noun	Carrier	Of course, our Army should be hard on its own if they have erred, but it needs a sense of proportion, which it has failed to show in the case of Sgt Nightingale.	
5									He			NP	Pronoun	Goal	He should be freed, re-instated in the Army, and have his home returned to him and his family.	

## D. Appendix

#	Outer Textual	Inner theme Pre-Head							Head			Text			
		Stance	Reader	Element	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant	Text	
1					As an independent candidate in the recent police and crime commissione r election	self mention		Role	I	self mention		NP	Pronoun I	Actor	SIR – As an independent candidate in the recent police and crime commissioner election, I can take my defeat at the ballot box relatively phlegmatically.
2									I	self mention		NP	Pronoun I	Senser	I feel the opposite when I consider the Government's abysmal handling of this election, and the appallingly low interest it stimulated in the electorate.
3									A lack of information, not apathy	attitude		NP+NP	NP+NP	Carrier	A lack of information, not apathy, seems to have been the problem.
4					Whatever people thought of the merits or otherwise of that election,	hedge		Contingency	Contingency- Clause			NP	Det Noun	Senser	Whatever people thought of the merits or otherwise of that election, the majority wanted to vote but knew nothing about the candidates – a particularly significant point when so many were concerned that the police should not become politicised.
5									A lame effort	attitude		NP	Det Adj Noun	Goal	A lame effort has been made by Downing Street to lay responsibility at the door of the media, but then what is ever the fault of politicians?.
6									I	self mention		NP	Pronoun I	Senser	I believe the blame lies with the Home Office, and more specifically with stubborn ministers and arrogant officials.
7									The idea that the electorate would be informed simply by visiting a Home Office website, even if they could ever be so persuaded or inclined	hedge		NP+App	NP+Clause e	Recipient	The idea that the electorate would be informed simply by visiting a Home Office website, even if they could ever be so persuaded or inclined, is ludicrous.
8					On countless occasions	booster		Time	Time-PP		inclusive pronoun	NP+App	NP+PP	Actor	On countless occasions a number of us sought to persuade the Home Office to include candidate details in the Electoral Commission booklet that was to be delivered to every household in England and Wales to explain the voting system.
9									This			NP	Pronoun this	Goal	This was refused on the grounds of cost – alleged to be £30 million.
10					This figure			Other	Other-NP	self mention		NP	Pronoun I	Senser	This figure I do not believe.
11									The additional cost			NP	Det Adj Noun	Carrier	The additional cost would have been in printing two or three pages more in each booklet, while the bulk of the cost for the booklet was in the delivery, which had already been committed to.
12	Even if				£30 million was the correct figure	hedge		Contingency	Contingency- clause			NP	Pronoun it	Carrier	Even if £30 million was the correct figure, it would seem a small price to have paid to help the electorate exercise its democratic right on a substantially more informed basis than was the case.
13		Sta	directive	Let us hope	for once	hedge		Time	Time-PP			NP	Noun	Scope	Let us hope that, for once, lessons really have been learnt by

## D.6 Annotated Letters to the Editor in English

#	Outer Textual	Inner theme					Head					Text				
		Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant			
1					Once again			Time-pp	the Synod of the Church of England			NP+App	Actor			
2									I	self mention		NP	Pronoun I Goal			
3					By approving the ordination of women priests 20 years ago	hedge		Manner-NonFinite	the Synod			NP	Det Noun Actor			
4									To deny this		directive	NonFinite	NonFinite Carrier			
5				Com	Surely				canon law			NP	Adj Noun Actor			

Native ash tree hope

#	Outer Textual	Inner theme					Head					Text				
		Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant			
1	While				the fungus that causes ash dieback may be windblown up to a point,	hedge		Time-Clause	the greatest threat to our native broadleaves	booster		NP+App	Actor			
2									The finger			NP	Det Noun Actor			
3					As a condition of tree planting grants;	hedge		Evidence-As	it			NP	Pronoun it			
4	But								it			NP	Pronoun it			
5									Tens of million of trees, grown from non-native seed			NP+App	NP+NonFinite Clause			
6									This			NP	Pronoun we			
7									We		Inclusive pronoun	NP	Actor			
8	If				as some scientists believe	hedge		Role	there		shared knowledge	NP	There			

## D. Appendix

No purple pride

#	Outer Textual	Inner theme PreHead					Head					Text				
		Reader	Stance	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant				
1								This country		Inclusive	NP	Det Noun	Carrier	SIR –		
2								English national teams			NP	Adj Noun	Actor	This country has no link with purple (Letters, November 19).		
3								These colours			NP	Det Noun	Carrier	English national teams play in white, and sometimes in blue or red.		
4								Other leading teams such as New Zealand and South Africa			NP+App	Det Noun	Senser	These colours have historical significance.		
5					Playing in purple,			we			NP	Pronoun	Actor	Other leading teams such as New Zealand and South Africa are proud of their national colours and play in the same strip every time.		
6							Manner-NonFinite	The Rugby Football Union			NP	Proper Noun	Actor	Playing in purple, we lost against a weak Australian team that was thrashed by France the weekend before.		
7		Eng		It is humiliating that				money-grubbing of this kind			NP	Pronoun	Carrier	The Rugby Football Union should focus on building a better team, not on alienating England's supporters.		
8								Our team		Inclusive	NP	Det Noun	Senser	It is humiliating that money-grubbing of this kind is allowed at a national level.		
														Our team will not have been inspired by its kit.		

Overburdened forces

#	Outer Textual	Inner theme PreHead					Head					Text				
		Reader	Stance	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant				
1								You		Inclusive pronoun	NP	Pronoun you	Sayer	SIR –		
														You report (November 15) Gen Sir David Richards, the Chief of the Defence Staff, as saying: "We have a whole load of tasks expected of us. Our political masters are quite happy to reduce the size of the Armed Forces, but their appetite to exercise influence on the world stage is, quite understandably, the same as it has always been... if you reduce your Armed Forces... something gives."		
2								Even a cursory glance at the position	attitude		NP+App	NP+PP	Actor	Even a cursory glance at the position since the Strategic Defence and Security Review of 2010 exposes the truth: too few "RAF and RN force enablers" to support even a much reduced Army.		
3	Therefore							the problem, to which the Chief of the Defence Staff alludes			NP+App	NP+Claus	Carrier	Therefore, the problem, to which the Chief of the Defence Staff alludes, is not so much the size of the Army but how, in an expeditionary age, other than in very small packets, expeditionary forces can safely and speedily be transported to where they are needed, and provided with adequate air and sea power support in-theatre.		

## Terrible cycle attacks in Israel: retaliation and recruitment to Hamas

#	Outer	Inner theme										Text
		PreHead					Head					
	Textual	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant		
1						There		shared knowledge	NP	Existent	There was a throwaway line in your report from Gaza yesterday about "Hamas, which is the elected government in Gaza" .	
2				In September 2005, as Israel completed its evacuation of the Gaza Strip,	Time	the Palestinian Authority			NP	Senser	In September 2005, as Israel completed its evacuation of the Gaza Strip, the Palestinian Authority was planning elections.	
3		The hope was that	Stative			Other-Clause			NP+AppP	Actor	The hope was that Palestinians, in Gaza, the West Bank and east Jerusalem, would be able to govern themselves, and a significant step would have been taken towards achieving the two-state solution.	
4						Hamas			NP	Actor	Hamas won 74 seats; the ruling Fatah 45.	
5						President Mahmoud Abbas			NP	Actor	President Mahmoud Abbas formed a national unity government, led by Ismail Haniyeh of Hamas as prime minister.	
6	But					sharing power with Fatah			NonFinite	Phenomenon	But sharing power with Fatah did not suit Hamas.	
7				In June 2007	Time	its "Executive Force"			NP	Actor	In June 2007 its "Executive Force" seized control of the Gaza Strip in a bloody coup.	
8						President Abbas			NP	Actor	President Abbas dissolved the national unity government and formed an emergency government based in the West Bank city of Ramallah.	
9						Forlorn efforts at reconciliation between the two power blocs within the Palestinian body politic			NP+App	Actor	Forlorn efforts at reconciliation between the two power blocs within the Palestinian body politic began as early as 2008.	

#### D. Appendix

10	Insofar	as Mahmoud Abbas has embraced the concept of the two- state solution to the Israeli- Palestinian conflict – and went so far in September 2010 as to sit down at the same table with Israel's prime minister and talk peace –	hedge				Time	Time-Clause	the Palestinian Authority				NP	Det Adj Noun	Actor	Insofar as Mahmoud Abbas has embraced the concept of the two-state solution to the Israeli-Palestinian conflict – and went so far in September 2010 as to sit down at the same table with Israel's prime minister and talk peace – the Palestinian Authority has placed itself beyond the pale in Hamas's eyes.
11	For								Hamas				NP	Proper Noun	Carrier	For Hamas remains an extreme Islamist and terrorist organisation committed to the destruction of Israel.

## The Arab spring has brought a new extremism to bear against Israel

#	Outer Textual	Inner theme							Text						
		Pre-Head			Head										
	Textual	Stance	Reader	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant			
1								I	self mention		NP	Pronoun I	Carrier	I am in Israel with my family, either glued to the television or reading Chaim Herzog's history, The Arab-Israeli Wars.	
2								The adage of history repeating itself	attitude		NonFinite	NonFinite	Carrier	The adage of history repeating itself is evident, with one new facet: the Arab Spring.	
3								This			NP	Pronoun this	Actor	This is producing a string of extremist Islamic governments, in Tunisia, Libya, Egypt and probably Syria, then most likely Jordan thereafter.	
4	When			the Western world wakes up to the fact that it is being challenged by extreme Islam,	booster		Time	Time-Clause	it		NP	Pronoun it	Goal	When the Western world wakes up to the fact that it is being challenged by extreme Islam, it will most probably be too late.	
5				Meanwhile			Time	Time-Advb	Israel, the only spearhead of Western ideological values,	attitude		NP+App	Proper Noun+NP	Target	Meanwhile, Israel, the only spearhead of Western ideological values, is demonised, and criticised for "disproportionate" reaction to its citizens being made to live half their lives in bomb-proof shelters.

## D.6 Annotated Letters to the Editor in English

6	My only constructive suggestion	is that if Egypt can be coerced (or bribed) into annexing Gaza, and likewise Jordan into annexing the West Bank, you would have a chance of some peace between three sovereign countries.	hedge	Contingency clause	you	Element	Stance	Reader	Syntactic Realization	Participant	Carrier	NP	Inclusive pronoun	My only constructive suggestion is that if Egypt can be coerced (or bribed) into annexing Gaza, and likewise Jordan into annexing the West Bank, you would have a chance of some peace between three sovereign countries.
---	---------------------------------	---	-------	--------------------	-----	---------	--------	--------	-----------------------	-------------	---------	----	-------------------	---

Deciding who benefits

#	Outer Textual	Inner theme										Head		Text
		Pre-Head	Reader	Stance	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant	
1									Universal benefits			NP	Adj Noun	Universal benefits are economically and morally dubious.
2									The welfare state			NP	Det Adj Noun	The welfare state exists to provide a safety net to those in genuine need of assistance.
3	It is wrong, and unjustifiably costly	to have benefits that	Eng		Contingency pp			Contingency	everybody			NP	Pronoun everybod y	It is wrong, and unjustifiably costly, to have benefits that everybody can claim.
4									Your correspondents (Letters, December 19)			NP	Pronoun Noun	Your correspondents (Letters, December 19) offer strange defences of universal benefits for pensioners.
5									Pensioners' bus passes			NP	Noun Noun noun	Pensioners' bus passes do not subsidise transport for children – this is the role of local councils.
6									Means testing			NP	Noun Adj	Means testing is not too difficult – surely it would be easy to declare your income on the form for claiming benefits.
7									The Chancellor			NP	Proper Noun	The Chancellor would not be amiss in removing many of the tax changes that have decimated private pension provision.
8	However								this			NP	Pronoun this	However, this does not justify giving benefits to those who have survived government attacks with a generous retirement income.

Can the police force be trusted to investigate his own failures?

#	Outer Textual	Inner theme										Head		Text
		Pre-Head	Reader	Stance	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant	
1									The most recent fiasco over Andrew Mitchell, the former Conservative chief whip.	attitude		NP+App	Np+Np	The most recent fiasco over Andrew Mitchell, the former Conservative chief whip, is just one of a long list of police failures.
2									Respect for the police			Nonfinite	Nonfinite Actor	Respect for the police has declined over the years, often because the documentation and reports have been doctored in some way, or the police have ignored minor criminal activities.

## D. Appendix

4					In view of the past failures by the police, can we trust them to investigate themselves?	we				Inclusive pronoun	Pronoun we	senser	In view of the past failures by the police, can we trust them to investigate themselves?
---	--	--	--	--	--	----	--	--	--	-------------------	------------	--------	--

Hospital consultants already carry our work commitments at weekends

#	Outer Textual	Inner theme										Text				
		Pre-Head					Head									
	Textual	Stance	Reader	Element	Stance	Reader	Syntactic Realization	Contingency	Advbp	Element	Stance	Reader	Syntactic Realizati	Participant		
1										Cristina Odone's views on doctors and weekend working ("Doctors should get off the golf course and into the wards", Comment, December 17)			NP+App	Actor	Cristina Odone's views on doctors and weekend working ("Doctors should get off the golf course and into the wards", Comment, December 17) would have brought a wry smile to the thousands of hospital consultants who, like me, spent the weekend working in their local hospital as part of their on-call commitment.	
2										Weekend working			NP	Noun Adj	Carrier	Weekend working has been part of my professional life since I qualified 35 years ago.
3					As I will have worked for 12 straight days by Friday,			Contingency	Contingency-Advbp	I	self mention		NP	pronoun I	senser	As I will have worked for 12 straight days by Friday, I do hope that no one will begrudge me a weekend off.
4										Improving provision of diagnostic services, such as scans, at weekends			NonFinite	NonFinite	Actor	Improving provision of diagnostic services, such as scans, at weekends does help speed diagnosis and patient throughput.
5										My own hospital, like many others, is already working on this.		inclusive pronoun	NP+App	NP+pp	Actor	My own hospital, like many others, is already working on this.
6	However									the question of carrying out planned non-emergency work at weekends			NonFinite	NonFinite	Actor	However, the question of carrying out planned non-emergency work at weekends comes down to three factors: available workforce, facilities and finance.
7					As I am already working full time, if I operate on a Saturday,			Contingency	Contingency-Advbp	someone else			NP	Pronoun someone else	Actor	As I am already working full time, if I operate on a Saturday, someone else will have to do my weekday operating list.
8	In addition									doctors			NP	Noun	Actor	In addition, doctors do not work in isolation, but are part of a highly trained team.
9										This			NP	Pronoun this	Carrier	This all costs money.
10										Tesco			NP	Proper Noun	Actor	Tesco doesn't open at weekends out of a sense of social beneficence, but because doing so generates profits.

Dementia screening

#	Outer	Inner theme										Text	
---	-------	-------------	--	--	--	--	--	--	--	--	--	------	--



## Memorable Christmas song are things of the past

#	Outer		Inner theme										Text	
	Textual		PreHead					Head						
	Textual		Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant	
1									I	self mention		np	senser	I know that Nigel Farndale (Opinion, December 16) is a Yorkshireman.
2	But if				he is not prepared to fork out 15p to stay in touch with old friends by buying a Christmas card	hedge		Contingency Advdp	that			np	Actor	But if he is not prepared to fork out 15p to stay in touch with old friends by buying a Christmas card – well, that is taking stereotypes too far.

## D. Appendix

#	Textual	PreHead				Head				Text
		Stance	Element	Reader	Syntactic Realization	Stance	Element	Reader	Syntactic Realization	
1							we	inclusive pronoun	np	We are told that a majority of people are opposed to fox hunting.
2		Com	Unfortunately		Contingency pp		the majority		np	Unfortunately for the fox, the majority do not understand the countryside and are too easily influenced by television programmes out to prove a point.
3		Com	All too often				I	self mention	np	All too often I find old, diseased foxes eaten alive by mange.
4	When		fox hunting was allowed,		time		these old and diseased foxes		np	When fox hunting was allowed, these old and diseased foxes would have been put out of their misery.
5				shared knowledge			The strong, healthy foxes		np	The strong, healthy foxes would outwit the hounds and breed to ensure the survival of a strong and healthy stock.
6			As it is now	shared knowledge	Manner-clause	attitude	we	inclusive pronoun	np	As it is now, we have a prime example of the law of unintended consequences, which is already resulting in the deterioration of a fine native species and making its long-term survival doubtful.

The Government has no mandate to redefine the meaning of marriage

#	Outer Textual	Inner theme PreHead				Head				Text
		Stance	Element	Reader	Syntactic Realization	Stance	Element	Reader	Syntactic Realization	
1			As parliamentarians from different political parties and none	self mention	Role		we	inclusive pronoun	np	As parliamentarians from different political parties and none, we are united in supporting the institution of marriage defined in law as a union between a man and a woman.
2							we	inclusive pronoun	np	We recognise the value of a loving and committed relationship and we respect civil partnership, but affirm the distinctive value of marriage reflecting the complementarity of a man and woman often evidenced in parenthood.
3			At the last election		time pp		none of the three main parties		np+pp	At the last election, none of the three main parties stood on a platform to redefine marriage.
4							It		np	It was not contained in any of their manifestos, nor did it feature in the Coalition's Programme for Government.
5							These facts		np	These facts alone should have led to extreme caution on the part of those calling for this change to be made.
6		Mod	instead				the Government		np	Instead the Government is ignoring the overwhelming public response against the plans.
7							The consultation		np	The consultation has ignored the views of 500,000 British residents in favour of anonymous submissions from anyone anywhere in the world.
8							we	inclusive pronoun	np	We believe that the Government does not have a mandate to redefine marriage.
9							we	inclusive pronoun	np	We recognise these are issues of conscience which will be given free votes in Parliament.
10							we	inclusive pronoun	np	We will be seeking legal guarantees of the same freedom of conscience for our constituents and religious organisations to teach, preach and express a traditional view of marriage.
11							us	inclusive pronoun	np	We are grateful that the proposed restrictions will protect the

12								The proposed redefinition of marriage		np+pp	Carrier	The proposed redefinition of marriage is unnecessary, given the legal rights established through civil partnerships.
13								we	inclusive pronoun	pronoun we	senser	We understand some parliamentarians support freedom for same sex couples to marry, but we support a freedom from the state being able to redefine the meaning of marriage.

## Women's premiums

#	Outer Textual	Inner theme						Head			Text				
		PreHead		Stance		Syntactic Realization		Element	Stance	Reader		Syntactic Realization	Participant		
1								The justification by Mark English of the European Commission (Letters, December 21) for outlawing gender discrimination in insurance		shared knowledge	Np+App np+pp	Carrier	The justification by Mark English of the European Commission (Letters, December 21) for outlawing gender discrimination in insurance is nonsense.		
2								he			np	pronoun he	sayer	He says “all safe drivers will pay less than drivers who are less safe”.	
3but								how		question	interrogative	attributive		But how are insurers to know who all the safe drivers are?	
4								accidents			np	noun	Carrier	Accidents are relatively uncommon (666 casualties per billion vehicle miles in 2011), so an accident-free driving record is a poor indicator of a safe driving style.	
5								New drivers,			np	adj noun	Carrier	New drivers, of course, have no record on which to assess their accident risk.	
6								rheme				rheme	None	Hence the need to use generalised risk factors such as gender, age, postcode and occupation.	
7								rheme				rheme	None	According to the commission, “male drivers will no longer pay more just because they are men” .	
8but								young drivers				np	adj noun	actor	But young drivers will continue to pay more just because they are young.
9								where		question	interrogative	attributive		Where is the logic in that?	

Traditional institution of marriage is central to the welfare of society

#	Outer		Inner theme										Text	
	Textual		PreHead			Head								
	Textual		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant
1										Charles Moore (Comment, December 15)			np	sayer
2										That			np	Carrier

3									Redefining marriage as simply a contract between individuals irrespective of their sex, without regard either to its procreative function or to the complementarity of the relationship between man and woman, would be an abuse of language.				
4	more important	Stra						it	Redefining marriage as attitude simply a contract between individuals irrespective of their sex, without regard either to its procreative function or to the complementarity of the relationship between man and woman		np	NonFinite	Carrier
5									That, and not homophobia, is why many people outside what Mr Moore calls the culturally dominant "minority" are opposed to the Government's proposal — and why more than 600,000 people have signed a petition against it.		np+app	nonFinite	Actor
6								The state	The state has the right to oversee the administration and legal aspects of marriage, but it has never been accepted that the state can dictate to individuals and society itself what marriage should mean to us.		np	det noun	Carrier
7	It is clear that	Eng						many problems	many problems would arise if the legislation as now tabled were to be implemented.		np	det noun	Actor
8			In the run-up to the last election,	time	time-clause			David Cameron	In the run-up to the last election, David Cameron led us to believe that the strengthening of marriage as an institution was one of his important objectives; and the Conservative Party's manifesto, which made no mention of "gay marriage", included a proposed tax break for married couples.		np	proper noun	Actor
9								nothing	Nothing has been heard of the latter proposal, and instead of action to strengthen marriage we have the proposal to abandon the traditional understanding of marriage on the basis of a "consultation" which explicitly excluded the possibility of a negative result.		np	pronoun nothing	verbiage
10								Protestations that this is all fundamentally "conservative"	Protestations that this is all fundamentally "conservative" ring a bit hollow.		np+app	np+claus e	actor
11	It is difficult	Eng	not to wonder	other	other- NonFinite	booster		how far the Prime Minister	It is difficult not to wonder how far the Prime Minister is someone whose steadiness of purpose can be relied on.		np	proper noun	Carrier

### Why Lord Heseltine was right in calling the Civil Service amateurish

#	Outer Textual	Inner theme PreHead						Text	
		Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Head	
	Textual	Element						Element	Participant
1								The letter from Sir Bob Kerslake, the head of the Civil Service, and Sir Jeremy Heywood, the Cabinet Secretary (December 15), establishes that, even at the highest level, distinguished civil servants do not understand the fundamental deficiencies in the Civil Service.	Sayer

Bullies and victims in the hoax call row

#	Outer Textual	Inner theme										Text	
		PreHead					Head						
	Textual	Element	Stance	Reader	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant	
1					eye for an eye		evidence	it				Carrier	Eye for an eye, or so it may seem.
2								The media and general public's reaction to the death of nurse Jacintha Saldanha		shared knowledge	NonFinite	Carrier	The media and general public's reaction to the death of nurse Jacintha Saldanha is predictable and equally deplorable.

### D. Appendix

3							Aside from the lack of attention given to the nurse herself, we have chosen to focus on demonising the two Australian DJs and their innocuous prank call as the sole reason for her death.	booster		Contingency	Contingency-nonfinite	we				inclusive pronoun	np	pronoun we	actor	Aside from the lack of attention given to the nurse herself, we have chosen to focus on demonising the two Australian DJs and their innocuous prank call as the sole reason for her death.
4	While						privacy laws may have been compromised in the telephone call,	hedge		Contingency	Contingency-clause	the act itself					np	det noun pronoun	Carrier	While privacy laws may have been compromised in the telephone call, the act itself is a staple of popular radio anywhere in the Western media.
5	so											why				question	verb	verb	attribute	So why do we find it perfectly acceptable to spew hatred, death threats, and unfair blame on two DJs who had never intended to harm anyone?
6												The hospital and Palace					np+np	det noun conj noun	sayer	The hospital and Palace have confirmed they offered their full support to the Duchess's nurse following the prank call.
7											other-clause	The same					np	det noun	verbiage	The same cannot be said about the DJs – Mel Greig and Michael Christian; the public has not recognised its own role in the mental well-being of these DJs. Mourn Jacintha, but let's also mourn the loss of livelthood and quality of life of the DJs, because if this is indeed the case, we need also to mourn the loss of human decency and compassion.

## Changing American's gun culture

#	Outer	Inner theme										Text	
		Textual					PreHead						
	Textual	Element	Stance	Reader	Syntactic Realization	Head	Element	Stance	Reader	Syntactic Realization	Participant		
1						In the past eight months or so,	there		shared knowledge	np	there	Existent	In the past eight months or so, there have been mass shootings in Seattle; in Aurora, Colorado; in Oak Creek, Wisconsin; in Minneapolis; in Portland, Oregon; and now in Newtown, Connecticut.
2							The Second Amendment			np	proper noun	sayer	The Second Amendment states, "A well-regulated Militia, being necessary to the security of a free state, the right of the people to keep and bear arms shall not be infringed".
3							The US Justice Department's official position		shared knowledge	np+app	np+np	Carrier	The US Justice Department's official position is that these rights are limited to the collective rights of states to regulate their own militia.
4							theme					None	<b>Good starting point.</b>
5	but						the National Rifle Association			np	proper noun	actor	But the National Rifle Association has hijacked the Second Amendment to assert that every individual has the right to be armed with guns.
6	but						how		question	interrogative	interrogative	attribute	But how do you control the possession of 300 million privately owned guns in your country?
7							Draconian laws			NP	Adj noun	Carrier	Draconian laws will be of limited effect because they will be flouted and probably unenforced.
8							A more effective route	attitude	question	np	det adj noun	Carrier	A more effective route might be to control the supply of ammunition?
9							The Second			np	proper	sayer	The Second Amendment says nothing about this.

11								Ammunition	np	noun	goal	Ammunition has to be replaced, so the federal control of ammunition would soon dwindle its availability.
12								This	np	pronoun this	actor	This will not solve the gun problem nor necessarily change the gun culture of its people, but something drastic has to be tried and pretty soon too.

Clegg's welcome word on drugs

#	Outer Textual	Inner theme PreHead					Head					Text
		Reader	Stance	Element	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant	
1							I	self mention		np	pronoun I carrier	I am a USMC Vietnam Veteran.
2							I	self mention		np	pronoun I Actor	I served in 1966-68.
3							All of the men in my 12-man squad	self mention		np+pp	Goal	All of the men in my 12-man squad were killed and replaced or wounded and replaced during my tour.
4							I	self mention		np	pronoun I beneficiary	I received three purple hearts and carried many dead and wounded brothers off the field of battle.
5				In June of 1968		time-pp	I	self mention		np	pronoun I actor	In June of 1968 I joined the NYPD in time for the Martin Luther King riots and the Black Panther killings of many police officers.
6				In all the experience I have had in dealing with death		evidence-pp	I	self mention		np	pronoun I senser	In all the experience I have had in dealing with death I have learned one thing: when you are dead, you are dead.
7							we		inclusive pronoun	np	pronoun we carrier	We have had guns in the United States for as long as this country has existed and we never saw the senseless killing that we are seeing today.
8							rheme		question		attribute	Why?
9							I	self mention		np	pronoun I senser	I believe that the answer to this new breed of senseless killer is violent video games, movies and television shows.
10							The children they			np	det noun actor	The children of today play video games from infancy on.
11							they			np	pronoun actor they	They shoot and kill invading aliens.
12							they			np	pronoun actor they	They shoot at and kill bad guys with the realism of red blood spurting out of their targets, heads and limbs being shot or cut off.
13	the more...the more						they			np	pronoun they	The more they kill, the more points they are rewarded with.
14	On the other hand if			the player is killed		Contingency clause	he or she			np+np	pronoun he beneficiary	On the other hand, if the player is killed he or she gets a second chance.
15							Our children		inclusive pronoun	np	pronoun noun goal	Our children are not taught that in real life there is no second chance.
16							what		question	the	interrogative attribute	What kind of video game is your child playing now?

## EBacc plans are unfair on the arts

## D. Appendix

Textual	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant	
1				As teachers and researchers who work at Goldsmiths, University of London,	self mention		role	we		Inclusive pronoun	np	senser	As teachers and researchers who work at Goldsmiths, University of London, we are deeply concerned about the Coalition Government's proposals to downgrade the creative arts and design in the secondary schools curriculum.
2								The Ebacc proposals			np	sayer	The Ebacc proposals discriminate unfairly against students with ability in the visual and creative arts.
3								why		question	interrogative	attribute	Why should the thousands of students with talent in these areas be denied the opportunity to develop and gain confidence in their abilities at school?
4								to do so			nonfinite	Carrier	To do so is blatantly discriminatory, and will undermine the state school system's responsibility to promote equality of opportunity.
5								Much evidence			np	sayer	Much evidence suggests that students who study creative arts do better at more traditional academic subjects such as mathematics and computer science.
6								Many of our competitors, including Singapore and China, where core academic subjects have had absolute pre-eminence in education, are now introducing the creative arts into their curricula.		Inclusive pronoun	np+app	actor	Many of our competitors, including Singapore and China, where core academic subjects have had absolute pre-eminence in education, are now introducing the creative arts into their curricula.
7				is			verb	it		question	NP	Carrier	Is it wise to abandon a curriculum in which academic and creative subjects are integrated, just at the point when our competitors have decided to copy it?
8								The new Ebacc			np	actor	The new Ebacc will cause long-term economic harm to Britain.
9								Marginalising the creative arts in the school curriculum will decimate the next generations of architects, designers, filmmakers, theatre-makers and composers.	attitude		nonfinite	actor	Marginalising the creative arts in the school curriculum will decimate the next generations of architects, designers, filmmakers, theatre-makers and composers.
10								Creativity			np	Carrier	Creativity is good for the economy, and intelligent governments should not seek to stifle it.
11 finally				quite apart from the economic arguments,	hedge		Manner	design and the creative arts			np+np	Carrier	Finally, quite apart from the economic arguments, design and the creative arts are simply good for Britain.
12								they			np	actor	They make us interesting and innovative and the envy of the world.
13								why		question	interrogative	attribute	Why would we cut ourselves off from that?

Gay wedding in church? Are you sure?

#	Outer Textual	Inner theme PreHead	Head	Text
---	---------------	---------------------	------	------



## D.6 Annotated Letters to the Editor in English

1										I		self mention		np	pronoun I	senser	I understand why Catholics, with their odd notion that marriage is a sacrament, would wish to refuse to conduct one for same-sex couples, but as an evangelical Christian my concern is less whether I should be compelled to do so than why on earth any same-sex couple would want me to.
2										Church weddings			np	noun noun	goal		Church weddings have to be conducted according to the "rites and ceremonies" of the church in question.
3										it	Contingency clause		np	pronoun it	carrier		Jumping over the inconvenience that no churches have or have ever had any rite or ceremony for such circumstances, it is without question that in any evangelical or reformed church a wedding includes a sermon on a topic of the minister's choice.
4										my sermon	Contingency Clause	self mention	np	det noun	verbiage		Were I ever to be obliged by the civil authorities to conduct such a wedding, my sermon would certainly include a forthright exposition of the Bible's teaching on sodomy and the evils of laws made by godless governments.
5	if									I	Contingency Clause	self mention	np	pronoun I	senser		If any homosexual couple and their guests were willing to sit through that for the 45 minutes or more of my average-length sermon, I suppose they would have earned the right to the piece of paper that entitles them to the very few and steadily diminishing privileges still conferred by the legal state of marriage in this country.

How did prankst get through?

#	Outer Textual	Inner theme					Head			Text				
		Stance	Element	Reader	Stance	Pre-Head	Syntactic Realization	Element	Stance	Reader	Syntactic Realization	Participant		
1								The tragic outcome of the Australian prank	attitude		np+app	actor	The tragic outcome of the Australian prank has spawned a deluge of fire and condemnation which now verges on a tasteless parade of moral indignation.	
2								The agony of the nurse who has ended her life, and of her family,	attitude		np+app	carrier	The agony of the nurse who has ended her life, and of her family, isn't respected by the endless churning of opinion and interviews.	

## D. Appendix

3	moreover					attitude	time	time-pp	there		shared knowledge	np	there	Existent	Moreover, in the midst of the furore about the broadcast there is no focus on the role of the hospital in ensuring the privacy of its patients.
4									Experience			np	noun	actor	Experience must have taught the management that there will always be attempts to breach privacy, and perpetrate pranks, not least when there is a royal patient in the building.
5									The key question	booster		np	det adj noun	token	The key question is what protocols were in place to ensure that all inquiries were handled with skill and discretion.
6									it			np	pronoun	Actor	It should not have been possible for a duty nurse to think that she had to deal with an incoming call about a royal patient.
7									Statements about supporting the nurse after the event			np+app	np+nonfinite	Carrier	Statements about supporting the nurse after the event are simply ducking the question of why she found herself in that position in the first place.
8									A question that some patients			np+app	np+pp	sayer	A question that some patients might well be asking also.

Limit access to guns but in wont end massacres

#	Outer Textual	Inner theme										Text			
		PreHead					Head								
		Reader	Stance	Element	Stance	Reader	Syntactic Realization	Contingency	Element	Stance	Reader	Syntactic Realization	Participant		
1			Sta	It would be naive			hedge		Contingency-Nonfinite			np	adj noun	carrier	It would be naive to assume that even if strict gun controls could be enforced, US schools would become safe overnight.
2															
3	yet											np	pronoun we	carrier	Yet we had another massacre in 2010 by a licensed firearms holder, who managed to kill 12 people and injure 11 with a shotgun and a bolt-action .22 rifle.
4							booster	place	place-clause	booster		np+app	np+pp	carrier	Where there are guns of any description available to the public (and I can't see us ever getting rid of sporting guns), the opportunity for mass shootings by mentally ill people remains.
5	Even if						hedge	Contingency	Contingency-clause			np	pronoun they	actor	Even if they can't get guns, they can still mow people down with their vehicle, as happened here earlier this year.

7						A much tougher challenge	booster	np	det adj adj noun	carrier	A much tougher challenge is for our social systems to reliably identify and contain potential homicidal maniacs before they can do catastrophic harm by any means at their disposal.
---	--	--	--	--	--	--------------------------	---------	----	---------------------	---------	--

Poor penalised for phoning HMRC

#	Outer Textual	Inner theme PreHead					Head				Text
		Element	Stance	Reader	Syntactic Realization	Pre-Head	Stance	Reader	Syntactic Realization	Participant	
1	Textual										Your front-page report "The great HMRC telephone rip-off" (18 December) was right to point out the high cost to HMRC customers of premium rate 0845 numbers.
2	but										But this practice goes on throughout government.
3											My research has shown that 148 of the Department for Work and Pensions' helplines – or 87 per cent – are 0845.
4											The most vulnerable, poor and distressed people have to call these numbers to claim, change or enquire about sickness and disability benefits, carers' support, jobs, pensions, child support and even crisis loans.
5	Well	Mod			Degree-pp						Well over 30 million people call just seven of the DWP's most-used helplines each year but many give up on getting through after being kept on hold and charged.
6	in one example										In one example, more than one in three calls to the employment and support allowance helpline are abandoned before they are answered, but on average over five minutes after they have been connected.
7											Callers to the incapacity benefit reassessment line wait nearly 13 minutes without being answered before they hang up.
8	That means	Eng									That means some people who can ill afford it are being charged more than £5 without even getting through.
9											The DWP and HMRC tell us they don't profit from having 0845 numbers, but someone is making money from these calls.
10	If					it's the Government,	hedg				If it's the Government, it's immoral.
11	If					it's only the telephone company, then	hedg				If it's only the telephone company, then it's a very bad deal for taxpayers.
12											Calls to such public service lines should be free.

Settle this Plebgate stand-off now

#	Outer	Inner theme	Text
---	-------	-------------	------

#### D. Appendix

1									it				np	pronoun it	Carrier	It is no one's interest (bar politicians and the Police Federation) for our governing political party to be swinging hatchets with the Metropolitan Police.
2									The public				NP	det noun	Carrier	The public need to have faith that police officers remain politically impartial and do not lie, and that politicians abide by the laws they create.
3									The "Plebgate" affair, which (in policing terms) amounts to little more than a routine police encounter with an angry and frustrated member of the public, has spiralled out of proportion, and is diverting too much political effort and attention away from the more pressing issues affecting our nation.	booster			np+app e	np+claus e	actor	The "Plebgate" affair, which (in policing terms) amounts to little more than a routine police encounter with an angry and frustrated member of the public, has spiralled out of proportion, and is diverting too much political effort and attention away from the more pressing issues affecting our nation.
4									we				np	pronoun we	Actor	We should not have to accept senior government politicians swearing at police, any more than they should accept any hint of serving police officers undermining or bullying politicians.
5									The stand-off that has now developed				np+app knowledge	np+claus e	actor	The stand-off that has now developed undermines the sacrosanct divide between politics and criminal justice, and goes well beyond what can be dealt with by "trial by proxy" in the press.
6									the Director of Public Prosecutions to draw a line under the affair, by charging Mr Mitchell, based on what evidence is available, and co-charging any officer who it is believed may have lied in the course of their duty.				np+app knowledge e	np+nonfinite	Actor	It is time for the Director of Public Prosecutions to draw a line under the affair, by charging Mr Mitchell, based on what evidence is available, and co-charging any officer who it is believed may have lied in the course of their duty.
7									The whole thing				np	det adj noun	goal	The whole thing should then be rolled up into a single trial, and the court can decide who lied and who has committed an offence.

The right to be a flintlock musket

#	Outer Textual	Inner theme						Text		
		PreHead			Head					
	Textual	Element	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Syntactic Realizati	Participant
1	I find it astonishing that	I find it astonishing that	Stia						np	senser
2	When	this amendment was introduced,			time	time-clause was	the only firearms available	shared knowledge	np	goal
3							Rifling such weapons	booster	nonfinite	actor

4						shared knowledge	time	time-clause	it		np	pronoun it	Carrier	In the environment of the day, when in the vastness of the new West no one was ever far from a wolf, mountain lion or unfriendly native American, it was obviously considered reasonable that people should be entitled to have and carry one of these weapons.
5	but								to construe the right to bear arms as an absolute when the nature of weapons has changed so much	shared knowledge	nonfinite	nonfinite	Carrier	But to construe the right to bear arms as an absolute when the nature of weapons has changed so much is just ridiculous.
6		Eng	It is interesting				place	place-pp	this curious interpretation	attitude	np	det adj noun	carrier	It is interesting in the USA that this curious interpretation exists alongside fundamentalist interpretations of religious books which are even more obsolete than the Constitution.
7	When					shared knowledge	time	time-clause	you	inclusive pronoun	np	pronoun you	actor	When the hardware and knowledge changes you have to change the rules.

Free labour to stack shelves

#	Outer Textual	Inner theme PreHead					Head				Text				
		Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance		Reader	Participant		
1									Iain Duncan Smith and many right-wing populists			Proper Noun+NP	senser	Iain Duncan Smith and many right-wing populists miss the point about graduates thinking they are "too good to stack shelves".	
2	When				I was at school and university			Time	I	self reference		NP	Pronoun	Actor	When I was at school and university I spent most holidays doing jobs like delivering milk, working on building and road construction sites and labouring in paper mills.
3									I	self reference		NP	Pronoun	Actor	I was doing it because I wanted to – mostly outdoor jobs and hard physical exercise, and I was being relatively well paid – not to qualify for job-seekers benefit.
4	In fact								It			NP	Pronoun	Actor	In fact it was several years into working as a highly qualified postdoctoral researcher before I was paid more per week than for my last summer job on a new power station construction site.
5									It			NP	Pronoun	Carrier	It is nothing to do with not wanting to do a job you are "too good for" – it is why you are doing it and the rewards that matter.
6									most Tories			NP	Adj Noun	Carrier	Just another example to add to the growing mountain of evidence that most Tories do not have an earthly of what life is like for most people.

## D. Appendix

Supermarkets must take blame for horse meat

#	Outer Textual	Inner theme PreHead							Head			Text		
		Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic Realization	Element	Stance	Reader	Participant		
	Textual													
1					So far			Time-Advvp	the supermarkets		NP	Det Noun	Actor	So far the supermarkets have managed to avoid much of the blame for the horse meat fiasco, casting themselves in the light of well-meaning victims of organised crime.
2									we		NP	Pronoun	Actor	We should not let them get away with it.
3									They		NP	Pronoun	Carrier	They are responsible for the food they sell, in the same way that an airline is responsible for passenger safety.
4					imagine			verb	imagine		Verb		process	Imagine an airline claiming that an air disaster was caused by a faulty component having been used by a rogue maintenance company.
5									That industry		NP	Det Noun	actor	That industry has achieved astonishing safety levels through a massive investment in the control of its supply chain.
6									Our supermarkets		NP	Det Noun	actor	Our supermarkets on the other hand prefer to invest in customer tracking and loyalty schemes, with the goal of improving margins.
7									They		NP	Pronoun	actor	They squeeze every last drop out of their suppliers' margins while failing to require and verify the application of rigorous standards of quality and traceability.
8	It is time for	Eng		directive					our supermarkets		NP	Det Noun	actor	It is time for our supermarkets to take responsibility for the products they are selling, and to build appropriate quality into their supply chains.
9									The good news in all of this		NP+App	NP+pp	carrier	The good news in all of this is for all those great local businesses and their suppliers who make a point of buying and selling locally with full traceability of their products.

Water companies call the shots

#	Outer		Inner theme					Head					Text		
	Textual		PreHead	Element	Stance	Reader	Syntactic Realization	Manner-Advb	Element	Stance	Reader	Participant			
1								Manner-Advb	another batch of companies, this time water suppliers				actor	So another batch of companies, this time water suppliers, has been using highly creative accounting and offshore subsidiaries to avoid tax (report, 15 February).	
2								verb	there		question	There	Existent	Is there no end to the list of companies engaging in these highly dubious practices?	
3									This			Pronoun	Phenomenon	This reminds me of a recent report in my local newspaper that the EU is to force German water companies to privatise.	
4	Yes, apparently	Com							an edict			NP	Det Noun	actor	Yes, apparently an edict has arrived from Brussels stipulating that privatisation should move ahead despite well-founded worries that investors' demands for returns will both drive prices up and cut funding for infrastructure and maintenance.
5			Now				Time		we		inclusive pronoun	NP	actor	Now we can add tricky tax accounting to the list of the disadvantages of privatisation.	
6			directive						we			NP	Pronoun	Carrier	We have to ask ourselves what the democratic mechanism behind all this really is?
7															When decided that privatisation is the main motivation?

## D.6 Annotated Letters to the Editor in English

8																		Interrogative Pronoun	Attribute	Who decided that we should not permit local authorities to run their own water service?
9									This									Pronoun	Actor	This bears the hallmark of the investment lobby which, looking for copper-bottomed deals, wants to park its cash in publicly underwritten investments such as water-works.
10	And							for me	self mention	Evidence	Evidence-pp	this						Pronoun	Actor	And for me this all goes far beyond the tax-avoidance issue, with economic life now seeming to have become a farcical game of monopoly, with the finance industry and corporations calling the shots and with both Brussels and national governments happily dancing to their tune.
11								Sta	directive			we						Pronoun	Carrier	For the sake of democracy and common sense we urgently need to call the lobbyists to heel and hold their masters to account.

Why Scotland should go it alone

#	Outer Textual	Inner theme										Head		Text						
		PreHead										Element	Stance	Reader	Participant	Reader	Stance	Participant	Text	
1												rheme							None	Your leader on 12 February "raises a query as to the point of independence".
2								Sta				how						Interrogative Pronoun	attribute	So kind to worry about how we poor dears will manage without you to tell us how to do everything.
3												is							process	Or is your concern perhaps the other way round?
4								Com				here						Noun	carrier	Whatever, here are few "points of independence" to mull over.
5									hedge		Contingency-NonFinite	it						Pronoun	Actor	Instead of Europe's largest nuclear weapons store being in the Clyde, it will reside, close to those who actually want the obscenity, beside an English river.
6												how						Interrogative Pronoun	attribute	How about the Thames?
7										Manner-NonFinite	Manner-NonFinite	we						Pronoun	Actor	Relieved from the burden of financing the monstrosity we shall be able properly to finance worthwhile projects such as a decent road and sensible rail system.
8												you						Pronoun	Goal	You are welcome to HS2.
9												You						Pronoun	Carrier	You shall have all your NHS privatisers back to reinforce those currently devastating English hospitals for the benefit of the USA and English shareholders of American culture medical companies.
10												Education						Noun	carrier	Education will remain an investment by our people in our future, with the threat removed of a Goveian machine turning it into a commodity to be bought and sold like parsnips, as in your nation of shopkeepers.
11												An end						Det Noun	Actor	An end will be put to the BBC banging on about Richard III "being

## D. Appendix

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Contingency- Clause	Element	Stance	Reader	Participant
1		Undoubtedly	Mod							Richard III		NP	Carrier
2										He		NP	Actor
3	But if				he had not murdered the princes in the Tower,	hedge		Contingency		why		question itive Pronoun	attribute itive Pronoun
4										The Richard III Society		NP	Actor

The Government must have known about the damaging culture in the National Health Service

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Contingency- Clause	Element	Stance	Reader	Participant
1										Jeremy Hunt, the Health Secretary,		NP+App Noun+NP	Actor
2					are		question verb			we		NP	senser
3					is		question verb			it		NP	Carrier

After the EU speech and gay marriage, what next for Cameron?

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Contingency- Clause	Element	Stance	Reader	Participant
1										Conservatives with whatever size "c"	attitude	NP+App NP+PP	Senser
2	But									this		NP	Carrier
3										rheme		rheme	None

If doctors and nurses can be struck off, why are managers immune?

#	Outer Textual	Inner theme PreHead						Head			Text		
		Stance	Reader	Stance	Pre-Head	Stance	Reader	Syntactic Realization	Contingency- Clause	Element	Stance	Reader	Participant
1										Conservatives with whatever size "c"	attitude	NP+App NP+PP	Senser
2	But									this		NP	Carrier
3										rheme		rheme	None



1				Throughout industry and business,					those whose methods lead to death or injury	attitude		question	NP+App interrogative	NP+clause interrogative	Goal	Throughout industry and business, those whose methods lead to death or injury are routinely prosecuted.
2									why						attribute	Why have no charges been brought against those responsible for practices in Mid Staffs, and Stafford Hospital in particular, even though their actions resulted in the deaths and suffering of hundreds of patients?
3									appropriate prosecutions				Np	Adj Noun	Actor	It may be that appropriate prosecutions would result in managers and frontline staff serving well-deserved prison sentences.
4									they				NP	Pronoun	Beneficiary	Instead, they have been promoted, retired with generous pensions, are still in their posts, or have moved to similar posts elsewhere.
5									justice				NP	Noun	actor	Let justice be done.
6									The victims, their families, and those in the NHS who truly care for their patients	attitude			NP+App NP+clause	NP+clause NP+clause	senser	The victims, their families, and those in the NHS who truly care for their patients deserve nothing less.

We don't want mangey foxes anywhere near u

#	Outer	Inner theme				Head				Text			
	Textual	Pre-Head	Reader	Stance	Element	Pre-Head	Reader	Stance	Element	Stance	Reader	Participant	
1									I	self mention		Pronoun	I abhor the spread of foxes, but I think the Toxocara risk from them and from dogs is exaggerated (Letters, February 13).
2									I	self mention		Pronoun	I can find no case of blindness caused by Toxocara recorded in the last 10 years.

Tory U-turn on inheritance tax pledge puts pressure on George Osborne

#	Outer	Inner theme				Head				Text			
	Textual	Pre-Head	Reader	Stance	Element	Pre-Head	Reader	Stance	Element	Stance	Reader	Participant	
1									George Osborne's pledge to raise the inheritance tax threshold to £1 million			NP+App NP+NonFinite	In 2007, George Osborne's pledge to raise the inheritance tax threshold to £1 million was greeted by a seven-point rise in the polls for the Tories.
2									Britain's aspirational classes			Proper Noun Noun Adj senser	Britain's aspirational classes appreciated the prospect of improving the lot of themselves and their children without having their pockets picked by the Government.
3									Some			Pronoun	Some called it a political masterstroke at the time.

D. Appendix

4	it is time	Eng	directive	Now that the proposed trebling of the limit has been replaced by a real-terms decrease of roughly 20 per cent between now and 2019,			Time	Time-Clause	to reassess				NonFinite	NonFinite	NonFinite carrier	Now that the proposed trebling of the limit has been replaced by a real-terms decrease of roughly 20 per cent between now and 2019, it is time to reassess Mr Osborne's aptitude for the post he has been filling.
5				With borrowing higher even than Labour planned, with growth almost non-existent, and with a tax code that penalises job creation and enterprise,			Manner	Manner-NonFinite	he				NP	Pronoun	Goal	With borrowing higher even than Labour planned, with growth almost non-existent, and with a tax code that penalises job creation and enterprise, he must be replaced.

## **D.7    Annotated Letters to the Editor in Spanish**

## D. Appendix

### LETTERS TO THE EDITOR SPANISH

¿Cómo están ustedes?

#	Outer				Inner theme				Head				Participants				Text
	Textual	Interpersonal	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Verb	Element	Stance	Reader	Syntactic	Verb	1p sing Pres Ind	Process	
1					No sé					1p sing Pres Ind	self mention						No sé, ustedes, pero yo me encuentro enriquecido y nostálgico.
2										La muerte de Miliki, más que la de una persona —eso también, por supuesto—, ha significado para mí, y supongo que para millones de jóvenes adultos de mi generación, la muerte de un pedazo de nuestra infancia.							La muerte de Miliki, más que la de una persona —eso también, por supuesto—, ha significado para mí, y supongo que para millones de jóvenes adultos de mi generación, la muerte de un pedazo de nuestra infancia.
3										Recuerdos preciados	attitude						Recuerdos preciados que hoy han vuelto y que el tiempo ha convertido en nostalgia, dolorosa nostalgia.
4	Luego, sin embargo,				entre los amigos al comentar la noticia han ido apareciendo			Time	Place-PP+NonFinite+Verb	3p plu PRT. Perfecto Compuesto							Luego, sin embargo, entre los amigos al comentar la noticia han ido apareciendo recuerdos compartidos que poco a poco nos han hecho sonreír, como no podía ser de otra manera al tratarse del gran Miliki, nuestro queridísimo Miliki.
5		Hasta	Com		en estos momentos			Time	Time-PP	nos							Hasta en estos momentos nos ha hecho sonreír, con ese extraño don que tenía para ello.
6										A Miliki							A Miliki se le harán tributos por su trabajo, seguramente todos enfocados a su arte como payaso, sus virtudes como compositor y músico, y también su faceta humana y su trabajo a favor de la infancia.
7	Pero									yo	self mention						Pero yo quiero añadir mi pequeño tributo a Miliki como maestro de la comunicación.
8					En un tiempo en el que no había Twitter ni Facebook, no estaban de moda conceptos como hashtag o SocialTV, Miliki logró que millones de niños contestaran a la tele: ¡Bleeeen!			Time	Time-Clause	Miliki							En un tiempo en el que no había Twitter ni Facebook, no estaban de moda conceptos como hashtag o SocialTV, Miliki logró que millones de niños contestaran a la tele: ¡Bleeeen!
9										Quién		question					¿Quién lo supera?

Contra la amnesia

#	Outer		Inner theme					Head					Participants					Text
	Textual	Interpersonal	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Goal	Carrier						
1							La Ley Hipotecaria			NP	Det Noun			La Ley Hipotecaria se reformó en los años noventa, creo que durante el Gobierno socialista de Felipe González.				
2			Hasta entonces,			Time	Time-PP	se	shared knowledge	NP	Pronoun se			Hasta entonces, no se podía hipotecar un bien si uno no era propietario del mismo.				
3							La modificación			NP	Det Noun			La modificación consistió en que a partir de esa fecha, fue posible hipotecar un bien en el momento mismo de su adquisición ante el notario.				
4							Eso			NP	Pronoun eso			Eso hizo que una cantidad enorme de españoles que habían vivido privados toda su vida, pudieran adquirir una vivienda propia.				
5			Desde entonces,			Time	Time-PP	la frase más escuchada	booster	NP+App	NP+AdvP			Desde entonces, la frase más escuchada era la de que alquilar era "tirar el dinero".				
6							La demanda de pisos y casas para comprar			NP+App	NP+NP			La demanda de pisos y casas para comprar fue creciendo vertiginosamente y los precios de las mismas multiplicándose cada año, pero creo sinceramente que a nadie le amenazaron para que comprara una casa solicitando una hipoteca.				
7	Y						me	self mention		NP	Pronoun me			Y me consta que los trámites para conseguirla no eran fáciles ni rápidos: avalés, nóminas, otras propiedades, etcétera.				
8			En cuanto a			Matter	Matter-PP	los bancos y los banqueros		NP	Det Noun con Det Noun			En cuanto a los bancos y los banqueros, han pasado de ser envidiados y adulados, a villanos.				
9	Parece que	Sta	hemos olvidado			Verb	1p plu PRT. Perfecto Compuesto			Verb	1p plu PRT. Perfecto Compuesto			Parece que hemos olvidado que su negocio siempre ha sido pagar poco por el dinero que les confiamos y cobrar mucho por el dinero que les pedimos prestado.				
10	Pero	pese a todo, Mod	prácticamente			Degree	Degree-Advb todos			NP	Pronoun todos			Pero pese a todo, prácticamente todos seguimos utilizando sus servicios, porque nadie quiere arriesgarse a guardar el dinero de su nómina o de su pensión en un cajón, ni hacer cosas interminables para pagar los innumerables recibos de todo tipo que generamos; y preferimos seguir utilizando las tarjetas de débito y de crédito para casi todo.				

11	ahora más que nunca	com	es				verb	verb	es	directive	Verb	3p sing Pres Ind	Process	Ahora más que nunca es necesario reformar la Ley Hipotecaria por otra más justa y controlar lo que los bancos nos cobran por sus servicios.
----	---------------------	-----	----	--	--	--	------	------	----	-----------	------	------------------	---------	---

Cooperación Internacional

#	Outer		Inner theme				Head		Text			
	Textual	Interpersonal	Pre-Head	Stance	Reader	Stance	Verb	Element	Stance	Reader	Syntactic	Participants
1			Leo				Verb	1 sing plnd		self mention	Verb	Process
2			Esta vez,				Time-NP	lo			NP	Actor
3	O sea	que con la está cayendo		Eng				se			NP	Actor
4			Afirmo			booster	Verb	-o		self mention	Verb	Process
5		Se puede afirmar que		Eng				el dinero facilitado			NP	Goal
6	Pues							ningún informe			NP	Actor
7								Me		self mention	NP	Senser
8			Hasta el momento solo son		shared knowledge		Time-PP	3 p plu Plnd			Verb	Process
9	Y							esto			NP	Carrier

euro por receta

#	Outer		Inner theme				Head		Text			
	Textual	Interpersonal	Pre-Head	Stance	Reader	Stance	Syntactic	Element	Stance	Reader	Syntactic	Participants
1								El Gobierno de la Comunidad de Madrid			NP	Sayer
2	Supongo que	Sta	se toman este tipo de medidas				Time-clause	hay			Verb	Existent
3	Pero en este caso considero	Sta	desde mi profesión de médico			self mention	Role	están dejando			Verb	Process
4								Esa forma de actuar			NP+App	Actor
5								Una patología de atención primaria			NP+NP	Carrier
6	Se sabe que	Eng	shared knowledge					los ingresos hospitalarios			NP	Carrier
7								¿Nadie		question	NP	Carrier
8								¿Nadie es consciente de esto?			None	None

La violencia en Gaza

#	Outer		Inner theme				Head		Text			
	Textual	Interpersonal	Pre-Head	Stance	Reader	Stance	Syntactic	Element	Stance	Reader	Syntactic	Participants
1								Gaza			NP	Carrier
2								¿Nadie		question	None	None
3								Con aquello que limita tu vida, que te marca y clasifica en una causa que nadie sabe.			None	None

## D. Appendix

4						Da	booster		Verb				3p sing Pres Ind	Verb		Process	Da lo mismo la fe que tengas ni en lo que creas si otro puede taparse la cabeza en vez del cuello.
5						Da	booster		Verb				3p sing Pres Ind	Verb		Process	Da lo mismo lo que significa la vida si otro puede dejarse patillas o rezar sentido.
6						Me ahi cuando	booster		Time-Clause				3p sing Pres Ind	Verb	shared knowledge	Process	Me ahi cuando se rompe con la ética de la coexistencia haciendo de esto más que un conflicto territorial, uno cultural y que, como se ha venido dando en los últimos siglos, no tiene ni tendrá solución alguna.
7	Pero					Hay			Verb-Hay				3p sing Pres Ind	Verb	shared knowledge	Existent	Pero hay algo de cada bandera que se ha quedado en el camino, esa ausencia de color pacífica que pinta a ambos géneros y que representa lo mismo en ambos territorios: la gente, gente que se cansa del dolor.
8													Eso	NP		Token	Eso es lo único que puede llevar al fin de todo esto; a la, hasta ahora, inabarcable paz en medio Oriente: el cambio de paradigma.
9													Un cambio de contratos sociales, no legales, que puedan llegar a reivindicar el concepto de vida en vez de acabar con ella.	NP+NP		Sayer	

La preocupación del ciudadano catalán

#	Outer Textual	Interpersonal	Inner theme	Head	Element	Stance	Reader	Syntactic	Stance	Reader	Syntactic	Participants	Text
1												Carrier	Parece ser que todos los catalanes debemos posicionarnos a favor o en contra de la independencia, o lo que es lo mismo, a favor del Gobierno de la Generalitat o contra él, como si no importase nada más.
2												Carrier	Creo que el ciudadano de a pie está mucho más preocupado por llegar a fin de mes, por no perder su trabajo, por dar una buena educación a sus hijos y por tantas otras cosas que no están alejadas de esa clase política que se cree por encima del bien y del mal.
3												Process	Lo importante ahora es sacar el país adelante, no dejar que el talento joven se vaya consiguiendo crear puestos de trabajo.
4												Process	Lo importante en realidad es parar los desahucios, jugar a los corruptos, evitar más recortes, recuperar aquel espíritu del 15-M que reunió a toda clase de personas en las principales plazas de España.
5												Actor	Un espíritu que consiguió unir a todos por una causa común, con pancartas que rezaban "Si en Barcelona no tienen miedo, a Madrid no tendrán por".
6												Process	Puede parecer nostalgia.
7												Carrier	No lo es.
8												Process	Se trata de un sentimiento reivindicativo de derechos, lejos de secundar cualquier idea política o sindical.

Las instituciones siguen copadas por los varones

#	Outer Textual	Interpersonal	Inner theme	Head	Element	Stance	Reader	Syntactic	Stance	Reader	Syntactic	Participants	Text
1												Senser	Una vez más, los académicos de la RAE no consideran la posibilidad de que una mujer sea candidata a ocupar un sillón.
2												None	Es, en este caso.
3												Process	Es fácil, sin embargo, encontrar periodistas, fotógrafos, cineastas, filósofos, juristas, historiadoras, novelistas, lingüistas, traductoras, médicas, científicas, arquitectas, etc. etc. con sobrados méritos.
4												Process	Resulta curioso comprobar cómo, desde 1982 —año en el que se licenciaron tantas mujeres como hombres— los únicos ámbitos profesionales que reflejan esa realidad son aquellos a los que se accede por estricta oposición (no un híbrido de oposición y pacto entre amiguetes como a menudo ocurre en la universidad).
5	Por el contrario,											Carrier	Por el contrario, las instituciones a las que se accede por cooptación siguen copadas por varones (en la RAE son 42 hombres y 4 mujeres).
6	Ellos defienden su "neutralidad".											Actor	Ellos defienden su "neutralidad".
7	Es decir, que, para empezar,											Process	Es decir, que, para empezar, no leyeron a Barthes y no han oído hablar del mundo "natural" de la ideología dominante.
8	Pero, además,											Process	Pero, además, debe ocurrirles una de estas cosas:
9												None	Que se crean de verdad que a nosotros se nos va toda la ciencia, el saber y la inteligencia cuando acabamos la carrera (antes no, porque los expedientes académicos prueban lo contrario).

10												Rheme			None		None	Que, en el fondo, las mujeres les sigamos pareciendo unas squafetas y piensen que con admitir a dos o tres por institución ya que han cumplido.
11												Rheme			None		None	Que no tengan ni idea de quién trabaja, investiga, crea, vale en los últimos campos.
12												Mode- Advb+Verb			Verb		Phenomenon	No sabría decir cuál de estas opciones me inquieta más.

Otras políticas sí son posibles

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Verb	Head	Element	Stance	Reader	Syntactic	Verb	Participants	Text
1											comprendo						Comprendo perfectamente a los miles de españoles que el pasado 14 de noviembre se han movilizado en España y han secundado la huelga general convocada por los dos sindicatos mayoritarios en contra de los recortes sociales.
2											La gente					Carrier	La gente está indignada y, lo que es peor, no ve ninguna esperanza.
3											Las previsiones de los organismos internacionales —tanto del FMI como del Banco Central Europeo— para nuestro país echan por tierra las previsiones optimistas del señor de Guindos y de la ministra de Empleo Editha Báñez.					Actor	Las previsiones de los organismos internacionales —tanto del FMI como del Banco Central Europeo— para nuestro país echan por tierra las previsiones optimistas del señor de Guindos y de la ministra de Empleo Editha Báñez.
4			Mod								Opalá se equivocan						Opalá se equivocan, pero mucho me temo que no va a ser así.
5											El PIB					Actor	El PIB caerá más de lo previsto y el número de desempleados puede llegar a superar los set millones.
6											3 p sing PInd						Con austeridad sólo no basta.
7	Mientras										Manner- Clause						Mientras no crezca la economía no se va a crear un solo puesto de trabajo en España.
8											Time						Todas las medidas anunciadas por el Gobierno exclusivamente a reducir el déficit público, abultado déficit público, nadie lo discute, pero muy pocas por no decir ninguna a relanzar el crecimiento económico.
9		Creo desde mi modesta opinión que	Sta								es compatible aplicar medidas dirigidas a reducir el déficit público						Creo desde mi modesta opinión que es compatible aplicar medidas dirigidas a reducir el déficit público como ya se está haciendo, eso sí con un calendario más flexible para cumplirlo, con políticas de crecimiento.
10											Las políticas de austeridad que nos marca Bruselas						Las políticas de austeridad que nos marca Bruselas nos están empobreciendo a muchas familias.
11											El próximo país en caer						El próximo país en caer puede ser Francia, que ya ha dado tasas negativas de crecimiento.
12		Hay que estar ciego para	eng								Manner- NonFinite						Hay que estar ciego para no darse cuenta de que esto nos conduce al abismo y a un punto sin retorno.

Quiénes son los morosos

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Verb	Head	Element	Stance	Reader	Syntactic	Verb	Participants	Text
1		Es interesante leer	eng								noticias						Es interesante leer noticias que afirman que la morosidad de la banca marca un nuevo récord y pasa del 10,7%.
2	Pero										mi pregunta						Pero mi pregunta es la siguiente: ¿quiénes son los "morosos"?
3											¿Aquellas personas sin empleo y sin vivienda						¿Aquellas personas sin empleo y sin vivienda que no pueden pagar sus deudas por culpa de un Gobierno, unas empresas y una banca; o aquellas culpables de la crisis y que se están lucrando de unos recortes e impuestos sin precedentes a las clases medias?
4		Posiblemente	Mod								los "morosos" de los que estas noticias hablan						Posiblemente los "morosos" de los que estas noticias hablan sean personas sin empleo debido a que nuestros políticos están más preocupados en hacer frente al déficit que en crear puestos de trabajo.
5	Y										estos "morosos"						Y sin empleo, posiblemente estos "morosos" sean personas sin casa.
6											Una vivienda que los propios bancos						Una vivienda que los propios bancos les han quitado.
7	Quizá		Mod								sp sing Cond						Quizá habría menos morosos si los culpables de la crisis dejaran de mirar tanto por sus intereses y se preocuparan, aunque solo fuera un poco, por la clase trabajadora, que no es que no quiera pagar, es que no puede.

Un dierre incomprendible

## D. Appendix

#	# Outer		Inner theme						Head						Participants		Text
	Textual	Interpersonal	Pre-Head	Reader	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Verb	Process				
1		Me quedo estupefacto al saber que, prácticamente sin aviso previo, cerrar el Instituto de Cardiología, del que soy (lera) el atorinado paciente.	cierran				Verb	cierran						Me quedo estupefacto al saber que, prácticamente sin aviso previo, cierran el Instituto de Cardiología, del que soy (lera) el atorinado paciente.			
2		Nunca podrá quedar			booster		Time	Time-Advb-Verb				Verb	Process	Nunca podrá quedar tan en evidencia la falta de sentido común como en este caso.			
3								Cerrar un centro, que ha conseguido, después de muchos años y con mucho esfuerzo y dedicación de su personal	attitude			NP+App	Carrier	Cerrar un centro que ha conseguido, después de muchos años y con mucho esfuerzo y dedicación de su personal, convertirse en una referencia en cardiología que atiende a muchos pacientes a los que aporta más que la vida: la calidad de vida.			
4								Dónde				Interrogative Pronoun	Attribute	¿Dónde iremos, los pacientes, para encontrar la misma atención, la misma relación de confianza —tan importante como la medicación para nuestra salud— la misma competencia y trato cariñoso de sus enfermeras y enfermeros y de todo el personal, aunque muchos estén en situación precaria?			
5		Me entero de que						este magnifico y eficaz servicio				NP	Actor	Me entero de que este magnifico y eficaz servicio cuesta 8,7 millones de euros al año.			
6								yo	self mention			NP	Senser	Yo invito encarecidamente a los responsables a que pasen revista a todo lo que, con nuestro dinero, se está costeando y que sale por más de 8,7 millones.			
7	si	A ver			hedge		Contingency if	Contingency-les				NP	Actor	A ver si la vergüenza les hace recapacitar y deciden cortar por otro lado menos fundamental para todos.			
8	si	Que no crean que			hedge		Contingency clause	Contingency-nos				NP	Actor	Que no crean que, si siguen adelante, nos limitaremos a quejas por escrito.			
9			Mantengan				Verb-Imperative	3p plu imp				Verb	Process	[Mantengan el Instituto de Cardiología			

el arte no entiende de razas ni de sexos

#	# Outer		Inner theme						Head				Participants				Text
	Textual	Interpersonal	Pre-Head	Reader	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Pronoun me	Sayer				
1								me	self mention				NP		Me dirijo a usted para mostrarle mi indignación por una entrevista que se publicó el domingo 8 de febrero en la sección de Cultura al cantaor flamenco Manuel Aguiñeta.		
2								Su frase más resaltada		shared knowledge	NP+App	NP+NonFinite	Token		Su frase más resaltada es la de que los blancos no pueden hacer jazz porque el jazz pertenece a la filosofía de los negros, al igual que el flamenco no lo pueden hacer más que gitanos, porque supuestamente el flamenco les «pertenece».		
3			Soy				Verb	1p sing PInd	self mention		Verb	1p sing Pres Ind	Process		Soy músico, estudiante concretamente en un conservatorio de Madrid, llevo más de 10 años tocando la trompeta (un instrumento usado habitualmente en el jazz), soy blanco, he tocado en una big band de jazz, tengo amigos músicos blancos que tocan jazz y conozco casos de músicos no negros que tocan jazz estupendamente como el trompetista Arturo Sandoval, al igual que conozco casos de músicos negros que tocan jazz y clásico de manera extraordinaria como el también trompetista Wynton Marsalis.		
4								¿Mi pregunta		question	NP	Det Noun	Token		¿Mi pregunta es la siguiente: si un compositor blanco de música culta aparece en una entrevista diciendo que los negros no pueden tocar música clásica porque todos los compositores son blancos, ¿ustedes publicarían dicha entrevista?		
5	0		peor aún, si dijese que una mujer no puede tocar música clásica porque todas las obras de la historia conocidas las han hecho hombres, ¿creen		self mention	contingency clause	contingency clause	3p plu PInd		question	Verb	3p plu Pres Ind	Process		O peor aún, si dijese que una mujer no puede tocar música clásica porque todas las obras de la historia conocidas las han hecho hombres, ¿creen ustedes que quedaría impune ante la opinión pública?		
6								El arte			NP	Det Noun	Senser		El arte no entiende de razas, culturas, sexos...		
7								Cualquier persona con capacidad expresiva y	booster		NP+App	NP+PP Conj PP	Actor		Cualquier persona con capacidad expresiva y creativa puede hacer una obra maestra del tipo que sea.		



El fracaso del independentismo

#	Outer				Inner theme				Head				Text			
	Textual	Interpersonal	Element	Stance	Reader	PreHead	Stance	Reader	Syntactic	Contingency	Element	Stance	Reader	Syntactic	Participants	Text
1											me	self mention		NP	Senser	Me gustara felicitar a los ciudadanos catalanes de los municipios que este fin de semana han tenido referéndum, y por que no, al resto, por ignorar esas consultas.
2	Creo que	Sta				por fin			Time		la mayoría de los catalanes	self mention		NP+App	Actor	Creo que por fin la mayoría de los catalanes está desemmascarando al independentismo.
3	Espero que	Sta		attitude		sigan			Verb		sigan	self mention		Verb	Actor	Espero que sigan así cuando llegue la sentencia del TC y las elecciones.
4						digo			Verb			self mention		Verb	Senser	Digo esto porque se confirma algo que ya sabíamos, aunque el nacionalismo lleva 35 años intentando enredar: que los lazos entre catalanes y españoles son fuertes, como corresponde a unos pueblos que llevan 2.000 años viviendo juntos y más de 500 siendo una misma nación.
5	Si					fueran débiles, ¿no			Contingency	question	se		question	NP	Actor	Si fueran débiles, ¿no se habrían separado después de machacar en ese sentido durante 40 años?
6	Espero que	Sta				no,			Contingency		los nacionalistas	boost		NP	Senser	Espero que los nacionalistas entiendan este mensaje y rectifiquen.
7						se			Contingency		se extinguirán	boost		NP	Goal	Si no, se extinguirán como los dinosaurios.
8	Pero					quienes insisten en defender una Cataluña independiente			Contingency		quienes insisten en defender una Cataluña independiente	boost		NP+App	Senser	Pro quienes insisten en defender una Cataluña independiente deben saber que ningún país oviado renuncia a su territorio histórico y legítimo. Y España no va a ser menos. Europa no va a apoyar jamás la autodeterminación de ninguna región de sus Estados miembros, muchos españoles no vanos a renunciar jamás a Cataluña, y no por enroqueamiento, sino por amor.
9											Los que han pensado que con el odio a lo español iban a provocar odio a lo catalán han fracasado	boost		Clause	Actor	Los que han pensado que con el odio a lo español iban a provocar odio a lo catalán han fracasado

El santuario de ETA y las FARC

#	Outer				Inner theme				Head				Text			
	Textual	Interpersonal	Element	Stance	Reader	PreHead	Stance	Reader	Syntactic	Contingency	Element	Stance	Reader	Syntactic	Participants	Text
1											La Audiencia Nacional			NP	Actor	La Audiencia Nacional ha procesado a varios miembros de ETA y de las FARC por colaborar para intentar atentar en España contra dos presidentes colombianos.
2						Al mismo tiempo indica			Time-PP+Verb		3p sing PInd			Verb	Process	Al mismo tiempo, indica que hay indicios de la «cooperación» del Gobierno venezolano en la alianza de estos grupos de terroristas y narcoterroristas.
3						Con esta acusación			Manner-PP		Chávez ex colapista y compadre del dictador Castro.	attitude		NP+App	Carrier	Con esta acusación, Chávez ex colapista y compadre del dictador Castro, ya tiene un motivo más de preocupación.
4						Gracias a la Justicia			Contingency		su verdadera imagen	attitude		NP	Carrier	Gracias a la Justicia, su verdadera imagen está siendo redescubierta por el resto del mundo.

Explorar la desgracia en televisión

#	Outer				Inner theme				Head				Text			
	Textual	Interpersonal	Element	Stance	Reader	PreHead	Stance	Reader	Syntactic	Contingency	Element	Stance	Reader	Syntactic	Participants	Text
1						Como voluntaria de una self asociación que todas las semanas ayuda a sin techo en las calles de Madrid y que realmente vive su realidad de cerca, me indigna que se utilice la miseria ajena para obtener audiencia y promocionar a pseudofamosos.			Role		me	self mention		NP	Senser	Como voluntaria de una asociación que todas las semanas ayuda a sin techo en las calles de Madrid y que realmente vive su realidad de cerca, me indigna que se utilice la miseria ajena para obtener audiencia y promocionar a pseudofamosos.
2											El reality show invisibles			NP	Token	El reality show invisibles es vergonzoso.
3						¿Cuántas			Interrogative		¿Cuántas historias personales de gente sin hogar escuche esta gente famosa a la semana?	question		NP	Attribute	¿Cuántas historias personales de gente sin hogar escuche esta gente famosa a la semana?
4											Rheme			None	None	Las suficientes para irse a casa a dormir tranquilo pensando que les van a pagar por hacer de indigente unas horas.
5	Que dicen de	Eng				quieren enseñar			Verb		3p plu Imp			Verb	None	Que dicen de decir que quieren enseñar la realidad social con estos programas.
6											Los sin techo			NP	Carrier	Los sin techo son para ellos pura carnsa.

## D. Appendix

7	En lugar de perder el tiempo	Eng	directive	participando o viendo programas como éste	directive	Manner-NonFinite	3p sing Cond		directive	Verb	3p sing Cond Ind	Process	En lugar de perder el tiempo participando o viendo programas como éste, convendría emprender acciones de verdadero voluntariado.
8							Yo		self mention	NP	Pronoun yo	Senser	No conozco la realidad de los sin techo y sé lo que les tiene que pasar. Me gustaría aprovechar de su condición para sacarlos minutos de audiencia en la tele.
9							Este programa			NP	Det Noun	Carrier	Este programa es un atentado a la ética, a la justicia y a la solidaridad.

La conciliación por encima del cargo

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Participants	Text
1									MI reconocimiento a Joaquín Manuel Montero, concejal del PSOE y segundo teniente de alcalde de Paradas (Sevilla)	self mention		NP+App	Actor	MI reconocimiento a Joaquín Manuel Montero, concejal del PSOE y segundo teniente de alcalde de Paradas (Sevilla), que solicitó su baja como millante el mismo día que se aprobó la nueva Ley del Aborto.
2		Me alegro ver que	Sta		hay	verb		Place-PP	hay	self mention		Verb	Existent	Me alegro ver que todavía hay personas con principios dentro del partido socialista.
3									En la carta que ha enviado Montero a Rodríguez Zapatero y a la secretaria de Organización Pajín alega que «jamás permitirá que su nombre aparezca junto al de una organización que legitima la muerte de inocentes, mediante la aprobación de leyes injustas».			3p sing Pres Ind	Process	En la carta que ha enviado Montero a Rodríguez Zapatero y a la secretaria de Organización Pajín alega que «jamás permitirá que su nombre aparezca junto al de una organización que legitima la muerte de inocentes, mediante la aprobación de leyes injustas».
4									Este gesto			NP	Token	Este gesto ha sido el mejor tributo que ha podido ofrecer a la defensa de la vida.
5	es que		Eng					Other-Clause	Montero, a diferencia de otros			Verb	Actor	Y es que Montero, a diferencia de otros, ha demostrado ser una persona decente que coloca su conciencia por encima de todo tipo de poder, cargo o fama.

la doble moral de la casta política

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Participants	Text
1									Se quitaron			Verb	Process	Se quitaron la careta el president Montilla y su esposa al decir que prefieren que sus hijos se eduquen en el Colegio Alemán, porque de esa forma tendrán seguro un trabajo en cuanto finalicen los estudios.
2					No es		shared knowledge	Verb	3p sing PInd			Verb	Process	No es el único dirigente político que condena a la sociedad a dedicar tiempo y recursos a aprender lenguas poco útiles pero al mismo tiempo preparan a sus hijos para el día de mañana.
3									Líderes como Pía (PSOE) en Valencia			NP+App	Actor	Líderes como Pía (PSOE) en Valencia hace lo mismo, y hasta algunos dirigentes de Añalar llevan a su prole en verano a Oxford.
4									Las Autonomías			NP	Carrier	Las Autonomías se han convertido en un negocio para los que dominan el castellano y la lengua autonómica.
5									Los que solamente hablamos castellano de las lenguas oficiales del Estado (mi caso, aunque domino inglés, alemán y francés)			Clause	Actor	Los que solamente hablamos castellano de las lenguas oficiales del Estado (mi caso, aunque domino inglés, alemán y francés) no podemos siquiera optar a un puesto público.
6	Pero								para quien realmente esto supone un gran negocio			Verb	Process	Pero para quien realmente esto supone un gran negocio es para la clase dirigente, que al mismo tiempo que distrae a la sociedad con dialécticas dialécticas, alecciona a sus cachorros para que les sucedan como pastores de un rebaño que en el fondo menosprecian.

La rebelión global de la naturaleza

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Participants	Text
1		Como todo el mundo	Eng						he advertido			Verb	Process	Como todo el mundo, he advertido que el clima ha cambiado bruscamente en poco tiempo.
2									El año pasado sentimos el frío hasta mayo y tras un verano prolongado casi hasta diciembre y varias nevadas desde entonces, padecemos lluvias históricas y temporales.			Verb	Process	El año pasado sentimos el frío hasta mayo y tras un verano prolongado casi hasta diciembre y varias nevadas desde entonces, padecemos lluvias históricas y temporales.
3									La fuerza de la naturaleza			NP+NP	Actor	La fuerza de la naturaleza ha dejado tragedias como la de Haití o Chile.

Las televisiones. un paso por detrás

protesta del embajador israelí

Willly Toledo y la muerte de Zapata

#	Outer		Inner Theme				Head				Participants				Text
	Textual	Textual	Interpersonal	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Carrier	
1										Guillermo Toledo, tan buen cómico como incoherente intolerante,	attitude		NP+App	Carrier	Guillermo Toledo, tan buen cómico como incoherente intolerante, debía tener la conciencia intranquila por no haberse manifestado en las últimas horas.
2										El actor, que parece pertenecer a una especie que sufre si está más de 48 horas sin manifestarse o firmar un manifiesto,	attitude		NP+Clause	Carrier	El actor, que parece pertenecer a una especie que sufre si está más de 48 horas sin manifestarse o firmar un manifiesto, no debía soportar saber que toda la sociedad estaba esperando su sentencia sobre el caso Zapata.
3										El actor			NP	Carrier	El actor estuvo acompañando a la saharaui Aminatu Haidar durante su huelga de hambre, y ha manifestado que mientras la saharaui «es una defensora y luchadora de los derechos humanos, pacífica, que ha cumplido años y años de agresiones y torturas», Zapata era «un delincuente común, si siquiera era un disidente político», coincidiendo con la teoría del diario Gramma y el régimen castrista.

[illegible]

Con todo respeto ¡esto es ridículo!

#	Outer		Inner theme						Head				Text	
	Textual	Interpersonal	Pre-Head	Reader	Stance	Reader	Syntactic	Verb	Element	Stance	Reader	Syntactic	Process	Participants
1	Textual	Element						Verb	1p sing Pnd	self mention		Verb	1p sing Pres Ind	
2									en su punto seis			Verb	3p sing Pres Ind	Sayer
3	Pero								esto			NP	Pronoun esto	Carrier
4			imaginar		hedge			Verb	1p sing Pnd	self mention		Verb	1p sing Pres Ind	Process
5		Con todos los respetos.							esto			NP	Pronoun esto	Carrier

## Contra el cambio de programa de Mas

#	Outer			Inner theme							Text	
	Textual	Element	Stance	Reader	Pre-Head	self mention	Stance	Reader	Syntactic	1p sing Pres Ind	Participants	Process
1					Como residente en Cataluña y, además, volante de la formación política que lidera el muy honorable Artur Mas rechazar							Como residente en Cataluña y, además, volante de la formación política que lidera el muy honorable Artur Mas, rechazo algunos de los acuerdos que previamente adoptará en materia político-fiscal el nuevo Gobierno surgido de las anticipadas elecciones autonómicas catalanas.
2		En concreto	Mod					sobre la recuperación del impuesto de sucesiones que, tras su cesión por el Estado a las comunidades autónomas	PP+App	PP+clause	Goal	En concreto, sobre la recuperación del impuesto de sucesiones que, tras su cesión por el Estado a las comunidades autónomas, fue suprimido por muchas de ellas y, entre otras, por la de Cataluña, según decisión del muy honorable Mas, dando cumplimiento a lo dispuesto en el programa electoral de su anterior legislatura y del que curiosamente ha sido beneficiado ya que, en marzo de 2012 (es decir, en el breve periodo de tiempo en que el impuesto estuvo suprimido), se causó a su favor una herencia por la que no tributaré, lo que de ninguna manera quiere decir que la supresión del impuesto fuera hecha simili modo berlusconiano, esto es, sin tratarse de una simple coincidencia cronológica.
3								La probable reinstauración del impuesto sobre sucesiones exigida por el nuevo socio político del muy honorable Mas	NonFinite	NonFinite	Actor	La probable reinstauración del impuesto sobre sucesiones exigida por el nuevo socio político del muy honorable Mas no hará sino reindicar en el agravio comparativo entre los españoles de distinta vecindad civil —y político—, fomentar la inmigración como protección del ahorro familiar para los hijos y disminuir en futuras elecciones la cuota de voto del muy honorable por deslealtad con su programa.
4		Es una falacia decir que	Eng					el impuesto de sucesiones	NonFinite	NonFinite	Actor	Es una falacia decir que el impuesto de sucesiones guardará a las claves altas (que, como es notorio, lo eluden por extraterritorialidad o por la constitución de sociedades patrimoniales), cuando realmente su sujeto pasivo lo constituye la clase media, clientela habitual del partido en el Gobierno.

Es preciso bucear otros fondos

[illegible]

1						booster	Time	Time-Clouse	se			NP	pronoun se	Carrier	En medio del inmenso dolor que aflige a los familiares de las víctimas inocentes de la masacre de la localidad norteamericana de Newtown se sitúa el ansia de toda una sociedad por encontrar una explicación racional que permita recuperar un mínimo consuelo.
2									se			NP	pronoun se	Actor	Se ha vuelto a abrir el debate sobre la necesidad de controlar con rigor el comercio libre de armas de fuego, aunque todo el mundo sea consciente de que el problema no reside tanto en la facilidad para comprarlas como en las personas que las manejan.
3						booster	Manner	Manner- Advb	basta achacar			Verb	3p sing Pres Ind	Process	Tampoco basta achacar este tipo de tragedias repetidas a la locura pasajera de asesinos como Adam Lanza.
4			Eng			bucuar	Verb		bucuar			Verb	3p sing Pres Ind	Process	Es preciso bucar también otros fondos que no siempre se quieren analizar con el rigor que merecen y que tienen mucho que ver con el relativismo y el individualismo exacerbado.

## Invertir en ciencia

#	Outer		Inner clause						Main clause				Text		
	Textual	Interpersonal	Pre-Head	Head	Post-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic		Participant	
1												NP+App	Det Noun+pp	Actor	La noticia sobre dos científicas españolas en Oxford (EL PAÍS, 25 de diciembre) ilustra perfectamente los problemas de la ciencia en España.
2												Verb	3p plu Pres Ind	Process	Cuentan su estancia en un centro de investigación creado en 2005 mediante la donación de un acaudalado ciudadano, James Martin, que pasará a la posteridad por este gesto.
3												Verb	3p plu Pres Ind	Process	En el mundo anglosajón es muy habitual hallar este tipo de mecenas, mientras que en el nuestro es extremadamente raro.
4												NP	Pronoun me	Senser	Me viene a la memoria el proyecto de Botín para construir un centro "de las artes y la cultura" en primera línea de mar (terreno público, por cierto), con arquitecto famoso pero sin objetivo claro ("conferencias, exposiciones...").
5	A pesar de que	se trata de un propósito muy loable										NP+App	NP+Clause	Actor	A pesar de que se trata de un propósito muy loable, este dinero (y el de muchos otros que podrían darse por aludidos) se podría invertir en un centro, por ejemplo para la investigación oncológica, tierra adentro y bien feo, para que el diseño no pase por delante del contenido.
6	Y los siete millones	(siete) de euros que costará su mantenimiento anual										NP+App	NP+Clause	Actor	Y los siete millones (siete) de euros que costará su mantenimiento anual podrían dedicarse a recuperar talentos, que los tenemos, como las dos doctoras de que hablaba al principio.
7												Verb	3p sing Contd Ind	Process	Sería el primer paso para evitar el desmantelamiento de la ciencia española, que conozco de primera mano como médico y profesor universitario, supliendo el impulso público, últimamente de capa caída.
8	A ver	si alguien se anima										NP	Pronoun se	Senser	A ver si alguien se anima, que en España sigue habiendo quien haga holgadamente a fin de mes y es una forma magnífica de que se recuerde su nombre durante siglos.

## La lotería y la salida de la crisis

#	# Outer		Inner theme										Head		Text			
	Textual	Interpersonal	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Contingency clause	Element	Stance	Reader	Syntactic	Pronoun se	Participants		
1	si					con 2.520 millones repartidos en la Lotería de Navidad gran parte del país ha salido a la calle a brindar con cava por ver resueltos sus problemas,	hedge				Contingency+ se		question	NP		Senser	Si con 2.520 millones repartidos en la Lotería de Navidad gran parte del país ha salido a la calle a brindar con cava por ver resueltos sus problemas, ¿se imaginan cuantos agujeros se podrían haber tapado con la lotería de los 60.000 millones íntegros que han tocado a las entidades financieras para salvarlas?	
2											se		question		NP		Senser	¿Se imaginan cuantos miles de familias podrían haber evitado entrar en mora por esa maldita cantidad de tal vez 300 o 500 euros que faltaban para pagar?
3						podrán			verb	verb	podrán		question	verb	Epi plu Pres Ind	Process	¿Podrán seguir durmiendo los responsables del robo mayúsculo de tantos miles de millones, cuando sabiendo que con tan solo 2.500 millones podrían haber sido felices tantas familias?	

## D. Appendix

4	por desgracia	Mod			Y para salir de esta pronto será	booster		Contingency	Contingency-PP+Verb	será		verb	3p sing Futuro ind	Process	Por desgracia y para salir de esta, pronto será más fácil que te toque la lotería que encontrar personas decentes que administren nuestras finanzas.
---	---------------	-----	--	--	----------------------------------	---------	--	-------------	---------------------	------	--	------	--------------------	---------	--

La nueva Constitución Egipcia

#	Outer Textual	Interpersonal Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Det Noun Adj	Participants	Text
1										El Egipto democrático		directive	NP		Actor	El Egipto democrático debe nacer respetando la diversidad, religiosa, política y étnica, apostando por la modernidad.
2										La nueva Constitución		directive	NP		Senser	La nueva Constitución debe inspirarse en el consenso y el acuerdo entre el Gobierno y la oposición.
3										El presidente Mohamed Morsi y los Hermanos Musulmanes			NP+NP		Actor	El presidente Mohamed Morsi y los Hermanos Musulmanes se juegan una nueva partida en este nuevo Egipto que debe afrontar los nuevos cambios sin complejos, haciendo una apuesta clara por separar la religión de la vida institucional y contando con las demás fuerzas políticas a través del diálogo como herramienta necesaria para garantizar la supervivencia de la democracia y las conquistas de revolución de la plaza Tahrir que derrocó a Mubarak.
4										Esta coyuntura			NP		Carrier	Esta coyuntura es un nuevo examen para las conquistas democráticas en Egipto.
5										Una nueva Constitución referendada en las urnas		directive	NP+App		Carrier	Una nueva Constitución referendada en las urnas debe ser el reflejo de todo el país, teniendo en cuenta sus minorías y mayorías, consolidando las libertades y acabando con cualquier tipo de mecanismo y poder absoluto.
6										La Constitución egipcia		directive	NP		Actor	La Constitución egipcia debe servir como modelo a otros países que aspiran a la democracia en el mundo árabe, de su éxito o fracaso dependerá el de toda la región.

Las causas del fracaso escolar

#	Outer			Inner theme										Text
	Textual	Interpersonal	Pre-Head	Stance	Reader	Syntactic	verb	Element	Stance	Reader	Syntactic	verb	Process	
1			cuesta			verb	cuesta				3p sing Pres Ind	Process	Cuesta mucho desde el aula difundir las causas y las soluciones del fracaso escolar vigente.	
2	Quando		docentes curtidors así lo intentan			time	Time-Clause	se			Pronoun se	Actor	Cuando docentes curtidors así lo intentan se impone por desgracia la versión platónica y alejada de teóricos que poco o nada pisan el aula.	
3			El pasado 30 de noviembre			time	Time-NP	ASPEPC-SPS			Proper Noun	Actor	El pasado 30 de noviembre ASPEPC-SPS inició en el Illustre Col·legi de Doctors i Llicenciats de Catalunya las I Jornadas de Secundaria sobre el fracaso escolar.	
4								Los acuerdos alcanzados entre todos los asistentes			NP+App	Actor	Los acuerdos alcanzados entre todos los asistentes mostraron una clara y dramática conclusión: la nueva ley de educación, la LOMCE, está diseñada por teóricos que viven a años luz de los docentes que sí salvan, arreglan y remiendan el actual fracaso escolar.	
5			En gran parte			place	place-Pp	este desacuerdo educativo			Det Noun Adj	Carrier	En gran parte este desastre educativo radica en leyes educativas anteriores, defectuosas y llenas de teorías nunca comprobadas.	
6	de hecho	Com	hace pocas semanas			time	time-pp	la OCDE			Proper Noun	Sayer	De hecho hace pocas semanas que la OCDE refregó a España ostentar el máximo fracaso escolar de Europa.	
7	Se hace obvio que	Eng						los docentes			Det Noun	Carrier	Se hace obvio que los docentes siguen siendo ignorados a pesar de los acuerdos alcanzados durante las Jornadas de Secundaria.	
8						Other- question	¿Que por qué				Interrogative Pronoun	Attribute	¿Que por qué actualmente nuestros estudiantes suspenden más?	
9	pues simplemente porque	Eng					estudian				3p plu Pres Ind	Process	Pues simplemente porque estudian menos.	
10							cómo				Interrogative Pronoun	Attribute	¿Cómo resolver entonces el fracaso escolar vigente?	
11	Primero		hay			verb	verb-hay	directive			verb-hay	Existent	Primero hay que evitar los pedagogos y teóricos que alejados de las aulas redactaron una LOGSE, una LOCE, una LOE y ahora una LOMCE que hacen imposible el correcto desarrollo de la actividad docente.	
12							es	booster			3p sing Pres Ind	Process	Es muy importante el ambiente de orden, silencio y concentración para facilitar la memorización y la comprensión de conceptos; las rutinas en clase y en casa de trabajo, estudio y descanso; los maestros con excelentes conocimientos en su especialidad y con un dominio rico, elegante y preciso de los idiomas oficiales; y, por último, mucho esfuerzo aprendizaje, enseñado y adquirido ya desde primaria...	

Por una educación para los alumnos

### Recortes también en residencias

#	Outer			Inner theme									Text		
	Textual	Interpersonal		Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic		Participants	
1									La Comunidad Autónoma de Murcia y en su nombre el IMAS			NP+Np	Proper Noun + Proper Noun	Sayer	La Comunidad Autónoma de Murcia y en su nombre el IMAS; ha equiparado a los ancianos de una residencia con los funcionarios.
2									me	self mention		np	pronoun me	sayer	Me explico.
3									A todos los residentes			pp	Prep Det Det Noun	Goal	A todos los residentes, hasta ahora se les retenía alrededor del 90% de su pensión.
4									Esta retención			NP	Det Noun	Carrier	Esta retención, era para ayudar de en los gastos de su alojamiento y manutención, el resto se supone que lo sumaba el IMAS, como ayuda social, con excepción de las pagas extras de verano y Navidad, de las que podían disponer en su totalidad.
5	Pues bien								los señores que nos gobiernan a nivel autonómico,			NP+App	Inclusive pronoun NP+clause	Senser	Pues bien, los señores que nos gobiernan a nivel autonómico, han decidido que a partir de ahora (han empezado en diciembre de 2012), ya no dispondrán de este pequeño ingreso semestral.
6									Estas pagas			NP	Det Noun	Carrier	Estas pagas eran empleadas, en la mayoría de los casos para pagarse el seguro de defunción, la peluquería, algún calzado, el podólogo, gafas, etc.
7									se		Interjection	NP	Pronoun se	senser	A partir de ahora que se aguanten y dejen de derrochar (tienen bastante con la pequeña cantidad que mencionaba al principio).
8		es lamentable que	Eng						se			NP	Pronoun se	Actor	Es lamentable que se celebren con el colectivo más débil y lo han hecho con hostilidad y alejosa, sin informar previamente ni a residentes ni a familiares.

## D. Appendix

Un trato vejatorio

#	Outer Textual	Interpersonal Element	Inner theme Pre-head	Reader	Stance	Head time-pp	Element	Stance	Reader question	Syntactic NP	Participants Actor	Text
1			En estos días navideños				se			Interrogative Pronoun	Actor	En estos días navideños se realizan atentados encubiertos y rastros en los bajos de una ciudad llamada Madrid.
2							rheme			Interrogative Pronoun	Attribute	¿Cómo?
3							rheme			rheme	None	A través de sus líneas del metro.
4							rheme			rheme	None	Dirigidos (como siempre) a la sufrida clase media.
5							rheme			rheme	None	A través de sus líneas del metro.
6	son incongruentes a todas luces	Eng					aquella			NP+App	Actor	Aquella que paga religiosamente sus impuestos.
7			Una vez más			time-pp	se			NP	Actor	Son incongruentes (a todas luces) los sistemas empleados para reivindicar derechos, que se van diluyendo, gracias a la poca capacidad organizativa y de previsión de sus líderes.
8							qué			Interrogative Pronoun	Attribute	Una vez más, se pasan por el forro las consecuencias de esas huelgas mediáticas que, no solucionando nada en sus reivindicaciones, si producirán un trato vejatorio e injusto a esa masa impersonal que por el momento no les interesa minar.
9							rheme			None	None	¿Qué les importa tener encerrados como ratas a cientos de personas durante más de cuarenta minutos?
										None	None	Decididamente, nada.

¿Balleas y balleos?

#	Outer Textual	Interpersonal Element	Inner theme Pre-head	Reader	Stance	Head	Element	Stance	Reader	Syntactic Interrogative Pronoun	Participants Attribute	Text
1							por qué			Interrogative Pronoun	Attribute	¿Por qué se empeñan nuestros representantes políticos y sindicales en hablar siempre de ciudadanos y ciudadanas, catalanes y catalanas, vascos y vascas, trabajadores y trabajadoras, remarcando así una diferencia de sexo innecesaria?
2							La repetición obsesiva de la alternancia masculino y femenino	attitude		NP+App	carrier	La repetición obsesiva de la alternancia masculino y femenino es, a mi oído, inútil, un latigallo molesto, mimetismo absurdo fruto de la poca preparación de nuestros representantes, pues creo, sin miedo a equivocarme, que a menudo es usado como recurso para preparar la siguiente frase cuando escasan los recursos lingüísticos.
3							Alguien		Inclusive pronoun	NP	sayer	Alguien debería decirles que los ciudadanos ya sabemos que cuando se refieren a ciudadanos, catalanes, vascos y trabajadoras, todos los sexos nos sentimos representados, pues se trata de un masculino de género epiceno, no marcado. Y que esta distinción es artificial, que no resuelve problemas de sexismo, sino que lo crea y agrava.
4							rheme		question	rheme	None	¿Balleas y balleos?

GUERRA DE IMPUESTOS

#	Outer Textual	Interpersonal Element	Inner theme Pre-head	Reader	Stance	Head	Element	Stance	Reader	Syntactic NP	Participants carrier	Text
1							La huida fiscal de Gérard Depardieu a Bélgica			NP+PP+PP	carrier	La huida fiscal de Gérard Depardieu a Bélgica fue el principio del conflicto.
2			Desde entonces,			time-pp	Paris y Bruselas			NP+NP	actor	Desde entonces, París y Bruselas intercambian golpes bajos a modo de impuestos contra las bebidas nacionales del vecino.
3	Primero		fue			verb	fue		shared knowledge	verb	Process	Primero fue Francia, que aumentó un 160% el gravamen de bebidas espirituosas y cervezas, y la más perjudicada es la cerveza belga que producen las brasseries.
4							La decisión			NP	carrier	La decisión supuso un ataque frontal a los productores belgas que provocó levantamientos contra la decisión francesa, Y también en España, por los 6.000 empleos directos que generan las cervezas de brasserie y los casi 70.000 empleos indirectos.
5			Molestos por el impuesto francés,			Manner	los belgas		shared knowledge	NP	Actor	Molestos por el impuesto francés, los belgas aprobaron una subida general de la tributación sobre bebidas alcohólicas excepto, claro está, sobre la cerveza. Aun no se sabe a ciencia cierta el impacto en precios sobre el vino francés, pero sí que los belgas consumen unos 280 millones de botellas y el 74% es de origen galo.



6						se	np	pronoun se	senser	No se descarta que continúe la guerra fiscal y se carguen nuevos y excesivos tributos sobre el chocolate belga o el foie gras.
7	al final					Contingency pp	Contingency pp	boost	verb	Al final, por simple cercanía física e interacción indirecta, sufrirá consecuencias el vino español.

Debarian sentir miedo

#	Outer			Inner Theme			Interpersonal						Process						Text
	Textual	Stance	Reader	Pre-head	Stance	Reader	Syntactic	verb	Head	Element	Stance	Reader	Syntactic	verb	Participants				
1				He llegado			verb		He llegado					verb	Process	He llegado a la conclusión que los poderosos que obran mal lo hacen porque no tienen miedo.			
2				saben			verb		saben					verb	Process	Saben que nos les pasará nada de nada.			
3				Repasando la prensa			time-nonFinite		una de las noticias que debería llamar la atención				NP+App		Actor	Repasando la prensa, una de las noticias que debería llamar la atención es el juicio que se está realizando en Islandia a los llamados vikingos financieros.			
4									La lista				NP		Carrier	La lista está compuesta por políticos, banqueros y empresarios islandeses.			
5				en lo sucesivo			time		tendrán				Verb		Process	Se supone que en lo sucesivo tendrán miedo antes de acometer negligentemente cualquier actitud, ya sea política o comercial, o sea, que el miedo hará guardar la vida.			
6	si			hacemos un pequeño resumen de todo lo acontecido últimamente, veremos			Contingency	Other-if	veremos				inclusive pronoun	Verb	Process	Si hacemos un pequeño resumen de todo lo acontecido últimamente, veremos que aquí de miedo, nada de nada, o no hay vida o lo tienen todo atado y bien atado.			
7				construyen			verb		construyen				shared knowledge	verb	Process	Construyen AVE minoritarios, aeropuertos sin pasajeros. Falifican empresas, muchos apartamentos son ridos preaviciados.			
8	en cambio								los políticos					NP	actor	Y en cambio, los políticos recorran y recorran en sanidad, enseñanza y asistencia social.			
9									los culpables					NP	carrier	los culpables son los otros.			
10	y mientras								el pueblo					NP	carrier	Y mientras el pueblo se ve cada día más asfixado, ellos son siempre los mismos, esos que se ven envueltos cada día en toda clase de tejemánjes.			
11				Solo una pregunta a políticos, banqueros y empresarios:			question	time-question	cuándo				question	Interrogative Pronoun	attribute	Sólo una pregunta a políticos, banqueros y empresarios: ¿cuándo pasará como en Islandia, que el miedo hará guardar la vida?			
12									este					NP	carrier	Este sería un buen punto de partida, que los políticos tuvieran miedo a...			

El frustrado puede matar

#	Outer			Inner theme			Head			Participants			Text
	Textual	Interpersonal	Pre-Head	Reader	Stance	Pre-Head	Reader	Syntactic	Element	Stance	Reader	Syntactic	
1								other-pp	Rheme			Det pronoun	Sólo una puntualización a todo lo que se ha publicado sobre la terrible matanza en la escuela Sandy Hook, de Newtown, a cargo de un joven de 20 años.
2								Place-pp-Verb	hemos hablado			1p plu Pft. Perfecto Compuesto	En los medios de comunicación hemos hablado de la falta de red de entorno no social, de fallos en la estructura matrimonial de los progenitores, del carácter reservado del asesino, pero... ¿cómo le iba su vida escolar?
3	porque							Verb	ovidemos			1p plu Pres Subj	Porque no ovidemos que también hay escuelas de tortura.
4	y quizá	Mod						Verb	hemos insistido			1p plu Pft. Perfecto Compuesto	Y quizá no hemos insistido demasiado en un detalle que como profesionales de la salud mental creo que tendríamos que destacar: la frustración continuada engendra agresividad, que, tarde o temprano, se transforma en violencia extrema.
5									frustración			Noun	Frustración que se da en las escuelas, ya que ellas han sido el marco explotario de la mayoría de los crímenes de esta brutal naturaleza acontecidos en los últimos años.
6	y								por qué			Interrogative pronoun	¿Y por qué mayoritariamente estos jóvenes vuelcan su violencia en las escuelas y no en las canchas deportivas o en los supermercados, por ejemplo?

[illegible]

midg nación social

#	Outer	Inner theme										Text			
		Textual		Interpersonal		Pre-Head		Head		Participants					
		Textual	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Det noun	Goal
1		Cada vez es mayor la indignación en relación a	Eng	Eng	shared knowledge					time-clause		los recortes	Np		Cada vez es mayor la indignación en relación a los recortes que se están produciendo en la sanidad, la educación, la justicia, la vivienda y las pensiones, sin olvidarnos de la sangría del paro, de los despidos colectivos, de los desahucios, de las preferentes, etcétera.
2										se			Np	pronoun se	se augura un 2013 cargado de protestas y movilizaciones sociales por la defensa de los derechos fundamentales y la dignidad de las personas.
3		ya está bien	Eng	Eng	interjection	se trata de culpabilizar a los mismos de siempre, a los inocentes que nada tienen que ver con la crisis, y		shared knowledge	Contingency	Contingency-clause		inclusive pronoun	NP	Pronoun nos	Ya está bien de que se trate de culpabilizar a los mismos de siempre, a los inocentes que nada tienen que ver con la crisis, y nos olvidemos de los verdaderos responsables: los especuladores, corruptos, responsables financieros y políticos de cualquier signo, que continúan con el discurso que más les interesa, tratando de responsabilizar de todo ello a las familias más humildes.
4										la indignación social			NP	Det adj noun	La indignación social está más que justificada cuando vemos que, a costa de la mayoría de ciudadanos, viven unos pocos privilegiados que son los primeros en aplaudir las políticas de recortes sociales y la pérdida de derechos adquiridos por parte de la clase media para que ellos puedan vivir al margen de la realidad, en urbanizaciones de lujo, en mansiones y con ingresos que superan los seis mil euros mensuales.

Lista de defraudadores

#	Order	Inner theme										Text					
		Textual		Interpersonal		Pre-Head											
		Textual	Textual	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	
1											España			directive	NP	proper noun	Carrier
2											El ministro del ramo, Montoro,				NP+App	NP+Proper Noun	sayer
3											Esa lista			directive	NP	Det Noun	Actor

### D.7 Annotated Letters to the Editor in Spanish

	No obstante			con la cantidad de casos de corrupción - una encuesta llevada a cabo en Catalunya muestra que el 80% de los ciudadanos piensa que está viviendo en un país de corruptos, y los casos que cada día saltan a los medios muestran que los ciudadanos no están equivocados- dada entre políticos y gestores de empresas dependientes, directa o indirectamente, de las administraciones, va a hacer poco creíble la medida y crea dudas sobre la eficacia.	Contingency PPI+App	Shared knowledge	Contingency	booster	verb	2p sing Pres Ind	actor	No obstante, con la cantidad de casos de corrupción -una encuesta llevada a cabo en Catalunya muestra que el 80% de los ciudadanos piensa que está viviendo en un país de corruptos, y los casos que cada día saltan a los medios muestran que los ciudadanos no están equivocados- dada entre políticos y gestores de empresas dependientes, directa o indirectamente, de las administraciones, va a hacer poco creíble la medida y crea dudas sobre la eficacia.
4				con la cantidad de casos de corrupción - una encuesta llevada a cabo en Catalunya muestra que el 80% de los ciudadanos piensa que está viviendo en un país de corruptos. Y los casos que cada día saltan a los medios muestran que los ciudadanos no están equivocados- dada entre políticos y gestores de empresas dependientes, directa o indirectamente, de las administraciones								

Rajoy dice la verdad

#	Outer		Inner theme										Text
	Textual	Interpersonal	Pre-Head		Head		Syntactic		Reader		Participant		
	Textual	Element	Stance	Reader	Stance	Reader	Syntactic	time-PP	Element	Stance	Reader	Syntactic	Participant
1				Hace unos días			time	time-PP	Rajoy			NP	sayer
2		lo curioso	Eng				verb	verb	es			verb	Process
3									Rajoy			NP	senser
4				creo			verb	verb	creo	self mention		verb	Process
5				Como lo creo, pues no es un mentiroso, llego			Manner	Manner-PP	llego	self mention		verb	Process
6									esto			NP	carrier

Matanza de inocentes

#	Quota		Interpersonal		Interloc theme						Text					
	Textual	Textual	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	
1											La opinión pública			NP	senser	
2	Si						hedge			Contingency Clause	habríamos actuado		inclusive pronoun	verb	Det noun adj 1p plu Pst. pluscumperfecto Subj	
3	Pero										el hecho	booster		NP	det noun Actor	
4	Sin embargo										nuestro gobierno		inclusive pronoun	NP	det noun actor	
5	Además					aquí			place	place-advb	el día de la muerte de los inocentes, el 28 de diciembre	attitude		NP+App	NP+pp Carrier	

lo todo vale

#	Outer Textual	Inner theme										Text	
		Interpersonal		Pre-Head		Head		Syntactic		Participants			
	Textual	Element	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Syntactic	Participants
1									la democracia			NP	det Noun Actor
													La democracia permite a todo ciudadano la libertad de expresary

## D. Appendix

#	Outer Textual	Interpersonal Element	Inner theme Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	Text
2																El portavoz del partido Ciudadanos, Jordi Cañas, ya nos tiene acostumbrados a su mal talante y a su cara de eterno enfado con todo y con todos –especialmente con todo lo que tiene que ver con Catalunya–, pero ha llevado su dureza (por decirlo de alguna manera) a un límite no aceptable cuando, supongo que ahorrando viejos tiempos y regímenes, ha amenazado al president Mas con la pena a los marineros que se rebelan.
3																Para el que lo desconozca, es la pena de muerte.
4																¿Hasta cuándo la libertad de expresión tiene que hacernos aceptar afirmaciones de este tipo o como las del exaltado de A Coruña comparando los alumnos castellanos en Catalunya con los judíos marcados por el nazismo?
5																Todo no es válido, aunque parece que para algunos políticos, sí.

Preitos con dinero

#	Outer Textual	Interpersonal Element	Inner theme Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	Text
1																Quisiera manifestar la indignación que me ha producido leer que el galardón Premi Internacional Catalunya, que concede anualmente la Generalitat desde 1989 en Barcelona, además de otorgar una escultura, está dotado con 80.000 euros ("Lula llama a no rendirse ante la crisis", Internacional, 14/XI/2012).
2																Con independencia del destinatario del galardón, que este año ha correspondido al expresidente brasileño Lula da Silva, que me parece muy bien merecido, lo sorprendente del caso es que la Generalitat se puede permitir el lujo de regalar un importe como el citado en la precaria situación económica en que se encuentra Catalunya.
3																Opino que si se quiere homenajear a alguien por su valía, no hace falta otorgarle un dinero que hubiera podido servir para cubrir la subsistencia de mucha gente que tiene que mendigar.

Voy a dejar de volar

#	Outer Textual	Interpersonal Element	Inner theme Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	Text
1																Soy un jubilado de 67 años, médico de profesión, que me he pasado la vida trabajando muchas horas para ganarme el pan.
2																Siempre he creído que la corrupción existía como algo humano, pero no como algo muy generalizado.
3																He votado siempre, desde que existe posibilidad de hacerlo.
4																No me esperaba este alud de corrupción que ha ido ahorrando últimamente por parte de todos los partidos políticos mayoritarios en Catalunya y en España.
5																Dada la situación, quiero decir a los políticos en general que o hay una limpieza a fondo, evidente y ejemplarizante de las persona corruptas, sean del partido que sean, o déjase de votar para siempre (al menos no me sentire ni responsable ni engañado).
6																Esta opinión, compartida por mi entorno, puede ir extendiéndose.

Cidistas anolados

#	Outer Textual	Interpersonal Element	Inner theme Pre-Head	Reader	Stance	Pre-Head	Stance	Reader	Syntactic	Head	Element	Stance	Reader	Syntactic	Participants	Text
1																Empieza a ser preocupante la frecuencia de atropellos mortales de ciclistas por parte de automóviles y camiones.
2																Yo también soy ciclista y he estado mediando sobre algunas de las causas que producen estos accidentes.

3						He observado	Verb			Self mention		Verb		Process	He observado que muchos de ellos llevan vestimenta de color negro o muy oscuro, lo cual hace que no destaquen contra el fondo negro del asfalto.
4						Imaginemos	Verb				Inclusive pronoun	Verb		Process	Imaginemos una situación muy frecuente, el coche tiene el parabrisas ligeramente polvoriento o sucio, la luz del sol viene de cara y, al dar en la superficie del parabrisas, provoca que el conductor no vea bien un ciclista que va de oscuro sobre el fondo negro del asfalto o del verde oscuro de la vegetación.
5										question		Interrogative Pronoun		Attribute	¿Que ocurre entonces?
6										question		Rheme		None	Pues un altísimo riesgo de que lo acabe atropellando.
7												Interrogative Pronoun		attribute	¿Cómo se podría, en parte, evitar ese riesgo?
8												Rheme		None	Aplicando la solución que me he impuesto a mí mismo: ponerme un chaleco de esos amarillo reflectante como los que son obligatorios para los conductores que sufren una avería o pinchazo.
9						Tendría que ser	Verb			directive		Verb		Process	Tendría que ser obligatorio para la seguridad de los ciclistas.
10						que canta	Verb					Verb		Process	¿Que canta un poco?
11												Rheme		None	Seguro que si, pero mejor eso que no que te tengan que cantar un responso.
12						Seguro que	Manner-Adverb					Verb		Process	Seguro que se evitarían muchos atropellos.

Todo un Pontífice

#	Outer Textual	Interpersonal	Inner rheme	Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Text
1				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Participants
1				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
1				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
2				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Senser
2				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Senser
3				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
3				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
4				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
4				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	Process
5				Pre-Head	Head	Element	Stance	Reader	Stance	Reader	None

Medidas preventivas

#	Outer		Inner rheme								Text			
	Textual	Interpersonal	Pre-Head	Stance	Reader	PreHead	Stance	Reader	Syntactic	Element		Stance	Reader	Participants
		Textual												
1														La corrupción lastra el progreso económico y social de un Estado, crea incertidumbre en los mercados e inseguridad jurídica.
2														Para frenar esta lastra social, el Gobierno está ofreciendo transparencia y judicialización.
3														Es indiscutible que los culpables de delitos de corrupción y detracción de fondos públicos deben ser perseguidos y condenados.
4														Sin embargo, estas acciones no aseguran el retorno de los fondos públicos derruidos a las arcas públicas.
5														Lo que los ciudadanos no saben y ningún grupo político parece estar interesado en contarla es que en la actualidad un funcionario puede ser disciplinado en un proceso que se lleva a cabo sin garantías procesales y sin la presencia de un abogado.
6														En estas circunstancias, ¿quién se atreve a decir que no? Los funcionarios no pueden convertirse en un ejército disciplinado al servicio de la corrupción política.
7														Urge aprobar una batería de medidas preventivas contra la corrupción que garanticen la independencia de la función pública respecto del poder político.
8														Entre ellas, el desarrollo reglamentario del Estatuto Básico del Empleado Público y la regulación por ley de su régimen disciplinario.

## D. Appendix

Chorizos británicos

#	Outer Textual	Interpersonal				Inner theme				Head				Participants				Text
		Element	Stance	Reader	Pre-Head	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Verb	Process	Process	Process	Process	
1						Dice			Verb	Dice								Dice el Financial Times que "en España no hay bastante pan para tanto chorizo".  Quizá sea verdad, pero me ofende que los británicos se atrevan a leerme la cartilla.  No tienen que mirar tan lejos, en Londres están los cerebros de la operación de fagotitar y vaciar Iberia en beneficio de British Airways y en breve se irán al paro 4.000 trabajadores españoles.  Si tenemos en cuenta que British Airways quebrada y lo estará aún más gracias al desastre de su fondo de pensiones, y que por el contrario, hasta la fusión con British, Iberia estaba sana y era rica, habrá que preguntarse dónde está la caja de Iberia, quién va a pagar las pensiones de los británicos y dónde están los chorizos, porque en Londres no habrá chorizo, pero chorizos hay un montón.
2		Quizás	Mod			sea			Verb	sea								
3						no tienen			Verb	tienen								
4	si					tenemos en cuenta que British Airways estaba quebrada y lo estará aún más gracias al desastre de su fondo de pensiones, y que por el contrario, hasta la fusión con British, Iberia estaba sana y era rica	hedge		Contingency	Habrà	directive		Existent	Existent				

Vigilar a los empresarios

#	Outer Textual	Interpersonal				Inner theme				Head				Participants				Text
		Element	Stance	Reader	Pre-Head	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	NP+App	NP+Proper Noun	Sayer	Sayer	Sayer	
1						Recientemente,			Time	el presidente de la CEOE, Juan Rosell,								Recientemente, el presidente de la CEOE, Juan Rosell, manifestaba que dado que "existe mucho empleo público que en este momento ni tiene competencias ni trabajo por hacer sin nada que hacer gastando bolígrafos y teléfono, saldría más barato para el país mantener a algunos funcionarios en sus casas cobrando subsidios que en sus puestos de trabajo".  Pero estaría bien que a la vista de las últimas noticias sobre supuestos pagos en negro a trabajadores de sus horas extras, evitando así la correspondiente cotización a la Seguridad Social, alguien preguntara al señor Rosell si no sería más conveniente que en lugar de dejar en casa a los funcionarios supuestamente ociosos los puséramos a vigilar a los empresarios supuestamente tramposos, para que cumplan con sus obligaciones para con los trabajadores y para con el resto de la ciudadanía.
2	Pero	estaba bien que	Sta			a la vista de las últimas noticias sobre supuestos pagos en negro a trabajadores de sus horas extras, evitando así la correspondiente cotización a la Seguridad Social			Contingency	alguien			NP	Pronoun	Sayer			

Habrà que regenerar la política

#	Outer Textual	Interpersonal				Inner theme				Head				Participants				Text
		Element	Stance	Reader	Pre-Head	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Verb	Process	Process	Process	Process	
1		Por una vez	Com			estoy			Verb	estoy	self mention							Por una vez estoy de acuerdo con Esperanza Aguirre: es necesaria la regeneración política.  Desde mi punto de vista, habría que actuar de forma contundente sobre tres niveles.  Cese de las personas que ejercen cargos directivos en todos los partidos políticos.  Tampoco podrían acceder a estos puestos aquellos que en algún momento los hayan desempeñado.  Así, la militancia de base tendría el camino libre para ocuparlos.  Imposibilitar presentarse en las próximas elecciones al Congreso, Senado y municipales a los políticos que anteriormente hayan sido electos, incluso también a los que hayan formado parte de listas electorales de cualquier partido político.  Los organismos públicos no podrán contratar a asesores políticos que ya hayan trabajado en la Administración.
2		Desde mi punto de vista	Sta			habría que actuar			Verb	habría								
3						cese			Verb	cese		directive						
4						tampoco podrían acceder	hedge		Manner	podrían acceder								
5						Así,			Manner	la militancia de base								Así, la militancia de base tendría el camino libre para ocuparlos.  Imposibilitar presentarse en las próximas elecciones al Congreso, Senado y municipales a los políticos que anteriormente hayan sido electos, incluso también a los que hayan formado parte de listas electorales de cualquier partido político.  Los organismos públicos no podrán contratar a asesores políticos que ya hayan trabajado en la Administración.
6									Manner	rheme			NP+App	Actor	None			
7										rheme	directive							Los organismos públicos no podrán contratar a asesores políticos que ya hayan trabajado en la Administración.
										Los organismos públicos	directive		NP	Actor				

Tasa Robin Hood

La gala de los Premios Goya

#	Outer	Inner theme								Head				Participants	Text	
		Textual		Interpersonal		Pre-Head		Stance		Syntactic		Element				
		Textual	Reader	Stance	Reader	Stance	Reader	Syntactic				Stance	Reader			
1									se			Np		Pronoun	Sayer	Se está hablando mucho del discurso de Gacela Peña en los días de anoche, y aunque fue contundente y certero, yo quiero dar eco al de Bardem al recibir su premio al mejor documental por Hijos de las nubes, dando voz al conflicto saharauí que dura más de 37 años.
2									Bardem			Np		Proper Noun	Phenomenon	Bardem no me cae especialmente bien, pero nadie puede negar su implicación con la causa saharauí.
3									Que un personaje público se signifique así			Clause			Carrier	Que un personaje público se signifique así creo que es algo meritorio, al margen de la simpatía que este nos despierte.
4				Sta				Time-Np	Sahara			Np		Proper Noun	Carrier	Me gustó que anoche Sahara fuera trending topic, pero también fui consciente del peñalísimo poder que tiene Twitter.
5				Eng				Verb	haya			Existent			Existent	Es estupendo que haya un sitio donde todos opinemos, pero.
6									Qué			Question		Interrogative pronoun	Attribute	¿Qué pasa cuando quien opina no tiene ni idea de lo que habla? Anoche pude leer tuits tan absurdos como "defender la causa saharauí es ser proetarra".
7								verb	sabrás			question		Verb	Process	¿Sabrá quien afirmó con tal rotundidad en qué consiste exactamente el conflicto saharauí?
8									lo que más me molestó			clause			Phenomenon	Sin embargo, lo que más me molestó fue la doble moral con la que algunos critican a Bardem, afirmando que es muy fácil protestar cuando tienes todas las comodidades a tu alcance.

#### D. Appendix

[illegible]

Cuestión de sentido común

#	Outer				Inner theme				Head				Participants				Text
	Textual	Interpersonal	Stance	Reader	Pre-Head	Stance	Reader	Syntactic	Element	Stance	Reader	Verb		Process			
1					tengo			verb		self mention				Process	Tengo 47 años.		
2					soy			verb		self mention		NP		Carrier	Soy madre de tres hijos y víctima de un ERE.		
3					estoy			verb		self mention				Process	Estoy en el paro y mi marido también, como casi seis millones de españoles.		
4									La gente de a pie		Inclusive pronoun	NP+App		Senser	La gente de a pie observamos atónita cómo unos y otros se echan los trastos a la cabeza, culpándose de quién originó esta catástrofe.		
5					A la sociedad ya no			Contingency pp			Inclusive pronoun	NP		Senser	A la sociedad ya no nos interesan estos debates.		
6	si				la crisis ya es un azote;			Contingency clause				NP		Actor	Si la crisis ya es un azote, esta se multiplica por cada miembro de una familia.		
7					Siempre			Time				NP		Actor	Siempre se ha luchado por su protección.		
8									Qué		question	Interrogative		Attribute	¿Que está pasando ahora?,		
9									por qué		question	Interrogative		Attribute	¿por qué ni se la nombra cuando la ONU y varios organismos internacionales están haciendo un llamamiento por defenderla?		
10		Si							los padres, profesionales como los que más, con cargas familiares no deben ser los primeros en engrosar una lista de despidos o estar de los primeros en un ERE.			NP+App+pp	Carrier	MI petición es que los padres, profesionales como los que más, con cargas familiares no deben ser los primeros en engrosar una lista de despidos o estar de los primeros en un ERE.			
11									Rheme		directive	Rheme		None	Es tiempo de recortes, pero recortes con sentido común.		
12					Hay			verb	Hay			Verb		Existent	Hay que fomentar e incentivar la contratación de estos padres con familia, puesto que coincidimos con el sector de población de mayores de 45 años.		
13					Somos			Verb	Somos		Inclusive pronoun	Verb		Process	Somos los que más difícil lo tenemos en el mercado laboral y el colectivo más empleado de la sociedad española.		
14									se			NP		Actor	Se ha luchado por el empleo joven, que lo aplaudo como madre también que soy de un joven, pero el verdadero problema es el empleo de los padres de esos jóvenes, que somos los que tenemos que alimentarlos, pagar la luz, el agua, el gas, la alimentación día a día y los gastos de familia.		
15					hay			verb			directive	Verb		Existent	Hay que aunar esfuerzos.		
16	si					hechge		Contingency	conseguimos el incentivo de la contratación de estos padres profesionales con cargas familiares y mayores de 45 años, habremos dado un paso de gigante en nuestro país.		Inclusive pronoun	Verb	Process	Si conseguimos el incentivo de la contratación de estos padres profesionales con cargas familiares y mayores de 45 años, habremos dado un paso de gigante en nuestro país.			





## Appendix

### **E.1 English Automatically Annotated Corpus (in CD attached)**

The full corpus automatically annotated according to the created rules for theme is contained in the CD attached

### **E.2 Spanish Semi-automatically Annotated Corpus (in CD attached)**

The full corpus semi-automatically annotated is contained in the CD attached