

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Departamento de Lingüística Aplicada



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



TESIS DOCTORAL

Estudio crítico lingüístico de publicaciones periódicas sobre la  
crisis financiera en España

Presentada por:

Dña. Lucía Gadea Boronat

Dirigida por:

Dr. Keith Douglas Charles Stuart

Dra. Ana Paloma Botella Trelis

Septiembre 2017



## AGRADECIMIENTOS

Escribir una noticia lleva implícito un proceso de investigación. El periodista busca, recaba y contrasta datos que *a posteriori* publica y con ello contribuye a ampliar el conocimiento de la sociedad. Con esta idea mis directores de tesis, Keith D. Stuart y Ana P. Botella, me hicieron ver que el ejercicio periodístico, que había realizado durante años en medios de comunicación como *El País* o en *Radio Alcoa-Cadena Ser*, podía ser una buena base sobre la que cimentar una investigación doctoral. Sobre ese punto de apoyo comenzó a moverse en diciembre de 2011 el complicado engranaje de lo que ha acabado siendo esta tesis doctoral. Las *piezas* de esta compleja maquinaria y cómo ponerla en movimiento de manera adecuada han sido los valiosísimos conocimientos académicos y lingüísticos que me han transmitido mis directores en estos cinco años de compromiso. Ellos han logrado hacer fácil lo difícil. Me dieron la confianza para conseguir que el estudio lingüístico de la prensa de economía, algo que *a priori* consideraba como una materia densa y complicada, resultase cercano.

El puntal sobre el que gira esta investigación es el *Corpus de la Crisis Financiera* con sus 3.254.499 palabras. Su elaboración no hubiera sido posible sin las incansables aportaciones e incontables horas de trabajo de Pilar Miquel, *Pili*. Desde el minuto cero mi cuñada se (y me) ilusionó con este proyecto, aunque por desgracia no haya podido verlo finalizado desde este mundo.

Este trabajo me ha dado la oportunidad de reencontrarme con la Universitat Politècnica de Valencia, que en 1994 me concedió la licenciatura de Periodismo, que cursé en el entonces CEU San Pablo. Me acerca además a uno de los deseos de mi hija Helena, de verme convertida en profesora. Y ha acabado por demostrar el convencimiento de mi esposo Carlos en mi empeño por alcanzar nuevos retos.

*“Las actitudes son más importantes que las aptitudes” (Winston Churchill, político y estadista británico)*



## RESUMEN

Esta tesis se centra en el estudio lingüístico de publicaciones sobre la crisis financiera en España. Se trata de un acontecimiento que ha impactado sobremanera en la sociedad española y que ha generado un aluvión de información que hemos aprovechado como fuente documental. La investigación desde sus inicios persiguió radiografiar el lenguaje utilizado en artículos sobre la crisis financiera. El primer objetivo ha sido realizar una recopilación de textos con los que se ha elaborado un corpus lingüístico, el Corpus de la Crisis Financiera (CCF), lo suficientemente representativo (3,2 millones de palabras) para desarrollar una metodología de estudio de muestras reales del lenguaje periodístico sobre la crisis financiera en España en un año natural (2012) en las dos publicaciones de referencia (reflejo del bipartidismo político) por aquel entonces, *El País* y *El Mundo*. Las posibilidades que aporta el avance de las nuevas tecnologías, con los nuevos formatos digitales han favorecido la compilación, agrupación vinculada, clasificación y análisis de los textos periodísticos. De este modo, se ha elaborado un corpus especializado, a partir de textos escritos, monolingüe, sincrónico (año 2012), cerrado, codificado mediante etiquetas y finito.

Revisamos las aportaciones de los teóricos de las ciencias de la comunicación para demostrar que el trabajo de redacción, que en ocasiones los periodistas realizan de manera instintiva, tiene detrás de sí toda una argumentación teórica. Comprobamos cómo los periodistas elaboran las informaciones, a través de distintos géneros, y cómo influyen en ellos las teorías de la comunicación. También cómo se han visto afectadas las informaciones por las nuevas formas de transmisión de las noticias, es decir, con la expansión de las redes sociales. Es el caso, por ejemplo, de la *posverdad*, en que se da más importancia a las emociones que genera una noticia que a los hechos en sí. Podremos comprobar cómo el lenguaje empleado es el vehículo de intercambio de conocimiento entre el Gobierno y la sociedad.

La Lingüística de Corpus ha proporcionado la base metodológica de trabajo: hemos revisado los conceptos teóricos más relevantes y comparado los diferentes criterios a la hora de recopilar, diseñar y analizar un corpus. El análisis del CCF comenzó con el estudio cuantitativo de sus textos (frecuencias, palabras clave, líneas de concordancia, colocaciones, patrones recurrentes y

clústeres) de manera automática con herramientas informáticas para, a continuación, desarrollar la interpretación cualitativa. Nuestro fin prioritario fue detectar la carga semántica y los recursos estilísticos contenidos en las publicaciones sobre la crisis financiera. Para ello, los *actores sociales* de la crisis (las doce personalidades del ámbito público con mayores índices de aparición en cada trimestre del año 2012) han constituido el hilo conductor del análisis del CCF. El método diseñado e implementado ha proporcionado evidencias sobre determinadas categorías semánticas que co-aparecen con estas entidades. Se han identificado matices en el mensaje a través de etiquetas evaluativas y se han detectado términos, frases y oraciones que confieren al texto gran expresividad.

Las últimas fases de la investigación han consistido en el análisis manual de textos completos (macroestructura y lenguaje evaluativo) para poder establecer una categorización de los sentimientos expresados mediante diferentes mecanismos y figuras retóricas. La creación de un corpus periodístico y su análisis informático y lingüístico nos permite detectar y analizar el lenguaje evaluativo presente en los textos de forma explícita o implícita y desenmascarar evidencias sobre pensamiento político e ideológico. Este último objetivo se alcanza ya al final de la investigación en que la lectura pormenorizada de textos completos (editoriales y cartas al director) muestra la intención y posición de los autores respecto a la crisis financiera.

**Palabras clave:** lingüística de corpus, periodismo, crisis financiera, análisis lingüístico, lenguaje evaluativo, recursos retóricos

## RESUM

Aquesta tesi se centra en l'estudi lingüístic de publicacions sobre la crisi financera a Espanya. És un esdeveniment que ha causat un gran impacte en la societat espanyola i ha generat una allau d'informació que hem aprofitat com a font documental. La investigació des dels seus inicis va tractar de radiografiar el llenguatge utilitzat en la premsa econòmica. El primer objectiu ha estat realitzar una recopilació de textos amb els quals s'ha elaborat un corpus lingüístic, el Corpus de la Crisi Financera (CCF), prou representatiu (3,2 milions de paraules) com per desenvolupar una metodologia d'estudi de mostres reals del llenguatge periodístic sobre la crisi financera a Espanya en un any natural (2012) en les dues publicacions de referència (reflex del bipartidisme polític) com han estat, *El País* i *El Mundo*. Les possibilitats que aporta el progrés de les noves tecnologies, amb els nous formats digitals, han afavorit la compilació, agrupació vinculada, classificació i anàlisi dels textos periodístics. D'aquesta manera, s'ha elaborat un corpus especialitzat, a partir de textos escrits, monolingüe, sincrònic (any 2012), tancat, codificat mitjançant etiquetes i finit.

Fem una revisió de les aportacions dels teòrics de les ciències de la comunicació per demostrar que el treball de redacció, que en ocasions els periodistes realitzen de manera instintiva, té com a rerefons tota una argumentació teòrica. Comprovem com els periodistes elaboren les informacions a través de diferents gèneres, i com influeixen en ells les teories de la comunicació. També com s'han vist afectades les informacions per les noves formes de transmissió de les notícies, és a dir, amb l'expansió de les xarxes socials. És el cas, per exemple, de la *posveritat*, on es dona més importància a les emocions que genera una notícia que als fets en si. Podrem comprovar com el llenguatge emprat és el vehicle d'intercanvi de coneixement entre el Govern i la societat.

La Lingüística de Corpus ha proporcionat la base metodològica de treball: hem revisat els conceptes teòrics més rellevants i comparat els diferents criteris a l'hora de recopilar, dissenyar i analitzar un corpus. L'anàlisi del CCF va començar amb l'estudi quantitatiu dels seus textos (freqüències, paraules clau, línies de concordança, col·locacions, patrons recurrents i clústers) de manera automàtica amb eines informàtiques per a, a continuació, desenvolupar la interpretació qualitativa. La finalitat primera va ser detectar la càrrega semàntica i els recursos

estilístics continguts en les publicacions sobre la crisi financera. Per aquesta raó, els *actors socials* de la crisi (les dotze personalitats de l'àmbit públic amb majors índexs d'aparició en cada trimestre de l'any 2012) han estat el fil conductor de l'anàlisi del CCF. El mètode dissenyat i implementat ha proporcionat evidències sobre determinades categories semàntiques que coapareixen amb aquestes entitats. S'han identificat matisos en el missatge a través d'etiquetes d'avaluació i s'han detectat termes, frases i oracions que confereixen al text gran expressivitat.

Les últimes fases de la recerca han consistit en l'anàlisi manual de textos complets (macroestructura i llenguatge avaluatiu) per tal de poder establir una categorització dels sentiments expressats mitjançant diferents mecanismes i figures retòriques. La creació d'un corpus periodístic i la seva anàlisi informàtica i lingüística ens permet detectar i analitzar el llenguatge avaluatiu present en els textos de manera explícita o implícita i desemmascarar evidències sobre pensament polític i ideològic. Aquest últim objectiu s'aconsegueix ja al final de la investigació quan la lectura detallada de textos complets (editorials i cartes al director) mostra la intenció i posició dels autors respecte a la crisi financera.

**Paraules clau:** lingüística de corpus, periodisme, crisi financera, anàlisi lingüística, llenguatge d'avaluació/llenguatge avaluatiu, recursos retòrics.

## ABSTRACT

This thesis focuses on the linguistic study of publications on the financial crisis in Spain. It is a socioeconomic event that has impacted greatly on Spanish society and has generated a barrage of information that we have used as a documentary source. Our research has sought to examine the language used in newspaper articles about the financial crisis in Spain. The first objective was the compilation of a corpus of texts (3.2 million words), the Corpus of the Financial Crisis (CCF). It is sufficiently representative to develop a methodology for studying real samples of journalistic language about the financial crisis in Spain in a natural year (2012). The corpus, based on newspaper articles from *El País* and *El Mundo*, is a reflection of political bipartisanship in Spain at that moment in time. The possibilities, offered by the improvement in new technologies, with new digital formats have favoured the compilation, linked grouping, classification and analysis of journalistic texts. In this way, a specialized written corpus of newspaper articles has been designed and developed. The corpus is written text (newspaper articles), specialized (economic domain), synchronous (year 2012), closed, tag-coded and finite.

We review the contributions of communication science theorists to show that the written work, which is sometimes instinctively carried out by journalists, has a theoretical argument behind it. We check how journalists elaborate information, across different genres, and how theories of communication influence them. We also show how information has been affected by the new forms of transmission of news, that is to say, with the expansion of social networks. It is often the case that more importance is given to the emotions generated by a story than to the facts themselves. We can verify how language is used as vehicle of knowledge exchange between government and society.

Corpus Linguistics has provided the methodological basis for this thesis: we have reviewed the most relevant theoretical concepts and compared the different criteria when compiling, designing and analyzing a corpus. Computational analysis of the CCF began with the quantitative study of texts (frequencies, keywords, concordance lines, placements, recurrent patterns and clusters) to then carry out manually a more qualitative approach through the analysis and interpretation of individual texts. Our priority was to detect the semantic load and the stylistic resources contained in the publications on the financial crisis. For this reason, the social actors of the crisis (the twelve personalities of the public sphere with the highest rates of occurrence in

each quarter of 2012) have been the thread of our corpus analysis. The method designed and implemented has provided evidence on certain semantic categories that co-appear with these entities. We have identified nuances in the message through evaluative labels and have detected terms, phrases and sentences that often realize rich ideologically loaded expressions in the texts.

The final stages of the research have consisted in the manual analysis of complete texts (macrostructure and evaluative language) in order to establish a categorization of the sentiments expressed through different rhetorical figures. The creation of a journalistic corpus and its computational and linguistic analysis allows us to detect and analyze the evaluative language present in the texts, which are realized explicitly or implicitly and unmask evidences about political and ideological thinking. Detailed readings of full texts (editorials and letters to the director) show the intention and ideological position of the authors regarding the financial crisis.

**Keywords:** corpus linguistics, journalism, financial crisis, linguistic analysis, evaluative language, rhetorical resources

## TABLA DE CONTENIDO

Página

|   |    |
|---|----|
| AGRADECIMIENTOS .....   | 3  |
| RESUMEN .....   | 5  |
| RESUM .....   | 7  |
| ABSTRACT.....   | 9  |
| LISTA DE TABLAS .....   | 13 |
| LISTA DE FIGURAS .....  | 15 |
| CAPÍTULO 1: ANÁLISIS LINGÜÍSTICO DE LA CRISIS FINANCIERA .....  | 17 |
| 1.1 Introducción.....   | 17 |
| 1.2 Breve contexto de la crisis financiera .....  | 19 |
| 1.3 Hipótesis .....   | 22 |
| 1.4 Definición de objetivos .....   | 23 |
| 1.5 Fuentes primarias y secundarias <i>versus</i> (o hacia) fuentes formales e<br>informales .....      | 25 |
| 1.6 Las fuentes para elaborar el CCF. Breve descripción de <i>El País</i> y <i>El Mundo</i> .....       | 28 |
| 1.7 Conclusión.....   | 30 |
| CAPÍTULO 2: LOS GÉNEROS PERIODÍSTICOS .....   | 33 |
| 2.1 ¿Qué es el periodismo? .....  | 33 |
| 2.1.1 Los códigos de los medios de comunicación de masas .....  | 36 |
| 2.1.2 Claves de la redacción periodística.....  | 38 |
| 2.1.3 Los principales rasgos del lenguaje periodístico en sus diferentes<br>estilos.....                | 41 |
| 2.2 Los géneros periodísticos.....  | 42 |
| 2.2.1 Los géneros informativos: la noticia, la crónica, el reportaje, la<br>entrevista.....             | 46 |
| 2.2.2 Características lingüísticas de los géneros informativos .....                                    | 54 |
| 2.2.3 La superestructura del discurso periodístico.....   | 55 |
| 2.2.4 Los géneros de opinión: editorial, columna, artículo o tribuna y<br>cartas al director .....      | 56 |
| 2.2.4.1 Características lingüísticas de los géneros de opinión .....                                    | 57 |
| 2.2.5 El periodismo especializado: origen, definición, tipos y ejemplos .....                           | 64 |
| 2.2.5.1 Información periodística especializada. Periodismo y<br>prensa especializada.....               | 65 |
| 2.2.5.2 Orígenes de la prensa económica.....  | 66 |
| 2.2.5.3 La prensa económica en España: antecedentes y<br>evolución.....                                 | 68 |
| 2.3 Libertad de información y de expresión y sus límites.....   | 70 |
| 2.3.1 Marco normativo .....   | 71 |
| 2.3.2 Conceptos de libertad de expresión y de libertad de información:<br>dos derechos distintos .....  | 72 |
| 2.3.3 Límites a la libertad de expresión y de información: el honor,<br>intimidad y propia imagen.....  | 73 |
| 2.3.4 ¿Qué es noticia en los medios de comunicación? <i>Agenda Setting</i> y<br><i>Gatekeeper</i> ..... | 73 |
| 2.4 El Periodismo en redes sociales .....   | 74 |

|   |     |
|---|-----|
| 2.5 Conclusión.....   | 79  |
| CAPÍTULO 3: LINGÜÍSTICA DE CORPUS .....   | 81  |
| 3.1 Introducción.....   | 81  |
| 3.2 Origen y evolución de la lingüística de corpus.....                             | 84  |
| 3.2.1 El tamaño importa .....   | 86  |
| 3.3 Lingüística computacional.....  | 89  |
| 3.4 Conceptos clave en la lingüística de corpus.....                                | 91  |
| 3.5 Tipología de corpus. ¿Para qué se utilizan? ¿Qué analizan? .....                | 101 |
| 3.5.1 Clasificación de los corpus .....   | 106 |
| 3.6 Diseño de corpus textuales.....   | 106 |
| 3.7 Herramientas para el análisis de corpus.....                                    | 109 |
| 3.8 Aplicaciones de la lingüística de corpus.....                                   | 110 |
| 3.8.1 Análisis de corpus especializados .....                                       | 111 |
| 3.9 Corpus y la Web .....   | 116 |
| 3.10 Ruido en el corpus y las redes sociales .....                                  | 120 |
| 3.11 Copyright, ética y legalidad.....  | 121 |
| 3.12 Lingüística de Corpus y la Sociedad del Conocimiento .....                     | 122 |
| 3.13 Organigrama con ejemplos de (posibles) corpus de textos orales y escritos..... | 126 |
| 3.14 Conclusión.....  | 129 |
| CAPÍTULO 4: METODOLOGÍA Y RESULTADOS .....  | 133 |
| 4.1. Corpus de la Crisis Financiera (CCF): Justificación y creación .....           | 133 |
| 4.2. Diseño del corpus.....   | 140 |
| 4.2.1 Diseño de una base de datos .....   | 143 |
| 4.2.2 Organización de los textos.....   | 146 |
| 4.2.2.1 Descripción del CCF.....  | 149 |
| 4.3 Análisis del corpus.....  | 152 |
| 4.3.1 Extracción de listas de palabras.....   | 155 |
| 4.3.2 Estudio comparativo de términos del CCF.....                                  | 158 |
| 4.3.3 Localización de los actores sociales en el CCF. ....                          | 162 |
| 4.3.3.1 Actores sociales. <i>Collocates</i> .....                                   | 164 |
| 4.3.3.2 Actores sociales: líneas de concordancia .....                              | 166 |
| 4.3.3.3 Sustantivos abstractos del CCF .....  | 171 |
| 4.3.3.4 Actores sociales: categorización semántica de adjetivos .....               | 173 |
| 4.4 Análisis manual: etiquetas de lenguaje evaluativo.....                          | 178 |
| 4.4.1 Evaluación explícita a nivel de palabra y frase .....                         | 182 |
| 4.4.2 Evaluación implícita en el CCF .....  | 185 |
| 4.5 Análisis de textos completos: macro estructura y lenguaje evaluativo .....      | 187 |
| 4.6 Análisis detallado de textos completos .....                                    | 195 |
| 4.7 Conclusión.....   | 232 |
| CAPÍTULO 5: CONCLUSIONES Y REFLEXIONES DE CIERRE.....                               | 235 |
| 5.1. Hipótesis .....  | 235 |
| 5.2 Consecución de Objetivos.....   | 236 |
| 5.3 Limitaciones y Perspectivas de Futuro.....                                      | 242 |
| 5.3.1 Limitaciones .....  | 242 |
| 5.3.2 Perspectivas de Futuro .....  | 244 |
| ANEXOS .....  | 249 |
| BIBLIOGRAFÍA.....   | 253 |

## LISTA DE TABLAS

| Tabla   | Página |
|---|--------|
| Tabla 1.1: Variación de los afiliados a la Seguridad Social en España entre 2008 y 2016.<br>Fuente INE .....  | 22     |
| Tabla 2.1: Estructura de la noticia .....   | 47     |
| Tabla 2.2: Las cinco (o seis) W .....   | 48     |
| Tabla 3.1: Corpus gestionados por Mark Davies<br>( <a href="http://corpus.byu.edu/corpora.asp?b=y">http://corpus.byu.edu/corpora.asp?b=y</a> )..... | 88     |
| Tabla 3.2: Patrones gramaticales.....   | 97     |
| Tabla 4.1: Metadatos del CCF .....  | 147    |
| Tabla 4.2: Metodología: Etapas de la investigación .....  | 148    |
| Tabla 4.3: Datos del CCF .....  | 150    |
| Tabla 4.4: Evolución por trimestres del CCF .....   | 151    |
| Tabla 4.5: Lista de palabras del primer trimestre.....  | 161    |
| Tabla 4.6: Los doce actores sociales que más se repiten en los cuatro trimestres de 2012 .....  | 163    |
| Tabla 4.7: Etiquetas con ejemplos .....   | 181    |
| Tabla 4.8: Macroestructura de los editoriales .....   | 189    |
| Tabla 4.9: Lista de palabras y frecuencias.....   | 191    |
| Tabla 4.10: Macroestructura del texto .....   | 194    |
| Tabla 4.11: Repetición de figuras retóricas .....   | 194    |
| Tabla 4.12: Textos analizados de manera detallada .....   | 196    |
| Tabla 5.1: Figuras retóricas 1T del CCF .....   | 237    |
| Tabla 5.2: Figuras retóricas 2T del CCF .....   | 238    |
| Tabla 5.3: Figuras retóricas 3T del CCF .....   | 239    |
| Tabla 5.4: Figuras retóricas 4T del CCF .....   | 240    |



## LISTA DE FIGURAS

| Figura   | Página |
|--|--------|
| Figura 2.1: Esquema básico de la comunicación .....  | 34     |
| Figura 2.2: Estructura de la noticia .....   | 45     |
| Figura 2.3: La superestructura de la noticia.....  | 56     |
| Figura 2.4: Derecho a la Información y lo que abarca .....   | 71     |
| Figura 2.5: Ejemplos de blogs de economía en <i>El País</i> .....                                    | 78     |
| Figura 2.6: Lanzamiento en España del diario <i>The Huffington Post</i> .....                        | 79     |
| Figura 3.1: Concordancias con el término (actor social) <i>Rajoy</i> .....                           | 93     |
| Figura 3.2: Organigrama de textos escritos.....  | 127    |
| Figura 3.3: Organigrama de textos orales.....  | 128    |
| Figura 4.1: Diagrama de confluencia realidad social-relato periodístico .....                        | 136    |
| Figura 4.2: Descripción esquemática del CCF .....  | 152    |
| Figura 4.3: Inicio página EXCEL de textos El País (EP) .....   | 153    |
| Figura 4.4: Inicio página EXCEL de textos de opinión de El País (OEP) .....                          | 153    |
| Figura 4.5: Inicio página EXCEL de textos de opinión de El Mundo (OEM).....                          | 153    |
| Figura 4.6: Lista de palabras más frecuentes correspondiente a textos del tercer trimestre 2012..... | 158    |
| Figura 4.7: Hápx legomena hallados en El País 3T .....   | 160    |
| Figura 4.8: Hápx legomena hallados en El Mundo Opinión 3T .....                                      | 160    |
| Figura 4.9: Hápx legomena hallados en El País Opinión 3T .....                                       | 160    |
| Figura 4.10: Ejemplo líneas de concordancia de Rajoy en 3T.....                                      | 166    |
| Figura 4.11: Rajoy y sus <i>collocates</i> .....   | 168    |
| Figura 4.12: Actores sociales y sus <i>collocates</i> : primer trimestre.....                        | 169    |
| Figura 4.13: Actores sociales y sus <i>collocates</i> : segundo trimestre .....                      | 170    |

|   |     |
|---|-----|
| Figura 4.14: Sustantivos abstractos 1T .....  | 171 |
| Figura 4.15: Documento sobre la relación de los actores sociales y sus adjetivos en el 1er trimestre..... | 173 |
| Figura 4.16: Esquema categorización semántica de los adjetivos, según <i>Appraisal Theory</i> .....       | 175 |
| Figura 4.17: Extracción de etiquetas a través de las líneas de concordancia de Montoro .....              | 177 |
| Figura 4.18: Metodología para el análisis de la crisis financiera en España .....                         | 188 |
| Figura 4.19: El ‘banco malo’ .....  | 230 |
| Figura 4.20: Dudoso .....   | 230 |
| Figura 4.21: El ‘banco malo’ es dudoso porque produce pérdidas, déficit, deuda .....                      | 231 |
| Figura 4.22: El ‘banco malo’ es dudoso porque es difícil gestionar.....                                   | 231 |
| Figura 4.23: Las dudas que suscita el ‘banco malo’ .....  | 232 |
| Figura 5.1: Modelización del lenguaje evaluativo.....   | 241 |

## CAPÍTULO 1: ANÁLISIS LINGÜÍSTICO DE LA CRISIS FINANCIERA

### 1.1 Introducción

"Estoy segura de que la crisis tiene su propio lenguaje y su utilización puede ser intencionada. Y no es lo mismo decir ayudas que rescate, recesión por crecimiento negativo o reestructuración en vez de recortes". La frase<sup>1</sup>, recogida en *El País* del 16 de mayo de 2013 la pronunció la entonces princesa Letizia con motivo de la inauguración del VIII Seminario Internacional de Lengua y Periodismo<sup>2</sup>, en San Millán de la Cogolla (La Rioja). El objetivo de nuestra investigación se centra en el estudio lingüístico de publicaciones sobre la crisis financiera, acerca del significado de las palabras utilizadas y el sentido que los distintos autores y periodistas han dado a quienes denominaremos en esta investigación como *actores sociales* de la crisis. Para este análisis lingüístico de la crisis financiera hemos recopilado y analizado textos periodísticos a lo largo de un año natural, el 2012. Todo ello con el objetivo de extraer conclusiones sobre el lenguaje de la crisis, su utilización y posible intencionalidad.

"Words, words, words..." No es la frase más popular de *La tragedia de Hamlet, Príncipe de Dinamarca*, de William Shakespeare (1605), pero encaja bien para definir el espíritu inicial de este trabajo. "Palabras, palabras, palabras..." es la respuesta del protagonista de la tragedia danesa a la pregunta de un entrometido consejero, Polonio, acerca de lo que estaba leyendo Hamlet en un momento de la obra. Es, además, una muestra del poder que Shakespeare ya concedía por aquel entonces a las palabras. Como erudito que es, Hamlet justifica sus acciones mediante las palabras y la reflexión. Utiliza la retórica como arma para confundir a sus semejantes y enemigos. Se defiende del acecho de las amenazas y traiciones con las trampas semánticas de la lengua. En esta investigación vamos a analizar ese poder oculto de las palabras y que aparece según decidan quienes las utilicen, con trampas o sin ellas, por parte de los distintos actores o protagonistas.

Hemos escogido como protagonistas a analizar a los actores sociales de la crisis económica. Para ello hemos seleccionado en esta investigación los nombres que más se repiten en los textos que

---

<sup>1</sup> [http://politica.elpais.com/politica/2013/05/16/actualidad/1368712226\\_024687.html](http://politica.elpais.com/politica/2013/05/16/actualidad/1368712226_024687.html)

<sup>2</sup> [http://cultura.elpais.com/cultura/2013/05/16/actualidad/1368730703\\_051922.html](http://cultura.elpais.com/cultura/2013/05/16/actualidad/1368730703_051922.html)

hemos recopilado de dos publicaciones periodísticas y con los que hemos creado del Corpus de la Crisis Financiera, el cual, una vez construido, también hemos analizado.

El motivo de la elección de textos periodísticos para elaborar un corpus de estas características responde, en primer lugar, a que no teníamos constancia de la existencia de un registro anterior de estas mismas características. En segundo lugar, nos pareció una manera óptima de unificar en esta investigación dos áreas en la que vemos nexos de unión. Hablamos de la lingüística y el periodismo. El análisis de las palabras lo hemos realizado tras la elaboración de un corpus formado por 3.254.499 palabras recogidas de 5.480 textos de *El País* y *El Mundo* sobre informaciones que tratan información económica en el año 2012. Las posibilidades que ofrecía el formato digital de ambas publicaciones han sido cruciales en la elección de ambos medios. Los escogimos además porque representaban el bipartidismo que, en el momento del inicio de la investigación, imperaba en España. A lo largo de este trabajo trataremos de analizar y comprender si en los inicios del siglo XXI se cumple aquello que Shakespeare ya escribió en 1601 como broche final a *Hamlet*. Sus últimas palabras en la tragedia fueron: "The rest is silence..." Esta frase, "El resto es silencio", lo interpretamos como una defensa de la utilidad del discurso.

Cerramos este apartado dedicado a introducir este trabajo de investigación con una nueva referencia periodística. Durante la elaboración de este trabajo, la intervención de Mariano Rajoy en un programa de televisión, el 2 de diciembre de 2015, *En tu casa o en la mía* de TVE, nos regala una magnífica confesión del presidente del Gobierno. Tres años después de inicio del trabajo viene a confirmar la hipótesis que en el origen buscábamos llegar a demostrar. El artículo, que firma Íñigo Domínguez en *El País* (3 de diciembre, 2015), recoge este episodio.

Se sinceró sobre los momentos más duros de la crisis: "Los primeros tiempos en el Gobierno yo no explicaba nada. Seguramente me explicaba mal, ni me expliqué. Tenía mil líos". Habló de "momentos enormemente duros, muchos, muchísimos" al borde del rescate. "Aquella época fue espantosa", concluyó, pero no había alternativa, él solo hizo lo normal, lo que hubiera hecho cualquiera.

Este episodio lo hemos visto recopilado también en nuestro corpus y se produjo en el ecuador del periodo analizado, en junio de 2012. Lo consideramos, por la información periodística que aporta, el momento culminante de la crisis.

## 1.2 Breve contexto de la crisis financiera

La crisis financiera internacional, que analizamos en esta investigación a través de las publicaciones periodísticas de 2012, tiene su origen en 2008 en Estados Unidos. Es importante conocer el contexto socio-económico que impera en el momento de elaboración de la información analizada a lo largo del año 2012 en la prensa nacional, para comprender bien cómo y por qué se producen las noticias que han elaborado los periodistas y autores de los textos de opinión que analizamos en esta investigación. Esta crisis económica que comenzó a dar signos de evidencia en 2008 tuvo su origen en Estados Unidos. La subida de los precios de las materias primas, la sobrevalorización de los productos, unida a la crisis alimentaria en todo el planeta, agravada por la crisis energética, la subida de la inflación, una creciente amenaza de recesión junto a una crisis crediticia, hipotecaria y de confianza en los mercados derivaron en la crítica situación socio-económica en los países más desarrollados del mundo.

Torrero (2008) establece una serie de detonantes y consecuencias de la crisis. Entre los detonantes habla de la caída de la bolsa, el alza de las materias primas, los cambios en el ordenamiento legal, la quiebra inesperada de la gran empresa. A la hora de establecer las consecuencias señala aspectos tan dispares como que puede darse el caso de que no suceda nada o bien que provoque un verdadero alud, que arrastre todo tras de sí. Para comprender bien el estallido de esta crisis debemos remontarnos a los primeros años del siglo XXI. La subida de las materias primas, alimentos y del petróleo comenzó a ser una seria amenaza para los países en vías de desarrollo. La especulación elevó el precio del petróleo a los 147 dólares estadounidenses y arrastró consigo nuevas subidas en metales industriales como el cobre, que alcanzó una cotización en la Bolsa de Metales de Londres los 8.940 dólares de USA por tonelada. Otros materiales básicos para la producción industrial como el ácido sulfúrico o la soda cáustica, también vieron crecer su precio hasta un 600%. Esta crisis se expandió rápidamente por los países más desarrollados: Japón, Australia o Nueva Zelanda, e hizo estallar la incertidumbre respecto a los países con economías emergentes como China, India, Argentina, Brasil y Sudáfrica, que ya en 2009 comenzaron a superar las bolsas de Estados Unidos y Europa. Torrero (2008) alude a apoyos sobre los que se basan los economistas para comprender el fenómeno de las crisis; los desequilibrios acumulativos, la lógica del proceso, el nulo control del proceso de una crisis, la intervención de los estados y los Bancos Centrales en situaciones de crisis y por

último el efecto rebaño, que el autor atribuye a observar el comportamiento de los demás y emular sus pautas.

La inflación comenzó a crecer para alcanzar a mediados de 2008 máximos históricos en los países exportadores de petróleo por el aumento de las reservas de divisas extranjeras. El problema se extendió como una plaga en los países desarrollados, aunque en menor medida. Un año más tarde el panorama económico derivó en deflación y generó un incremento histórico en el número de desempleados de 205 millones en todo el mundo. Las monedas de los países desarrollados (dólar, euro, yen y yuan) comenzaron su particular batalla en 2010, para alcanzar ventajas competitivas, de cara a una mejora en las exportaciones y así facilitar la salida de la crisis. Lejos de acercarse al final del túnel, el enfrentamiento derivó en una guerra comercial que alejaba la recuperación. Todas las miradas apuntaron a China que podría haber solucionado el problema apreciando su moneda.

El año 2008 fue crítico para Estados Unidos, que comenzó a acusar los síntomas de la crisis crediticia e hipotecaria que afectó a la burbuja inmobiliaria, hinchada en los últimos años del siglo XX. Ese año se produjo la bancarrota del banco de inversión Lehman Brothers<sup>3</sup> y otras compañías hipotecarias como Fannie Mae y Freddy Mac. El estallido de esta crisis se sitúa en agosto de 2007 cuando los bancos centrales se vieron abocados a intervenir para dar liquidez al sistema bancario. La quiebra, sin embargo, fue inevitable. Cayeron medio centenar de bancos y entidades financieras y con ellas la bolsa, la capacidad de consumo y el ahorro de la población. En 2011 la situación de crisis tocó techo al alcanzar la deuda soberana del país 14,4 billones de dólares. Torrero (2008:13) atribuye el origen de la crisis a la fuerte expansión del crédito hipotecario en Estados Unidos:

La expansión del crédito hipotecario requería un aumento de la demanda lo cual implicaba relajar las condiciones de los préstamos, concediéndolos a personas de dudosa reputación financiera, pero dispuestas a aceptar –por eso mismo- un tipo de interés más elevado que hacía extraordinariamente rentable estas operaciones. Todo descansaba en la confianza de que el continuo aumento del precio de las viviendas garantizaba el buen fin de los créditos concedidos.

El autor del estudio detalla cómo la rueda de la crisis pronto adquirió velocidad y proporcionó beneficios a bancos e instituciones financieras. Para aumentar la velocidad las entidades

---

<sup>3</sup> [http://news.bbc.co.uk/hi/spanish/business/newsid\\_7615000/7615910.stm](http://news.bbc.co.uk/hi/spanish/business/newsid_7615000/7615910.stm) [26-04-2016]

concesionarias de los créditos hipotecarios podían sacarlos de su balance sin consumir recursos propios. Este fue el origen de las conocidas como hipotecas subprime<sup>4</sup>, que han tenido un fuerte impacto en los países europeos y que en esta investigación veremos como una de las palabras que más popularidad alcanza junto a otros vocablos hasta la fecha desconocidos para la mayoría como *prima de riesgo* o *fondos buitres*.

La mala experiencia vivida en Estados Unidos no sirvió de señal de alarma en el viejo continente. El Banco Central Europeo fue incapaz de prevenir la crisis que arremetía. Las medidas llegaron con retraso y apuntaron principalmente a la austeridad y contención del gasto público. La consecuencia fueron créditos escasos y la dificultad de acceso a la financiación de consumidores y productores. Dinamarca entró en recesión y otros países como Francia y Alemania comenzaron a ver como su economía se contraía.

La primera consecuencia en España fue la subida del desempleo y la escalada de los Expedientes de Regulación de Empleo, los lamentablemente cada vez más comunes EREs. 2012, el año analizado en esta investigación, terminó en España con una tasa de desempleo de 26%<sup>5</sup>. El PIB español del cuarto trimestre de 2012 se redujo 0,7% con respecto al trimestre anterior y 1,8% con respecto al último trimestre de 2011<sup>6</sup>. En el año natural 2012<sup>7</sup> se perdieron casi 800.000 afiliados a la seguridad social, una cifra sólo superada en 2009.

El INE certifica una caída del PIB del 0,5% en el primer trimestre del año siguiente 2013. En términos interanuales la caída fue del 2% respecto a 2012 y el desempleo rebalsó el 27,3% con más de 6.202.700 parados<sup>8</sup>. En el año 2013 el desempleo descendió muy lentamente.

Las cifras de la variación del número de afiliados en enero de 2016 eran claramente inferiores a las existentes en enero de 2011. El desglose de estos datos en la tabla 1.1 muestra la variación

---

<sup>4</sup>[http://web.archive.org/web/20070928111114/http://serviciodeestudios.bbva.com/TLBB/fbin/EUES\\_070310\\_PrensaEspana\\_24\\_tcm208-89011.pdf](http://web.archive.org/web/20070928111114/http://serviciodeestudios.bbva.com/TLBB/fbin/EUES_070310_PrensaEspana_24_tcm208-89011.pdf) [26-04-2016]

<sup>5</sup> <http://www.dw.com/es/m%C3%A1s-de-26-de-desempleo-en-espa%C3%B1a/a-16547227> [27-04-2016]

<sup>6</sup> <http://www.gurusblog.com/archives/caida-pib-espana-2012/30/01/2013/> [27-04-2016]

<sup>7</sup> <http://www.publico.es/actualidad/ine-certifica-pib-cayo-primer.html> [27-04-2016]

<sup>8</sup> <http://www.publico.es/actualidad/dia-internacional-trabajadores-202-700.html> [27-04-2016]

absoluta y la porcentual de los afiliados a la seguridad social en un ciclo de 8 años; entre 2008, año del estallido de la crisis financiera con la quiebra de Lehman Brothers y el año 2016.

| <b>AÑO</b>       | <b>VARIACIÓN ABSOLUTA</b> | <b>VARIACIÓN PORCENTUAL</b> |
|------------------|---------------------------|-----------------------------|
| <b>2008-2009</b> | <b>-980.108</b>           | <b>-5,11%</b>               |
| <b>2009-2010</b> | <b>-635.735</b>           | <b>-3,50%</b>               |
| <b>2010-2011</b> | <b>-184.173</b>           | <b>-1,05%</b>               |
| <b>2011-2012</b> | <b>-403.571</b>           | <b>-2,32%</b>               |
| <b>2012-2013</b> | <b>-778.829</b>           | <b>-4,59%</b>               |
| <b>2013-2014</b> | <b>-5.829</b>             | <b>-0,04%</b>               |
| <b>2014-2015</b> | <b>401.703</b>            | <b>2,48%</b>                |
| <b>2015-2016</b> | <b>529.045</b>            | <b>3,19%</b>                |

Tabla 1.1: Variación de los afiliados a la Seguridad Social en España entre 2008 y 2016. Fuente INE

Pero, si un país se ha visto duramente afectado por la crisis, ese es Grecia. La deuda pública emitida por el gobierno griego para financiar sus déficits agravó la situación en este periodo. La mala gestión gubernamental derivó en recortes que generaron manifestaciones públicas y disturbios. Fue el primer país en solicitar un rescate para sostener su deuda pública. Rescate es una de las palabras de la crisis y que analizaremos en esta investigación. La solicitud de un segundo rescate en 2011 tensó las relaciones del país con la Unión Europea y el FMI. Siguió los pasos de Grecia Irlanda, Portugal y España. En este contexto arranca esta investigación. 2012 fue un año especialmente azotado por la crisis financiera.

### **1.3 Hipótesis**

En esta tesis partimos de las siguientes hipótesis:

1. La evaluación en el discurso periodístico es un fenómeno complejo, difícil de determinar y sistematizar tanto mediante el análisis cuantitativo como a través de la interpretación cualitativa de muestras de texto.
2. El periodista recurre a una gran variedad de mecanismos y recursos estilísticos para presentar y opinar sobre los acontecimientos de un determinado momento.
3. El lenguaje evaluativo en los artículos periodísticos está presente en diferentes niveles: palabra, frase, oración, e incluso distribuido por todo un texto (hasta la intertextualidad puede tener una función evaluativa).
4. La recopilación y el diseño de un corpus de artículos de prensa escrita ofrece al investigador una buena base de datos para llevar a cabo un estudio lingüístico sobre los acontecimientos económicos/financieros en un periodo relevante para la sociedad.
5. La Lingüística de Corpus se convierte así en una metodología empírica de trabajo para el análisis minucioso de las estrategias utilizadas para evaluar las instituciones, los actores sociales y los procesos implicados y afectados por la crisis financiera en España.
6. El análisis del Corpus de la Crisis Financiera en sus diferentes etapas proporcionará evidencias sobre la complejidad del lenguaje evaluativo, cargado de recursos estilísticos y matices, y permitirá radiografiar la cobertura periodística sobre esta crisis social que tanto ha afectado a la sociedad española en general.

#### **1.4 Definición de objetivos**

El objeto de la investigación es la construcción de un corpus elaborado a partir de textos periodísticos relacionados con las noticias de economía que permitan analizar su carga semántica. Los hemos recopilado de las secciones de opinión de *El País* y *El Mundo* y de la sección de economía de *El País* a lo largo del año natural 2012. El corpus se ha denominado *Corpus de la Crisis Financiera* (CCF en adelante) y en él vamos a comprobar una serie de aspectos de los artículos periodísticos, qué palabras son las más frecuentes y con qué actores se relacionan para extraer conclusiones. Estos son, en esencia, nuestros objetivos:

1. Objetivo general del estudio: Realizar una radiografía lingüística de la crisis financiera.

2. Objetivos específicos:

- Recopilar y diseñar una base de datos de artículos periodísticos sobre la crisis financiera en España del año 2012
- Recopilar datos de cada trimestre por periódico (*El País* y *El Mundo*) y sección (Economía y Opinión)
- Extraer listas de palabras de cada trimestre por periódico y sección
- Diseñar tabla comparativa de las palabras más frecuentes por trimestres
- Determinar los 12 actores sociales más frecuentes por trimestres y extraer sus colocaciones y líneas de concordancia. Diseñar redes ilustrativas de las relaciones
- Extraer los sustantivos abstractos del CCF
- Seleccionar los adjetivos que acompañan a cada actor social
- Establecer categorías semánticas de los adjetivos
- Analizar manualmente las líneas de concordancia y asignar etiquetas de ejemplos de lenguaje evaluativo, recopilar etiquetas más frecuentes y transferir a un documento Excel las etiquetas propuestas con ejemplos
- Analizar en textos completos (editoriales): su macroestructura y lenguaje evaluativo

Tras presentar el objetivo general y enumerar los diez objetivos específicos de esta investigación recurrimos a Eco (2001: 43) para tratar de que nuestro trabajo alcance los siguientes requisitos genéricos que el autor establece como fundamentales en toda investigación científica:

- La investigación versa sobre un objeto reconocible y definido de tal modo que también sea reconocible por los demás
- La investigación tiene que decir sobre este objeto cosas que todavía no han sido dichas o bien revisar con óptica diferente las cosas que ya han sido dichas
- La investigación tiene que ser útil a los demás, ya que se considera útil un artículo que presente un nuevo descubrimiento sobre el comportamiento de las partículas elementales
- La investigación debe suministrar elementos para la verificación y la refutación de las hipótesis que presenta, y por tanto tiene que suministrar los elementos necesarios para su seguimiento público

El objetivo inicial es cumplir el máximo de requisitos, para lo cual nuestro primer trabajo sería radiografiar las publicaciones de referencia en 2012, un año especialmente condicionado por la crisis económica. El CCF que nos va a permitir este análisis es diacrónico. Lo comenzamos a

analizar por trimestres para poder extraer conclusiones a través de su evolución. Las disciplinas de la lingüística y el periodismo van a ser los ejes que guíen esta investigación. Vamos a exponer, a continuación, a lo largo de los dos capítulos siguientes (capítulos 2 y 3) la referencia teórica a ambas materias motoras del presente trabajo. Eco (2001: 58) establece que:

Se puede hacer de modo científico una tesis que otros definirían, en cuanto al tema, como puramente periodística. Y se puede hacer de modo puramente periodístico una tesis que, a juzgar por el título, tendría todo lo necesario para parecer científica.

Veamos a continuación los nexos de unión entre el periodismo y la investigación científica a través del análisis de fuentes primarias, secundarias, formales e informales.

### **1.5 Fuentes primarias y secundarias *versus* (o hacia) fuentes formales e informales**

Del mismo modo que un investigador recurre a las fuentes primarias y secundarias para desarrollar su trabajo, un periodista ha de informarse y contrastar en las fuentes oficiales y extraoficiales un determinado hecho antes de publicarlo. Eco (2001: 18) a la hora de delimitar las características que ha de tener una investigación para una tesis doctoral señala que:

...se trata precisamente de una investigación original, hay que conocer lo que han dicho sobre el tema los demás estudiosos y, sobre todo, es preciso *descubrir* algo que los demás no hayan dicho todavía. Cuando se habla de *descubrimiento*, especialmente en humanidades, no se trata de inventos revolucionarios como el descubrimiento de la escisión del átomo, la teoría de la relatividad o un medicamento que cure el cáncer: también puede haber descubrimientos modestos, y se considera *científico* incluso un nuevo modo de leer y comprender un texto clásico, la localización de un manuscrito que arroja nuevas luces sobre la biografía de un autor, una reorganización y relectura de estudios precedentes que lleva a madurar y sistematizar ideas que vagaban dispersas por otros textos variados. En cualquier caso, el estudioso ha de producir un trabajo que, teóricamente, los demás estudiosos del ramo no deberían ignorar, pues dice algo nuevo.

La verificación de los hechos a través de sus fuentes es un principio inalterable para el periodismo y que ratifica aquello que ha sido un lugar común en la historia del periodismo: *un periodista vale lo que valen sus fuentes*. Pero, ¿cuáles son el tipo de fuentes habitualmente utilizadas por el periodista? El informe *El periodista en la encrucijada* (Diezhandino et al., 2012: 57) cita las siguientes fuentes periodísticas con sus respectivos porcentajes de utilización por parte de los informadores:

- Contactos personales: 81'67%
- Notas de prensa y gabinetes de comunicación: 73'33%
- Enlaces en red: 38'33%
- Blogs: 11'67%
- Redes Sociales: 28'33%
- Otros: 11'67%

Las notas de prensa y comunicados de entidades públicas, así como las personas que los representan, formarían parte de las fuentes oficiales a las que recurre el periodista. Las redes sociales carecen de esta oficialidad, así como los informadores anónimos que no representan a colectivos. Las fuentes periodísticas, al igual que las de la investigación académica, responden a una clasificación genérica, según la relación al objeto de conocimiento. Esta tipología distingue entre:

- Fuentes primarias
- Fuentes secundarias

La diferencia entre fuentes primarias y secundarias es la relación que mantienen con el objeto de conocimiento. Para un periodista o para un investigador una fuente primaria es la que está informada por haber participado o presenciado en un hecho noticiable o conoce de primera mano la materia a investigar. En la elaboración de las noticias los periodistas recurren a los testigos de un accidente o afectados por una causa para redactar o narrar sus crónicas. Los investigadores que, por ejemplo, están elaborando una tesis doctoral, bucean en los manuales teóricos para construir su trabajo, consultan y adaptan a su investigación a través de lo publicado previamente. En ambos casos la identificación de la fuente es la que concede valor a la investigación, pero no siempre es posible, sobre todo en periodismo.

Dentro de las fuentes primarias se encuentran las de atribución directa. Son las conocidas como formales, las que se pueden identificar claramente con nombres y apellidos. Es el caso de los testigos de un hecho, los cargos o entidades públicas o privadas, gabinetes de comunicación. Aportan información para ser publicada.

Pero existen otras fuentes tan o más importantes que no desean ser identificadas directamente. Sin embargo, se pueden citar en determinados contextos y transmitir sus palabras. Se trata de las fuentes informales. Se identifican en las noticias como una fuente judicial, o una fuente próxima al presidente, o testigos presenciales. Se da la circunstancia de que el informador puede reclamar al periodista que omita parte de su testimonio. Esta atribución *off the record* es la información que se le da a quien elabora una noticia sólo para que comprenda de manera global algún hecho, pero, que en ningún caso debe ser publicada.

La fuente secundaria es la que interpreta, pone en contacto y analiza las fuentes primarias. En los periódicos esta interpretación de los hechos la encontramos en la sección de opinión. Sus autores ofrecen una versión de una realidad que en la mayoría de las ocasiones no han vivido directamente. Su función es encaminar su pensamiento. El investigador en una tesis recurre a unos u otros teóricos de las materias que conforman las fuentes para orientar el sentido de una investigación académica.

Para ejemplarizar lo que en una investigación académica serían las fuentes primarias y secundarias Eco (2001: 61) detalla los siguientes instrumentos de los que se vale un investigador en una tesis para estudiar un determinado objeto:

Muchas veces el objeto es un libro y los instrumentos son otros libros. Tal es el caso de una tesis, supongamos, sobre El pensamiento económico de Adam Smith, en la cual el objeto está constituido por los libros de Adam Smith mientras que los instrumentos son otros libros sobre Adam Smith. En tal caso diremos que los escritos de Adam Smith constituyen las fuentes primarias y los libros sobre Adam Smith constituyen las fuentes secundarias o la literatura crítica. Naturalmente, si el tema fuera Las fuentes del pensamiento económico de Adam Smith, las fuentes primarias serían los libros o escritos en que se inspiró Smith. En realidad, las fuentes de un autor pueden haber sido también acontecimientos históricos (ciertas discusiones acaecidas en su tiempo sobre ciertos fenómenos concretos), pero estos acontecimientos siempre son accesibles en forma de material escrito, esto es, de otros textos.

En nuestro caso, el objeto a estudiar son las publicaciones sobre la crisis financiera para su posterior análisis y la extracción de conclusiones. En el inicio de la investigación, la prensa de referencia en España se centraba en dos medios de comunicación escritos con los que elaborar un corpus lingüístico: *El País* y *El Mundo*, ambos medios eran exponentes del entonces bipartidismo. Vamos a continuación a analizar las fuentes de nuestro corpus.

### 1.6 Las fuentes para elaborar el CCF. Breve descripción de *El País* y *El Mundo*

Veamos en este punto algunos datos de interés de los periódicos que han servido para la construcción del CCF. Aportamos la información de fundación, crecimiento, tendencia y evolución de *El País* y *El Mundo*, con especial atención al periodo (el año 2012) analizado en esta investigación.

- a. ***El País***. Se fundó en 1976. Desde 2014 su director es Antonio Caño, pero el primero fue Juan Luís Cebrián (1976-1988). Le han sucedido Joaquín Estefanía (1988-1993), Jesús Ceberio (1993-2006) y Javier Moreno (2006-2014). La Oficina de Justificación de la Difusión (OJD) certifica en el periodo analizado en esta investigación -de enero a diciembre de 2012- que el promedio de tirada<sup>9</sup> de *El País* fue de 400.212 ejemplares y el promedio de difusión<sup>10</sup> de ejemplares vendidos 324.814. Según el Estudio General de Medios (EGM), en julio de 2013 alcanzó 1.929.000 lectores en su edición impresa.

*El País* tiene también una edición global o internacional. El 4 de mayo de 1996 se convirtió en el segundo periódico de España (después del periódico en catalán *Avui*) en ofrecer una edición electrónica en internet. También fue el primer periódico español que, el 18 de noviembre de 2002, impuso el sistema de pago para acceder a los contenidos informativos de su edición digital. La versión para internet alcanzó en abril de 2013 los 15,6 millones de lectores únicos, según la agencia Comscore de marketing en internet. *El País* está también presente en las redes sociales, con más de 4 millones y medio de seguidores en Twitter y casi 1.900.000 en Facebook.

El Grupo PRISA, al que pertenece *El País*, es el mayor grupo mediático de España. Su principal accionista es, a partir de 2010, Liberty Acquisition Holding (los principales accionistas de este fondo de inversiones son, a su vez, los estadounidenses Nicolas Berggruen y Martin E. Franklin). PRISA es también propietario de *Cadena SER* (radio), *Cinco Días* (prensa económica), Grupo Santillana, con las editoriales *Santillana*, *Alfaguara*, *Taurus*, *Aguilar* y otras; *Diario As* (prensa deportiva), *Los 40 Principales*, *Máxima FM*, *M80 Radio*, *Radiolé*, *Cadena Dial* (radiofórmula), Prisa TV, Localia, Digital+ (televisión), entre otros medios.

---

<sup>9</sup>[http://web.archive.org/web/20120531092423/http://www.ojd.es/Introl/Portal/publicacion\\_detalle/\\_LX1O\\_Pr-rb\\_siYZO\\_A6y71P4ZANY9Afh](http://web.archive.org/web/20120531092423/http://www.ojd.es/Introl/Portal/publicacion_detalle/_LX1O_Pr-rb_siYZO_A6y71P4ZANY9Afh) [13-04-2016]

<sup>10</sup> <http://prnoticias.com/hemeroteca/20119354-ojd-2012-las-ventas-de-el-pais-caen-un-16-en-un-ano-y-las-de-la-gaceta-un-38> [13-04-2016]

La crisis también afectó a este grupo y, en 2011, el Grupo PRISA registró las mayores pérdidas<sup>11</sup> de su historia, de 451 millones de euros, una cifra seis veces mayor que la recogida en el año 2010.

En su fundación *El País* se definió como un periódico independiente, de calidad, con vocación europea y defensor de la democracia pluralista. Fue referente del centro-izquierda española o tendencias socialdemócratas y su línea editorial cercana al PSOE. Tras sucesivos cambios accionariales<sup>13</sup>, sus lectores notaron una evolución<sup>14</sup> al liberalismo o al neoconservadurismo<sup>15</sup>.

**b. *El Mundo*.** En sus orígenes se denominaba de manera formal *El Mundo del siglo XXI*. Desde 2015 su director es David Jiménez, que sucedió a Casimiro García-Abadillo (2014-2015). El director desde su fundación, el 23 de octubre de 1989, hasta el 30 de enero de 2014 fue Pedro J. Ramírez.

En el año analizado en esta investigación, el 2012, según la Oficina de Justificación de la Difusión el promedio de difusión<sup>16</sup> de *El Mundo* fue de 206.179 ejemplares. La versión digital de la publicación impresa tuvo un año más tarde, en 2013, el promedio de difusión de 28.731 ejemplares. A esta versión se accede a través de la plataforma digital Orbyt de contenidos de pago con productos editoriales del Grupo Unidad Editorial General S.L.U. propiedad de Unidad Editorial, S.A., propietaria también del Grupo Recoletos, las cuales están participadas mayoritariamente por el grupo RCS MediaGroup (Rizzoli), que controla el diario italiano de mayor tirada, *Corriere della Sera*. La plataforma además permite consultar la hemeroteca de todas las publicaciones, opinar e interactuar con la redacción. También cuenta con edición abierta en internet denominada *elmundo.es*. Los contenidos no son exactamente los

---

<sup>11</sup> [http://www.nytimes.com/2012/10/24/business/global/el-pais-leading-newspaper-in-spain-announces-plans-for-layoffs.html?pagewanted=1&\\_r=2&smid=tw-share](http://www.nytimes.com/2012/10/24/business/global/el-pais-leading-newspaper-in-spain-announces-plans-for-layoffs.html?pagewanted=1&_r=2&smid=tw-share) [13-04-2016]

<sup>12</sup> <http://prnoticias.com/hemeroteca/20112471> [13-04-2016]

<sup>13</sup> <http://www.publico.es/actualidad/pais-cambia-director-y-redaccion.html> [13-04-2016]

<sup>14</sup> [http://www.eldiario.es/rastreador/testamento-Pais-periodicos\\_6\\_232786742.html](http://www.eldiario.es/rastreador/testamento-Pais-periodicos_6_232786742.html)

<sup>15</sup> <http://tiempoargentino.com/nota/55768> [13-04-2016]

<sup>16</sup> <http://prnoticias.com/prensapr/519-ojd-1/20119354-ojd-2012-las-ventas-de-el-pais-caen-un-16-en-un-ano-y-las-de-la-gaceta-un-38> [13-04-2016]

mismos que en la edición impresa, aunque desde el portal se puede acceder a una versión digital de la edición impresa a través de un quiosco virtual específico.

Unidad Editorial también cerró el ejercicio de 2011 con pérdidas<sup>17</sup> de 322 millones de euros. Fue el quinto año consecutivo en números rojos que dejó la empresa editora en una situación de quiebra técnica<sup>18</sup>. El 28 de mayo de 2012 la editora *El Mundo* firma un ERE, ante el fuerte descenso de los ingresos publicitarios.

*El Mundo* define su línea editorial como liberal. Es habitualmente crítico con el PSOE y los nacionalismos periféricos y próximo (sobre todo en política española) al PP, aunque se muestra alejado de principios puramente conservadores. La periodista Elena de Regoyos (2006) defiende la teoría<sup>19</sup> de que el objetivo de *El Mundo* es convertirse en el medio escrito de referencia de la derecha española desplazando al diario conservador ABC.

## 1.7 Conclusión

La crisis financiera, que comenzó en 2008, por sus características y por tratarse de un acontecimiento novedoso para la sociedad de la época, generó un aluvión de información en los medios de comunicación con el objetivo de responder a la incertidumbre de una población que se disponía a afrontar una situación inédita hasta la fecha. Conscientes de que el relato de estas informaciones sobre la crisis ha contado con su propio lenguaje, surgió esta investigación. El primer objetivo fijado fue la construcción de un corpus lingüístico a partir de artículos periodísticos para analizar el lenguaje de la crisis financiera en España en 2012, un año especialmente crítico. El estudio de las publicaciones sobre la crisis financiera se ha centrado en los dos medios de comunicación de referencia escritos de la época; *El Mundo* y *El País*. Este trabajo se ha organizado siguiendo las pautas que ha de seguir un investigador, con el objetivo de darle el mayor carácter académico posible, tratando de que aporte algo novedoso a la comunidad científica y sin olvidar la finalidad de poder darle divulgación al trabajo y que resulte

---

<sup>17</sup> [http://economia.elpais.com/economia/2012/03/16/actualidad/1331914134\\_344710.html](http://economia.elpais.com/economia/2012/03/16/actualidad/1331914134_344710.html) [13-04-2016]

<sup>18</sup> [http://www.periodistadigital.com/periodismo/prensa/2010/06/23/unidad-editorial-elmundo-perdidas-quiebra-cuentas.shtml?sms\\_ss=meneame](http://www.periodistadigital.com/periodismo/prensa/2010/06/23/unidad-editorial-elmundo-perdidas-quiebra-cuentas.shtml?sms_ss=meneame) [13-04-2016]

<sup>19</sup> [http://blogs.periodistadigital.com/periodismo.php/2006/09/14/el\\_plan\\_secreto\\_para\\_acabar\\_con\\_abc](http://blogs.periodistadigital.com/periodismo.php/2006/09/14/el_plan_secreto_para_acabar_con_abc) [13-04\_2016]

interesante de cara a futuras investigaciones. En esta tesis hemos recurrido a fuentes del periodismo para construir el CCF y a los métodos de la lingüística para analizarlo. Los dos siguientes capítulos los dedicamos a analizar estas dos disciplinas. En el capítulo 4 abordaremos la metodología y resultados que darán paso a la exposición de las conclusiones en el apartado final.



## CAPÍTULO 2: LOS GÉNEROS PERIODÍSTICOS

### 2.1 ¿Qué es el periodismo?

El periodismo es el oficio o la profesión que tiene por finalidad ofrecer la información para que la sociedad conozca lo que ocurre a su alrededor. Los medios de comunicación de masas son el canal que utilizan los periodistas para desarrollar esta labor informativa que, en ocasiones, se combina con el entretenimiento sobre todo en los medios audiovisuales. La importancia de los periodistas radica en que cumplen una función primordial en la sociedad actual como difusores de la información y en el desarrollo del deber de vigilancia del poder. Sin las aportaciones a la sociedad de los medios de comunicación y sus profesionales sería muy complicado entender lo que pasa a nuestro alrededor. Saperas (1987: 10) establece los fines y orientación de los medios de comunicación:

Los medios de comunicación actúan como instituciones mediadoras entre la población y la realidad, entre la población y aquellas instituciones que protagonizan los procesos de decisión pública.

Los medios de comunicación de masas son aquellos que dirigen su mensaje a un colectivo, es decir, a un colectivo de lectores, de oyentes, espectadores o internautas (Figura 2.1). Abordan una función algo más compleja que el simple procedimiento comunicativo entre un emisor y un receptor. Esto último sería lo que ocurre, por ejemplo, en una conversación telefónica, donde la relación comunicativa se lleva a cabo entre dos personas. En el caso de los medios de comunicación de masas el mensaje llega a los receptores a través de la prensa escrita, la radio, la televisión o los ordenadores. Entre finales del siglo XX e inicios del XXI, se han incorporado con fuerza los dispositivos móviles que han ido ganando terreno a los medios tradicionales no sólo en la difusión, sino también en la creación y elaboración de las noticias. Ponemos como ejemplo los medios digitales y los blogs, que han adaptado la inmediatez que caracterizaba la emisión radiofónica, la posibilidad de entretenimiento de la televisión y el espacio para el análisis propio de la prensa escrita. Todo esto lo veremos con más detalle cuando analicemos los códigos que caracterizan los distintos medios de comunicación. Pero retomamos el tema de los canales para establecer que son las vías que utilizan los periodistas o técnicamente emisores para comunicar. El mensaje es la información que llega al receptor y por eso hablamos de medios de comunicación de masas.

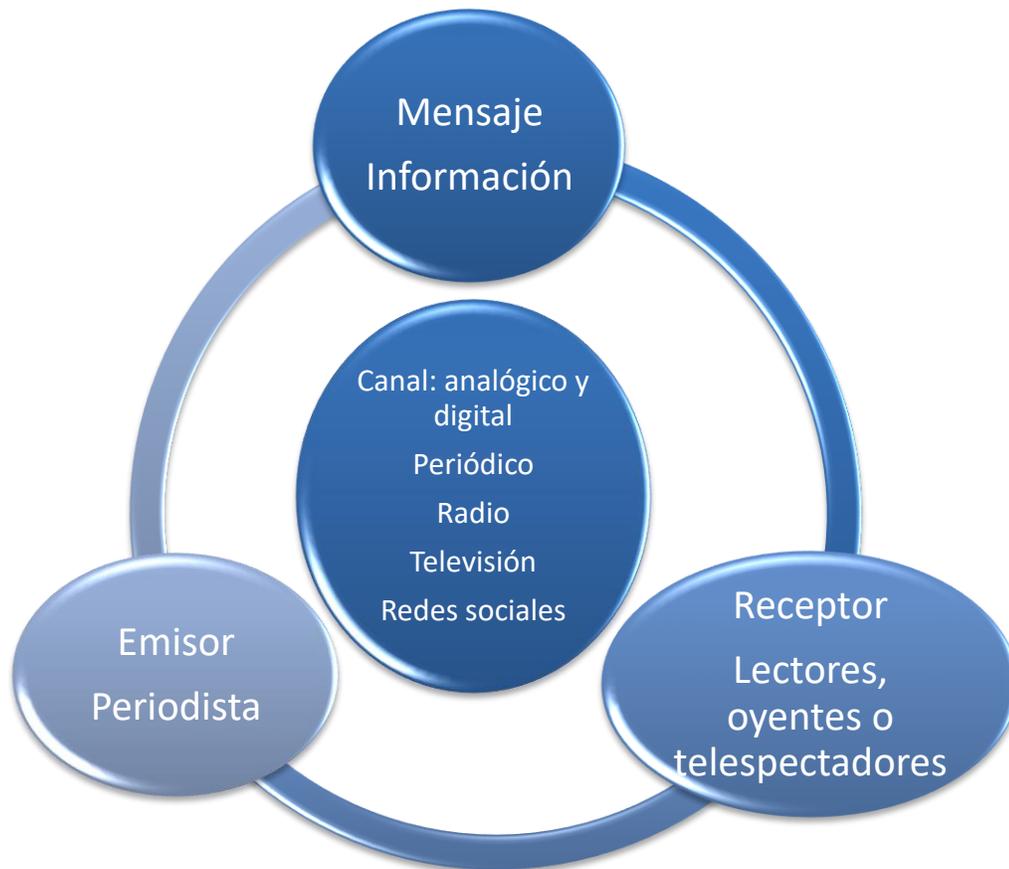


Figura 2.1: Esquema básico de la comunicación

Ya hemos señalado anteriormente que los medios de comunicación han aportado un gran avance en la difusión de la cultura y de la información y en la posibilidad de que el receptor pueda participar en la vida pública. Por eso cumplen un papel destacado en la sociedad. Hay que tener en cuenta, sin embargo, algunos aspectos negativos en su concepción tales como la simplificación de los contenidos para adaptarlos al público medio. El concepto del receptor como una masa homogénea y no individualizada puede afectar a la elaboración y resultado final del mensaje y eso lo vamos a ver en esta investigación, ya que el mensaje de la información sobre economía en la etapa de crisis analizada ha de transformarse en algo cercano al lector para servir a una población poco acostumbrada a este tipo de datos. Además, tenemos que tener en cuenta que el sometimiento a la sociedad de consumo de los medios o la capacidad que tienen para imponer una determinada ideología puede también condicionar el resultado final de la información. De ahí que hayamos optado por analizar dos medios que presentan planteamientos opuestos en cuanto a su ideología. Ahora veremos el sentido de la célebre

máxima acuñada por el profesor de literatura inglesa, crítica literaria y teoría de la comunicación Marshall McLuhan; *el medio es el mensaje*.

Nos detenemos en el trabajo de McLuhan (1996: 28) que analiza los medios de comunicación en su vertiente social, vinculada al ser humano. En la definición, en la que establece que el medio es el mensaje, quiere mostrar la importancia social del canal de información a través del cual conocemos las noticias.

En una cultura como la nuestra, con una larga tradición de fraccionar y dividir para controlar, puede ser un choque que le recuerden a uno que, operativa o prácticamente, el medio es el mensaje. Esto significa que las consecuencias individuales y sociales de cualquier medio, es decir, de cualquiera de nuestras extensiones, resultan de la nueva escala que introduce en nuestros asuntos cualquier extensión o tecnología nueva.

La naturaleza del canal, del emisor y el receptor determinará los fines de los diferentes medios de comunicación. El término medio de comunicación engloba todo procedimiento que un emisor emplea para establecer una relación comunicativa con un receptor. El acto comunicativo se puede establecer entre dos personas, micro grupos, macro grupos y entre canales naturales o artificiales. Ya hemos dicho que los medios de comunicación son aquellos que se dirigen a un colectivo y que tradicionalmente el receptor del mensaje es un colectivo donde la comunicación es unilateral. Esta última característica ha evolucionado en los últimos años con el desarrollo de internet, que ha permitido a la audiencia, al público o los lectores interactuar con los medios de comunicación ante la publicación de una noticia, una entrevista o un reportaje. A la radio, la prensa escrita y la televisión de siempre se han sumado en los últimos años no sólo los periódicos digitales con elementos multimedia, sino también las redes sociales, que ya comienzan a reivindicarse como canales generadores de información.

Los medios de comunicación han evolucionado a la velocidad que lo ha hecho la tecnología y han crecido según se ha producido el desarrollo económico. La prensa escrita alcanzó su edad de oro entre finales del siglo XIX y principios del XX en Europa y sobre todo en Estados Unidos, con el auge de la sociedad industrial. Sin embargo, los periódicos escritos iniciaron su declive no con la invención y desarrollo de la radio, ni con la aparición de la televisión, sino con la llegada de internet.

Retomamos la idea de que la finalidad primordial de los medios de comunicación es formar e informar para presentar las principales características que diferencian a cada uno de los medios o canales por los que transcurre la información. La irrupción de la televisión a principios del siglo XX agregó al objetivo de la formación e información un nuevo fin: el entretenimiento<sup>20</sup>. A estos tres podemos añadir el de la creación de opinión, más propia de la prensa escrita, que ofrece un formato más adecuado para el desarrollo y el análisis. Esta capacidad de influir en el público es la razón por la que los medios de comunicación han pasado a ser considerados el cuarto poder<sup>21</sup>. El cuarto poder es la última pata de los tres poderes que tradicionalmente caracterizan el constitucionalismo: el poder Ejecutivo, el Legislativo y el Judicial. Esta atribución inicialmente vinculada a la prensa escrita surgió en los años previos a la Revolución Francesa por la influencia extraordinaria que adquirieron los textos escritos en la sociedad del siglo XVIII. Se estaba gestando el influjo que con el tiempo llegarían a tener la televisión, la radio e internet.

### 2.1.1 Los códigos de los medios de comunicación de masas

Los medios de comunicación de masas con tradición periodística son la radio, la prensa escrita y digital y la televisión. Ya hemos dicho que las redes sociales y plataformas digitales han aprovechado el camino hecho para conectar con el público. Vamos a analizar sus principales características y los códigos de comunicación que emplean.

**La radio.** Tradicionalmente, ha sido el medio de comunicación más inmediato ya que utiliza códigos auditivos. Gracias a la tecnología y la posibilidad de retransmisión simultánea a bajo coste su característica ha sido y sigue siendo la inmediatez. El avance de las redes sociales a través de los dispositivos móviles ha ganado terreno a la radio tradicional en su capacidad de

---

<sup>20</sup> Las primeras emisiones públicas de la televisión las realizó la BBC en 1927. Le siguieron la CBS y la NBC en Estados Unidos en 1930. La primera emisora con programación y horario regular fue creada en 1930 en Berlín.

<sup>21</sup> La creación de la expresión “cuarto poder” se atribuye al escritor, político y orador anglo-irlandés, Edmund Burke (1729-1797). Con ella da una prueba, casi profética, de perspicacia política ya que en aquel momento la prensa no había logrado, ni siquiera en Inglaterra, el extraordinario poder que alcanzaría más tarde en todos los países libres. En tiempos modernos muchos críticos consideran que la prensa no se limita a reflejar la opinión pública, sino que puede crear esa misma opinión pública, proporcionando la casi totalidad de la información con la que ésta cuenta en cualquier momento dado, siendo uno de los generadores de *mainstream* (lo traducimos como corriente generalizada o intereses comunes).

informar de manera inmediata. Retomaremos este asunto, más adelante, cuando veamos el papel de los nuevos medios.

**La televisión.** Combina los códigos auditivos con los visuales. Éste último es un elemento fundamental. La inmediatez que caracteriza tanto a la radio en el caso de la televisión pasa a un segundo plano. Aunque son posibles las retransmisiones en directo<sup>22</sup>, como es el caso de eventos deportivos o de carácter social, lo normal es que la información esté previamente editada a través de los vídeos. Si algo caracteriza a este medio es su capacidad para entretener.

**La prensa.** La prensa escrita ha utilizado tradicionalmente el soporte papel, sobre el que se ha desarrollado el código escrito para presentar las informaciones. La invención de la imprenta moderna en 1440 a cargo de Johannes Gutenberg supuso un salto decisivo en el posterior desarrollo y difusión de los periódicos. Frente a la radio (inmediatez) y la televisión (entretenimiento) la prensa escrita ofrece mayor capacidad de análisis y reflexión sobre la información transmitida por los medios audiovisuales.

La evolución de la imprenta de Gutenberg a la imprenta electrónica y la digital ha permitido la aparición de nuevos medios de comunicación mucho más rápidos y que permiten la difusión de la información con una mayor rapidez e inmediatez, sólo asociada a la radio, y la posibilidad de ofrecer contenidos visuales editados, que únicamente podían dar las televisiones. Los nuevos medios de comunicación con soporte digital han aparecido en un momento de cambio en el que el lector solicitaba una comunicación más veloz. Los periódicos digitales se caracterizan además por no borrar por completo los anteriores sistemas y estructuras, sino que las superponen. Esta última evolución de la prensa escrita a la digital ha introducido la inmediatez y la simultaneidad a la retransmisión de las noticias escritas al introducir videos. Internet ha aportado a todos los medios la globalización como sinónimo de expansión de proceso creciente de redes complejas entre sociedades, culturas, instituciones e individuos en el ámbito internacional.

La evolución de las nuevas tecnologías y la investigación de nuevos campos han ido cambiando la forma de los mensajes que se transmiten hacia receptores. Las tradicionales fronteras desaparecen ante el crecimiento de los imperios globales de los medios de comunicación. Aquí

---

<sup>22</sup> por ejemplo, Ana Blanco se enfrentó al telediario más largo de la historia de TVE, el 11 de septiembre del 2001.

cobra importancia el concepto de “aldea global”<sup>23</sup> acuñado por el sociólogo canadiense Marshall McLuhan y que era totalmente premonitorio a mediados del siglo pasado.

### 2.1.2 Claves de la redacción periodística

Volvemos a los clásicos para comprender los conceptos básicos del Periodismo como profesión y como método de transmitir la información a la sociedad. Vamos a analizar a continuación qué es noticia y porqué se escoge un asunto determinado como información de interés. Vamos a mirar de inicio como han venido funcionando los medios tradicionales ya que consideramos los nuevos medios como una consecuencia o derivación de la prensa escrita, la radio o la televisión. Y el primer concepto a esclarecer es el de redacción periodística. Martínez Albertos (1983: 27) cita a Janowitz y a Schulze (1969) para definir la redacción periodística como una rama de las Ciencias de la Comunicación. En su definición provisional la establece como:

La ciencia que se ocupa del estudio de unos determinados signos –naturales y técnicos– ordenados en una unidad de pensamiento con el fin de transmitir datos e ideas de interés general a través del periódico o de cualquier otro medio de comunicación de masas; o bien, de otra manera, utilizando procedimientos técnicos: Prensa, Radio, Cine, etcétera.

Mucho antes del desarrollo de la era digital Martínez Albertos ya establece los periódicos, la radio, la televisión como medios de comunicación de masas y por ello son considerados una especie de espejo que refleja la información sobre cualquier parte del mundo, aunque tengan más relevancia las noticias de interés en los países del primer mundo, que otras acontecidas en los países subdesarrollados. Pero, ¿qué convierte un hecho en noticia de interés para la sociedad? Aquí cobra importancia el término *Agenda Setting*<sup>24</sup> de los medios de comunicación y

---

<sup>23</sup> El término aldea global se remonta a la década de los 60 del siglo XX y aparece en algunos de los libros de Marshall McLuhan, *The Gutenberg Galaxy: The Making of Typographic Man* (1962) y *Understanding Media* (1964). Posteriormente dio título a otro de sus libros *Guerra y Paz en la Aldea Global* (1968).

<sup>24</sup> La teoría de la *Agenda Setting* de McCombs y Shaw hace referencia a cómo los medios influyen en el público directa o indirectamente dando la relevancia o el espacio informativo a temas o cuestiones que los medios eligen. "Su nombre metafórico proviene de la noción de que los mass-media son capaces de transferir la relevancia de una noticia en su agenda a la de la sociedad" (McCombs, 1996: p.17). <https://es.scribd.com/doc/284445/TEORIA-DE-LA-AGENDA> [15-04-2016]

el concepto de *gatekeeper*<sup>25</sup> (el que filtra). En el caso que hemos analizado, la crisis financiera en el año 2012, vemos cómo cualquier información que tiene que ver con la economía se convierte en tema de agenda en los medios, es decir, cobra una presencia destacada de manera diaria y prácticamente obligada. Los editores de los diarios estudiados son los que ejercen como filtros de las noticias de este ámbito para suministrar los datos a la audiencia. En el caso concreto de la información financiera en los medios de información general, estamos hablando de un tipo de noticias o artículos de opinión que precisan de un trabajo extra por parte de sus autores, que han de hacer comprensible la información y para ello filtran los datos para dar forma a sus crónicas, editoriales, entrevistas, etcétera. Más adelante en el punto 2.3.4 analizaremos las funciones de la *Agenda Setting* y el *Gatekeeper* como mecanismos de selección de noticias en los medios de comunicación. Seguir a diario o de manera periódica la actualidad en los medios de comunicación es la forma más práctica que tiene el público de conocer lo que sucede y de buscar o encontrar respuestas o explicaciones. Dovifat (1959) y Fattorello (1969) coinciden en que la información periodística debe de contener las características básicas de actualidad, periodicidad y difusión. Las informaciones periodísticas ayudan al lector a conocer los hechos y que cuanto más informado esté un lector, un oyente o un espectador más capaz se sentirá para tomar decisiones, de seleccionar la variada oferta informativa y de escoger a sus representantes. La comunicación periodística ha de contener:

- novedad de los hechos
- actualidad
- diversidad e interés o narrar algo atípico

En lo referente a la forma de expresión, la información puede llegar al público a través de las secciones de la información general y de las páginas de opinión. En el primer caso, en el de la información general, los hechos deben ser narrados con objetividad. El autor de las noticias ha de abstenerse de emitir juicios de valor. En cambio, en la sección de opinión los medios de comunicación aprovechan esta plataforma para establecer su línea editorial a través de las informaciones que difunden.

---

<sup>25</sup> El término *gatekeeper* acuñado por Kurt Lewin en 1947 hace referencia al emisor o al autor o editor que se encarga de filtrar los datos del entorno y dar forma a las noticias.  
<http://www.nap.edu/read/9566/chapter/8> [15-04-2016]

La función valorativa o de opinión exige profundidad en el tratamiento de los contenidos y la utilización de códigos extralingüísticos como la tipografía, gráficos, fotos, etcétera, además de los puramente lingüísticos que presuponen en el lector referencias y conocimientos previos y abordan temas diversos y exigen un conocimiento léxico amplio.

Dentro de cada una de las secciones mencionadas, la de la información general y la de opinión, existe otra subdivisión en géneros periodísticos, en la que profundizaremos en el punto 2.2 e iremos al detalle de los géneros de información general y los de opinión.

Añadimos otra división de tipos de periodismo, en relación al tema que se aborde:

- Cultural
- Deportivo
- Político
- Social
- Económico
- Internacional
- Entretenimiento

Hay tantas clases de periodismo como temas en la sociedad. Un ejemplo claro de esto son las secciones en las que está dividido un periódico. Los periodistas que trabajan en cada una de estas secciones tienen una intención clara, tratan de reflejar al máximo la sociedad del momento.

- Economía
- Nacional
- Internacional
- Política
- Sociedad
- Deportes

Martínez Albertos (1983) define los géneros periodísticos como las diferentes modalidades de creación lingüística destinadas a ser canalizadas a través de cualquier medio de difusión colectivo y con el ánimo de atender a los dos grandes objetivos de la información de actualidad: el relato de los acontecimientos y el juicio valorativo que provocan tales acontecimientos.

Dovifat (1959) contabiliza hasta tres modalidades distintas que pueden ser consideradas como estilo diferenciado de los otros. Según este autor, hay aspectos en la presentación de los temas periodísticos que prácticamente tienen una expresión propia, determinada siempre por la obligación de hacer la lectura interesante y cautivadora. Las clasifica como el estilo informativo, el estilo de sollicitación de opinión y el estilo ameno.

No podemos olvidar que los medios de comunicación tradicionales se han caracterizado por la unilateralidad. Al dirigirse a un macro grupo no permitían la retroalimentación. El desarrollo de las redes sociales ha permitido una evolución en esta función, con la aparición de comentarios o *retweets* sobre las informaciones publicadas, que sirven a los medios para conocer los gustos o preferencias del lector, oyente o espectador.

### **2.1.3 Los principales rasgos del lenguaje periodístico en sus diferentes estilos**

La comunicación periodística se ha visto influida por una serie de factores. La cada vez mayor presencia de lenguas extranjeras en la sociedad, la influencia de términos utilizados en el ámbito político, judicial o económico o el avance de las nuevas tecnologías se nota y, mucho, en los textos periodísticos. A estos factores se añade la más que posible intencionalidad del mensaje o el descuido del periodista emisor en los usos lingüísticos. Teniendo en cuenta todo esto, podemos dividir los principales rasgos del lenguaje periodístico en tres: rasgos morfosintácticos, léxicos y retóricos.

- a. Rasgos morfosintácticos.** Son aquellos que muestran en el lenguaje periodístico una tendencia al alargamiento de las oraciones a través de perífrasis, aposiciones, incisos, frases explicativas, locuciones adverbiales, preposicionales y conjuntivas. Es más abundante la voz pasiva y el sujeto tiende a colocarse al final. El estilo directo se mezcla con el indirecto y es habitual el empleo de barbarismos. Veamos algunos ejemplos extraídos de nuestro corpus:

- *El Gobierno está gestionando este problema* (perífrasis)
- *Sería una decisión errónea revalorizar las pensiones* (condicional con valor de posibilidad)
- *Gobierno Rajoy* (supresión de preposiciones)

**b. Rasgos léxicos.** Hacen referencia al uso de anglicismos, de calcos semánticos cuando se utiliza por ejemplo *crédito* en lugar de rótulos o firmas, y al utilizar siglas y acrónimos:

- *UNESCO*
- *OTAN*
- *SIDA*

Eufemismos:

- *Medidas de ajuste* en lugar de *recortes*

Formación de nuevas palabras por derivación o composición:

- *Eurozona*

**c. Rasgos retóricos.** Las metáforas, metonimias, personificaciones, hipérbolos y otras figuras retóricas son bastante comunes en los textos periodísticos. Veamos algunos ejemplos más de nuestro corpus:

- *Estrangulamiento del crédito*
- *Rajoy movió bien sus piezas*
- *Hasta que el BCE se convierta en supervisor*

## 2.2 Los géneros periodísticos

Ya hemos avanzado que los medios de comunicación utilizan distintos géneros para hacer llegar al lector o al público los acontecimientos. Los géneros periodísticos se subdividen a su vez en informativos y de opinión. El CCF que hemos elaborado para esta investigación cuenta con ejemplos de cada uno de estos géneros periodísticos que, a continuación, pasamos a enumerar.

Las noticias generalistas o de carácter informativo aparecen en los diferentes medios de comunicación a través de los siguientes géneros periodísticos:

- Noticia
- Reportaje
- Crónica
- Entrevista

En la sección de opinión la información se transmite a través de:

- Editorial
- Artículo de opinión o de fondo
- Columna
- Cartas al director

Los periodistas que trabajan para la sección de opinión, como los que lo hacen en la de la información general cuentan, por tanto, con diversos géneros a través de los cuales transmiten un hecho noticiable. En los periódicos de referencia, en las cadenas de televisión de ámbito nacional o las emisoras de radio suele haber profesionales especializados en géneros concretos. Sin embargo, no todo el monte es orégano y en la mayoría de medios, de menor tamaño, es mucho más frecuente la figura del periodista todo-terreno capaz de firmar un día una crónica, al día siguiente elaborar un artículo de opinión y ser capaz de realizar una entrevista, por ejemplo.

Martínez Albertos (1997) y Foncuberta (1993) coinciden a la hora de definir los elementos básicos del estilo informativo: concisión, claridad y una construcción lingüística que cautive la atención del lector. Pero, ¿qué convierte un hecho en noticioso? La lista que propone Warren (1975) establece que ha de contener los siguientes ingredientes:

- Actualidad
- Proximidad
- Consecuencias
- Relevancia personal
- Suspense
- Rareza
- Conflicto

- Sexo
- Emoción
- Progreso

Ismael Herraiz<sup>26</sup> recoge estas diez características de la escuela norteamericana a la que pertenece Warren y añade el factor subjetivo en función de los lectores particulares de cada publicación y que varía a la hora de comparar el tratamiento concreto que una u otra publicación da a los acontecimientos de un mismo día. Es por esta razón por la que encontramos ciertas noticias que abordan determinados temas en unos medios y total silencio en otros medios de comunicación.

Marro y Dellamea (1993) detallan que el análisis de los textos permite evaluar qué características de los textos pueden ayudar al lector a desentrañar los significados que quiere dar el autor y a recordarlos después de la lectura. Ciertos aspectos de las estructuras textuales influyen tanto en la cantidad como en el tipo de información que el lector recuerda y permiten predecir dónde ocurrirán posibles distorsiones, omisiones, adiciones, sustituciones y reestructuraciones sobre la lectura, según las investigaciones realizadas en el análisis de los textos. Las investigadoras sostienen que la estructura de los textos es fundamental para facilitar su comprensión ya que el lector deposita ciertas expectativas sobre la organización de la información para distintos tipos de textos que influyen, a su vez, en el procesamiento de la información y en la representación del significado de un texto. Es como, podríamos decir, ocurre muchas veces en la cocina moderna donde la comida entra por la vista. Las lentejas son lentejas, pero elaborarlas y presentarlas de manera apetecible puede garantizar que el plato se acabe o, por el contrario, quede intacto.

Los lectores de periódicos utilizan esquemas que guían sus expectativas sobre la organización de una noticia y esperan encontrar un resumen en la parte superior, el título, la entrada y el texto que

---

<sup>26</sup> Ismael Herraiz Crespo (Madrid, 1913-1969) fue periodista español alumno de la Escuela de Periodismo del diario *El Debate* (1932). Redactor de la *Agencia Logos* y del *Diario Ya* (1935-1936), de *Arriba* del que llegó a ser corresponsal en Alemania e Italia (1940-1941), subdirector (1944-1948) y director (1948-1956). Corresponsal de *La Vanguardia* en Viena (1960) y Jefe de Sección de *El Alcázar* (1968-1969). Recibió los Premios Nacionales de Periodismo José Antonio Primo de Rivera (1946) y Francisco Franco (1966). Escribió también bajo los pseudónimos de Carlos Crespo y Gaspar Ledesma.

contesten las cinco W (qué, quien, cuando, cómo, dónde y porqué). Más adelante abordaremos las cinco W, que caracterizan un texto periodístico bien redactado, y cuál es su origen. Ahora simplemente nos vamos a quedar con que Marro y Delamea (1993) consideran que la información debe presentarse en orden decreciente en cuanto a la importancia de los datos y contener la respuesta a estas cinco W, o preguntas, en las primeras líneas. Esta estructura es lo que se conoce como pirámide invertida en la que el interés de la información va de más a menos. Los datos más importantes ocupan las primeras líneas y según avanza el relato la información aportada cae en importancia. Existe otra fórmula del relato, la del yunque o martillo, en la que la estructura sigue la narración de los hechos por su mismo orden.

Para Marro y Dellamea (1993), la estructura de las historias y las expectativas y conocimientos que el lector tiene sobre ellas son cruciales para su comprensión. Argumentan que las historias desordenadas y las que omiten información afectan a la lectura y recuperación de la información. Si el autor del texto periodístico presenta la información de manera ordenada y se coloca cerca del lector a través del uso de la segunda persona del plural, facilita la comprensión. En cambio, con temas en los que el lector tiene dominio no apelará demasiado a la estructura para reconstruir la información. No obstante, cuando el tema es nuevo, la estructura de la información está destinada a desempeñar un papel fundamental para la comprensión final del texto.

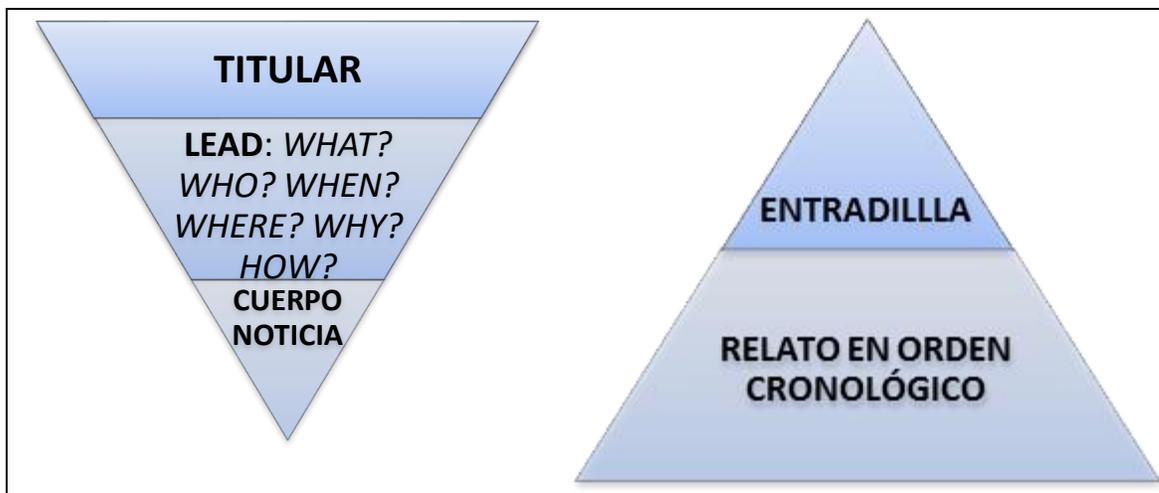


Figura 2.2: Estructura de la noticia

### 2.2.1 Los géneros informativos: la noticia, la crónica, el reportaje, la entrevista

Ya hemos visto que los géneros periodísticos se dividen en informativos y de opinión. Estos a su vez tienen sus propias clasificaciones. Veamos en qué consisten y qué caracteriza a cada uno de los géneros informativos.

**Noticia.** Es la esencia o la base de muchos de los géneros periodísticos. Sus principios básicos son brevedad, objetividad y claridad. Para que se produzcan otros géneros periodísticos como la entrevista, el reportaje, el artículo de opinión o el editorial primero tiene que producirse una noticia.

Para Martín Vivaldi (1973) la noticia es el género periodístico por excelencia que da cuenta de un hecho actual o actualizado, digno de ser conocido y divulgado, y de innegable repercusión humana. La noticia informa sobre los hechos de manera breve y directa. El periodista utiliza en su elaboración un lenguaje escueto, objetivo, ágil y conciso. Pero, ¿qué es lo que determina que un hecho se convierta en noticia? En las aulas de periodismo esta pregunta se contesta a través del siguiente ejemplo:

*Si un perro muerde a un hombre, no hay noticia. En cambio, si es el hombre el que muerde al perro, sí la hay.*

La noticia es algo que ha ocurrido, pero ha de ser nuevo e interesante. No suele ser habitual que un ser humano muerda a un animal y sí lo es, al contrario. Eso lo convierte en un hecho nuevo y, por tanto, noticiable. Es interesante además porque el lector quiere saber las causas de esta acción que podría llegar a afectarle.

Ya hemos analizado la conveniencia de presentar las informaciones de acuerdo a una estructura que facilite su comprensión (Marro y Dellamea, 1993) y la noticia tiene, por esta razón, una estructura determinada. Puede ser habitual que comience con un antetítulo inicial, aunque éste llegue en algunos casos a ser opcional. A continuación, figura el título de la noticia, que resume en poco más de una línea y de manera impactante la información. En tercer lugar, aparece un subtítulo, como complemento al titular. Después va la entradilla, o *lead*, que desarrolla de una

manera breve aquellos datos interesantes que no contiene el titular. Por último, el cuerpo de la noticia, o texto en el que el periodista narra lo acontecido con la crónica, reportaje o entrevista.

| ESTRUCTURA DE LA NOTICIA |   |
|--------------------------|---|
| Antetítulo               | Opcional  |
| Título                   | Obligatorio                                     |
| Subtítulo                | Opcional  |
| Entradilla o <i>Lead</i> | Debe contener las 5 W's                         |
| Cuerpo de la noticia     | Debe presentar estructura de pirámide invertida |

Tabla 2.1: Estructura de la noticia

Presentada ya la estructura de la noticia que mejor facilita la comprensión de la información volvemos a recordar aquello que planteamos en el apartado 2.1 que, si bien los medios de comunicación de masas han contribuido a la difusión de la cultura y de la información en la sociedad, la simplificación de los contenidos con esta estructura puede ser considerado como aspecto negativo en la concepción global de la información.

Siguiendo la estructura de la noticia, Martínez Albertos (1983: 302) divide en dos partes las informaciones. Hemos visto que la noticia está compuesta por un titular, un antetítulo o subtítulo, la entrada o *lead* y el cuerpo de la noticia. Para Martínez Albertos la primera parte la conforman el *lead* o párrafo inicial y la segunda el cuerpo de la información:

En el caso de la información, el párrafo inicial busca condensar sinópticamente toda la noticia en aquellos datos esenciales para una cabal comprensión de la misma. En el *lead* informativo se destaca la esencia o los datos más sobresalientes del acontecimiento del que se quiere dar noticia. Ese *lead* también se conoce como el *lead* del sumario y su técnica de realización aparece vinculada a la rutina profesional del quehacer periodístico que se designa con el nombre de la fórmula de las 5W's.

Para que una noticia incluya siempre los datos fundamentales, el periodista que la elabora debe seguir la técnica de las cinco uves dobles que ya establece Warren (1951), que recordamos es de la escuela norteamericana y que hemos introducido anteriormente. La w es la letra con la que

comienzan en inglés las preguntas que el texto debe de contestar en sus primeras líneas. Aquí ya hemos incluido el *How*, el *cómo*. Entraremos a detallarlo un poco más adelante.

|              |  |
|--------------|--|
| <b>WHO</b>   | ¿Quién/quienes es/son el/los protagonista/s de la noticia? Los sujetos |
| <b>WHAT</b>  | ¿Qué ha pasado? Los hechos   |
| <b>WHEN</b>  | ¿Cuándo ha pasado? El tiempo   |
| <b>WHERE</b> | ¿Dónde ha pasado? El lugar   |
| <b>WHY</b>   | ¿Por qué ha pasado? La causa   |
| <b>HOW</b>   | ¿Cómo? El modo   |

Tabla 2.2: Las cinco (o seis) W

Pero las cinco W, que se atribuyen a la escuela tradicional norteamericana de periodismo, tienen su verdadero origen en la retórica clásica. Según Foncuberta (1993: 69), ya en sus *Institutiones Oratoriae* Quintiliano se plantea las mismas preguntas que el periodista actual como método para responder a toda causa posible de un discurso.

Según la retórica, las fases de elaboración, producción y representación en público discurso comprenden: la *inventio* (o proceso productivo-creador); *dispositio* (el orden de las ideas y pensamientos que hemos encontrado gracias a la *inventio*); *elocutio* (traslada al lenguaje las ideas halladas en la *inventio*); *memoria* (memorizar el discurso); y *pronuntatio* (ejercitarse en la oratoria)

Foncuberta (1993: 69) considera que la estructura de la noticia directa ha sido concebida para suministrar la máxima información en el menor tiempo o espacio posible al lector o receptor de la información:

La mayoría de los acontecimientos son complejos, se producen dentro de un determinado contexto y pueden enfocarse de muchas formas diferentes. Además, no es lo mismo informar sobre el descubrimiento de un satélite artificial a un público masivo que a un sector especializado, ni se describe un accidente ferroviario del mismo modo para una agencia de noticias internacional (donde la noticia ocupará pocas líneas) que para el diario de la localidad en que se ha producido (donde posiblemente merecerá los titulares de portada y un amplio despliegue informativo).

En lo referente a esta estructura, enfatiza que *lead* y el cuerpo han de responder a:

- ¿Qué ha sucedido?

- ¿Quiénes son sus protagonistas?
- ¿Dónde, cuándo y por qué ha sucedido?

En cambio, la noticia de creación no sigue, según Foncuberta (1993), estas pautas porque incorpora nuevas formas de narración y lenguaje en periodismo. Las preguntas claves a la hora de redactar una noticia son:

- ¿Qué quiero decir?
- ¿A quién?

Algunos teóricos como Martín Vivaldi (1964) introducen el *how* como sexto elemento que se suma a las cinco preguntas que comienzan por *w* en inglés como asuntos fundamentales a tratar en la noticia. Martínez Albertos (1983) recoge esta matización ya que cree que el *cómo* puede aportar alguna claridad y precisión al típico esquema de las escuelas norteamericanas, que se apoyan básicamente en las cinco preguntas. Lo más importante es que el periodista conteste en las primeras líneas a las preguntas, si es en la entradilla o *lead*, mejor. De este modo, el informador, volvemos a insistir, sigue la estructura de la pirámide invertida en la que lo más interesante figura al principio y según avance el cuerpo de la noticia su atractivo va decayendo.

*El Libro de Estilo de El País* (1996) establece que, en un texto informativo, el uso de la técnica de la pirámide invertida (de mayor a menor interés) es conveniente pero no obligatorio. Siempre se ha de comenzar por el hecho más importante, que estará recogido a su vez en el título. No obstante, el párrafo siguiente puede constituirlo una frase que explique la entradilla o contenga los antecedentes necesarios para comprender el resto del artículo, rompiendo así la relación de los hechos. Lo mismo puede ocurrir con párrafos sucesivos. Esta técnica de la pirámide invertida para la elaboración de textos informativos, autores como Warren (1975), Foncuberta (1993) y Martínez Albertos (1983), la identifican como *el diagrama de la información*.

La mayor parte de los lectores se informan a través de la entradilla o *lead* o incluso del titular. Por eso, en ocasiones, elegir un buen titular que refleje de manera tan resumida la noticia es una tarea complicada que queda relegada a los jefes de sección. En su *Libro de Estilo El País* (1996: 40) determina que la entradilla es el primer párrafo de la noticia y añade acerca de su formato:

Esté diferenciada tipográficamente o no, contendrá lo principal del cuerpo informativo, pero no deberá constituir un resumen o un sumario de todo el artículo. Ha de ser lo suficientemente completa y autónoma para que el lector conozca lo fundamental de la noticia sólo con leer el primer párrafo.

Respecto al cuerpo de la noticia especifica el *Libro de Estilo* de *El País* que éste desarrolla la información con todo tipo de elementos complementarios; incluye los datos que no figuran en la entrada, explica los antecedentes y apunta las posibles consecuencias. El manual de *El País* añade que esto no significa que se puedan por ello incluir opiniones partidistas o juicios de valor sobre lo que se narra, más propios de las páginas de opinión.

En cualquier caso, Foncuberta (1993) recomienda recurrir a la contextualización de la noticia para ayudar a explicar el hecho que se está narrando. La profesora establece dos tipos de contextualización: la diacrónica que explica la sucesión de hechos anteriores a la noticia relacionados con ella, y la sincrónica que explica las circunstancias geográficas, políticas, sociales en las que se produce la noticia.

**Crónica.** Es la información interpretada sobre hechos actuales donde se narra un suceso. Son reportajes que se elaboran fuera de la redacción. Martínez Albertos (1983: 360) fija el sentido de la crónica en la siguiente definición:

Narración directa e inmediata de una noticia con ciertos elementos valorativos, que siempre deben ser secundarios respecto a la narración del hecho en sí. Intenta reflejar lo acaecido entre dos fechas: de ahí le viene su origen etimológico en la Historia de la Literatura.

En el periodismo español la crónica es un género híbrido entre el estilo informativo y el de sollicitación o estilo *editorializante*, a juicio de Martínez Albertos. La crónica se diferencia de la noticia por su intención valorativa y del reportaje por la carga objetiva que éste último debe mantener. Martín Vivaldi (1973: 128) recoge esta doble finalidad en la definición de crónica.

La crónica periodística es, en esencia, una información interpretativa y valorativa de los hechos noticiosos, actuales o actualizados, donde se narra algo al propio tiempo que se juzga lo narrado.

Como hemos dicho, las crónicas suelen escribirse en el lugar de los hechos a donde el periodista se ha desplazado para vivirlos en primera persona. Los medios de comunicación y las agencias de información tienen sus periodistas desplazados o corresponsales que viven en las principales

ciudades del mundo y que informan a través de sus crónicas de todo lo que acontece. *El Libro de Estilo* de *El País* (1996: 49) define la crónica como un estilo situado a medio camino entre la noticia, la opinión, y el reportaje. Este género puede emplearse:

Cuando se trate de informaciones amplias escritas por especialistas del periódico en la materia de que se trate, corresponsales en el extranjero, enviados especiales a un acontecimiento o comentaristas deportivos, taurinos o artísticos.

Este género se utiliza de manera asidua en las secciones de deportes, tribunales o sucesos. En este caso, son los periodistas especializados en un tema determinado quienes, al igual que los corresponsales en un país, elaboran las crónicas para su publicación. Gomis (1974: 51-52) traza una línea que marca la diferencia entre la crónica que cubre un lugar y la crónica que cubre un tema:

El cronista en ambos casos está puesto por el periódico para que cuente lo que pasa en el sector que tiene asignado. En una misma crónica caben diversos temas, porque la unidad de la crónica no es de contenido, sino más bien de marco espacio-temporal...la crónica está en el centro equidistante de la noticia y del editorial de la información pura y el puro comentario.

Martínez Albertos (1983) diferencia hasta diez tipos de crónicas entre las que incluye los relatos escritos en países o localidades distintas a la ubicación de la redacción de un periódico o cadena de radio o televisión. En su clasificación engloba:

- Crónica de sucesos
- Crónica judicial
- Crónica deportiva
- Crónica taurina
- Crónica local
- Crónica del corresponsal fijo en el extranjero
- Crónica de los corresponsales en provincias
- Crónica de enviado especial
- Crónica política
- Crónica de sociedad
- Crónica viajera

En lo referente a su estructura o su construcción, la crónica debe de contener, como bien recoge el *Libro de Estilo de El País*, (1996: 49) los elementos noticiosos.

Será titulada por regla general como una información y puede incluir análisis (por tanto, cierta opinión o interpretación). El autor debe, no obstante, explicar y razonar las interpretaciones que exprese, y construir su texto de modo que la información prime sobre la interpretación.

**Reportaje.** A diferencia de la noticia, un reportaje traslada no únicamente datos, sino que también olor, color y sabor. Y, ¿qué se conoce en periodismo como *color*? Recurrimos al Libro de Estilo de *El Mundo* (2002: 3-4):

¿Qué es el llamado *color*? Como su propio nombre indica, se trata de las pinceladas que dan vida a una información, y en particular a un reportaje de mayor calado, diferenciándolo de una fría y seca relación de hechos y datos. Esas pinceladas dan forma humana a los personajes y agregan elementos de atmósfera sobre el lugar, el momento, las voces, los ruidos o los olores que enmarcan la información. Es el mejor modo de proporcionar vigor y atractivo literario a un relato periodístico, enriqueciéndolo con el mayor número posible de elementos descriptivos.

El reportaje es la ampliación de una noticia, es mucho más extenso y no necesariamente ha de cumplir el requisito de actualidad que se exige a las noticias. Se elabora cerca de los protagonistas, con la colaboración de testigos y gente afectada. El Libro de Estilo de *El País* lo define como el género que combina la información con las descripciones e interpretaciones de estilo literario. Martin Vivaldi (1973: 348) alude al origen etimológico del francés *reportatge* para definir este género periodístico informativo:

Relato periodístico informativo libre en cuanto al tema, objetivo en cuanto al modo y redactado preferentemente en estilo directo. El reportaje es el género periodístico por excelencia, ya que todo lo que no sea comentario, crónica o artículo es reportaje que, en sentido lato, equivale a información. La mayoría de lo que se publica en un periódico es reportaje, salvo los géneros arriba indicados, y la estricta noticia que, casi siempre, se amplía y se desarrolla en crónicas y reportajes.

La escuela americana, con Warren (1979) a la cabeza, introduce el término *stories* y que une la información al reportaje. Su estructura más habitual es la del titular, seguida del *lead* o entradilla y el cuerpo.

En la elaboración de un reportaje para la televisión son muy importantes las imágenes captadas, en radio los sonidos y en los periódicos los testimonios recogidos por el periodista. El autor del

reportaje puede hacer uso de recursos estilísticos y de imprimir su estilo. Esta licencia se hace mucho más evidente a la hora de titular ya que el arranque de un reportaje puede ser desde divertido, a ilustrativo o sugerente. Puede reproducir frases hechas, refranes o títulos de películas. Estas circunstancias son las que permiten diferenciar el reportaje de la información pura y dura, aunque la estructura del reportaje sea la misma que la de la noticia. En el reportaje, la información debe de ir narrada de mayor a menor importancia, siguiendo la fórmula de la pirámide invertida y también debe de contestar las cinco o seis uves dobles.

**Entrevista.** Al margen de ser el instrumento básico que utiliza el periodista para elaborar sus informaciones en cualquiera de los géneros anteriores, es el género que permite conocer la personalidad, opiniones o manera de expresión del entrevistado. Su formato habitual es el de pregunta y respuesta. Es la mejor manera para obtener datos directos de la fuente. En esta modalidad se alternan las palabras textuales del entrevistado con descripciones o narraciones del periodista. Puede adoptar la forma de pregunta-respuesta. El Libro de Estilo de *El País* (1996: 50) matiza:

Las declaraciones mediante el diálogo con un personaje no siempre han de adquirir la forma de entrevista. La presentación con preguntas y respuestas debe reservarse para las entrevistas extensas y a fondo. En los demás supuestos, su presentación será la de un reportaje o, si tiene interés como información de actualidad, la de una noticia, en ambos casos con sus correspondientes entrecomillados.

*El País* diferencia tres tipos de entrevistas: la entrevista de declaraciones, la entrevista-perfil y la entrevista en suplementos. Vamos a ver en qué se diferencian y/o se parecen.

- I. **La entrevista de declaraciones.** Debe contar, según este manual, con una presentación del entrevistado en la que se refleje su personalidad y los motivos de porque se le interroga y a continuación las preguntas y respuestas señalizadas.
- II. **La entrevista perfil.** Admite una mayor libertad formal ya que no precisa la fórmula pregunta-respuesta y puede incluir comentarios, descripciones y datos bibliográficos del entrevistado.

- III. **La entrevista en suplementos.** Ya sean dominicales o diarios, es una mezcla de los anteriores.

### **2.2.2 Características lingüísticas de los géneros informativos**

La noticia, la crónica, el reportaje y la entrevista son géneros periodísticos informativos que deben cumplir una serie de características lingüísticas en el ámbito léxico-semántico y en el morfosintáctico. Desde el punto de vista periodístico, son muy importantes las aportaciones a los géneros informativos que han realizado autores como Warren (1975), Foncuberta (1993), Casasús y Nuñez (1991) y Martínez Albertos (1983). Otros autores como van Dijk (1990) han centrado sus esfuerzos en el acercamiento de los géneros periodísticos desde una perspectiva semiótica-estructural a través del análisis de los segmentos de la noticia y los procesos meta-cognitivos relacionados con la producción y comprensión de la información.

En el plano léxico, las palabras empleadas deben ser claras, concisas y correctas. El léxico en general ha de ser objetivo, aunque en ocasiones nos encontremos en las noticias publicadas un léxico valorativo y otros recursos selectivos de la información y entrecorillados y que hacen referencia a una declaración realizada por un protagonista y queda en ella su opinión. El redactor utiliza un léxico especializado, en ocasiones condicionado por la materia que trata sobre todo en informaciones concretas de las secciones de Política, Economía y Deportes. Para evitar repeticiones, el periodista recurre a los sinónimos.

En el plano morfosintáctico, en la redacción de las noticias predominan los sintagmas nominales que se construyen con adyacentes de intención aclaratoria. En los artículos determinantes es más común el uso de demostrativos y posesivos. En lo referente al tiempo verbal de los géneros informativos, es mucho más común el pretérito perfecto simple de indicativo. Son también abundantes los predicados transitivos con extensiones circunstanciales de lugar, de tiempo y de modo. Es además frecuente el uso de la subordinación adjetiva.

Los géneros periodísticos informativos suelen emplear el estilo directo para respetar la objetividad de la información. Y en cuanto a su construcción, es habitual encontrar errores en la construcción dada la rapidez con la que en ocasiones son elaboradas las noticias.

### 2.2.3 La superestructura del discurso periodístico

Hemos visto ya la estructura habitual que suelen presentar las noticias. El antetítulo es la pieza opcional que encabeza la información. A continuación, sigue el título, que es obligatorio en toda noticia, el subtítulo, también opcional, seguido de la entradilla, que ha de contener las respuestas a las seis W y por último el desarrollo de la información en el cuerpo de la noticia.

A esta estructura van Dijk (1990) introduce una serie de variables que denomina *superestructura del discurso periodístico*. El autor establece que cada ítem informativo está compuesto por un titular, aunque no muchos tengan un encabezamiento. Juntos preceden al resto del ítem informativo con la función de expresar los principales temas de la noticia. Funcionan como un resumen inicial. El lingüista concreta que la restricción semántica es obvia: el titular y el encabezamiento resumen el texto periodístico y expresan la macro estructura semántica.

Respecto al contexto y sus antecedentes, van Dijk (1990) explica que podemos tomar estas categorías como constituyentes de los esquemas periodísticos, aunque no sean exclusivas de las noticias. Los antecedentes en un discurso periodístico aparecen más tarde, después de la parte que se ocupa de los acontecimientos principales, que conforman otra categoría. Esta puede asimilarse en lo que previamente denominamos contexto, que se encuentra a menudo señalado por indicadores como *mientras*, *durante* o expresiones parecidas de simultaneidad.

Semánticamente, la información de contexto debe denotar la situación actual, que consiste en otros sucesos informativos concretos. Difiere de los antecedentes que tienen naturaleza histórica. La categoría *sucesos previos* se utiliza para recordar al lector lo que ocurrió previamente y es parte de las actuales circunstancias dentro de las que se incluye el contexto con una dimensión histórica y que, por consiguiente, no es el suceso principal.

Van Dijk (1990) advierte de que las consecuencias vienen determinadas por el valor informativo de los acontecimientos sociales y políticos. Mediante la discusión real o posible de las consecuencias, un discurso periodístico puede otorgar coherencia casual a los acontecimientos informativos. Las consecuencias, en ocasiones, son más importantes que los acontecimientos informativos principales y el hecho de llegar a la misma posición jerárquica o convertirse en asunto destacado queda reflejado en los titulares. Por ejemplo, en el caso concreto de la crisis

financiera, en el periodo analizado de 2012, hemos visto como España se planteaba pedir un rescate bancario a la Unión Europea. Tal rescate no llegó a solicitarse, pero sus consecuencias fueron motivo de análisis en numerosas ocasiones por la prensa de ese año.

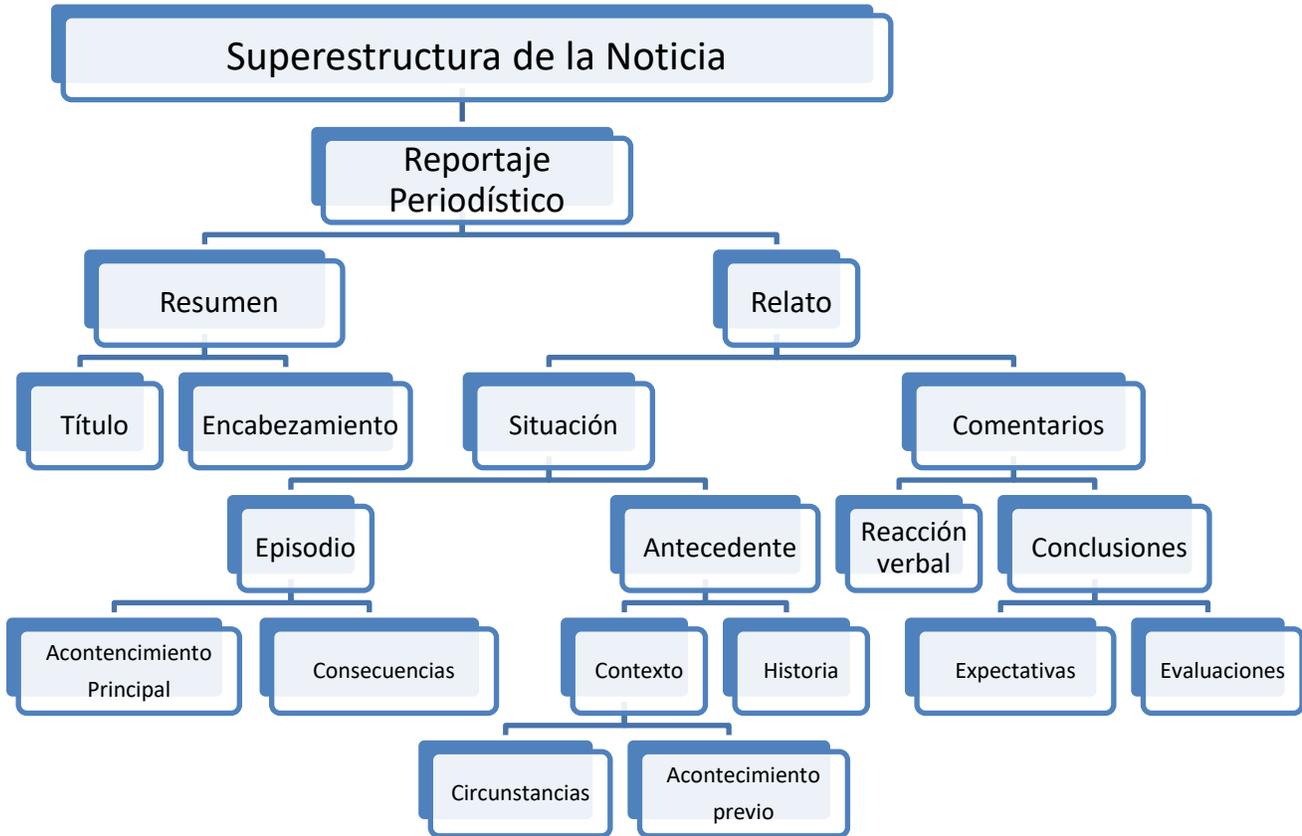


Figura 2.3: La superestructura de la noticia

#### 2.2.4 Los géneros de opinión: editorial, columna, artículo o tribuna y cartas al director

Hay informaciones que por su trascendencia traspasan las páginas de la información general y se cuelan en la sección de opinión. Son textos que afectan al hombre y a los temas de actualidad. Son sucesos que son debatidos y sometidos a la opinión de los medios de comunicación. El lector debe discernir entre los artículos de opinión y los exclusivamente informativos y de este modo saber si lo que está leyendo son datos o son opiniones. Los géneros de opinión que los medios de comunicación recogen de manera habitual son el editorial, la columna, el artículo o tribuna y las cartas al director.

### 2.2.4.1 Características lingüísticas de los géneros de opinión

Al igual que ocurre con los géneros informativos, los géneros de opinión presentan una serie de rasgos lingüísticos comunes en el plano léxico-semántico y en el morfosintáctico.

En el plano léxico-semántico se nota la utilización de un léxico culto, aparecen numerosos sustantivos abstractos y es notable el uso de recursos literarios como metáforas e hipérbolos. Se desarrolla así la función conativa y emotiva.

En el plano morfosintáctico comprobamos que suele ser abundante en los textos el empleo de proposiciones para apoyar su carácter argumentativo. Esto se consigue con las proposiciones causales, consecutivas y condicionales. La aparición de los verbos en el modo indicativo, condicional e imperativo dependerá del carácter expositivo.

Autores como Martínez Albertos (1983) establecen que el estilo de sollicitación de opinión, o estilo *editorializante*, no tiene, como ocurre con el estilo informativo, solera y personalidad autónoma dentro del mundo de las letras, sino que conecta directamente con la retórica tradicional, de la que deduce y aplica un buen paquete de recursos intelectuales mediante los cuales se consigue persuadir y conmover el ánimo de los lectores.

El informe *Committee on Modern Journalism* (1967: 605-606) detalla la finalidad del editorial de lograr el convencimiento de los lectores como función primordial:

El editorial es el celoso guardián contra la violación del derecho que tiene el público a saber. Aunque los editorialistas suelen trabajar en el anonimato, en todo el mundo periodístico no hay puesto más importante o respetado que el suyo. Es él quien fustiga el engaño, la injusticia, el desatino; es él quien trata de inspirar a los hombres y a las mujeres para que den lo mejor de sí mismos. El periodista ayuda a los lectores a saber. El editorialista les ayuda a entender.

Sobre la labor de convencimiento, Dovifat (1959: 129) añade que la labor de convencimiento con vistas a la formación de opinión se efectúa por medio de la fuerza probatoria del pensamiento y los hechos. Añade Martínez Albertos que este conjunto de formas de expresión periodísticas destinadas a conseguir el convencimiento de los lectores es lo que Dovifat (1959) denomina estilo de sollicitación de opinión. A juicio de Martínez Albertos (1983: 244), los géneros periodísticos de la sección de opinión conforman la conciencia del periódico.

Es obligación de los editorialistas pensar y escribir como si fueran la conciencia del periódico. El editorial –y por extensión todas las secciones del periódico que participan de la misma finalidad, es decir, lograr el convencimiento de los lectores- tiene como una de sus funciones más importantes.

Otros autores como van Dijk (1990), al hilo de la superestructura de la noticia antes detallada, establece que la categoría permite a los periodistas formular opiniones que no necesariamente son las suyas propias pero que han sido formuladas. La categoría de las reacciones verbales viene señalada por los nombres y los roles de los participantes periodísticos y por citas directas o indirectas de declaraciones verbales. Concluye con que suele situarse a continuación de la de sucesos principales, contexto y antecedente, al final del discurso periodístico. En relación a los comentarios, las opiniones y las evaluaciones del periodista y del medio de comunicación van Dijk (1990: 88) señala que:

Aun cuando muchos productores de noticias comparten la visión ideológica de que el hecho y la opinión no deben mezclarse, esta categoría aparece frecuentemente, a veces de forma indirecta. La evaluación caracteriza las opiniones evaluativas sobre los acontecimientos informativos actuales; la categoría de las expectativas formula consecuencias políticas o de otro tipo sobre los sucesos actuales y la situación. Puede predecir acontecimientos futuros.

Ya hemos visto que en los medios de comunicación los principales géneros de opinión se dividen en: editorial, columna, tribuna y cartas al director. Veamos a continuación cuáles son sus características informativas.

**Editorial.** Es el artículo que expresa la opinión del medio de comunicación en el que sale publicado y muestra su línea ideológica. El editorial<sup>27</sup> establece la postura del medio ante los acontecimientos. Se caracteriza por tratar temas de actualidad. Ocupa un lugar preferente en el periódico y define la línea ideológica de la empresa editora del medio. Nunca va firmado y su elaboración corre a cargo de un equipo de editorialistas. El texto que finalmente publicará el medio ha de ser el resultado del consenso del equipo completo, aunque haya sido escrito por un

---

<sup>27</sup> El editorial como género de opinión es masculino y lleva el artículo el delante. Si fuera femenino y llevara el artículo la delante tendría otro significado y sería la empresa encargada de la publicación y distribución de escritos.

solo autor, que puede ser o no el director del medio de comunicación. Éste es en última instancia responsable de todo lo que publique el medio en cuestión.

*El País* en su *Libro de Estilo* recoge la vinculación entre director y editorial. También lo hace Martínez Albertos (1983) al determinar que su elaboración es una labor delicada y comprometida para el medio de comunicación porque se confía a personas perfectamente identificadas con la línea política de la dirección o los propietarios de la empresa. Los editoriales son los textos que con mayor interés leen aquellos que quieren conocer el estado de la opinión pública de un país. Los gobernantes de un país encuentran en este género un punto de información muy importante acerca de la interpretación del medio sobre sus decisiones. Martínez Albertos (1983) se fija en Joseph Pulitzer<sup>28</sup>, célebre director de *The New World*, para afirmar que el lector debe conocer el punto de vista del periódico y considera inmoral cobijarse detrás de la neutralidad de las noticias. El *Committee on Modern Journalism* (1967) plantea que mientras el periodista ayuda a los lectores a saber, el editorialista les ayuda a entender.

Acerca de su estilo, Martínez Albertos (1983) aconseja que el editorial sea digno e incluso tenga tono algo mayestático. Con giros desenfadados, toques humoristas o desgarrados que pueden tener cabida en un comentario, un ensayo, una crítica, difícilmente pueden admitirse en el editorial. El editorial es para el profesor de periodismo la manifestación de la conciencia del periódico acerca de un problema particular. Si el periódico quiere mantener su prestigio y autoridad moral ante el público, debe presentar el editorial con esas notas exteriores y formales de dignidad y seriedad lingüística.

Mostaza (1966) y Martín Vivaldi (1972) establecen la importancia de un editorial para comprender mejor la realidad. El primero, en su *Enciclopedia del periodismo* detalla que si bien la noticia da el parte diario de lo sucedido, el editorial interpreta el sentido de ese parte o

---

<sup>28</sup> El editor estadounidense Joseph Pulitzer nació en Hungría en 1847. Férreo competidor de otro magnate de la prensa, William Randolph Hearst, dio origen a la llamada prensa amarilla y los periodísticos Premios Pulitzer. Fue pionero del estilo que mezcla la información con el entretenimiento. En 1883 adquirió el *New York World* con el que dirigió la atención hacia historias de interés humano, escándalos y sensacionalismo. En 1889 introdujo la tira cómica *The Yellow Kid* en su periódico. De ahí lo de prensa amarilla, como sinónimo del estilo periodístico que practicaba en su diario y que pasó a ser el primer periódico nacional en difusión, tras pasar de 15.000 a 600.000 ejemplares.

previene lo que en las profundidades de la colectividad humana se está fraguando y va estallar de un momento a otro. Por otro lado, Martín Vivaldi (1972) define el editorial como un “artículo periodístico, normalmente sin firma, que explica, valora e interpreta un hecho noticioso de especial trascendencia o relevante importancia, según una convicción de orden superior representativa de la postura ideológica del periódico”.

Respecto a su estructura, Martínez Albertos (1983) prefiere el inicio del editorial con una exposición de los hechos que van a dar pie al escrito y quedan la oportunidad al lector a conocer un juicio que puede orientarle. A continuación, se detallan los principios generales aplicables al caso expuesto y por último la conclusión que debe emitirse a la vista de los principios generales y los teóricos. Otros autores optan por el esquema del silogismo lógico con una premisa mayor general, caso concreto y la conclusión deducible.

La enunciación del tema, el planteamiento, es lo primero. Lo segundo es exponer sus implicaciones y sus consecuencias, o sea, operar con él, manipularlo, desmontarlo en piezas. Lo tercero es emitir un fallo o censura y adoptar, en consecuencia, una conducta, ofrecer una solución, trazar un rumbo. En sobradas ocasiones, esas tres faenas se embarullan y mezclan y el editorial degenera en enigma. (Mostaza, 1966: 184)

Aunque su estructura no suele ser fija, sí que responde a un esquema piramidal que tiene el momento más importante al final. En resumen, los editoriales están escritos de acuerdo con la siguiente estructura:

- Introducción: en ella se plantea el estado del problema.
- Cuerpo: el autor expresa una opinión.
- Solución: el texto aporta una solución al problema.
- Conclusión: el editorial hace hincapié en la cuestión principal.

El editorial tiene aspiraciones de objetividad, su tendencia ideológica introduce elementos subjetivos que acaban con esta objetividad. Se caracteriza además porque su título suele ser un sintagma nominal que muestra tendencias valorativas o creativas.

Veamos otros consejos adicionales a la hora de estructurar la opinión en un editorial:

- Iniciar el texto con una explicación objetiva. Incluir en esta explicación las cinco (o seis) *W* (*who, what, when, where, why* (y *how*)) en los hechos expuestos y utilizar las citas de las fuentes que sean pertinentes.
- Presentar el primer enfrentamiento. Al estar el autor en desacuerdo con ciertos puntos de vista, se pueden identificar a las personas a las que se opone o no representa. Para ello hay que utilizar hechos y citas que permitan exponer de manera objetiva las diferentes opiniones. Darle de este modo una posición fuerte a la oposición. No se gana nada en la refutación de una posición débil.
- Refutar las creencias de la oposición. Hacer uso de hechos y citas de personas que apoyan la misma línea ya que conceder validez a la oposición hará parecer racional al haber considerado otras opciones.
- Dar otras razones, originales. Defender su posición con razones de peso e incluso utilizar una alusión literaria o cultural que dé mayor credibilidad.
- Concluir con un golpe de efecto. Dar soluciones al problema o desafiar al lector a estar informado. Puede ser efectivo si se trata de una fuente esperada. Una pregunta retórica como conclusión puede ser efectiva.

El editorial ha de estar ligado a la actualidad. Eso es lo que le aporta carácter periodístico.

Santamaría (1990) le asigna las funciones de explicar los hechos, dar antecedentes, predecir el futuro y formular juicios. En el editorial se explica el tema al ofrecer una presentación articulada y relacionada de los acontecimientos con vistas a la comprensión del sentido que tienen. Los juicios de valor que incluyen sobre lo que está bien o está mal los convierte en guardianes de la conciencia pública.

Los manuales de periodismo establecen distintas clasificaciones de los editoriales en atención a la acción persuasiva que ejercen llevada a efecto con cierta carga de ideología y argumentos. De este modo, hay editoriales expositivos, explicativos, combativos, críticos, apologéticos, admonitorios y predictivos, principalmente.

*El Mundo* en su *Manual de estilo* (2002: 11-12) considera estos textos como los más importantes de cuantos se publican:

Los editoriales deben ser el género más cuidadosamente trabajado del periódico, tanto desde el punto de vista documental como desde el literario. Su objetivo es profundizar argumentalmente en un problema, bien tratando de defender un punto de vista, bien analizando diferentes puntos de vista y dejando el veredicto en manos del lector. El editorial debe apelar a la razón y no a las emociones; convencer, no enardecer; partir, más a menudo, de la duda que de la certeza. El editorial es como una bandera que el periódico despliega cada día, y ello obliga mucho. Los datos que se incluyan deberán estar inapelablemente confirmados. Las imágenes literarias deben ser el producto de una reflexión y de una reescritura concienzudas.

**Columna.** Este género periodístico de opinión se llama así porque regularmente adquiere este formato en la prensa escrita. Sus autores, por lo general, se denominan columnistas y no suelen ser especialistas en un determinado asunto, sino más bien periodistas o escritores que hablan de cualquier tema como harían en una conversación de amigos. Debe aparecer firmado y no necesariamente ha de tratar temas de actualidad. Martínez Albertos (1983: 389) define el comentario o columna como:

Artículo razonador, orientador, analítico, enjuiciativo, valorativo y con una finalidad idéntica a la del editorial. Se diferencia básicamente en que el comentario es un artículo firmado y su responsabilidad se liga tan sólo al autor del trabajo.

Las columnas suelen ir fuera de las páginas de opinión y es frecuente que cada sección del periódico cuente con su columnista que comente los acontecimientos más destacados. Al respecto Mostaza (1966: 181) afirma:

Lo que escribe el columnista vale por lo que valga su firma. Es una opinión individual que usa el periódico para expresarse. Además, la columna no siempre tiene finalidad orientadora: el *columnismo* se caracteriza por la variedad de contenido.

Es usual además que los periódicos publiquen su columna en la última página. El autor de la columna, a diferencia del editorialista, no tiene porqué adoptar una posición ante los hechos que narra. Tiene además mayor margen para expresarse, lejos del estilo mayestático del editorial. Puede utilizar en la columna giros y expresiones coloquiales, pero sin abandonar el estilo decoroso. El *Manual de Estilo de El Mundo* (2000: 10) resalta la creatividad literaria que aportan los columnistas a los periódicos:

Otra distinción importante reside en la posibilidad de que una columna fija, y sólo una columna fija, tenga un carácter claramente literario y creativo, a diferencia de otras que comentan más directamente la actualidad política, económica, cultural o social. Sólo muy excepcionalmente, en cambio, será puramente creativo el artículo firmado de carácter no regular. Por su propio

formato y por sus limitaciones, un periódico solamente puede dedicar un espacio limitado a la creación literaria, sin por ello renunciar del todo a esa tradición española tan arraigada, la del escritor de periódicos.

**Tribuna.** Son artículos estrictamente de opinión y que responden al estilo propio del autor. A diferencia de un editorial, los artículos de opinión son más extensos y siempre van firmados. Pueden estar escritos por algún periodista de la redacción, aunque lo más normal es que firme las tribunas o artículos de opinión alguna persona ajena a la redacción.

Una vez más recurrimos a *El País* y a su Libro de Estilo (1996: 53) para dejar claros algunos criterios de publicación de estos textos y otras colaboraciones de las páginas de opinión.

Todos los artículos de opinión llevarán, tras la última línea, un pie de autor –por desconocido que sea- donde se indique el cargo, título, militancia política (en su caso) u ocupación principal, o la que esté o estuvo relacionada con el tema abordado.

Desgranamos a continuación algunos de los pasos básicos a seguir para escribir un artículo de opinión:

- El autor debe plantear el artículo de opinión como un ensayo persuasivo. Tiene una opinión o un punto de vista sobre un tema y quiere convencer al lector acerca de ello
- Antes de escribir el artículo de opinión el autor debe investigar el tema a tratar, averiguar lo que está sucediendo y lo que sucedió en el pasado
- Conocer bien los hechos y ser capaz de referirse a ellos en su argumento
- Como ocurre con un abogado ante un jurado, el escritor ha de convencer a los miembros del jurado a creer que su cliente tiene razón. Por lo que es necesario presentar pruebas de todo lo que pueda para demostrar sus argumentos
- La estructura de un artículo de opinión es muy similar a la del editorial, pero no debe confundirse ya que el editorial carece de firma personal, mientras que el artículo de opinión o tribuna queda identificado con su autor. Al respecto el libro de Estilo de *El Mundo* (2000: 9) aclara que han de estar plenamente identificados de cara al lector:

Su clara identificación tipográfica reviste, claro está, particular importancia, recurriéndose a los títulos en letra cursiva y a la colocación de etiquetas fáciles de localizar, como son las palabras Tribuna libre u Opinión.

**Cartas al director.** Este género periodístico es la única oportunidad que tiene cualquier persona de hacer pública su denuncia o de expresar su opinión. Habitualmente llegan multitud de cartas a un medio de comunicación y es su cometido seleccionar aquellas que van a ser publicadas. En ocasiones, de las cartas al director surge también la posibilidad de obtener una noticia. En otras, las cartas atacan a un individuo o a una empresa por lo que es pertinente comprobar la veracidad del texto antes de publicarlo. En su *Libro de Estilo* (1996: 54) *El País* considera como cartas al director para publicar en la sección correspondiente:

Aquellas cartas firmadas y avaladas por el carné de identidad del autor, domicilio, teléfono, etcétera, de modo que puedan ser autenticadas.

Acerca de la veracidad de los hechos desarrollados en una carta, este medio siempre los comprueba antes de publicarla. Por regla general, las cartas suelen ser breves, en torno a 300 palabras y deben estar firmadas con el nombre y apellido o apellidos del remitente o al menos sus iniciales separadas por puntos.

*El Mundo* en su *Manual de Estilo* (2000: 11) abre al público a comentar en esta sección lo publicado en el medio y a asuntos políticos y sociales. En el caso de la crisis financiera, este tema generó también en los medios analizados respuesta en los lectores a través de las cartas al director.

### **2.2.5 El periodismo especializado: origen, definición, tipos y ejemplos**

El periodismo especializado surge de la necesidad del lector de satisfacer su interés por los avances en el conocimiento científico, técnico y cultural. Este tipo de publicaciones cumplen una función social, divulgativa y de mediación entre la ciencia y el resto de la sociedad. El avance del periodismo especializado coincide con la profesionalización de la labor periodística y al hecho de que esta actividad esté cada vez más ligada a un proceso productivo de elaboración de mensajes. El periodista-escritor de la información de carácter generalista va abriendo camino al profesional preparado que compete en un mercado en el que el público es cada vez más exigente. La era Internet y la irrupción de las nuevas tecnologías supone un destacado avance en la consolidación del periodismo especializado. Los medios de comunicación interactivos contribuyen a que los lectores y usuarios puedan satisfacer de una manera más amplia su

interés por profundizar más en una determinada materia. Los suplementos especializados de la prensa diaria compiten de este modo con las revistas especializadas.

El salto de la información general a la especializada transcurre dentro de un cambio de la sociedad de masas a la era en la que conviven distintas formas culturales y diferentes modelos de comunicación. El lector, el oyente, el espectador y últimamente el internauta, han pasado de informarse simplemente, a querer satisfacer la necesidad de contribuir mejor a la sociedad y participar en la resolución de problemas. Este tipo de periodismo surge para satisfacer esta necesidad. Es el caso de las publicaciones que abordan información de economía y que en esta investigación abordamos y como al inicio de este trabajo ya avanzamos han crecido de manera considerable durante la crisis financiera para satisfacer la necesidad de información por parte de la sociedad acerca de un tema con el que puede estar poco familiarizada. Analizamos a continuación las derivaciones de la prensa e información especializada.

#### **2.2.5.1 Información periodística especializada. Periodismo y prensa especializada.**

Recurrimos a los teóricos del periodismo (Fernández del Moral, 2004; Borrat, 1993; Martínez Albertos, 1983) para profundizar y así entender mejor las diferencias que existen entre:

- Prensa especializada
- Periodismo especializado
- Información periodística especializada

Martínez Albertos (1983) distingue entre la prensa especializada, - que es la que forman las revistas dirigidas a profesionales, especialistas, comunidad científica, técnica o industrial-, y el periodismo especializado. La prensa especializada puede tener o no periodicidad fija, está hecha por técnicos especialistas en una materia y está dirigida a especialistas. El periodismo especializado es el que encontramos en los diarios de información general en las secciones específicas y que va dirigido a todo tipo de público. Fernández del Moral (2004: 24) considera que es un periodismo que propone modelos comunicativos que hagan de las especialidades del saber o actividades profesionales materia periodística.

Borrat (1993: 81-82) a la hora de establecer el significado de información periodística especializada señala dos enfoques; el que identifica a la prensa especializada o de información selectiva (la de economía, deportes, etcétera) por oposición a la prensa de información general y el enfoque de los periodistas especialistas o elitistas, los científicos, los de investigación y los columnistas o articulistas que tienen rango de estrellas mediáticas. El autor define la información periodística especializada como la producción de textos periodísticos “caracterizada por la coherencia interna de esos textos, la correspondencia de sus afirmaciones con la realidad y la pertinencia de los conceptos, las categorías y los modelos de análisis aplicados”.

Atendiendo la función social de este tipo de publicaciones, que aludíamos antes, podemos establecer que la información periodística especializada es la disciplina que abarca la elaboración y difusión de textos informativos que persiguen la divulgación de la especializada del conocimiento experimental, social y humanístico en relación a la ciencia y de su interés, al gran público principalmente de manera narrativa, a través de los medios de comunicación.

Fontcuberta y Borrat (2006: 77-86, 220) conciben el periodismo especializado en función de del eje temático:

Al tratamiento en profundidad en los medios de comunicación de un determinado campo de conocimiento. Se refiere al periodismo político, económico, cultural, deportivo, etcétera. El concepto de periodismo especializado tiene un referente temático (la geografía, la segmentación demográfica, los medios especializan). Son, componentes complementarios y fundamentales de la auténtica razón de ser de la especialización; el ámbito temático.

El periodismo generalista está sujeto a rutinas productivas y la sumisión de la actualidad, las declaraciones, las fuentes y versiones oficiales, lo descriptivo. Es quizá, según Miralles Castellano (1998), el tipo de periodismo más impregnado del negocio periodístico sujeto al modo liberal. Sin embargo, el periodismo especializado tiene otros ritmos productivos, distinta temática no tan supeditada a la actualidad informativa y tiene mayor responsabilidad social y servicio público de información en profundidad.

### **2.2.5.2 Orígenes de la prensa económica**

El objetivo de esta investigación es la elaboración de un corpus sobre la crisis financiera con textos recopilados de publicaciones que abordan temas de economía. Conocemos por prensa económica aquellas publicaciones de actualidad cuyos contenidos informan prioritariamente de

temas y acontecimientos económicos en general. Arrese (2002) añade que este tipo de periodismo está dirigido mayoritariamente a personas con responsabilidad o interés en la toma de decisiones en ámbitos variados como la política económica, las finanzas, el comercio y los negocios. En el análisis de este tipo de publicaciones, surgidas a mediados del XVI, considera Arrese (2002) que alcanzan su consolidación en la última década del siglo XX, a través de los grandes conglomerados mediáticos. Fija el origen de esta prensa en las primeras redes de información privadas, las hojas de precios corrientes, publicación de noticias comerciales y de negocios en los periódicos de información general.

La depresión de 1873, la primera de las crisis cíclicas de carácter industrial, y la necesidad de la sociedad de la época de recurrir a información económica y la previsión de los mercados financieros fue el origen, en 1889, de *The Wall Street Journal*<sup>29</sup>. Es en la actualidad el diario económico y financiero de mayor tirada en Estados Unidos con una tirada de dos millones de ejemplares diarios. En el siglo XIX surgieron otras publicaciones económicas como *The Economist* y *Financial Times* que derivaron en la posterior aparición de las revistas especializadas de negocios como *Forbes*, *Business Week* y *Fortune*. Este modelo de periodismo especializado surge en Estados Unidos y se expande por países con menor tradición en el periodismo económico. Los cimientos de este nuevo modo de elaborar noticias y que se denomina periodismo especializado se consolidan a lo largo del siglo XX. Su consolidación estuvo acompañada de arduas jornadas de debate entre los profesionales de la información y los gobiernos por fijar en el nuevo modelo informativo especializado el cumplimiento de la responsabilidad social de los medios de comunicación, en relación con la política democrática, todo ello cuidando su independencia y comportándose como un servicio público, más que como una empresa privada (McQuail, 2002). En el debate (Orive y Fagoaga, 1974; Fernández del Moral

---

<sup>29</sup> Periódico internacional escrito en inglés que principalmente cubre temas de la economía estadounidense, noticias internacionales y problemas financieros. Su periodicidad de publicación en la ciudad de Nueva York es de seis días a la semana. Cuenta con varias ediciones internacionales y el periódico en circulación más grande de Estados Unidos. Comenzó a editarse en 1889 y ha alcanzado 2.4 millones de copias que incluyen unas 900 000 suscripciones digitales (marzo de 2013). Su rival principal en el sector de negocios es el periódico *Financial Times* con sede en Londres, el cual también publica varias ediciones internacionales.

Su nombre deriva de Wall Street, considerado el corazón del distrito financiero de Manhattan. Sus editores son Charles Dow, Edward Jones (que dan nombre al índice bursátil) y Charles Bergstresser. Ha obtenido el Premio Pulitzer treinta y cuatro veces. En 2007 lo hizo por su reportaje sobre opciones en acciones retroactivas y los efectos negativos del auge de la economía China.

y Esteve, 1993; Fontcuberta, 1997) para fijar las características de la nueva manera de informar del periodismo especializado frente al periodismo generalista, surgen aportaciones destacadas como la de Herrero (2002: 90) quien establece que:

Para clasificar un texto como especializado, no basta con que tenga como referente un área de actualidad; ni que su lenguaje sea más o menos legible o tenga una u otra distribución de elementos lingüísticos, icónicos o paralingüísticos, y tampoco que la audiencia a la que va dirigido esté más o menos segmentada. Lo especializado viene dado por el rigor de la estrategia en la construcción del texto, que es clave para su validez y eficacia.

En el caso de este trabajo, -que analiza los textos de un año concreto, 2012, con unas características socioeconómicas específicas de dos publicaciones con características concretas que hemos descrito en el capítulo 1-, hemos de tener en cuenta lo que Quesada (1998: 18-19) afirma sobre el periodismo especializado:

Es el que resulta de la aplicación de la metodología periodística a los múltiples ámbitos temáticos que conforman la realidad social condicionada siempre por el medio de comunicación que se utilice como canal, para dar respuesta a los intereses y necesidades de las nuevas audiencias sectoriales.

En lo referente a la prensa especializada de carácter económico, los periodistas que la elaboran requieren, según Muñoz Torres (1997: 171-172), un nivel de conocimiento y de valoración social suficiente para la divulgación de las noticias.

### **2.2.5.3 La prensa económica en España: antecedentes y evolución**

La información económica en España tardó en despegar. La situación política que imponían unos mercados limitados y ajenos a la internacionalización provocó que hasta 1985 fuera prácticamente inexistente. Para Portela (1999), las empresas no valoraban hasta esta fecha la información que estaba disponible y restringían el alcance público de sus propios informes internos al considerarlos un bien privativo. Destaca la autora la paradoja de que, por ejemplo, un centro de documentación de un gran banco podía tener datos sobre cualquier empresa noruega, pero no de una española.

En su análisis sobre los antecedentes y evolución de la prensa económica en España, Portela (1999) detalla que las empresas españolas empujadas por la regulación de la administración que

se adecuaba a las normas de la CEE y por la competitividad se han ido haciendo transparentes y ofreciendo los datos sobre su actividad al tiempo que en los medios de comunicación ha ido cobrando forma la industria de información económica y empresarial.

En esta evolución Portela destaca lo que han aportado la implantación de intranets, la aplicación de sistemas de gestión del conocimiento, el control electrónico de documentos, una mayor atención a los datos de los mercados exteriores la incorporación de los activos intangibles y el capital intelectual como elementos de la gestión empresarial. Todos estos condicionantes, según la autora, están transformado el esqueleto corporativo español y las políticas de fomento de las distintas administraciones e impulsando el desarrollo de una industria específica del sector de la información económica.

A la hora de analizar la prensa escrita habla de tres periódicos: *Expansión*, *Cinco días* y la *Gaceta de los negocios*. A estos suma las secciones de economía de los diarios generales o las páginas salmón que suelen incluir los domingos.

En su análisis señala que este tipo de prensa diversifica su presencia en el mercado ya que ha salido con fuerza en internet, hasta ahora con carácter gratuito. Las webs no sólo presentan el periódico digital sino también una amplia oferta con acceso a informes de empresas, gráficos de bolsa, mercados de divisas, tipos de interés, etcétera. Una de las publicaciones a la que hemos recurrido para la elaboración y análisis del CCF, *El País*, cumple esta premisa ya que el desarrollo de la crisis financiera ha obligado a la empresa editora a redoblar su esfuerzo con un incremento del espacio dedicado a cubrir esta información no sólo desde los géneros informativos sino también los de opinión.

Pese a todo, despegue económico y desarrollo tecnológico, la autora sostiene que España sigue estando en los puestos de cola europeos. Considera que el primer país productor y consumidor de servicios de información financiera es Gran Bretaña, seguido de Francia, Alemania, Países Nórdicos e Italia.

### 2.3 Libertad de información y de expresión y sus límites

La libertad de expresión puede plantear problemas en los medios de comunicación como vía para comunicarnos con los demás. Desde los orígenes de la escritura primaria en la civilización mesopotámica hasta la actualidad el hombre ha dispuesto de distintos medios para no sólo comunicarse, sino para fijar la intención de su mensaje. Al inicio de este capítulo sobre periodismo ya nombramos a Marshall McLuhan<sup>30</sup> (1996) y cómo el autor estableció la máxima “el medio es el mensaje”, con la que quería expresar la importancia de la forma que adopta la comunicación por encima del contenido. Retomamos la idea que pretende mostrar la importancia que en la sociedad han adquirido los medios de comunicación de masas, como por ejemplo en la difusión de las noticias. Entran en escena términos como *libertad de expresión*, que ejercen los medios para difundir la información y también sus límites. Al respecto Faúndez (2004: 170) destaca la fuerza que desde hace siglos ha tenido la expresión escrita y los controles y regulaciones de los que ha sido objeto:

Las garantías constitucionales relativas a la libertad de prensa, que obviamente estaban concebidas primordialmente en términos de expresiones escritas, han sido interpretadas en un sentido amplio, incluyendo en ella las expresiones habladas que se transmiten por la radio, al igual que los mensajes hablados y las imágenes transmitidas por la televisión.

El derecho a la información es un avance de la sociedad democrática y va más allá de la simple garantía que tiene un informador de poder transmitir noticias veraces. Contempla además el derecho de la sociedad a poder acceder a las noticias para estar informado y poder escoger libremente. En esencia, el esquema de la figura 2.4 resume todo lo que representa el derecho constitucional a la información, que va más allá de las publicaciones periodísticas.

---

<sup>30</sup> La teoría de Herbert Marshall McLuhan (educador, filósofo y teórico canadiense de la *Escuela de Toronto*) “el medio es el mensaje”, se convirtió en el lema de la contracultura de la década de los 60 del siglo XX. Sostenía que los medios electrónicos, en concreto la televisión, producen un impacto que supera el material comunicado.



Figura 2.4: Derecho a la Información y lo que abarca

### 2.3.1 Marco normativo

La libertad de expresión es un derecho universal. Ya hemos dicho que abarca más allá de la libertad de prensa y del ejercicio del periodismo. En España está regulada mediante la Constitución Española de 1978, en los artículos 18 y 20, la Ley Orgánica 5/1982 de Protección al honor, intimidad personal y familiar y el Código Penal: los delitos contra el honor.

El derecho a la información y la libertad de expresión es una conquista de las sociedades libres. No hay democracia si no se reconoce este derecho. Es algo evidente hoy en día. La Constitución española de 1978 no sólo defiende el derecho a la libertad de expresión, sino que también recoge en su artículo 20 el derecho a comunicar o a recibir libremente información veraz: Se reconocen y protegen (en el artículo 20) los derechos:

- A expresar y difundir libremente los pensamientos, ideas y opiniones mediante la palabra, el escrito o cualquier otro medio de reproducción
- A la producción y creación literaria, artística, científica y técnica
- A la libertad de cátedra

- A comunicar o recibir libremente información veraz por cualquier medio de difusión. La ley regulará el derecho a la cláusula de conciencia y al secreto profesional en el ejercicio de estas libertades

El ejercicio de estos derechos no puede restringirse mediante ningún tipo de censura previa. La ley regulará la organización y el control parlamentario de los medios de comunicación social dependientes del Estado o de cualquier ente público y garantizará el acceso a dichos medios de los grupos sociales y políticos significativos, respetando el pluralismo de la sociedad y de las diversas lenguas de España.

Estas libertades tienen su límite en el respeto a los derechos reconocidos en este Título, en los preceptos de las leyes que lo desarrollen y, especialmente, en el derecho al honor, a la intimidad, a la propia imagen y a la protección de la juventud y de la infancia.

### **2.3.2 Conceptos de libertad de expresión y de libertad de información: dos derechos distintos**

El derecho a recibir y dar información viene recogido en el artículo 20 de la Constitución española de 1978. El texto no sólo recoge el derecho a comunicar informaciones veraces sino el de recibirlas del mismo modo. La libertad de expresión es un derecho reconocido en el ámbito internacional. Sin embargo, es preocupante el silencio que guardan algunos gobiernos respecto a determinados asuntos, privando de este modo del derecho a la información mundialmente reconocido. La labor de un periodista es la defensa de estos derechos en contra de la censura a fin de facilitar el ejercicio de su profesión y que lleva inherente una amplia responsabilidad.

El filósofo José Luis López Aranguren (1909-1996) estableció como primer deber de un periodista la vigilancia del poder. Los periodistas son los testigos de los hechos que acontecen y que tienen que trasladar de una manera responsable, seria, exacta y elaborada. Cuando los periodistas abusan de estos derechos reconocidos por la Constitución y ponen por encima de su deber de informar otros intereses que se alejan de servir a la audiencia o los lectores, provocan la pérdida de credibilidad en los medios.

### **2.3.3 Límites a la libertad de expresión y de información: el honor, intimidad y propia imagen**

El derecho a estar informado y a poder difundir una información puede generar la vulneración de otros derechos inherentes a la persona en una sociedad democrática como el honor, la intimidad personal y familiar y a la propia imagen. Desde la invención de la imprenta al desarrollo de los medios de comunicación y el reciente avance de la información digital y las redes sociales, ha crecido la posibilidad de entablar comunicación con distintas partes del mundo y trasladar opiniones. Las nuevas tecnologías de la información han variado nuestra manera de comunicarnos, de pensar y de relacionarnos. Ya hemos visto que en España la regulación y el control y garantías de la libertad de expresión queda establecida a través de la constitución española y de los artículos 18 y 20. Sin embargo, Faúndez (2004: 201) advierte que este control precisa una revisión a la vista de gran avance de las redes sociales:

El contenido de los mensajes que se difunden en la red, al igual que el contenido de los mensajes que se difunden a través de otros medios, hace surgir los mismos conflictos que, en cualquier contexto, imponen la tarea de ponderar los límites de la libertad de expresión frente a otros derecho intereses igualmente legítimos; sin embargo, las características técnicas de la red, que permiten a cualquier persona que haga uso de ella se vea expuesta a que muchos de sus datos personales puedan ser descubiertos, almacenados y procesados, hacen mucho más difícil preservar la intimidad de las personas, e impedir que esa información sea utilizada indebidamente.

El desarrollo de internet y las redes sociales ha trascendido, de alguna manera, ajeno a las garantías que los límites a la libertad de expresión han tenido en los medios de comunicación tradicionales y como de nuevo veremos en el capítulo dedicado a este nuevo tipo de periodismo emergente.

### **2.3.4 ¿Qué es noticia en los medios de comunicación? *Agenda Setting* y *Gatekeeper***

En el capítulo 2.12 en el que analizamos las claves de la redacción periodística y los condicionantes que convierten un hecho en noticia de interés para la sociedad ya introdujimos el término *Agenda Setting* de los medios de comunicación y el concepto de *Gatekeeper* que vamos a detallar de la mano de expertos en la comunicación de masas.

Saperas (1987) y McQuail (1992) establecen la definición de la *Agenda Setting* como jerarquización de las noticias y la amplia a la capacidad para el establecimiento de la agenda temática o capacidad de la agenda temática. Según esta agenda, los medios de comunicación

escogen los temas sobre los que se habla y se discute, su importancia, su orden y la manera de transmitirlos. Los medios elaboran con antelación una agenda de trabajo sobre la información que van a difundir, determinan la importancia de los hechos del país y del mundo y les asignan un orden planeado y pensado racionalmente con el objetivo de lograr mayor audiencia, un mayor impacto y una determinada conciencia sobre un tema en concreto.

La idea básica de la *Agenda Setting* afirma la existencia de una relación directa y casual entre contenido de la agenda de los medios y la subsiguiente percepción pública de cuáles son los temas importantes del día (McCombs, 1996: 201). Saperas (1987: 60-63) deriva la función del *Gatekeeper* en los medios de comunicación de masas en:

Quienes determinan qué acontecimientos son periodísticamente interesantes y cuáles no y asigna una diferente relevancia en términos de diversas variables como la extensión, la importancia y el grado de conflictividad de todos los ítems que deben pasar por el tamiz.

El *Gatekeeper* adquiere una especial relevancia en la investigación sobre la capacidad de establecer la agenda temática al realizar la selección de los temas, al determinar el grado de relevancia del tema y, por ello, al iniciar el establecimiento de los media. Al mismo tiempo determinará cuál es el periodo de permanencia de un tema en los media y destacará cuales son los conflictos de mayor presencia pública.

En esta investigación hemos podido comprobar cómo un tema como es el de la economía y de las finanzas pasa de ser un asunto minoritario en los medios de comunicación a ocupar las portadas de los diarios y puestos destacados de los telediarios. El alcance de la crisis financiera lo convierte en asunto relevante, de agenda de los medios, cuyos editores abren la puerta a las informaciones que provengan de este ámbito. La crisis financiera y el aluvión de información que genera, como ya avanzamos en la introducción y veremos en el desarrollo de la metodología de este trabajo, deriva en un lenguaje propio.

#### **2.4 El Periodismo en redes sociales**

El desarrollo científico y tecnológico ha multiplicado las posibilidades de comunicación. La ventana al mundo que ha abierto internet ha ampliado el flujo de información entre las personas y ha hecho crecer la influencia de los medios de difusión más allá de la tradición de la original prensa escrita, la evolución de las ondas hertzianas o la señal de la televisión. A nadie se le escapa que los nuevos medios de comunicación son una evolución de los tradicionales. El

nuevo periodismo no es nada sin los elementos multimedia o las redes sociales que acrecientan el impacto de su trabajo y lo hacen llegar a todo el planeta. Ocurre también al contrario. Los nuevos soportes digitales de información precisan de la credibilidad que con el paso de los años se han labrado los medios tradicionales como garantes de la libertad de información y expresión que veíamos en el capítulo 2.3. Las nuevas tecnologías han variado la forma de comunicarnos, de expresarnos, de relacionarnos y de pensar. Sin embargo, la difusión y los contenidos de los nuevos medios necesitan del apoyo de los medios de comunicación tradicionales. En este sentido se expresa Daniel Moltó (*comunicación personal*, 2015), periodista de *elmundo.es*:

El periodismo digital, entendido como periodismo riguroso, no dista tanto del tradicional. Simplemente se adapta a otro canal distinto para difundir el mensaje y ese medio sí tiene unas características diferentes al periódico de papel. Ofrece ventajas como la inmediatez y la ruptura de barreras entre el lector y el redactor, favoreciendo al mismo tiempo la retroalimentación y la crítica -que debe ser aceptada y gestionada correctamente-. Otra de las cualidades del periodismo en la red es la posibilidad de convertir en global una información local, ya que las noticias se pueden leer desde cualquier punto del planeta y en cualquier dispositivo. Entre las desventajas, destacaría la falta de tiempo para la reflexión, derivada de la premura con la que se trabaja y que obliga al periodista a ser conciso y ágil a la hora de confirmar los datos antes de publicar cualquier información.

Ya hemos dicho que el nuevo periodismo comenzó siendo una copia en formato digital de la versión en papel editada. Ese fue el primer paso, pero la versión digital ha ido evolucionando y distanciándose cada vez más de la versión en papel. En el documento editado por FUNDEU BBVA, *Escribir en Internet. Guía para los nuevos medios y redes sociales*, Tascón y Cabrera (2012: 25-28) aproxima la siguiente definición:

Son los medios que proporcionan acceso a la información a través de las nuevas tecnologías de la información y comunicación. Los nuevos medios permiten llegar a la información en cualquier momento y desde cualquier lugar, siempre que tengamos acceso a internet, a través de cualquier dispositivo digital. Se apoyan en las tecnologías digitales y permiten la interactividad. Se caracterizan por la democratización en la creación, publicación, distribución y consumo de contenidos multimedia. Cualquier usuario puede generar información, así como comentarla y compartirla.

Atendiendo a estas dos apreciaciones, la de Moltó y la de Tascón y Cabrera, sobre los nuevos modos que tienen de ejercer los medios de comunicación social y de transmitir informaciones, es conveniente parar para destacar la importancia que han adquirido en este proceso las redes

sociales entre ellas Facebook<sup>31</sup> y Twitter<sup>32</sup> para la divulgación de la información. Las redes sociales incrementan la difusión de las noticias gracias a la acción de los comentarios, las interacciones manifestadas a través de *like* o *me gusta* y sus derivaciones<sup>33</sup> en el caso de Facebook o realizando *retweet*<sup>34</sup> en Twitter, marcando el favorito. La información de los medios llega más allá gracias a las nuevas fórmulas que constituyen las redes sociales tejidas entre amigos, profesionales de distintos ámbitos o desconocidos que comparten pensamientos, imágenes, opiniones, música, etcétera. Levinson (2013: 2) prefiere hablar de nuevos medios y no de medios sociales al determinar que “The media that allow consumers to become producers are undoubtedly social, and far more interactive than older, one way media such as television”.

---

<sup>31</sup> Facebook es un sitio web de redes sociales creado por Mark Zuckerberg y que lo fundó junto a Eduardo Saverin, Chris Hughes y Dustin Moskovitz. Originalmente era un sitio para estudiantes de la Universidad de Harvard, en el que pudieran intercambiar una comunicación fluida y compartir contenido de forma sencilla a través de Internet. Fue un proyecto tan innovador que se extendió hasta estar disponible para cualquier usuario de la red. A mediados de 2007 Facebook lanzó las versiones en francés, alemán y español para impulsar su expansión fuera de Estados Unidos, país en el que se concentraban sus usuarios junto a Canadá y Reino Unido.

En 2014 Facebook alcanzó 1.350 millones de usuarios y traducciones a 70 idiomas. Las principales críticas a la red social y la empresa se han centrado en la supuesta falta de privacidad que sufren sus millones de usuarios y que se acrecentaron en 2013, al descubrirse que la Agencia de Seguridad Nacional de Estados Unidos y otras agencias de inteligencia vigilan los perfiles de millones de usuarios y sus relaciones con amigos y compañeros de trabajo. El 9 de abril de 2012 Facebook anunció la adquisición de otra red social, Instagram, por 1.000 millones de dólares y en febrero de 2014 dio a conocer la compra del servicio de mensajería móvil WhatsApp por 16.000 millones de dólares.

<sup>32</sup> Twitter es un servicio de mensajes cortos, o tweets de un máximo de 140 caracteres, creado por Jack Dorsey en California 2006 y bajo la jurisdicción de Delaware desde 2007 que lo lanzó en julio del mismo año. Twitter ha alcanzado más de 500 millones de usuarios, generado 65 millones de tweets al día y manejado más de 800 000 peticiones de búsqueda diarias. Los usuarios pueden suscribirse a los tweets o mensajes de otros usuarios, a esto se le llama seguir y a los usuarios abonados se les llama seguidores o *followers* o *tweeps* de *Twitter + peeps*, seguidores novatos que aún no han hecho muchos tweets. Los mensajes son públicos y se pueden difundir de manera privada mostrándolos únicamente a unos seguidores determinados. El dominio fue registrado en 2000 pero no ha trascendido la fecha exacta de presentación pública del proyecto, si los tweets pueden ser o no redirigidos y si Twitter comparte los mensajes internos entre usuarios con la Agencia de Seguridad Nacional de Estados Unidos.

<sup>33</sup> [www.bbc.com/mundo/noticias/2015/10/151009\\_facebook\\_emojis\\_reacciones\\_jm](http://www.bbc.com/mundo/noticias/2015/10/151009_facebook_emojis_reacciones_jm)

<sup>34</sup> *Retweet* es un término inventado por los usuarios de la red social Twitter y que forma parte del sistema de su comunicación. Permite a los usuarios compartir sus mejores publicaciones o tweets. Funciona al mencionar lo que otra persona ha dicho para agregarlo o no un comentario.

El periodismo multimedia es la integración de medios y servicios. Cebrián (2003) habla de sistemas expresivos que superan la concepción de lo audiovisual al incorporar la interactividad, navegación y vínculos. Generan un cambio en el modelo de comunicación que repercute en todos los componentes del proceso: en los usuarios que intercambian continuamente sus papeles, en los medios ante la constante tendencia a la convergencia, en los contenidos y tratamientos que aparecen abiertos y potenciales para que cada usuario combine como considere oportuno y, sobre todo, en la recepción en la que cada participante se convierte en un usuario con capacidad para recibir y para ofrecer dentro de un proceso de diálogo e intercambio real de informaciones.

Los periódicos digitales emulan la función de la prensa tradicional, pero utilizando los canales que les permiten hacer útiles un conjunto de complementos inimaginables. El ciberperiodismo tiene la voluntad de mediación entre los hechos y el público y utiliza criterios y habilidades de los periodistas como los medios interactivos e hipertextos (Salaverria, 2005). Los nuevos complementos multimedia abren en los medios digitales una nueva dimensión que es necesario analizar frente a los géneros periodísticos clásicos (Fondevila Gascón, 2009). El periodismo digital crea un nuevo modelo informativo que combina el estilo de la prensa tradicional, de la radio y la televisión al tiempo que incorpora la imagen y sonidos a través de videos.

En los medios de comunicación online ha comenzado la proliferación de blogs de manera paralela a la información periodística. El blog, que también puede conocerse como weblog, es una página en la que se publican anotaciones al estilo de un diario y que puede contener enlaces a otras páginas, imágenes y textos. Hay blogs que se elaboran por un colectivo de personas que comparten inquietudes e interés por un tema. Pero también puede elaborarse por parte de un particular. Landow (2009) destaca la importancia que puede tener el blog para quien esté interesado en publicar hipertextos. Los blogs toman forma de periódico online, que suele escribir un autor y que se diferencian de los textos editados en papel en que presentan un orden cronológico inverso. Los periodistas suelen utilizarlos para promocionarse ya que suponen una plataforma para darse a conocer y en la que pueden incluir sus reportajes y su curriculum.

Destacamos casos como el del diario *El País*, el cual, desde que dio un impulso decisivo a su edición online en 2012, ha ido incorporando blogs de autores destacados como complemento a la información periodística.



Figura 2.5: Ejemplos de blogs de economía en *El País*

No es este un caso aislado. Otros medios online han aprovechado las posibilidades que aporta esta nueva tecnología para ampliar sus contenidos, con hipertextos y otros elementos multimedia a través de blogs. En España hay casos de prensa exclusivamente digital sin un soporte papel previo (*Libertad Digital, El Plural, Vilaweb*, etcétera). Esos casos han servido de reflexión a la prensa escrita sobre la cantidad y el tipo de uso que plantean con los recursos multimedia.

En el desarrollo de la prensa digital, a la hora de hablar de blogs, queremos resaltar el caso del periódico online y blog agregador de noticias fundado por Arianna Huffington y que fue lanzado el 9 de mayo de 2005. *The Huffington Post* ha logrado hacerse un destacado hueco entre la prensa online como colector de comentarios sobre política, medios, negocios, entretenimiento, moda, el movimiento verde, noticias internacionales y humor. La versión en castellano de *The Huffington Post* se puso en marcha en España en junio de 2012<sup>35</sup>. Antes se habían lanzado las

<sup>35</sup> Según Technorati Authority, al 15 de junio de 2012 *The Huffington Post* ocupaba el primer lugar en los Top 100 Blogs.

locales de Chicago, Nueva York, Denver y Los Ángeles. *The Huffington Post* tiene una comunidad activa de más de un millón de comentarios por mes. Las últimas versiones en incorporarse han sido las de Francia y Gran Bretaña, en septiembre de 2012, y Alemania, octubre de 2013.



Figura 2.6: Lanzamiento en España del diario *The Huffington Post*

## 2.5 Conclusión

“The truth shall set you free”. Así reza el epitafio escrito en la lápida de Marshall McLuhan. El erudito canadiense tomó prestada la frase bíblica como un alegato al conocimiento frente a la esclavitud que representa la ignorancia. El capítulo que aquí concluye empieza formulando la pregunta ¿qué es el periodismo? Una de las finalidades del periodismo es, precisamente, acercar el conocimiento o la verdad de las cosas a la sociedad con el objetivo de que cada cual tenga libertad de elegir cómo vestir, qué comer, qué leer o, por ejemplo, quién quiere que le

gobierno. Hemos analizado los códigos de los medios de comunicación de masas, las claves de la redacción periodística y los principales rasgos del lenguaje periodístico en sus diferentes estilos. Asimismo, hemos desentrañado cómo los medios de comunicación construyen la verdad que sirven a la sociedad a través de los distintos géneros periodísticos objetivos: noticia, crónica, reportaje, entrevista y los géneros periodísticos subjetivos, que llevan implícita la opinión del medio: editorial, columna, artículo, cartas al director. También se han analizado las características lingüísticas de los géneros informativos y las de los géneros de opinión, así como la superestructura del discurso periodístico.

Han quedado patentes desde las cuestiones más científicas, relacionadas con la construcción del mensaje informativo que aportan autores como José Luís Martínez Albertos, a las teorías más visionarias como la de la aldea global del ya mencionado Marshall McLuhan, de mediados del siglo pasado. Todo ello sin perder de vista la aparente evolución de la manera de contar las noticias en los medios de comunicación, desde la incipiente prensa escrita, la irrupción de la radio en los hogares, la llegada de la televisión, la presencia global de los medios digitales en la actual sociedad y la aportación de las redes sociales a su expansión, que han transformado y revolucionado las rutinas productivas del periodista y las empresas periodísticas.

El capítulo abunda en la prensa especializada, en concreto la de la sección de economía, ya que esta investigación se basa en el análisis del relato periodístico de la crisis financiera en 2012 y sus particularidades en los dos medios de referencia de prensa escrita en España en ese año (los dos medios que conforman nuestro corpus). En alusión, de nuevo, a McLuhan y su también célebre frase “el mensaje es el medio” descubrimos cómo el contenido de cualquier mensaje puede llegar a resultar menos importante que el medio (de comunicación) en sí mismo en que se publique. El capítulo abarca también el marco normativo que ampara el ejercicio periodístico como un derecho a ofrecer y recibir información y sus límites.

En el próximo capítulo, abordaremos la lingüística de corpus como metodología para analizar textos periodísticos, en los que hemos basado nuestra investigación.

## CAPÍTULO 3: LINGÜÍSTICA DE CORPUS

### 3.1 Introducción

La lingüística de corpus es la parte de la lingüística que investiga y analiza los datos obtenidos de un corpus elaborado con muestras reales del uso de la lengua. Para elaborar un corpus lo fundamental es recurrir a las fuentes para la recopilación de datos. Las fuentes a las que podemos acudir pueden ser orales o escritas. Orales como, por ejemplo, conversaciones grabadas o bien fuentes escritas como, por ejemplo, textos periodísticos. Éste es el caso de la presente investigación, la elaboración de un corpus a partir de publicaciones periodísticas sobre la crisis financiera para su posterior análisis.

La lingüística de corpus ha experimentado una evolución a lo largo de la historia. La literatura escrita sobre esta disciplina considera su concepción como muy antigua. Existen múltiples referencias a trabajos clásicos, como *Early Corpus Linguistics* (McEnery y Wilson, 1996), basados en muestras reales de la lengua. Los autores realizan su revisión desde los estudios anteriores a Chomsky, antes de 1950. La siguiente fase abarca trabajos entre 1960 y 1980. La tercera comprende los considerados como *Modern Corpus Linguistics*. Pero, regresamos a los considerados como el origen de la actual lingüística de corpus (Kennedy, 1998: 13):

Corpus-based research is often assumed to have begun in the early 1960s with the availability of electronic, machine-readable corpora. However, before then there was a considerable tradition of corpus-based linguistic analysis of various kinds occurring in five main fields of scholarship:

- 1 biblical and literary studies
- 2 lexicography
- 3 dialect studies
- 4 language education studies
- 5 grammatical studies

La lingüística de corpus cobra nuevas perspectivas con la implantación de la informática como método de trabajo. Ofrece grandes posibilidades de procesar y gestionar conjuntos de textos cada vez de mayor tamaño. Es el momento de la aparición de los grandes corpus de las lenguas modernas, que superan los 100 millones o incluso alcanzan los 200 millones de palabras. Es el caso del *British National Corpus* o el *Corpus de Referencia del Español Actual* elaborado por la Real Academia Española.

La tarea de este capítulo va ser el análisis de la definición de corpus, en qué consiste y su finalidad. Todo ello a través de los trabajos desarrollados por los estudiosos en la materia. Pero, antes de adentrarnos en el marco histórico de la cuestión, sirva como punto de arranque de la presente investigación la siguiente definición de Sinclair (1991: 171).

A corpus is a collection of naturally occurring language text, chosen to characterize a state or variety of a language. In modern computational linguistics, a corpus typically contains many millions of words: this is because it is recognized that the creativity of natural language leads to such immense variety of expression that it is difficult to isolate the recurrent patterns that are the clues to the lexical structure of the language.

Sinclair hace hincapié en que los corpus deben ser vistos como colecciones de millones de palabras. El principal objetivo de su investigación es la compilación de diccionarios de uso general y gramáticas, que por su naturaleza requiera el examen de millones de palabras para llegar a una descripción del sistema lo más completa como sea posible. Si atendemos la regla general de cuanto más grande sea el corpus mejores resultados obtendremos, la cuestión del tamaño apropiado para un corpus dependerá de la finalidad del investigador y del propósito de la investigación.

La actual concepción de un corpus es diferente a la que utilizaban los primeros lingüistas como recurso para sus investigaciones. Autores como McEnery y Wilson (2001: 2) señalan que a menor frecuencia de la función que se desea investigar, mayor debe ser el corpus. Esto se aplicaría a los sustantivos, adjetivos, adverbios, etcétera -es decir, palabras de contenido- que tienden a tener una frecuencia mucho más baja que las palabras gramaticales de un determinado corpus. Amplían su marco de actuación a toda la lingüística: "Corpus linguistics is a methodology that may be used in almost any area of linguistics, but it does not truly delimit an area of linguistics itself".

A la inversa, se puede argumentar que los corpus más pequeños pueden ser útiles en una investigación sobre las características más comunes del lenguaje, como pueden ser los elementos gramaticales. De hecho, Biber (1988) señala que los corpus más pequeños son perfectamente adecuados para este cometido.

Otros lingüistas como Hunston (2006) ven en la lingüística de corpus una metodología o conjunto de metodologías, más que una teoría de descripción de la lengua. Es comparable a cualquier enfoque teórico sobre la lengua y de ahí surge la gran teoría lingüística. Parte de la base de que el significado no reside en las palabras de forma aislada, sino en las unidades de significado. La autora desgrana cuatro ideas básicas en la lingüística de corpus:

- La observación de la lengua natural
- La observación de una cantidad relativamente elevada de muestras de esa lengua
- La observación de las frecuencias relativas, en su forma natural o modificada mediante operaciones estadísticas
- La observación de los patrones de asociación, bien entre un rasgo y un tipo de texto, bien entre grupos de palabras

Los grandes investigadores de la lingüística de corpus, como Sinclair (1991), Stubbs (1996), Biber et al. (1998) o Hunston (2002), coinciden en señalar que dicha disciplina es la recopilación de palabras escritas o habladas, creada para un propósito y que, por lo general, suele ser de carácter puramente lingüístico, aunque también puede ser de naturaleza socio-pragmática.

En la actualidad un corpus ya se concibe como un conjunto de textos a los que se accede informáticamente. Hunston (2002: 2) afirma: “Collections of texts (or parts of texts) that are stored and accessed electronically”.

Esta definición es más general en el ámbito anglosajón, ya que en el mundo hispánico la definición de corpus que aparece en el DRAE (2014) es la de un conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etcétera, que pueden servir de base a una investigación.

Leech (1992) en su estudio destaca la habilidad de los ordenadores de procesar y ordenar esta información, por lo que recomienda que se les conozca como *Computer Corpus Linguistics*, debido a que el término lingüística de corpus, sin más, se asocia al periodo anterior al desarrollo de la era digital. En la actualidad los datos de un corpus son accesibles de forma electrónica para su análisis y pueden ser leídos por un ordenador. Éste es el caso de esta investigación que ha

consistido en la creación de un corpus de textos de periódicos digitales para su posterior análisis lingüístico.

Seguimos con los criterios básicos, aunque no exclusivos, y a tener en cuenta a la hora de definir un corpus:

- La finalidad concreta para la que hemos creado un corpus
- La inclusión de datos auténticos
- Su elaboración conforme a criterios de diseño explícitos
- Que es representativo de un lenguaje o género particular

Más adelante abundaremos en estos criterios, pero ahora nos fijamos en la definición que Kennedy (1998: 1-3) hace de la lingüística de corpus. Veremos que esta disciplina no tiene un fin en sí mismo, sino que es una fuente de evidencia para descubrir cómo funcionan las lenguas (su estructura y su uso). Otras aplicaciones de la lingüística de corpus serían el procesamiento del lenguaje natural de forma mecánica, el aprendizaje y la enseñanza de lenguas. Este autor destaca como habilidad más importante la capacidad que tiene la lingüística de corpus de resolver una serie de preguntas clave a base de los estudios teóricos, descriptivos y aplicados del lenguaje.

### **3.2 Origen y evolución de la lingüística de corpus**

La lingüística de corpus tiene un origen muy remoto. Ya hemos apuntado que las primeras referencias de recopilación de muestras de la lengua se remontan a los textos clásicos. La lingüística de corpus se ocupa de la recopilación de datos lingüísticos, ya sean textos escritos o una transcripción del discurso grabado, que se pueden utilizar como punto de partida para la descripción lingüística o como medio para verificar las hipótesis sobre el lenguaje. Es una metodología y un enfoque del análisis del lenguaje natural que se basa en generalizaciones hechas a partir de datos. La lingüística de corpus estudia y analiza todos los aspectos del idioma.

La lingüística de corpus aborda la búsqueda de patrones. Una parte importante de esta disciplina es la localización de los patrones de uso de la lengua en los datos empíricos de un corpus y comprobar *a posteriori* la información que nos revela acerca del comportamiento del

lenguaje. En el caso de esta investigación, el cometido será la recopilación de textos periodísticos para, finalmente, llegar a conocer el sentimiento de sus autores. Esta información nos llega a través del conteo de palabras, partes de la oración, las secuencias gramaticales o léxicas, oraciones, párrafos, encabezados, títulos, expresiones, giros, figuras retóricas, etcétera. La lingüística de corpus explica las posibles opciones que tienen quienes utilizan una lengua. Es además una materia descriptiva, no prescriptiva. El lenguaje prescriptivo es aquel que se utiliza para indicarle al lector lo que tiene que hacer. En la lingüística de corpus destacan dos principios:

- Los aspectos del lenguaje se estudian empíricamente mediante el análisis de datos de lenguaje natural en un corpus
- A más datos, mejor conteo y mejor búsqueda de patrones en el lenguaje

El proceso de acumulación de datos lingüísticos que da origen a la lingüística de corpus tal y como la concebimos en la actualidad surge a mediados del siglo XX. Teuber (2007: 50-58) describe este hecho al final del capítulo primero *A brief history of Corpus Linguistics* y cómo arranca con el proyecto *Survey of English Usage* de Randolph Quirk en 1959 y cuyos datos fueron la base de la primera gramática estándar del inglés *A Comprehensive Grammar of the English Language* (Quirk et al., 1985). Este trabajo fue el inicio de la recopilación informática y fue un punto de referencia de otros estudios. Nelson Francis y Henry Kucera (1964) crearon el Corpus de la Universidad de Brown en Providence (Rhode Island) con un millón de palabras en las que se ejemplificaban los usos de 2.000 palabras escogidas de 500 textos americanos de 15 categorías textuales distintas. El *Lancaster-Oslo/Bergen Corpus* (LOB Corpus) (1961) es un corpus similar al Corpus de Brown. Contiene 500 textos de 2.000 palabras, distribuidas en 15 categorías de texto: 9 informativas y 6 imaginativas. Hay dos versiones de este corpus, la original (1970-1978), elaborado por Geoffrey Leech de la Universidad de Lancaster y Stig Johansson de la Universidad de Oslo, ambos líderes del proyecto, y Knut Hofland de la Universidad de Bergen, como jefe de informática. La otra es la versión del *LOB Corpus POS-tagged* (1981-1986), elaborado por Geoffrey Leech de la Universidad de Lancaster, Stig Johanson de la Universidad de Oslo, ambos líderes del proyecto, y Roger Garside de la Universidad de Lancaster y Knut Hofland de la Universidad de Bergen, como jefes de informática. Estos trabajos generaron investigaciones gramaticales y de frecuencia de aparición de palabras.

Con un millón de palabras tomadas de discursos orales y textos escritos, por su tamaño y repercusión, el corpus que dio pie al informe *English Lexical Studies*, el *OSTI Report*, llegó a ser el más destacado. Comenzó a elaborarse en 1963 en Edimburgo y se completó en Birmingham con John Sinclair a la cabeza de los investigadores. En este proyecto se investigaron las unidades léxicas. Concluimos este punto sobre el origen y evolución de la lingüística de corpus con los 10 principios básicos que Sinclair (2005: 1-21) describe para la elaboración de un corpus:

- The contents of a corpus should be selected without regard for the language they contain, but according to their communicative function in the community in which they arise.
- Corpus builders should strive to make their corpus as representative as possible of the language from which it is chosen.
- Only those components of corpora which have been designed to be independently contrastive should be contrasted.
- Criteria for determining the structure of a corpus should be small in number, clearly separate from each other, and efficient as a group in delineating a corpus that is representative of the language or variety under examination.
- Any information about a text other than the alphanumeric string of its words and punctuation should be stored separately from the plain text and merged when required in applications.
- Samples of language for a corpus should wherever possible consist of entire documents or transcriptions of complete speech events, or should get as close to this target as possible. This means that samples will differ substantially in size.
- The design and composition of a corpus should be documented fully with information about the contents and arguments in justification of the decisions taken.
- The corpus builder should retain, as target notions, representativeness and balance. While these are not precisely definable and attainable goals, they must be used to guide the design of a corpus and the selection of its components.
- Any control of subject matter in a corpus should be imposed by the use of external, and not internal, criteria.
- A corpus should aim for homogeneity in its components while maintaining adequate coverage, and rogue texts should be avoided.

### 3.2.1 El tamaño importa

Una cuestión destacada a la hora de completar la definición de un corpus es su tamaño. Este aspecto, su tamaño, está también relacionado con los cuatro criterios básicos señalados en el punto anterior: su finalidad, la inclusión de datos auténticos, elaboración de acuerdo a criterios de diseño explícitos y que sea representativo. De manera general, los corpus elaborados a gran escala que han tenido una finalidad generalizada han albergado entre 100 y 500 millones de

palabras, pero en la actualidad esa cifra puede ser de billones de palabras. Los más especializados, relativos a un género alcanzarían desde 50.000 hasta 5 millones de palabras.

Sobre el tamaño de los corpus también hay diversidad de opiniones. Ya hemos visto que Sinclair (1991: 171) es partidario de un corpus grande con millones de palabras como herramienta de trabajo que proporcione una representación adecuada del lenguaje y suficientes y múltiples ocurrencias de los ítems objeto de la investigación. Sampson (2001: 124) se decanta por que dispongan de una muestra importante de uso del lenguaje de la vida real para asegurar que existe evidencia suficiente para generar o probar hipótesis acerca del lenguaje.

We do not care how a scientist dreams up the hypotheses he puts forward in the attempt to account of the facts – he will usually need to use imagination in formulating hypotheses, they will not emerge mechanically from scanning the data. What is crucial is that any hypothesis which is challenged should be tested against interpersonally observable, objective data.

Otro factor que influye en el tamaño del corpus es el grado de variación interna en el idioma en estudio. Cuanto mayor sea la variación, se requieren más muestras y un corpus más grande para garantizar la representatividad y, por lo tanto, la validez de los datos (Meyer, 2002). Sin embargo, aquí de nuevo vemos que el tamaño tiene que ser equilibrado con el nivel de la investigación, un problema abordado por Kennedy (1998), quien advierte que a mayor tamaño mayor dificultad encontraremos en el manejo de datos.

El tamaño del corpus es increíblemente importante en términos de la riqueza de los datos. Por ejemplo, un corpus muy pequeño, digamos de un millón de palabras, es extremadamente limitado en cuanto a los fenómenos que puede abordar –comparado con un corpus de 500 millones de palabras–, donde podría haber 500 veces más datos. Sin embargo, los primeros corpus contenían alrededor de un millón de palabras (por ejemplo, los corpus *Brown*, *LOB*, *LLC*, *Kolhapur*, *Wellington*).

Las siguientes generaciones de corpus eran, al menos, 100 veces de mayor tamaño (por ejemplo, el *BNC*, que contiene 100 millones de palabras y el *Collins COBUILD*<sup>36</sup> corpus desarrollado en los años ochenta contenía más de 500 millones de palabras). En la actualidad,

---

<sup>36</sup> El *Collins Corpus* hoy en día contiene más de 4,5 billones de palabras.

con el auge de Big Data, existen corpus que ascienden a billones de palabras (véase <http://corpus.byu.edu/>).

| English  | # words      | language/dialect   | time period        |
|--|--------------|--------------------|--------------------|
| NOW Corpus                                     | 2.8 billion+ | 20 countries / Web | 2010-<br>yesterday |
| Global Web-Based English (GloWbE)              | 1.9 billion  | 20 countries / Web | 2012-13            |
| Wikipedia Corpus                               | 1.9 billion  | English            | -2014              |
| Hansard Corpus (British Parliament)            | 1.6 billion  | British            | 1803-2005          |
| Corpus of Contemporary American English (COCA) | 520 million  | American           | 1990-2015          |
| Corpus of Historical American English (COHA)   | 400 million  | American           | 1810-2009          |
| TIME Magazine Corpus                           | 100 million  | American           | 1923-2006          |
| Corpus of American Soap Operas                 | 100 million  | American           | 2001-2012          |
| British National Corpus (BYU-BNC)*             | 100 million  | British            | 1980s-1993         |
| Strathy Corpus (Canada)                        | 50 million   | Canadian           | 1970s-2000s        |
| CORE Corpus                                    | 50 million   | Web registers      | -2014              |
| <b>Other languages</b>                         |              |                    |                    |
| Corpus del Español                             | 100 million  | Spanish            | 1200s-1900s        |
| Corpus do Português                            | 45 million   | Portuguese         | 1300s-1900s        |
| <b>N-grams</b>                                 |              |                    |                    |
| Google Books: American English                 | 155 billion  | American           | 1500s-2000s        |
| Google Books: British English                  | 34 billion   | British            | 1500s-2000s        |
| Google Books: One Million Books                | 89 billion   | Am/Br              | 1500s-2000s        |
| Google Books: Spanish                          | 45 billion   | Spanish            | 1500s-2000s        |

Tabla 3.1: Corpus gestionados por Mark Davies (<http://corpus.byu.edu/corpora.asp?b=y>)

### 3.3 Lingüística computacional

El avance de la informática ha abierto nuevas posibilidades al estudio de la lingüística de corpus, sus patrones y frecuencias. La lingüística de corpus original se ha visto gratamente enriquecida con las aportaciones de la lingüística computacional (ya que realizan labores informáticas de diseño de etiquetadores o lematizadores, tales como el *POS-tagger*). Se trata de una metodología dentro del estudio del lenguaje y que en ocasiones llega a confundirse con la lingüística de corpus, pero hay algunas diferencias a tener en cuenta. Mientras que la lingüística computacional diseña y desarrolla métodos informáticos para analizar datos, la lingüística de corpus se centra en la explotación e interpretación de datos lingüísticos en formato electrónico para poder acometer su investigación (algunas veces los estudiosos de la lingüística de corpus desarrollan sus propios programas pero lo normal es que utilicen programas informáticos desarrollados por otros). Ahí radica la diferencia entre ambas, en que en la lingüística de corpus el objeto del estudio son los datos. Sin embargo, en la lingüística computacional, el enfoque está en el desarrollo de sistemas informáticos para profundizar nuestros conocimientos sobre el lenguaje humano.

Vayamos al origen de la lingüística computacional. Surgió en 1950 y se ha centrado en tres áreas de actuación:

- La traducción automática
- La extracción de información
- La interfaz hombre-ordenador para traducir textos al inglés de manera automática

En este ámbito destacamos que los nuevos recursos lingüísticos, desarrollados en la década de los 80 del siglo XX para procesar el lenguaje a gran escala, permitieron el desarrollo de innovadoras aplicaciones que favorecieron la traducción automática. Pero la lingüística computacional no es solo esto. Otro ejemplo de sus aplicaciones tradicionales es la extracción de información, a través de módulos de software que analizan el lenguaje en lo referente a la morfología, la sintaxis, o la semántica. Grishman (1986) define la lingüística computacional como el estudio de los sistemas informáticos para la comprensión y la generación de lenguaje natural.

Volvemos a los requisitos fundamentales que ha de cumplir un corpus hoy en día. McEnery y Wilson (2001: 21) fijan los siguientes requisitos:

**a. Textos en formato electrónico.** Para ser una herramienta útil al lingüista debe estar informatizado. Los textos de que consta tienen que estar en formato electrónico. La elaboración de los primeros corpus no contaba con ordenadores, lo que motivó ciertas críticas, en el sentido de que el procesamiento de los datos debía efectuarse de forma manual, con los errores y problemas que ello acarrea. Más adelante el uso del ordenador ha permitido automatizar tareas como:

- Búsqueda de información rápida. Localización ágil de una palabra o una secuencia de palabras o incluso una categoría gramatical.
- Recuperación de información. Al estar informatizado, un corpus permite obtener todos los casos de una palabra, secuencia de palabras, etcétera, registrados en el corpus, con su contexto inmediato anterior y posterior. Es lo que conocemos como concordancia.
- Cómputo de la frecuencia. De aparición de una palabra, secuencia de palabras, etcétera.
- Clasificación de los datos contenidos en el corpus, según diferentes criterios: por orden alfabético, frecuencia de aparición, autor, procedencia geográfica, tema, medio de publicación, etcétera.

**b. Autenticidad de los datos.** Los textos recogidos en el corpus deben ser muestras reales de uso de la lengua objeto de estudio. A partir de ellas se construyen o verifican, de forma empírica, las teorías que tratan de explicar el funcionamiento de la lengua o las aplicaciones computacionales.

**c. Criterios de selección.** Los textos que forman parte del corpus deben haber sido elegidos de acuerdo con unos determinados criterios –lingüísticos y/o extralingüísticos- para la finalidad concreta que persiga el corpus.

**d. Representatividad.** La selección de los textos ha de responder, además de a unos criterios adecuados, a parámetros estadísticos que garanticen que los textos son representativos de la

variedad de la lengua objeto de estudio. La variedad puede referirse a la obra de un autor determinado, a un período de tiempo, a un género, etcétera. Si lo que nos interesa es la lengua en su conjunto, la opción de reunir en un corpus todas las muestras parece algo impracticable, a diferencia de si lo que queremos es recoger todas las obras de un autor determinado. La solución que se presenta es tomar una muestra pequeña de la lengua que refleje a pequeña escala el funcionamiento del todo que es la lengua. Chomsky (2002) alertó de que los corpus puedan parecer por eso documentos sesgados, por lo que recurrimos a la selección, según criterios estadísticos, de diversos géneros, tipologías, temas o medios de publicación.

**e. Tamaño.** Por lo general, los corpus suelen medirse por millones de palabras o formas, algo que queda fijado antes de iniciar la recopilación de textos. Al alcanzar la cifra deseada o preestablecida se da por terminada la recopilación del corpus, que no es más que el primer paso de todo el proceso. Existen otros corpus abiertos o monitor, como el del proyecto COBUILD de la Universidad de Birmingham, de especial interés para la lexicografía. En el pasado se daba importancia al tamaño; a mayor corpus, más posibilidades de reflejar el funcionamiento real de la lengua en todas sus variedades. En la actualidad priman otros criterios de diseño.

### 3.4 Conceptos clave en la lingüística de corpus

El decálogo de conceptos que enumeramos a continuación es fundamental para entender la lingüística de corpus como metodología para el análisis de texto y para comprender el análisis lingüístico de los textos realizado para la presente investigación y la elaboración del CCF o *Corpus de la Crisis Financiera*.

**Wordlist/Lista de palabras.** Para definir este concepto recurrimos a la *Aproximación a ciertas perspectivas en la Lingüística de Corpus*, de Cristina Martín Herrero (2009). La autora de este artículo describe en su trabajo las listas de palabras como el conjunto de todas las palabras del corpus con la frecuencia de aparición que tienen (número de concordancias en las que aparecen).

En la clasificación que realiza Barnbrook (1996: 43-64) de diversos tipos de listas de palabras diferencia entre:

- Simples
- De formas canónicas
- De lemas
- De dos o más palabras
- De partes de la oración

Pueden ordenarse las listas de palabras de manera alfabética, por frecuencias, en orden alfabético inverso, por orden de aparición, por longitud de palabras, por categorías gramaticales. Según este autor, las listas de palabras cualquiera sea su tipo o su orden, no nos permitirán identificar los diferentes sentidos de las palabras. Para esto recurriremos a las concordancias. Una de las herramientas más utilizadas para la elaboración de listas de palabras es el *Wordlister* que pertenece a un conjunto de herramientas llamadas *Wordsmith Tools*, desarrollado por Mike Scott (2012) de la Universidad de Aston. Permite producir una única lista de todas las palabras en todos los archivos seleccionados o múltiples listas de palabras. La información procesada ofrece datos a través de análisis estadísticos, listas de palabras por índice de frecuencias y una lista de palabras por orden alfabético.

**Concordance/Concordancia.** Barnbrook (1996) define las concordancias como listas de palabras de un corpus, ordenadas dentro de un contexto. Son también conocidas como *Key Word In Context (KWIC)*. Suelen ser el paso posterior a las listas de palabras, ya que de estas últimas se selecciona la palabra para traer en concordancia. La palabra que se analiza viene en una columna centrada. Una concordancia es una colección que recoge todas las apariciones de una palabra en un texto o conjunto de textos con un número determinado de caracteres de un co-texto y posterior. Es la lista alfabética de las palabras principales usadas en un corpus, con sus contextos inmediatos. Pérez y Moreno (2009) sostienen que de esta manera es posible visualizar a la vez gran cantidad de ejemplos de uso de una palabra o un grupo de palabras y que los diversos tipos de ordenación permiten centrar la atención en el co-texto inmediatamente anterior o posterior de la palabra y así poder estudiar tipos comunes de sujetos y complementos en el caso de un verbo, la modificación adjetival de un sustantivo o el tipo de sustantivos que acompaña un adjetivo determinado. También podemos identificar la fuente original de una línea de concordancia determinada, ampliar el co-texto o acceder al texto original, a través de los distintos programas informáticos. La siguiente figura extraída de nuestro corpus muestra

ejemplos de concordancias realizadas para esta investigación con la palabra Rajoy, uno de los doce personajes analizados en los textos que componen el corpus de la crisis financiera.

|  |  |
|--|--|
| laborar en un proyecto que, si existe, no es fiable? [Compadeczo a Rajoy: ganó para perder]. No sólo porque esté mal rodeado, sino porque era su...<br>le noviembre: «Si esa fuera la herencia envenenada que dejáis a Rajoy, el líder popular no podrá cumplir con la parte sustancial de su programa<br>rejubilaciones ante el Gobierno. Los agentes sociales reclaman a Rajoy que tenga en cuenta su utilidad en ciertos casos. Uno de los acuerdos a<br>y, como afirmamos hoy, un primer ministro de la UE reconoció a Rajoy el jueves que mantener ese objetivo del 4,4% «mataría a España». La única | <b>Comentario [C2]:</b> Antítesis. No se ha obtenido el resultado esperado: ganar para perder.         |
| ...Schroder. Las declaraciones de Merkel suponen un espaldarazo a Rajoy para sacar adelante la reforma. Juan Rosell, presidente de CEOE, dijo a...   | <b>Comentario [C3]:</b> Valoración del periodista de apoyo a Rajoy                                     |
| ...a. Hay razones para manifestarse. Hay motivos para discutir a Rajoy los extremos de un plan de reformas que mantuvo oculto en su programa y qu...<br>ce que llega a su fin. Los mismos que hasta hace poco aplaudían a Rajoy por la valentía de las medidas y las reformas que iba tomando, contrarías  | <b>Comentario [C4]:</b> Crítica a Rajoy. Acusa al presidente de "ocultar" la reforma                   |
| ...s. Todo un gesto. [Quiero desde aquí enviar mis agradecimientos a Rajoy].   |  |
| ...spués de la asunción del Gobierno con mayoría absoluta, Rajoy puede exhibir al menos tres grandes reformas económicas: la laboral, la fi...<br>ondo y en la forma, distinto a lo que nos tiene acostumbrados. Rajoy es un hombre poco dado a apelar a las emociones y a los sentimientos en su  | <b>Comentario [C5]:</b> Valoración. Apoyo a las actuaciones de Rajoy.                                  |
| ...sobre la renta y la mayor tributación sobre el ahorro que activó Rajoy al llegar a La Moncloa. Esa subida de impuestos ya aplicada deja menos diñ...  | <b>Comentario [C6]:</b> Valoración   |
| ...ersonal y político. Merkel respaldó las medidas que ha adoptado Rajoy, pero también puso el énfasis en la importancia de la reforma laboral. Ent...<br>o evitar duplicidades y conflictos de competencias", ha afirmado Rajoy. "De ahí la reforma de nuestros organismos reguladores aprobada por el Cons...  | <b>Comentario [C7]:</b> Valoración. Destaca el verbo "exhibir", no es el apropiado. Mejor presentar... |
| ...frente a la desviación del déficit público. Afortunadamente, Rajoy cuenta con una ventaja de salida: los españoles son perfectamente conscien...<br>nstantia, el crecimiento de nuestra economía", ha agregado. Rajoy, que ayer ya aseguró que no habrá flexibilidad en el recorte exigido a las...   | <b>Comentario [C8]:</b> Evaluación negativa de la imagen.  |
| ...por la huelga general DESPUES de la tormenta de comentarios desencadenada por las elecciones andaluzas en los medios de información, M...<br>o capaces de bajar el déficit conjunto del 2,94% del PIB, y ahora Rajoy pretende que se reduzca a la mitad en un año. El derrumbe de los ingresos  | <b>Comentario [C9]:</b> valoración. Comentario positivo.   |
| ...atero dijo que bajar los impuestos era de izquierdas, ahora Rajoy nos dice por boca de Montoro no sólo que subirlos es de derechas sino que...<br>recuperación probablemente va a seguir siendo lenta. Ahora Rajoy se enfrenta a un momento decisivo. Puede seguir el ejemplo del primer mini   |  |

Figura 3.1: Concordancias con el término (actor social) Rajoy

En el Corpus de la Crisis Financiera (CCF) hemos extraído las concordancias de las palabras que se han considerado como más interesantes, como suele ser habitual en los corpus grandes. El método utilizado es una herramienta informática en la que se introduce la palabra o el patrón a localizar en la concordancia. La palabra aparece en una columna central con las palabras que van antes y después en cada línea. Este es, en ocasiones, el paso posterior a extracción y filtrado y la elaboración de las listas de palabras.

**Collocation/Colocación.** El término *collocation* procede del inglés, de la obra de Palmer y Hornby de 1933, *Second Interim Report on English Collocations*. Sinclair acabó definiendo el término colocación como la tendencia de dos o más palabras a aparecer en la misma concordancia, a combinarse unas junto a otras. En la lingüística de corpus es la secuencia de palabras o términos que ocurren de manera coordinada, o co-ocurren, con mayor frecuencia de lo que puede ocurrir por azar. Existen seis tipos principales de colocaciones:

- Adjetivo + sustantivo
- Sustantivo + sustantivo

- Verbo + sustantivo
- Adverbio + adjetivo
- Verbos + sintagma preposicional
- Verbo + adverbio

Las colocaciones son expresiones parcialmente o totalmente fijas, que se materializan a través del uso repetido dependiente del contexto. Veamos ejemplos de parejas de palabras yuxtapuestas:

- *Punto álgido*
- *Éxito rotundo*
- *Cirugía estética*

La colocación puede corresponder a una relación sintáctica (como verbo-objeto: tomar y decisión), una relación léxica (por ejemplo, antonimia), o puede no existir ninguna relación lingüística definida. El conocimiento de las colocaciones es vital para el uso competente de la lengua: una oración gramaticalmente correcta resultará extraña si no se ajusta a las preferencias colocacionales.

El procesamiento de las colocaciones implica una serie de parámetros. El más importante es la medida de la asociación, que evalúa si la co-ocurrencia es puramente por casualidad o estadísticamente significativa. Debido a la naturaleza no aleatoria de la lengua, la mayoría de las colocaciones se clasifican como estadísticamente significativas y los resultados de la asociación se utilizan simplemente para clasificar los resultados. Las medidas de asociación de uso común incluyen el test de información mutua, el test *T-Score* y el algoritmo de la verosimilitud (*Log likelihood*).

Un corpus lingüístico es una plataforma indispensable para la observación de patrones colocacionales, visible a través de la pantalla vertical de la palabra nodo (es decir, la forma de la palabra o lema objeto de estudio, conocido comúnmente como KWIC, palabra clave en contexto) y su *collocate*, es decir palabra o palabras con la que se produce co-ocurrencia. En

general, se acepta que el comportamiento *colocacional* de naturaleza léxica se puede manifestar hasta una línea de cuatro o cinco palabras a la izquierda o a la derecha de la palabra nodo.

El *Span* sería el número de palabras situadas a la derecha y a la izquierda.

**Tokens, types.** La primera actividad cuando comenzamos a analizar un corpus es la de elaborar una lista de todas las palabras que lo conforman. En la elaboración de las listas de palabras hay que tener bien claros los conceptos *token* (palabra) y *type* (forma). Estas son sus definiciones a partir de la clasificación que establece Barnbrook (1996):

- *Tokens* son el número de palabras totales que aparecen en el corpus, sin importar cuantas veces aparecen en el texto. El número total definirá el tamaño del corpus
- *Types*. Son cada una de las palabras diferentes de un corpus. La suma de las frecuencias de todos los *types* será la suma de todos los *tokens* de un corpus

Bowker y Pearson (2002) en el glosario de su obra proporcionan estas definiciones:

Token: each Word in a corpus. The total number of words in a corpus is referred to as the number of tokens, so if a corpus contains 1 million words, then it has 1 million tokens.

Type: each different word in a corpus. Language is repetitive, which means that some words will appear in the corpus more than once. When calculating the number of types in the corpus you do not count each word (tokens), you only count each different word. So, in a sentence such as 'the cat sat the mat' you have six tokens but only five types.

**Keywords/Palabras clave.** *Keywords* o palabras clave son aquellas que resumen un texto. En la búsqueda de información en la red, por ejemplo, las palabras clave o *Keywords* se utilizan para localizar enlaces de páginas web y se muestran como resultados. Un caso claro, conocido y muy común es el del buscador Google en el que después de teclear una serie de dos o más palabras es capaz de localizar enlaces que nos dan acceso a la información que almacena internet. Es decir, a partir de una serie de palabras clave el ordenador busca los archivos que contienen las palabras que el usuario teclea. En la lingüística de corpus, las palabras clave son las que aparecen con una frecuencia estadísticamente inusual en un texto o corpus de textos. Se identifican por medio de *software* que compara una lista de palabras de los textos con una lista de palabras basada en un corpus de referencia más amplio. Existen diversas herramientas para el tratamiento de los corpus y que ofrecen la posibilidad de realizar listados ordenados de diferentes maneras:

- Por orden alfabético
- Por número de palabras
- Por oraciones, párrafos y longitud

Estos listados pueden ser útiles en el ámbito lexicográfico. Pérez y Moreno (2009) señalan que ayudan a decidir la lista de voces que han de incluirse en un diccionario teniendo en cuenta su frecuencia de uso. También pueden ofrecernos índices de frecuencia en los que se muestre la ratio palabras/formas o el número total de palabras de un texto frente al número de palabras diferentes que aparecen en el mismo, o comparar los índices de varios ficheros de texto. En el análisis del corpus que Pérez y Moreno (2009) realizan vemos una diferenciación de análisis complementarios. El análisis cualitativo es el que ofrece una descripción detallada y completa de un fenómeno lingüístico o del comportamiento de una palabra o grupo de palabras. En el cuantitativo se asignan índices de frecuencia a los fenómenos lingüísticos observados en el corpus y estos pueden servir para construir modelos estadísticos más complejos, que expliquen la evidencia hallada en el texto. Pérez y Moreno (2009) detallan que en el mercado existe un importante número de programas - comerciales y gratuitos para fines académicos- con interfaces de usuario de fácil manejo versátil y sofisticado. Enumera algunos de estos programas de manejo de corpus para fines académicos:

- *Free text Browser* de la Universidad de Michigan
- *TACT* del *Department of Computing in the Humanities and Social Sciences* de la Universidad de Toronto

Respecto a los comerciales señala como los más usados:

- *Oxford Concordancing Program*
- *MicroConcord*
- *Word Cruncher*
- *WordSmith*

El procedimiento utilizado por *WordSmith* es el que hemos empleado para la presente investigación. El *software* creado por Scott (2012) ofrece la posibilidad de crear y desarrollar las

listas de las palabras clave. Sinclair (1991: 4) destaca la importancia de las herramientas informáticas como ésta que acabamos de presentar, para el análisis de grandes textos: “The ability to examine large text corpora in a systematic manner allows access to quality of evidence that has not been available before”.

Añadimos a la anterior cita la apreciación de Leech (1991: 13) sobre las herramientas informáticas para el análisis de textos: “The availability of vastly increasing computer corpus resources has enabled syntactic and lexical phenomena of a language to be open to empirical investigation on a scale previously unimagined”.

**Colligation/Coligación.** Un fenómeno relacionado con la función de proximidad es la conocida como coligación. Hoey (1998) definió el término *colligation* como la compañía gramatical que una palabra mantiene y el lugar en una secuencia que una palabra prefiere. En el caso de un vecindario sería la relación que esa palabra mantiene con sus respectivos vecinos con los que convive. Veamos, por ejemplo, la siguiente secuencia extraída del *Corpus de la Crisis Financiera* con el sustantivo *políticas* seguido de preposición (*de*) más sustantivos abstractos con cargas negativas y positivas.

| EVALUACIÓN | POLÍTICAS+PREPOSICIÓN+SUSTANTIVO ABSTRACTO |
|------------|--|
| Positiva   | POLÍTICAS DE DESARROLLO                    |
| Positiva   | POLÍTICAS DE CONCILIACIÓN                  |
| Positiva   | POLÍTICAS DE CRECIMIENTO                   |
| Positiva   | POLÍTICAS DE CONCILIACIÓN                  |
| Positiva   | POLÍTICAS DE PROTECCIÓN                    |
| Negativa   | POLÍTICAS DE RESTRICCIÓN                   |
| Negativa   | POLITICAS DE RECORTES                      |
| Negativa   | POLITICAS DE REDUCCIÓN                     |
| Negativa   | POLITICAS DE EXPATRIACIÓN                  |
| Negativa   | POLITICAS DE EMPOBRECIMIENTO               |

Tabla 3.2: Patrones gramaticales

Otros ejemplos de patrones gramaticales en castellano serían los que se forman con los verbos:

- *consistir* + preposición *de* + sustantivo
- *depender* + preposición *de* + sustantivo

**Frequency/Frecuencia.** Las listas de palabras vienen acompañadas de información acerca de su frecuencia en el corpus analizado. Una lista de frecuencias de un corpus muestra las palabras que se repiten y la proporción relativa en que cada una de ellas contribuye en el mencionado corpus.

Las listas básicas de frecuencias de palabras son muy útiles en las etapas preliminares de análisis, ya que proporcionan un buen punto de partida para una investigación al detalle de las palabras gramaticales o de contenido del texto.

El ejemplo más básico de la lista de frecuencias muestra el número de *types* (ocurrencias individuales de cualquier forma de la palabra) y sus *tokens* (es decir, frecuencia) en el corpus. Las listas de *types* y *tokens* están normalmente ordenadas alfabéticamente o por índice de frecuencia. Se pueden encontrar resúmenes de diferentes tipos de listas de frecuencia en los trabajos realizados por Bowker y Pearson (2002) y Sinclair (1991).

Una lista de palabras será buena si lo es el corpus del cual deriva. Por este detalle es importante el tamaño del corpus ya que según sea éste, generará los datos de frecuencias. En un texto muy largo la mitad del vocabulario habrá ocurrido una sola vez. Un corpus debe contener muchos millones de palabras con el fin de generar un número suficiente de ocurrencias de los diferentes usos en los que basar enunciados descriptivos. A la hora de fijar el tamaño del corpus tendremos que tener en cuenta el objeto de la investigación y si ésta es de carácter general o especializado. En los corpus más bien pequeños las listas de frecuencias pueden arrojar datos más útiles ya que queda mejor concentrada la información de una determinada área.

Es importante aclarar qué conocemos por frecuencia, ya que es difícil establecer en relación a qué. Por lo general son archivos de texto que ofrecen información cuantitativa y estadística. A nivel particular, la frecuencia ofrece información acerca de cómo las palabras se distribuyen en el texto. Por lo tanto, y a modo de resumen, podemos establecer que existen:

- Frecuencias de ocurrencia de los elementos lingüísticos. Son las frecuencias con morfemas, palabras, patrones gramaticales etcétera, que ocurren en un corpus o en parte de un corpus.
- Frecuencias de co-ocurrencia de estos elementos. Con qué frecuencia se producen los morfemas con determinadas palabras y con qué frecuencia se producen determinadas palabras en una determinada construcción gramatical.

La frecuencia no es sólo un fenómeno relevante que se da en un corpus de 100 a 1.000 ocasiones. Va algo más allá. Incluye palabras que ocurren una sola vez, o *hápax legomena*. La lingüística de corpus tiene que ver con si un elemento individual, o la ocurrencia de más de un elemento individual, se atestigua en los corpus o si la frecuencia observada de ocurrencia o co-ocurrencia es menor o mayor.

**Semantic prosody/Prosodia semántica.** Es un fenómeno colocacional particular. Algunas palabras tienen colocaciones negativas, otras tienen colocaciones positivas y otras son simplemente neutras. El término prosodia semántica se utiliza a menudo para describir los casos en los que una colocación típica de una palabra le confiere un determinado significado evaluativo. Es el caso de palabras como *muerte* o *enfermedad* que vienen acompañadas de verbos con connotaciones negativas, mientras que palabras como *oportunidad* o *ayuda* suelen estar acompañadas de verbos con connotaciones positivas.

Sinclair (2004: 19-20) utiliza la prosodia semántica en el ámbito general para referirse al fenómeno que comparten las palabras y las frases. Sinclair argumenta que las palabras no constituyen selecciones independientes y que el significado de las palabras unidas puede ser diferente al significado de esas mismas palabras por separado. Sinclair (1987: 319) señala que esto conduce a una *delexicalización* de las palabras: “A language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments”.

**Genre analysis/Análisis de género.** Uno de los grandes exponentes del estudio del género en el corpus es Biber (1998), que adoptó un enfoque de análisis multidimensional con el objeto de realizar distinciones entre los distintos géneros y descubrir sus características lingüísticas. El

enfoque multidimensional se basa en la idea de que si se producen algunos rasgos lingüísticos con frecuencia en un texto, aparecerán con menor frecuencia otras características en el mencionado texto. Los textos de conversaciones tienen una parte importante interactiva y mayor implicación por parte del autor que los textos académicos, que se caracterizan por la objetividad y el distanciamiento.

La investigación de Stuart (2006) es un ejemplo de estudio basado en un corpus de un género. En este trabajo, analiza tanto los propósitos comunicativos y rasgos lingüísticos formales de páginas web o de blogs académicos. La hipótesis de partida es que los blogs académicos son mucho menos formales y favorecerían un mayor uso de estructuras léxico-gramaticales de función interpersonal del lenguaje. Para cuantificar la posible variación entre géneros académicos tradicionales y blogs académicos, se investigaron algunas características lingüísticas típicas del discurso académico como la forma pasiva sin agente.

Después de examinar los resultados de 26 rasgos lingüísticos, se encontró con menos sorpresas que las que inicialmente imaginaba. Las posibilidades del idioma en páginas web tienen una mayor interactividad, aunque mantienen una gran influencia del discurso de la prosa académica tradicional.

***Dispersion/Dispersión.*** Aunque las frecuencias de co-ocurrencia son fundamentales y ubicuas en la lingüística del corpus, son potencialmente engañosas a menos que también se dé una medida de dispersión, es decir, la propagación o variabilidad de una distribución de valores. Las frecuencias de palabras u otros elementos lingüísticos se usan a menudo para indicar la importancia en un corpus, pero consideremos un corpus en el que dos palabras tienen el mismo número, sólo la primera palabra ocurre en el 99% de los documentos del corpus, mientras que la segunda palabra se concentra en sólo el 5%. ¿Qué palabra es "más importante"? ¿Y cómo debemos interpretar estadísticas basadas en estas frecuencias si el alto valor de la segunda palabra no es representativo de la mayoría del corpus?

Una forma común, aunque no totalmente ideal, de explicar la dispersión en la lingüística del corpus es calcular la frecuencia ajustada de las palabras, que a menudo es simplemente la frecuencia multiplicada por la dispersión. Tales frecuencias ajustadas son, por definición, alguna

fracción de la frecuencia absoluta, y las palabras con baja dispersión son penalizadas más que aquellas con alta dispersión.

A modo de cierre de este decálogo, es importante dejar claro que la intención ha sido tratar de poner luz en el conocimiento de los conceptos clave y que son básicos para utilizar y comprender la lingüística de corpus. Podemos concluir este apartado con que la lingüística de corpus se encarga de la recogida de datos, escritos o transcritos de un discurso grabado y que se pueden utilizar como punto de partida para la descripción lingüística o como un medio de verificación de hipótesis sobre el lenguaje. La lingüística de corpus aborda la búsqueda de patrones que se usan en la lengua como datos empíricos de los corpus y que nos aportan información lingüística. El análisis empírico de la lengua se realiza mediante el estudio de los datos del lenguaje natural de un corpus.

### **3.5 Tipología de corpus. ¿Para qué se utilizan? ¿Qué analizan?**

Ya hemos visto que la lingüística de corpus es una materia que ha ido variando a lo largo de su historia. La evolución de los corpus lingüísticos nos obliga también a tener en cuenta varios aspectos importantes como son, por ejemplo, su tamaño, su tipología o nivel o tipo de marcaje. En cuanto al tamaño, es mucho lo que se ha avanzado desde los primeros corpus. Entre el *Brown Corpus* (de un millón de palabras) a los actuales de cientos de millones de palabras. En este aspecto, volvemos a insistir en el papel que ha tenido el desarrollo de la informática y su contribución en el almacenamiento de datos y la creación de los megacorpora, un concepto que Leech incorpora en 1992.

La otra variable es la tipología, la cual vamos a tratar de aproximar en este apartado. La tendencia a seguir sería establecer un equilibrio entre los corpus orales y los escritos. La ayuda de las nuevas tecnologías es aquí, de nuevo, un aspecto clave y que puede contribuir a la conversión automática de habla al texto y viceversa.

Y en lo referente al nivel de marcaje, volvemos a recurrir a Leech (1992) para establecer que el corpus ha de superar la sintaxis para abordar aspectos pragmáticos y discursivos de los textos que conforman el corpus. Si la elaboración de un corpus va a depender de su finalidad última,

vamos a desgranar la utilidad final de algunos de los principales corpus que se conocen, su magnitud y características, para comprender cómo y para qué han sido creados.

- I. **Corpus diacrónico español.** El Corpus Diacrónico del Español, CORDE, recopila información de todas las épocas y lugares en que se habló el idioma español o castellano hasta 1975. Es monolingüe y su diseño se ha realizado de manera que de él se puede extraer información para estudiar las palabras, sus significados, la gramática y su utilización a lo largo del tiempo. Ha llegado a ser una fuente obligada para cualquier estudio diacrónico relacionado con la lengua española y ha servido de material básico para la confección del Diccionario Histórico. Contiene textos escritos en prosa y verso y dentro de éstos los narrativos, líricos, dramáticos, científico técnicos, históricos, jurídicos, religiosos, periodísticos en todas las variedades geográficas, históricas y genéricas. Su origen se remonta a 1994 y cuenta con más de 260 millones de palabras, lo que constituye el mayor volumen de conjunto de registros de la historia de la lengua española. Se hizo porque la Real Academia Española se planteó utilizar herramientas informáticas para elaborar un almacén de información que mejorase la calidad de sus materiales de trabajo y su acceso. Se utiliza para documentar palabras, clasificarlas de anticuadas, para conocer el origen de algunos términos, su tradición y primeras apariciones.
  
- II. **Corpus de referencia del Español Actual (CREA).** La Real Academia Española constituyó este banco de datos del español contemporáneo para la obtención de un corpus de referencia del español actual. Están almacenados en soporte informático que permite el estudio de las palabras, sus significados y contextos.

Es un corpus general y de referencia compuesto por una amplia variedad de textos orales y escritos que provienen de países de habla hispana. El CREA comenzó a elaborarse en 1975 con registros de procedencia tan diversa como periódicos y revistas, y transcripciones de documentos sonoros de la radio y la televisión. Es la única herramienta lingüística de gran magnitud para la lengua española. Es de gran utilidad para investigaciones de índole tan diversa como las lingüísticas o las de la publicidad, la terminología o la sociología y para la elaboración de productos derivados como gramáticas, diccionarios, tesauros, correctores ortográficos,

métodos de didáctica del español, desarrollos informáticos de traducción automática, entre otros.

- III. **Corpus del Español mexicano Contemporáneo (CEMC).** Es también un corpus de referencia. Se elaboró en 1973 para la realización del Diccionario del Español de México. Su finalidad ha sido deslindar y reconocer las características del español mexicano contemporáneo, sobre todo en cuanto a su léxico. Quedó formado por 996 textos, de 200 palabras cada uno, escritos por autores mexicanos desde 1921 a 1974, de transcripciones de conversaciones grabadas del *Atlas Lingüístico de México* y otros bienes lingüístico- etnográficos de la biblioteca del Colegio de México.
- IV. **Corpus del Español de Mark Davies.** El corpus elaborado por Mark Davies, profesor de la *Brigham Young University*, contiene 100 millones de palabras del español de los siglos XIII al XX. Cuenta con un buscador que permite localizaciones complejas.
- V. **Corpus del Grupo de Ingeniería Lingüística (GIL).** El Grupo de Ingeniería Lingüística del Instituto de Ingeniería de la Universidad Nacional Autónoma de México, UNAM, cuenta con tres corpus: el Corpus Lingüístico en Ingeniería (CLI), el Corpus Histórico del Español de México (CHEM) y el Corpus de las Sexualidades en México (CSMX).
- VI. **Archivo de textos hispánicos (ARTHUS).** El Archivo de los textos hispánicos de la Universidad de Santiago de Compostela está elaborado con textos de diferentes etapas de la historia del español, introducidos en ordenador por escáner y programas de reconocimiento óptico de caracteres en formato ASCII (acrónimo inglés de *American Standard Code for Information Interchange-Código Estándar Estadounidense para el Intercambio de Información*), con una codificación mínima en formato COCOA (*Count and Concordance on Atlas*), para conocer textos, página y línea de los ejemplos buscados, mediante programas de recuperación adecuados. Incluye textos de distintos períodos de la historia de la lengua española de diversos géneros literarios y periodísticos y transcripciones orales. Es un corpus textual y específico, con textos históricos que sólo abarcan el español. Su parte

contemporánea cuenta con treinta y cuatro textos narrativos, teatrales, ensayos, de periódicos y orales de España e Hispanoamérica.

- VII. Proyecto *Corpus Resources and Terminology Extraction (CRATER)*.** Este corpus multilingüe (español, francés e inglés) lo componen textos técnicos de la *ITU (International Telecommunications Union)*. Destaca por su etiquetado morfológico, su alineado de frases con equivalentes en francés e inglés. Para este proyecto fue creado un etiquetador morfosintáctico *part-of-speech tagger* en español. Contiene un millón de palabras y es muy útil para proyectos de traducción automática, lingüística computacional y corpus en general. Es un corpus textual y específico ya que fue creado para la ITU.
- VIII. Corpus textual especializado plurilingüe.** Es un proyecto de investigación prioritario del Instituto Universitario de Lingüística Aplicada (IULA) de la Universidad Pompeu Fabra de Barcelona. Es su principal soporte de las actividades de investigación y docencia. Está formado por textos escritos en cinco idiomas (catalán, castellano, inglés, francés y alemán). Las áreas que abarca son economía, derecho, medio ambiente, medicina e informática. Con este corpus se intentan inferir las leyes que rigen el comportamiento de cada lengua en cada área. Sus investigaciones van encaminadas a la detección de neologismos y términos, estudios de variación lingüística, análisis sintáctico parcial, alineación de textos, extracción de datos para la enseñanza de segundas lenguas, extracción de datos para la construcción de diccionarios electrónicos o elaboración de tesauros, entre otros. En su elaboración los textos son seleccionados por especialistas de cada área y agrupados en una clasificación temática de uso propuesta por los especialistas en cada una de las áreas. A continuación, los textos son marcados siguiendo el modelo de *Standard Generalized Markup Language*-Estándar de Lenguaje de Marcado Generalizado (SGML) y las directrices del *Corpus Encode Standar (CES)* de *EAGLES*.
- IX. Archivo Gramatical de la Lengua Española (AGLE).** Este está orientado hacia los estudios gramaticales. Lo forman más de 100.000 fichas de texto escrito y oral que recogió el gramático español Salvador Fernández Ramírez. Ilustra construcciones

sintácticas del español y sus fichas están ordenadas como una Gramática. En su primera entrega consta de 75 ficheros y puede ser consultado en el Instituto Cervantes, donde se está anotando y editando.

- X. Base de datos ETDEWEB (*Energy Technology Data Exchange World Energy Base*).** Es un corpus textual, especializado o específico. Cuenta con más de 3'8 millones de archivos, lo que lo convierte en la base de datos con la mayor colección de literatura sobre energía. Estos datos contienen referencias bibliográficas, artículos de prensa, informes, conferencias, libros y otros tipos de documentos. Cubre además, aspectos medioambientales sobre el uso y protección de la energía y las ciencias básicas que apoyan la investigación en la energía y su desarrollo. Entre sus usuarios están los científicos, ingenieros, bibliotecarios, industriales y estudiantes. Está a la disposición de miembros de ETDEWEB de México, Estados Unidos, Japón, a través de Internet, en otros formatos on line o en CD-ROM.
- XI. María.** Susana López Ornat, de la Universidad Complutense de Madrid, transcribió este primer corpus formado por las interacciones de una niña desde el año y medio a los cuatro con sus padres. Contiene información lingüística con etiquetas morfológicas y psicolingüísticas. Se codificó con información ortográfica.
- XII. Corpus de Derecho Ambiental Mexicano (CDAM).** Es una compilación de leyes, reglamentos y Normas Oficiales Mexicanas de ecología de 576 palabras clasificadas atendiendo a las fuentes. Lo recopiló Elena Bogomilova Lozanova del Centro de Estudios Lingüísticos y Literarios.
- XIII. *British National Corpus (BNC)*.** Está formado por 100 millones de palabras. El 90% de este corpus lo conforman textos escritos en inglés británico y el 10% restante son textos hablados. Para estos últimos se utilizaron conversaciones informales de voluntarios de diversas edades, sexo, región y clase social. En las muestras había lenguaje hablado en reuniones gubernamentales, programas de radio o llamadas telefónicas. Se elaboró entre 1991 y 1994 a cargo de un consorcio de editoriales,

centros de investigación, universidades, *The British Library* y patrocinadores de la industria.

### 3.5.1 Clasificación de los corpus

Hunston (2002) establece una clasificación de los principales corpus. Algunos de ellos los acabamos de nombrar y describir. Veamos a continuación esta reagrupación, según las características de los corpus:

- a. Corpus generales o de referencia**<sup>37</sup>. Son aquellos que contienen vocabulario estándar de una lengua, como el *Brown Corpus*, el *Bank of English* o el CREA<sup>38</sup>.
- b. Corpus especializados**. Son los formados por textos especializados en algún ámbito, como el Corpus del DICTER<sup>39</sup>.
- c. Corpus diacrónicos**. Son los compuestos por textos de distintas épocas, como el ARCHER<sup>40</sup> o el CORDE<sup>41</sup>.
- d. Corpus monitor**<sup>42</sup>. Son los que están abiertos, continuamente actualizándose. Es el caso del *Bank of English*.
- e. Corpus paralelos**<sup>43</sup>. Son útiles para el estudio de los mecanismos de traducción.
- f. Web and Corpus**. El uso de internet como un gran corpus, como es el caso del buscador de Google.

### 3.6 Diseño de corpus textuales

A la hora de diseñar un corpus hemos de tener en cuenta cuestiones como la aplicación que le vayamos a dar, qué tipo de corpus queremos elaborar, si hablado o escrito, si ha de ser estático o dinámico, general o especializado, monolingüe o multilingüe, comparable o paralelo. El objeto

---

<sup>37</sup> "Reference corpora contain the standard vocabulary of a language. They are the corpus linguist's main resource to learn about meaning" (Teubert, 2007: 67)

<sup>38</sup> Corpus de Referencia del Español Actual

<sup>39</sup> Diccionario de la Ciencia y la Técnica en el Renacimiento

<sup>40</sup> Ancient Corpus Historical English Register

<sup>41</sup> Corpus Diacrónico del Español

<sup>42</sup> "The monitor corpus is a corpus that monitors language change. It is, in principle, regularly updated and open" (Teubert, 2007: 71)

<sup>43</sup> "A parallel corpus, sometimes also called a translation corpus, is a corpus of original texts in one language and their translations into another (or several other languages)" (Teubert, 2007: 73)

de una determinada investigación a través de un corpus requerirá un método de elaboración concreto. Por ejemplo, puede ir desde el análisis transcripciones policiales, el discurso político de los cargos públicos, el discurso de los sospechosos de terrorismo, la autoría de un libro o un análisis para detectar el plagio, o el vocabulario que utilizan los profesionales de la medicina, profesores, abogados, el discurso de la publicidad o el televisivo. Una vez sepamos el tipo de corpus, su tamaño, la cantidad de textos a recopilar, trabajaremos en aspectos como la posible proporción en relación a otros *subcorpus*<sup>44</sup>, la extensión de los textos o fragmentos de texto completo, los temas, áreas temáticas, dominios que se incluirán, los géneros y tipos de texto que se incluirán.

En lingüística aplicada se trabaja con datos reales que permitan reproducir las características del objeto de estudio, lo que implica que hay que recopilar muestras más o menos grandes de lo que se quiere observar. El progreso de la informática ha ayudado al investigador en las tareas mecánicas de recopilación y organización en formato electrónico de los textos. Torruella y Llisterri (1999) se refieren a Svartvik (1992) para destacar que fue este autor quien señaló que la lingüística basada en los corpus hacía posibles aproximaciones a viejos problemas y convertir las ideas en afirmaciones fundadas. Es decir, los corpus están compuestos por datos reales y sus resultados son empíricos. Son empíricos porque su función principal es restablecer la relación entre la teoría y sus datos. En opinión de Torruella y Llisterri (1999), es necesario que un corpus sea neutro y que recoja muestras proporcionadas de todos sus aspectos. Los autores llegan a la conclusión de que un corpus nunca puede ser la realidad, sino un modelo de ésta y mostrar sus aspectos más característicos. El corpus ha de ser selectivo ya que no puede recopilar todo lo escrito y/o hablado de una lengua. Este carácter selectivo puede limitar sus conclusiones.

Para que los corpus faciliten la extracción de datos homogéneos y cuantificables consideran Torruella y Llisterri (1999) necesario restringir las ocurrencias léxicas a formales, las variantes a

---

<sup>44</sup> Un corpus puede dividirse en *subcorpus*. Un *subcorpus* tiene todas las propiedades de un corpus, pero pasa a ser parte de un corpus más grande. Corpus y *subcorpus* se dividen en componentes. Un componente no es necesariamente una muestra adecuada de un lenguaje y de esa manera es distinto de un corpus y un *subcorpus*. Se trata de una colección de piezas de la lengua que son seleccionados y ordenados de acuerdo a un conjunto de criterios lingüísticos que sirven para caracterizar su homogeneidad lingüística. Mientras que un corpus puede ilustrar la heterogeneidad, y también un *subcorpus* en cierta medida, el componente ilustra un tipo particular de lenguaje.

invariantes. Las variantes comprenden las gráficas, fonéticas, de naturaleza diatópica y las producidas por la polisemia de las lenguas. Defienden que la creación, mantenimiento o actualización de los corpus es una ciencia interdisciplinar en la que intervienen historiadores, sociolingüistas, matemáticos, informáticos y teóricos de la literatura en la que se debe definir las etapas diacrónicas, la variedad temática o la conveniencia de trabajar con documentos enteros o fragmentos. A la hora de analizar la variedad de corpus resaltan que son unas herramientas muy buenas en la investigación y que facilitan bases provechosas para comparar variedades de una lengua o explotar sus aspectos cuantitativos y probabilísticos. Acerca de los corpus informatizados defienden Torruella y Llisterra (1999) que han dado un nuevo impulso al estudio descriptivo de la lengua en diversos aspectos como la prosodia, el léxico, morfología, sintaxis e historia de la lengua. Según estos autores, la informática nos ayuda a recopilar en menos tiempo y de manera más ordenada un corpus. Los avances más significativos se han producido en el área de la creación de modelos probabilísticos de la lengua y como pruebas para verificar estos modelos, lo que ha permitido avanzar en el análisis gramatical automático de textos en aspectos morfológicos o *tagging*<sup>45</sup> (etiquetado) y en sus aspectos sintácticos o *parsing*<sup>46</sup> (análisis sintáctico). Concluyen con que la ventaja de trabajar con corpus informatizados es tal que los lingüistas tradicionales se están viendo obligados a trabajar conjuntamente con lingüistas computacionales. Ambos comparten la finalidad de entender mejor cómo funciona el lenguaje humano, aunque la primera tarea sea la obtención de datos.

En la planificación de la construcción del corpus, en el caso de que la fuente sean los periódicos, como es el caso de esta investigación, hemos de tener en cuenta otras variables como:

- ¿Cuáles son los objetivos de la construcción de un corpus?
- Criterios de diseño (periódico, número de textos y recuento de palabras)

---

<sup>45</sup> En lingüística computacional el etiquetado gramatical (conocido también por su nombre en inglés, *part-of-speech tagging*, *POS tagging* o *POST*) es el proceso de asignar (o etiquetar) a cada una de las palabras de un texto su categoría gramatical. Este proceso se puede realizar de acuerdo con la definición de la palabra o el contexto en que aparece, por ejemplo su relación con las palabras adyacentes en una frase, oración o en un párrafo.

<sup>46</sup> En la lingüística computacional, sistema de procesamiento de datos para reconocer y analizar una frase escrita en un lenguaje natural. Cualquier procedimiento para obtener la información en una base de datos no estructurados originalmente.

- Subgéneros en el corpus (De información general: noticia, crónica, entrevista, reportaje. De opinión: editorial, artículo, análisis, columna, cartas al director)
- ¿Cómo se recogen los textos?
- ¿Qué información necesitamos almacenar acerca de los textos?

Otro aspecto a tener en cuenta es el concepto de *reusability (replicability)* (McEnery y Wilson, 2001: 32), y que traducimos como *reutilización (replicabilidad)*.

The advantage of a widely available corpus is that it provides a yardstick by which successive studies may be measured and that a continuous base of data is being used and thus variation between studies may be less likely to be attributed to differences in the data being used, and more to the adequacy of the assumptions and methodologies contained in the study.

### 3.7 Herramientas para el análisis de corpus

La evolución de los corpus ha hecho necesaria la introducción de herramientas para su elaboración y análisis. El análisis cualitativo de un corpus nos permite una descripción detallada y completa de un determinado fenómeno lingüístico o del comportamiento de una palabra o grupo de palabras. Este tipo de análisis establece índices de frecuencia de los fenómenos lingüísticos observados en el corpus.

En la actualidad hay diversos programas para analizar un corpus. Las consultas pueden realizarse por el tipo de interfaz, que permite la conexión física y funcional entre dos sistemas o dispositivos y por el nivel de marcaje del corpus. En el mercado hay disponibles programas informáticos que permiten extraer concordancias y que operan sobre textos planos sin necesidad de ningún tipo de marcaje. *Antcoc*, por ejemplo, es un programa de análisis de concordancias y que dispone de versiones para Windows (3.4.3), Macintosh OS X (3.4.3.), Linux (3.4.3) y versiones más antiguas.

Existen herramientas que realizan el cómputo de palabras de manera informatizada para el procesamiento lingüístico del corpus. Una herramienta destacada son los etiquetadores morfológicos<sup>47</sup> que marcan gramatical y morfológicamente todas las palabras de un texto. Estos

---

<sup>47</sup> Existen varios tipos de etiquetadores morfológicos con diferentes repertorios de etiquetas. Algunos están basados en reglas, como el programa *TAGGIT* de la Universidad de Brown. Otros en cálculos probabilísticos, como el programa *CLAWS*, de la Universidad de Lancaster.

sistemas se desarrollan con la interacción de expertos que analizan los errores y retroalimentan el programa de modo que los resultados se vuelvan cada vez más precisos. Otro tipo de herramienta son los analizadores sintácticos (*syntax parser*)<sup>48</sup> que a través de programas informáticos identifican, analizan y agrupan (*syntax chunkers*)<sup>49</sup> las unidades sintácticas de un corpus (Jurafsky y Martin, 2000). Los etiquetadores morfológicos y los analizadores sintácticos utilizan subsistemas o subprogramas que permiten obtener corpus etiquetados gramaticalmente (*tagged corpora*)<sup>50</sup> o analizados sintácticamente (*parsed corpora*)<sup>51</sup> con un grado de precisión y confianza bastante alto.

### 3.8 Aplicaciones de la lingüística de corpus

Un corpus, especialmente si es de gran tamaño y elaborado en soporte informático, es una importante fuente de información del uso de una lengua, acerca de la gramática, la semántica, el discurso. La lingüística de corpus ha llegado a ser la base de redacción de diccionarios como el *COBUILD* para estudiantes avanzados de inglés (Sinclair, 1987a) y también un elemento clave en el avance de la lingüística computacional, para detectar estructuras lingüísticas habituales. Biber (1998) aporta otros fines como la traducción, a través del empleo de corpus bilingües y el procesamiento de la voz. La enseñanza es también un campo de aplicación de corpus como herramienta de investigación para detectar y reconocer usos y estructuras más frecuentes. McEnery y Wilson (1996) añaden la posibilidad de utilizar esta disciplina en el aula como herramienta de apoyo para el aprendizaje de una segunda lengua. Son distintos los usos y aplicaciones de la lingüística de corpus, de ahí la creación de corpus especializados.

---

<sup>48</sup> Es el proceso de analizar una serie de símbolos, ya sea en lenguaje natural o en lenguajes de programación, conforme a las reglas de una gramática formal.

<sup>49</sup> También conocidos como analizadores de poca profundidad (Abney, 1991). Proporcionan un análisis sintáctico parcial y simplificado.

<sup>50</sup> Un ejemplo es *The Brown University Standard Corpus* (o Brown Corpus) compilado en 1960 por Henry Kucera y W. Nelson Francis. Contiene 500 muestras de texto, por un total de aproximadamente un millón de palabras, de obras publicadas en EEUU en 1961.

<sup>51</sup> *TreeBank* o *Penn Treebank* es un ejemplo de corpus lingüístico que ha sido anotado (*parsed*) con superestructura sintáctica que se ha representado generalmente como una estructura arbórea que recibe la denominación de *TreeBank*.

En la industria de la lengua y las áreas de investigación relacionadas destacan las siguientes aplicaciones lingüísticas:

- La investigación lingüística. Léxico, sintaxis, morfología, semántica, análisis del discurso, pragmática, género y registro de análisis, estilística, sociolingüística, estudios culturales, dialectología, investigación del habla, desarrollo de modelos de lenguaje
- Enseñanza de lenguas: plan de estudios, diseño de materiales y aprendizaje basado en datos
- Edición de libros de referencia: diccionarios, gramáticas, materiales didácticos, guías de uso, tesauros, glosarios, ontologías
- Traducción automática
- Procesamiento del lenguaje natural / Inteligencia Artificial
- Reconocimiento de voz y síntesis

### **3.8.1 Análisis de corpus especializados**

El corpus que vayamos a trabajar en una determinada investigación tendrá que adaptarse a las necesidades u objetivos que se persigan. Un corpus puede estar elaborado a partir de textos periodísticos, conversaciones informales, bases de datos, redacciones escolares, informes, anuncios, etcétera. Sea cual sea su elaboración, ha de ser representativo del uso lingüístico del grupo específico a analizar. Será entonces un corpus especializado.

En la construcción del corpus lingüístico se han de tener en cuenta una serie de variables como el espacio temporal del cual se han recopilado textos, el contexto social y el tema central a analizar. En este sentido, los textos orales o escritos que recojamos en un corpus han de estar enmarcados dentro de un período de tiempo concreto, al seleccionarlos debemos que tener en cuenta el entorno social de donde los hemos extraído, si es más o menos distendido, si se trata de una fuente oficial y en tercer lugar ha de compartir una misma temática, por ejemplo, artículos publicados por la prensa española sobre un asunto en concreto.

¿Cuáles son los criterios a tener en cuenta en la elaboración de un corpus para que éste sea representativo? En la selección de textos para la elaboración de un corpus hemos de atender las

palabras de Bowker y Pearson (2002: 45-47) con las directrices generales para la recopilación de datos.

- I. **Tamaño del corpus.** Aunque hemos visto que un corpus es una colección grande de textos, el adjetivo *grande* resulta un tanto vago. ¿Qué normas o factores delimitan el tamaño del corpus? ¿Cuál es el tamaño ideal de un corpus? Estas cuestiones encuentran su respuesta cuando el investigador define cuáles son sus necesidades, los objetivos de su proyecto, la disponibilidad de material y el tiempo que dispone para su ejecución. No podemos asumir que cuanto más grande es el corpus mejores resultados se van a obtener. En ocasiones, se puede extraer información más útil de un corpus pequeño, pero bien diseñado, que de uno grande que no está adaptado para satisfacer nuestras necesidades. Los primeros corpus electrónicos de lenguaje para fines generales, como *The Brown Corpus* y *Lancaster-Oslo Bergen (LOB) Corpus*, contenían aproximadamente un millón de palabras. En la actualidad, ante la creciente disponibilidad de textos en formato electrónico, no es inusual encontrar corpus generales que contienen cientos de millones de palabras (por ejemplo, *The British National Corpus*, o *Bank of English*). Está reconocido, no obstante, que los corpus especializados sean de menor tamaño que los generales. Las razones para ello son, por una parte, que resulta más difícil encontrar muestras de textos especializados y, por otra, el lenguaje para fines específicos representa un subconjunto más restringido del lenguaje natural.

Sin embargo, los corpus pequeños presentan algunos inconvenientes que cabe mencionar. Si el corpus es demasiado pequeño, puede que no contenga todos los conceptos, términos o patrones lingüísticos que son relevantes al lenguaje especializado que se está investigando. Incluso, si el corpus es demasiado pequeño, puede que no sea posible hacer generalizaciones válidas. Es importante disponer de un corpus sustancial si queremos hacer afirmaciones basadas en frecuencia estadística.

- II. **Finito o infinito.** Otro criterio que se deberá adoptar y que está relacionado con el tamaño del corpus es si queremos que éste sea abierto o cerrado. Esto es, si

queremos que recoja aspectos de uso del lenguaje en un momento determinado (corpus finito o cerrado) o si, por el contrario, nuestro interés se centra en un corpus abierto o monitor, una entidad más flexible a la que podremos añadir o quitar textos para reflejar los cambios que se producen en la lengua. El lenguaje de especialidad es dinámico, pues los conceptos en las disciplinas de especialidad están constantemente evolucionando y los términos que se utilizan para describir esos conceptos también cambian. Dada la naturaleza dinámica de las lenguas de especialidad, un corpus abierto que se pueda actualizar de forma regular será probablemente más adecuado para muchas de nuestras necesidades en el estudio de los lenguajes especializados. A continuación, recogemos ejemplos de estudios de corpus lingüísticos. Tres tipos principales de corpus que se basan en el estudio lingüístico:

- *Léxico: por ejemplo, el uso de palabras, expresiones idiomáticas, colocación, léxico valorativo, prefijos, sufijos, antónimos, sinónimos, etcétera*
- *Sintáctico: características en el nivel de oraciones, por ejemplo, frases preposicionales, formas verbales, modales, deixis, grupos nominales, adverbiales, etcétera*
- *Discursivo: la estructura del texto, por ejemplo, la cohesión por encima del nivel de la oración*

**III. Extracto de textos versus textos completos.** Superada la cuestión del tamaño del corpus en su conjunto, a continuación deberemos plantearnos el tamaño de cada uno de los textos que componen nuestros corpus. En algunos corpus generales, el tamaño de cualquiera de las muestras de los textos puede ser limitado. Por ejemplo, el LOB (*Lancaster-Oslo-Bergen Corpus*) contiene extractos de 2000 palabras y el LLELC (*Longman-Lancaster English Language Corpus*) está formado por extractos de alrededor de 40.000 palabras. En los estudios de lenguajes de especialidad, sin embargo, los conceptos, términos, patrones y contextos que son de nuestro interés pueden aparecer en cualquiera de las secciones del texto. De hecho, su aparición en una u otra sección puede ser relevante.

Un concepto importante puede que esté definido en el párrafo introductorio o un término completo, como *read only memory*, puede que se utilice a principio de un texto mientras que su forma abreviada *ROM* puede aparecer más adelante. Si decidimos elegir un extracto al azar, podríamos eliminar de forma accidental una parte del texto que podría ser muy interesante en nuestro estudio. Por todo lo expuesto, es una buena práctica utilizar textos completos al recopilar un corpus especializado.

- IV. Número de textos.** Analizados los criterios sobre el tamaño de los corpus en cuanto al número de palabras en el corpus y el tamaño de los textos individuales y si esos textos deben ser completos o no, es importante considerar cuántos textos diferentes vamos a incluir en nuestro corpus, así como cuántos de esos textos han sido escritos por autores diferentes. Existe una gran diferencia entre un corpus de 20.000 palabras, que está formado por 20 textos diferentes, escritos por 20 autores diferentes y un corpus de 20.000 palabras formado por tan sólo dos textos largos escritos por el mismo autor. Dependiendo del tipo de estudio que queramos realizar, nuestra elección se decantará por un criterio u otro.
- V. Escritos versus orales.** Este criterio guarda relación con si cada uno de los textos fue originalmente preparado en forma escrita o si es una transcripción de un texto oral. Algunos usuarios de lenguas para fines específicos (traductores, terminólogos) trabajan frecuentemente con textos escritos, mientras que otros usuarios de los lenguajes de especialidad (expertos en un área de conocimiento) se suelen comunicar tanto de forma oral como escrita. Una vez más, nuestra decisión (recopilación de un corpus escrito, oral, o mixto) dependerá de lo que queramos estudiar.
- VI. Tema.** Los textos que vamos a incluir en nuestro corpus tratarán de temas de nuestra área de estudio. No obstante, en ocasiones nos encontramos ante el hecho de que un área de especialización es multidisciplinar (por ejemplo, la bioquímica), lo que significa que es difícil delimitar dónde termina un área y dónde empieza otra. Si nuestro estudio pretende analizar una misma cuestión en diferentes áreas,

tendremos que delimitar cada una de ellas. Si, por el contrario, examina un aspecto particular en artículos de investigación científica en general, el corpus consistirá en textos de diferentes áreas, siempre que todos reúnan la condición de ser artículos de investigación científica. La anotación de los textos nos facilitará esta tarea.

- VII. Tipo de texto.** Dentro de áreas temáticas muy especializadas se pueden encontrar distintos tipos de textos: aquellos escritos por expertos dirigidos a otros expertos (por ejemplo, artículos de investigación), textos diferentes -en cuanto a estilo y vocabulario- de aquellos escritos por expertos dirigidos a un público no experto (por ejemplo, revistas de ciencia de carácter popular, libros de texto). El tipo de textos que escojamos dependerá del estudio que llevemos a cabo.
- VIII. Autoría.** Los autores deben tener una formación académica y/o profesional adecuada en la disciplina en la que escriben y deberían ser reconocidos entre sus compañeros como que tienen el nivel de conocimientos requerido para escribir sobre la materia en cuestión.
- IX. Lengua.** Tanto si nuestro corpus es monolingüe o multilingüe, los textos deberán estar escritos en la/s lengua/s objeto de la investigación. Otro aspecto que podríamos considerar si estamos recopilando un corpus monolingüe es la inclusión de textos que están o no en su lengua original (a diferencia de los que son traducciones) y si han sido escritos por un nativo o no. En ocasiones resulta difícil determinar la lengua materna de una persona simplemente mirando un texto. En un corpus multilingüe puede resultar importante que todos los textos en la lengua A sean en la lengua original, mientras que los textos en la lengua B sean traducciones o, puede que se prefiera que tener una mezcla de textos originales y traducidos en cada una de las lenguas. De nuevo, el tipo de estudio que asumamos determinará la inclusión o no de traducciones sólo, de textos originales exclusivamente o una combinación de ambos en cada lengua.
- X. Fecha de publicación del texto.** Para una investigación de un área temática concreta probablemente necesitemos averiguar el estado de esa área en ese momento, a

nivel lingüístico y conceptual. En este caso decidiremos recopilar textos de fecha muy reciente. Textos más antiguos o de épocas diferentes pueden ser de gran utilidad para el estudio de la evolución de términos o conceptos, para aprender sobre cuestiones que nos resultan nuevas y queremos saber su desarrollo con el tiempo, etc. Las necesidades del proyecto serán la consideración primordial al diseñar el corpus.

### 3.9 Corpus y la Web

La irrupción de internet en los medios de comunicación ha provocado un cambio decisivo que ha generado que investigadores y responsables de medios de comunicación se muestren inquietos ante los cambios que afectan directamente a los nuevos modos de comunicación digital y de información. Estos cambios han afectado de manera especial a la industria periodística que, como otras, se ha visto afectada por la crisis económica. En el caso español, el informe de la *Open Society Foundations, Mapping Digital Media: Spain2* (2012: 7) resalta el hecho de que los mayores proveedores de noticias hayan adoptado los medios digitales en la producción y transmisión de noticias y las redacciones, por ello, hayan sufrido una reestructuración significativa por medio de la convergencia.

Las condiciones laborales de los periodistas digitales siguen siendo distintas de, y considerablemente inferiores a, la de los periodistas tradicionales...En estas condiciones no sorprende que el periodismo de investigación no se haya beneficiado de un renacimiento post-digitalización, o que las noticias *online* dependan cada vez más de servicios de agencia y de editoriales.

El informe añade que la crisis económica global ha provocado una caída marcada y sostenida en los ingresos por publicidad de prensa, televisión y radio desde 2008. Especialmente el sector prensa, que perdió un 28% de ingresos entre 2005 y 2010. El informe predice que las plataformas de internet móvil podrían extender el acceso a la diversidad de noticias ofrecida *online* y expandir la esfera pública, lo que podría compensar en cierto modo el déficit de pluralidad causado por la concentración y los cierres en el sector de la prensa escrita.

Los periódicos digitales emulan la función de la prensa tradicional, pero utilizan canales que les permitirá construir el uso de un conjunto de complementos impensables. Los *cibermedios* o los medios de comunicación de internet actualizados tienen voluntad de mediación entre hechos y

público y los criterios de uso y niveles de cualificación, multimedia, interactividad, hipertexto (Salaverría, 2005).

Ante el crecimiento de los contenidos publicados *online* por el estancamiento de los medios tradicionales por la crisis económica, la web puede ser una magnífica fuente para la elaboración de un corpus. Utilizar la web puede proporcionarnos un corpus colosal y podremos consultarlo directamente de internet. Eso sí, precisamos una conexión *online* para trabajar. Si escogemos esta opción hemos de tener en cuenta que existe software y herramientas disponibles en la web para construirlo, para su análisis y consulta del corpus. Vamos a desarrollar esta opción a través de cinco ideas fundamentales sobre la utilización de la web como un corpus:

- World Wide Web como un corpus colosal (un corpus multimodal: texto, imágenes, sonido y vídeo)
- Directamente consultar la web
- Acceso online (con facilidades de búsqueda)
- Web como fuente para un corpus de creación (géneros emergentes)
- Software y herramientas que son de libre acceso con el fin de construir y consultar el propio corpus

#### **I. World Wide Web como un corpus colosal: texto, imágenes, sonido y vídeo**

Comenzamos con las ventajas de la obtención de datos de la web para la construcción de un corpus. Los textos que vayamos recuperando para su inclusión en el corpus ya son legibles a través de ordenador, no los tenemos que transcribir de manera manual. Este método automatizado de recuperación de texto nos permite que el corpus pueda ser construido en un tiempo más breve. También para el investigador la web es una fuente adecuada para recuperar una gran cantidad de datos en un tiempo razonable y con un esfuerzo razonable.

La web nos proporciona más y mejores datos. El gran tamaño de la web, como fuente de datos lingüísticos, es una de las grandes ventajas para la construcción de un corpus. La información que recopilamos de la web está constantemente renovada. Si utilizamos la web para la elaboración de un corpus podemos además analizar géneros textuales emergentes. Es el caso de los blogs, wikis o foros de discusión. En el caso de los blogs, contienen grandes cantidades de

texto, se refieren a una amplia gama de temas y desde el punto de vista de un lingüista de corpus, se producen de forma espontánea por los usuarios web, cuyas características demográficas, como pueden ser la edad o la profesión, pueden variar los contenidos en gran medida.

No sólo los blogs pueden aportar variedad de temas. La web es el único recurso disponible para el estudio de los sub-dominios lingüísticos especializados de vocabulario científico y técnico, como el campo de la extracción de terminología. De este modo, la web puede ser usada también para construir corpus especializados o para construir corpus paralelos formados por gran número de bitextos<sup>52</sup>.

La mayor parte de la web está escrita inglés, esto podría ser una limitación de su uso para otros idiomas. También en su mayoría la web está en formato escrito, no oral, lo que limita también la creación de corpus orales.

Los problemas de la web de ofrecer datos sin control son aparentemente compensados por las ventajas de su enorme tamaño y la fácil disponibilidad de texto que permite a (casi) cualquier persona llevar a cabo su propia investigación lingüística de una manera más sencilla.

Díaz Noci y Salaverría (2003) establecen que cinco son las características que configuran el discurso cibernético:

- Elementos multimedia (la escritura se mezcla con la imagen del sonido en movimiento)
- El hipertexto (navegación de nodos)
- La instantaneidad (el presente adquiere una nueva dimensión debido a que puede ser sustitutivo acumulativo y no sucesivo y secuencial como en los otros medios de comunicación)
- La interactividad (el usuario se convierte en un elemento activo del proceso comunicativo)
- La universalidad (en internet todo está en una décima de segundo, ya que la información circula a la velocidad de la luz)

---

<sup>52</sup> <http://www.statmt.org/euoparl/>[11/04/2016]

## II. Consultas directas a la web

Este enfoque considera la web como un gran corpus. Los sistemas que implementan este enfoque, por lo general, ofrecen una interfaz para la consulta y la búsqueda de concordancias en la web. Básicamente, añaden una capa a los motores de búsqueda tradicionales como *Google* o *Microsoft Live Search* y ofrecen diversas opciones de visualización que son útiles para el trabajo lingüístico (concordancias, extracción de colocaciones, listas de palabras, etcétera).

Existen distintos ejemplos:

- *El WebCorp*: <http://www.webcorp.org.uk> (Renouf, 2003)
- *Key Word in Context Concordances from the Web*:  
<http://kwicfinder.com/KWiCFinder.html> (Fletcher, 2007)
- *CorpEus*: <http://corpeus.elhuyar.eus/cgi-bin/kontsulta.py> (Leturia et al., 2007)

## III. Corpus a los que se puede acceder online (con facilidades de búsqueda)

- *BNC*: <http://www.natcorp.ox.ac.uk/>
- <http://corpus.byu.edu/bnc/x.asp>
- *Collins*: <http://wordbanks.harpercollins.co.uk/>
- *Corpus de Inglés Americano*: <http://www.americancorpus.org/>

Otros corpus:

- <http://corpus.byu.edu/> (<http://www.corpusdelespanol.org>).
- <http://corpus.rae.es/creanet.html>
- <http://corpus.rae.es/cordenet.html>

## IV. La web como fuente para la creación de corpus (géneros emergentes)

Otro enfoque consiste en utilizar la web como una fuente para la extracción de textos que serán recogidos (con el uso de rastreadores), filtrados y registrados en un corpus no *online*. El proyecto *Wacky* es un ejemplo bien conocido de esta opción (Baroni y Bernardini, 2009).

- <http://wacky.sslmit.unibo.it/>

Otro ejemplo similar es: corpus *USENET* (2005-2007)

- <http://www.psych.ualberta.ca/~westburylab/downloads/usenetcorpus.download.html>

#### V. Software y herramientas que son de libre acceso para construir y consultar corpus

- *Sketch Engine* (Adam Kilgarriff): <http://www.sketchengine.co.uk/>
- *Developing Linguistic Corpora: a Guide to Good Practice* (Martin Wynne): <http://ota.ox.ac.uk/documents/creating/dlc/index.htm>
- *Tools, Ideas & Resources for Linguistics* (Martin Weisser): <http://martinweisser.org/index.html>

#### 3.10 Ruido en el corpus y las redes sociales

Los textos extraídos de las redes sociales pueden constituir una interesante base de elaboración de un corpus lingüístico por su variedad de discurso y lo heterogéneo de sus registros y estilos. La recopilación de datos puede llevar a explorar, a través del análisis cualitativo comparativo, las semejanzas y diferencias de la ortografía o el léxico en las distintas redes sociales. La evolución de la manera de entender los corpus y, en concreto, el avance experimentado en el caso de los textos digitales publicados en la web así como su explotación obliga a tener en cuenta análisis desde conceptos y dimensiones diversas. Es, por ejemplo, el caso de la publicación de mensajes en las redes sociales que han cobrado relevancia en el inicio del siglo XXI tales como Facebook o Twitter<sup>53</sup>. En ellas pueden comprobarse entornos semióticos complejos con fotografías, videos, etiquetas o enlaces relacionados a enunciados que determinan el sentido global de los mensajes publicados. Esta variedad ha dado pie a diversas investigaciones de registros y géneros para establecer comparaciones entre variedades distintas de una misma lengua, cuestión que es más evidente cuando se utilizan páginas de la web para constituir un corpus de estilo lingüístico y discursivo (Rundell, 2000; Resnik y Smith, 2003; Fletcher, 2004). Como señalan Kilgarriff y Grefenstette (2003: 33):

---

<sup>53</sup> *La 6ª Oleada del Observatorio de redes sociales* (octubre 2014) analiza las redes más utilizadas y con el mayor número de usuarios en España. Facebook es líder, con un 83% de los entrevistados que tiene una cuenta personal y la utiliza, frente a Twitter (42%). La media de edad de los usuarios es de 31 años en Facebook y de 28 en Twitter.

Language scientists and technologists are increasingly turning to the Web as a source of language data, because it is so big, because it is the only available source for the type of language in which they are interested, or simply because it is free and instantly available.

Los datos extraídos de la web son muy heterogéneos, lo cual dificulta la obtención de información homogénea, estructurada y libre de ruido (Martínez Santiago et al., 2001). A la hora de construir un corpus, cuando consideramos aspectos fundamentales como la extensión del mismo, la representatividad, o la ética en la recolección de datos, Claridge (2007) y Beibwenger y Storrer (2007) advierten que necesario distinguir entre los datos que genera el sistema como las líneas de entrada y salida de un canal de chat, o la fecha de envío de un tuit y los que genera el usuario, que pueden incluir texto, enlaces, emoticonos e imágenes.

### **3.11 Copyright, ética y legalidad**

En una visión simple del mundo podemos decir que la ley establece lo que está permitido y lo que no. Sin embargo, en el caso de la creación de un corpus y sus aplicaciones en el ámbito de la lingüística de corpus, la cuestión ya no es tanto el estar o actuar dentro de la legalidad, sino cómo gestionamos los riesgos de un eventual incumplimiento. En otras palabras, si no existe una oportunidad de negocio, el riesgo de ser penalizado por cometer una posible ilegalidad es menor, ya que nadie (o casi nadie) gastaría su dinero en interponer denuncias ante los tribunales cuando no ve la oportunidad de ganar. Pero, si el sujeto en cuestión es el responsable de una gran organización el hecho acudir a la justicia puede tener como compensación la obtención de grandes sumas de dinero, en caso de que prospere y se gane la demanda. Aquello que cada uno esté dispuesto a denunciar ante los tribunales, como persona física o jurídica, va a depender de su motivación por el interés final y por la cantidad de capital en juego. Es muy diferente si se trata de una organización, o una universidad, que cuentan con mayor respaldo económico para financiar una defensa jurídica más fuerte y solvente. Por lo tanto, aunque no valga la pena demandar porque el autor de la denuncia sea una persona sin importancia, hay que tener en cuenta que la demanda puede estar dirigida a la universidad que respalda una determinada investigación por usar sus servidores y publicar un corpus donde hay infracciones de derechos de autor.

Técnicamente hablando, las listas de palabras y perfiles de frecuencia parecen estar seguros en el ámbito del uso justo ([http://en.wikipedia.org/wiki/Fair\\_use](http://en.wikipedia.org/wiki/Fair_use)). Los datos de *N-grams* de Google

*Books*, distribuidos hasta cadenas de 12 palabras por Google, son ejemplos de una institución que distribuye cadenas de palabras de una extensión considerable con bastante asiduidad. Esto puede provocar que los representantes legales de las universidades se sientan incómodos si ven la posibilidad de que un determinado autor pueda llevarlos ante los tribunales, aunque no exista garantía de evadir en primera instancia el tribunal y pueda esto tener consecuencias económicas. Si sólo está distribuyendo listas de frecuencias, no hay rastro de una obra protegida por derechos de autor registrada. Incluso en el nivel de 5 gramas, es muy difícil encontrar cualquier infracción accionable.

Hay ejemplos como los listados de frecuencias de palabras de *COCA/COHA* (<http://www.wordfrequency.info>) y datos de (<http://www.ngrams.info>). Mark Davies, profesor de Lingüística de la Universidad Brigham Young, ha estado distribuyendo datos de “texto completo” del corpus *COCA* de 450 millones de palabras, y del corpus *GloWbE*, de 1,9 billones de palabras (<http://corpus.byu.edu/glowbe>) por un tiempo, y no ha tenido problemas. Sin embargo, hay un cambio de dirección en términos de cómo los datos de texto completo han sido ligeramente alterados para evitar problemas de copyright (<http://corpus.byu.edu/full-text/limitations.asp>).

Por lo tanto, parece que las listas de frecuencias de palabras y listas de *N-grams* no son un problema. El gran problema que los lingüistas de corpus están encontrando es que actualmente no hay ninguna guía o base legal establecida sobre si los corpus pueden ser compartidos. Como resultado, casi todos los corpus siguen siendo inaccesibles, lo que significa que cada investigador tiene que recoger su propio corpus. Esto supone un enorme trabajo que significa que poco trabajo acumulativo se puede hacer. Este es el verdadero problema de los derechos de autor. A todo hecho cabe añadir el hecho de que muchos corpus no son lo suficientemente profesionales como para ser distribuidos.

### **3.12 Lingüística de Corpus y la Sociedad del Conocimiento**

La lingüística basada en el análisis del corpus ha evolucionado de la misma manera que lo ha hecho la tecnología. Los avances informáticos han aportado precisión y fiabilidad al procesamiento de los datos (Vargas, 2012). La sociedad del conocimiento es un concepto que engloba las transformaciones sociales que se están produciendo en la sociedad moderna. El

término ofrece una visión del futuro, pero ha tenido una adaptación desigual en las diferentes áreas lingüísticas al concurrir con términos como sociedad de la información. En la década de los 90 del siglo XX el concepto sociedad del conocimiento resurgió en el ámbito de la política y las ciencias sociales (Reich, 1992; Drucker, 1994; Lundvall y Johnson, 1994; Stehr, 1994; OECD, 1996; Krohn, 1997; Knorr-Cetina, 1998; Willke, 1998 y Weingart, 2001). Sin embargo, la sociedad de la información no guarda relación con la base tecnológica para caracterizar la sociedad actual y para resaltar las diferencias con la sociedad industrial. Heidenreich (2003) distingue cuatro definiciones:

- Al igual que la sociedad de la información, la noción sociedad de conocimiento indica la importancia de las tecnologías de la información y la comunicación (TIC's) y su utilización en los procesos económicos.
- Resalta las nuevas formas de producir conocimiento, considerado como uno de los principales causantes del crecimiento junto con los factores capital y trabajo. En este sentido, se concede una relevancia crucial a la producción de productos intensivos en conocimiento y a los servicios basados en el conocimiento.
- Destaca la importancia de los procesos educativos y formativos, tanto en su vertiente de educación y formación inicial como a lo largo de la vida.
- Destaca la creciente importancia de los servicios intensivos en conocimiento y comunicación, que generalmente se denominan trabajo de conocimiento.

La sociedad del conocimiento hace referencia a cambios en las áreas tecnológicas y económicas estrechamente relacionadas con las TICs, en el ámbito de planificación de la educación y formación, en el ámbito de la organización (gestión de conocimiento) y del trabajo (trabajo de conocimiento). Teniendo en cuenta cómo se han ido adaptando los cambios tecnológicos a las distintas áreas, vamos a comprobar cómo la sociedad del conocimiento también ha podido marcar una evolución en la lingüística de corpus.

Ya hemos abordado la evolución de la lingüística de corpus desde sus orígenes hasta la actualidad. Los cambios se han notado mucho en apartados tales como la obtención de textos y su análisis mediante métodos electrónicos y las TIC's, que han permitido el análisis del lenguaje hablado y escrito de una manera más simple. La evolución de la lingüística de corpus está

directamente relacionada con la disponibilidad de programas informáticos para el análisis y explotación del corpus.

Veamos a continuación algunas de las tareas a realizar para la elaboración de un corpus:

- Escaneo de textos originales
- Recopilación de textos electrónicos de un corpus existente
- Recopilación de datos de la *World Wide Web*. Por ejemplo, utilizando Google, a través de los motores de búsqueda, bibliotecas electrónicas o revistas electrónicas
- Recopilación de periódicos o libros en formato electrónico
- Transcripciones de las grabaciones de un discurso hablado o conversaciones

Algunos de los métodos:

- Grabación y digitalización
- Transcripción ortográfica
- Transcripción fonética
- Anotación prosódica
- Fenómenos paralingüísticos

A continuación, vemos algunas de las tareas a realizar en la elaboración de un corpus (tanto de géneros orales como escritos):

- Lematización
- POS tagging
- Anotación sintáctica
- Anotación semántica
- Anotación discurso
- Anotación pragmática
- Representación de las divisiones textuales

Vargas (2012) establece que las herramientas informáticas empleadas en el tratamiento y análisis de un corpus se dividen en etiquetadores morfosintácticos y programas de análisis

textual o programas de concordancias. El etiquetado más habitual es el morfosintáctico (*part of speech tagging*). El etiquetado de un corpus a este nivel suele comprender cuatro fases:

- Segmentación. Recibe el nombre de tokenización y consiste en la segmentación del texto en cadenas de caracteres y cifras que se encuentran entre espacios
- Lematización o especificación de la forma no marcada de cada palabra
- Análisis morfológico y de las categorías gramaticales
- Etiquetado morfosintáctico o desambiguación de las categorías gramaticales dudosas

En lo referente a los programas de concordancias, Vargas (2012) los define como un conjunto de aplicaciones informáticas capaces de analizar una base de datos textual. La herramienta *WordSmith Tools* es capaz de generar tipos de listas de los resultados tales como:

- Listas con estadísticas
- Listado monoléxico
- Listado poliléxico de las palabras del corpus o de una selección. Estos listados pueden ofrecerse a través de agrupamientos técnicos o clusters o con el cálculo de información mutua o cualquier otra medida estadística (*Z Score*, *MI3*, *Log Likelihood*, etcétera)
- Líneas de concordancia o listados de aparición de una palabra en concreto que puede denominarse de búsqueda, base clave, y que puede estar formada por una unidad, varias o parte de ésta y que está acompañada del texto que la rodea o co-texto. El más común es el la palabra clave en contexto, *KWICK* o *Key Word In Context*. Este tipo de lista agrupa las apariciones de la palabra de búsqueda y aparece destacada en el centro para analizar y detecta con rapidez sus colocaciones o palabras de su entorno.
- Listado de palabras clave (*keywords*). Con ella se obtiene una lista de palabras cuyas referencias son estadísticamente diferentes en el corpus de estudio respecto al de referencia.
- Gráfico de distribución de la palabra de búsqueda o *plot*. Permite apreciar de manera visual la posición en la que figura una palabra en los textos en los que aparece. El resultado es similar a un código de barras y el investigador lo interpreta por la relevancia de la posición de mayor o menor frecuencia de aparición de una parte de texto determinada.

- Listado de *collocates* o de palabras que aparecen alrededor de la palabra base, en posiciones determinadas.

### **3.13 Organigrama con ejemplos de (posibles) corpus de textos orales y escritos**

Las fuentes para la elaboración de un corpus pueden ser muy diversas. La acumulación de ítems puede ser procedente de conversaciones convencionales a escritos académicos y oficiales o prensa escrita como es el caso de esta investigación para la que se ha creado el CCF. Los siguientes organigramas tratan de mostrar la variedad de posibilidades que se presentan para construir un corpus y su posterior análisis.

Por ejemplo, si la decisión es el análisis de textos escritos nos podemos decantar por dos ámbitos diferentes como son la investigación o el ámbito educativo. Los medios de comunicación también presentan posibilidades de análisis de textos y cuentan con la novedosa incorporación de las redes sociales y su característica manera de ofrecer información. La cultura, el ocio, los documentos de la esfera privada y la pública también presentan posibilidades de análisis de textos. Veamos en el siguiente organigrama algunos ejemplos sobre las posibilidades de creación de corpus de textos escritos.

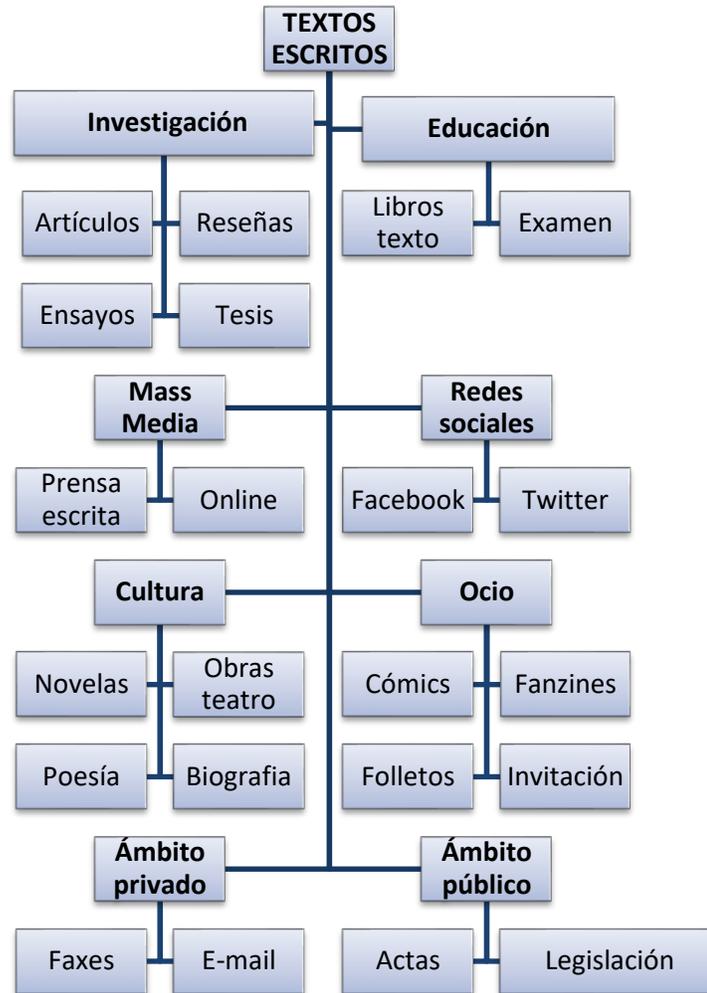


Figura 3.2: Organigrama de textos escritos

También los textos orales presentan un amplio abanico de posibilidades de análisis lingüístico, tanto en el ámbito de lo privado como lo público. Determinadas actividades de ocio o de la cultura, la educación escolar o las desarrolladas en medios de comunicación pueden ser también una fuente de recopilación para su posterior análisis lingüístico.

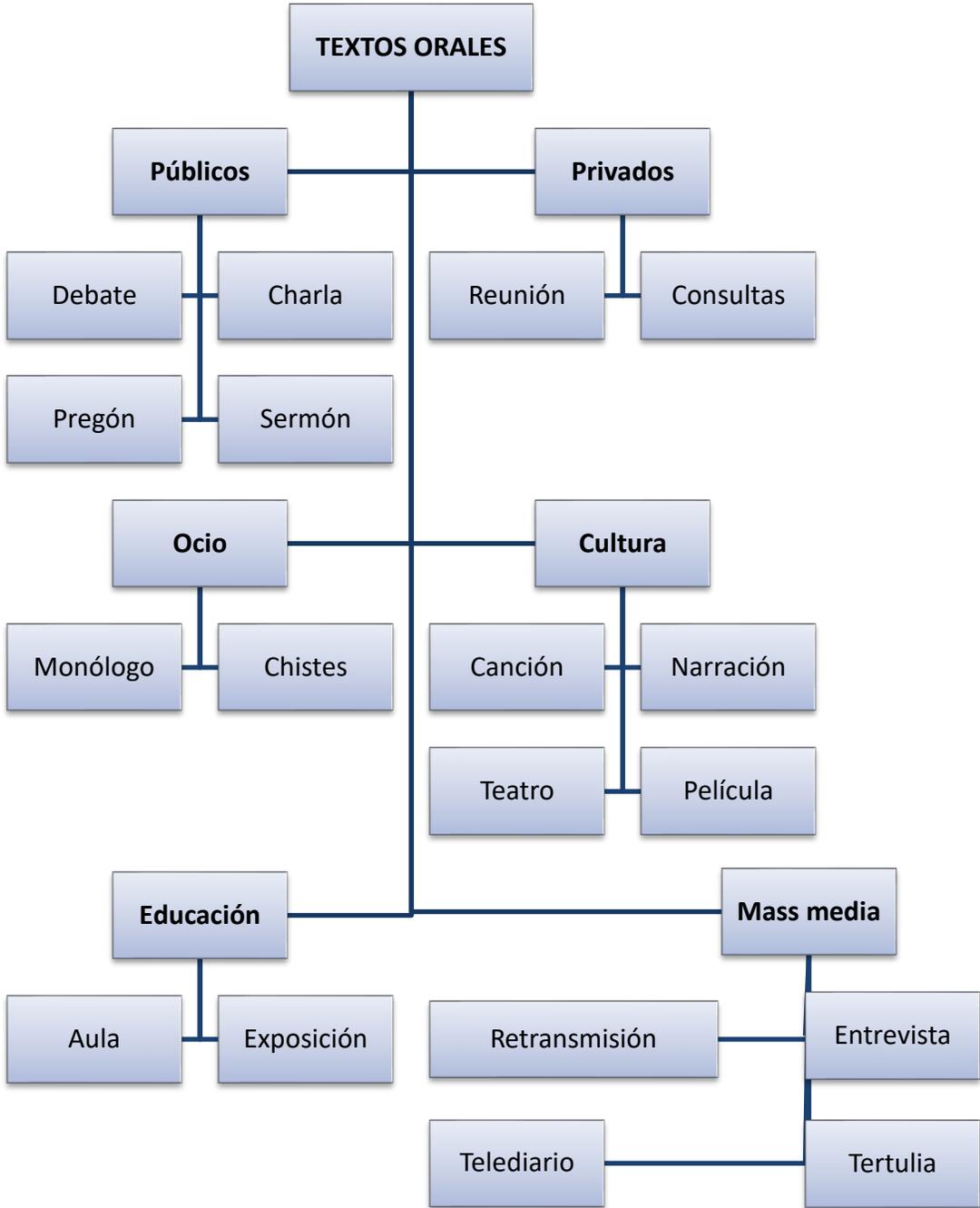


Figura 3.3: Organigrama de textos orales

### 3.14 Conclusión

En este capítulo, hemos discutido algunas de las premisas fundamentales de la lingüística de corpus. Entre estas ideas centrales, podemos mencionar las siguientes:

- las diferencias formales de distribución reflejan diferencias funcionales. En otras palabras, una diferencia en forma sintáctica produce una diferencia en el significado.
- los corpus deben diseñarse teniendo en cuenta los siguientes principios: representatividad, equilibrio, marcado (¿mantenemos información sobre la apariencia física de los documentos originales?) y anotación (¿vamos a añadir etiquetas morfosintácticas o semánticas al corpus?)
- existen algunos métodos básicos de la lingüística de corpus (listas de frecuencias, palabras clave, concordancias, colocaciones)

Podemos ir más allá del concepto de frecuencia diciendo que inicialmente todo lo que se puede obtener de los corpus, que son archivos (generalmente de texto), es información distributiva (o cuantitativa y estadística). Esta información estadística puede ser la siguiente:

- frecuencias de aparición de elementos lingüísticos, por ejemplo, con qué frecuencia aparecen morfemas, palabras, patrones gramaticales en un corpus (tal y como aparecen en las listas de frecuencias)
- frecuencias de co-ocurrencia de estos elementos, por ejemplo, con qué frecuencia aparecen morfemas con palabras particulares, con qué frecuencia aparecen palabras particulares en una determinada estructura gramatical (como figuran en las líneas de concordancias)

La noción de frecuencia no presupone que un fenómeno lingüístico relevante se produzca en un corpus 100, 1.000 o 1.000.000 de veces. La noción de frecuencia también incluye fenómenos que ocurren sólo una vez (*hapax legomena*). De hecho, la lingüística de corpus ha demostrado que un pequeño número de elementos lingüísticos se repiten con frecuencia mientras que hay un número mucho mayor de ítems que ocurren muy infrecuentemente. Por lo tanto, la lingüística del corpus se ocupa de descubrir si:

- aparece en el corpus un elemento lingüístico individual o la co-ocurrencia de más de un elemento lingüístico (si la frecuencia observada de ocurrencia o co-ocurrencia es 0 o mayor)
- un elemento lingüístico aparece en el corpus con más frecuencia que otro elemento lingüístico (si una frecuencia observada es mayor que la frecuencia de otro elemento lingüístico)
- se detecta un elemento lingüístico más o menos frecuentemente de lo que se esperaría por casualidad

Por lo tanto, se asume que las diferentes frecuencias de (co) ocurrencias de elementos formales (morfemas, palabras, patrones gramaticales) reflejan regularidades funcionales, entendiendo "funcional" en un sentido muy amplio, por ejemplo como semántico o discursivo-pragmático, destinado a realizar una función comunicativa particular. En otras palabras, los patrones gramaticales de la evaluación (como veremos en el siguiente capítulo) realizan funciones comunicativas típicamente expresadas en editoriales de periódicos.

En este capítulo hemos abordado los criterios para construir un corpus, tales como la representatividad y el equilibrio. Hay un último criterio para que un conjunto de textos sea un corpus. Los textos que componen el corpus deben haber sido producidos en un entorno comunicativo natural. Esto significa que los textos fueron emitidos de forma oral o escrita para algún propósito comunicativo auténtico, pero no con el fin de ponerlos en un corpus. Por ejemplo, nuestro corpus consiste en artículos periodísticos. Nuestro corpus cumple el criterio de haber sido producido en un entorno natural porque los periodistas escriben el artículo para ser publicado en periódicos, revistas y blogs para comunicar noticias y opiniones a sus lectores, pero no porque quieran llenar el corpus de un lingüista.

Un corpus de artículos periodísticos sobre la economía española es un corpus especializado y monolingüe. Los corpus pueden clasificarse de varias maneras:

- corpus general y corpus especializado
- corpus diacrónico y corpus sincrónico
- corpus monolingüe y corpus paralelo

- corpus estático (tamaño fijo) y corpus dinámico/monitor (en constante cambio, ampliado con nuevos textos)

Volviendo a la cuestión de la frecuencia, podemos vincular la frecuencia con la probabilidad a través de la descripción de Halliday del lenguaje como un sistema semiótico que consiste en una serie de redes entrelazadas de opciones (Halliday, 1978). Cada rasgo lingüístico (elemento lingüístico) tiene una probabilidad de ser elegido. Nosotros, como usuarios del lenguaje, optamos por utilizar un elemento lingüístico con mayor o menor frecuencia que decide la probabilidad de esa selección del elemento lingüístico. En la teoría sistémica de Halliday, las redes representan series de elecciones que pueden hacerse en el lenguaje (por ejemplo, entre los modos interrogativo, imperativo y declarativo, o entre tiempos presentes y pasados). A cada opción se le puede asignar una probabilidad de ocurrencia basada en su frecuencia real de aparición en un corpus grande. Por ejemplo, Halliday y James (1993) calculan que en un corpus general grande, los tiempos verbales en presente y en pasado aparecen aproximadamente con igual frecuencia, mientras que las oraciones positivas ocurren extraordinariamente con más frecuencia que las negativas, en una proporción de aproximadamente 9: 1. Otro argumento es que los textos y cada registro diferente tienen un conjunto de frecuencias y, por tanto, de probabilidades, que son diferentes de la probabilidad global general. Más exactamente, la probabilidad global es una amalgama de las probabilidades de todos los registros que constituyen el conjunto. Además, las probabilidades de ocurrencia en cada registro son en sí mismas el resultado de la frecuencia de una característica dada en cada uno de los diferentes textos.

Para resumir, esencialmente, la lingüística del corpus implica:

- examinar el lenguaje natural;
- examinar cantidades relativamente grandes de dicho lenguaje;
- observar frecuencias relativas, ya sea en forma de datos brutos o procesados a través de operaciones estadísticas;
- observar patrones de asociación, bien entre un rasgo y un tipo de texto o entre grupos de palabras.

En el capítulo 4, en el análisis de nuestro corpus (CCF), trataremos de observar patrones de asociación haciendo especial hincapié en el lenguaje evaluativo.

La lingüística de corpus es una metodología de estudio de textos recopilados de muestras reales de una lengua. Un corpus lingüístico puede proporcionar una lectura no lineal, por ejemplo cuando se utiliza líneas de concordancia para obtener datos sobre el uso de la lengua en múltiples textos de diferentes géneros. Un mismo grupo de textos recopilados en un corpus puede aportar datos diferentes según el método de análisis que utilicemos, por ejemplo, listados de palabras, índices de frecuencias, palabras clave, colocaciones, líneas de concordancia, *clusters*, entre otras.

Hemos repasado las aportaciones a la lingüística de corpus de autores como Stubbs (1996), Biber et al. (1998), Kennedy (1998), McEnery y Wilson (2001), Hunston (2002) y Sinclair (2005) que muestran las diferentes posibilidades del análisis no lineal de los textos compilados. De esta manera, podemos alcanzar los diferentes matices de una lengua a través del seguimiento de unos determinados patrones. En el aspecto científico y metodológico, esta parte de la investigación desgrana las características de los principales corpus, su aportación, las herramientas para analizarlos o la influencia de las nuevas tecnologías. Todo ello sin perder de vista la finalidad de esta investigación: el análisis lingüístico de textos periodísticos.

## CAPÍTULO 4: METODOLOGÍA Y RESULTADOS

### 4.1. Corpus de la Crisis Financiera (CCF): Justificación y creación

Este capítulo describe la metodología a través de la cual se ha diseñado e implementado esta investigación cuya finalidad es evaluar un corpus de textos periodísticos. El objetivo del presente capítulo es delinear las etapas en el proceso de análisis de nuestro corpus a través de una selección de ejemplos que ponen de relieve las construcciones lingüísticas utilizadas por los periodistas y autores de los textos para entablar, mantener e intensificar sus relaciones con los lectores. Veremos cómo la combinación de procedimientos cuantitativos y cualitativos, junto con técnicas de lingüística de corpus, proporcionan la base para la descripción de distintas categorías de evaluación en el discurso de los diarios y los resultados que muestran.

Una vez quedó definido el objetivo de la investigación, el siguiente paso fue la creación del Corpus de la Crisis Financiera (CCF) para posteriormente analizarlo. La fecha de arranque del trabajo, enero de 2012, era un momento de cambio por la coyuntura de inestabilidad que se presentaba por el avance de la crisis financiera cuyo origen se fijó en Estados Unidos con la quiebra en 2008 de Lehman Brothers<sup>54</sup> y otras compañías hipotecarias, como ya vimos en el capítulo 1. A este acontecimiento y posteriores consecuencias en la economía mundial se sumó el hecho de que no tuviésemos constancia de la existencia de algún otro corpus con estas mismas características sobre publicaciones periódicas que abordasen la crisis financiera en España. Al inicio de esta investigación pretendíamos alcanzar el cumplimiento los requisitos que Eco (2001) establece como esenciales en una tesis doctoral. Los vimos en el Capítulo 1: una investigación ha de ser sobre un objeto reconocible y estar definida, sobre cosas que todavía no han sido dichas, que sea útil a los demás y que suministre elementos para la verificación y la

---

<sup>54</sup> Lehman Brothers Holdings Inc, fundada en 1850, fue una compañía global de servicios financieros de Estados Unidos. Destacaba en banca de inversión, gestión de activos financieros e inversiones en renta fija, banca comercial, gestión de inversiones y servicios bancarios en general. Tenía su sede social en la ciudad de Nueva York, con sedes regionales en Londres y Tokio, y oficinas por todo el mundo. El 15 de septiembre de 2008 la compañía anunció la presentación de quiebra gravemente afectada por la crisis financiera provocada por los créditos *subprime*. Acumuló enormes pérdidas por títulos respaldados por las hipotecas a lo largo de 2008. En el segundo trimestre fiscal declaró pérdidas por 2.800 millones de dólares y la empresa se vio obligada a vender 6000 millones de dólares en activos. En el primer semestre de 2008 había perdido el 73% de su valor en bolsa.

refutación de las hipótesis que presente. El objetivo marcado era llegar a demostrar cómo la crisis financiera ha podido haber influido en el lenguaje de los medios de comunicación. No teníamos constancia de una investigación similar, por lo que sería algo novedoso. El tema de debate de las primeras sesiones de este trabajo fue fijar el objeto a investigar y en qué podría derivar. Decidimos construir un corpus sobre publicaciones de economía en que se pudieran extraer conclusiones sobre la aportación semántica de la crisis financiera a partir del sentimiento expresado por los autores en los textos. Fijados los objetivos de arranque, las consideraciones preliminares a aclarar fueron las siguientes:

- Marcar los criterios de selección de los textos para el CCF
- Establecer el tamaño del corpus a analizar para que fuese representativo

El primer objetivo que nos marcamos era que el corpus que íbamos a construir fuese representativo del lenguaje objeto de estudio y que el posterior análisis de textos de ese corpus resultase fructífero y, por tanto, concluyente. En el debate de las fuentes de las que íbamos a obtener los textos del CCF entraron a debate el análisis de distintos medios de comunicación. Si algo teníamos claro era que la materia a analizar tenía que ser periodística. En un primer momento, con un objetivo mucho más ambicioso, en el que incluso planteamos la posibilidad de analizar junto a la prensa española la inglesa<sup>55</sup>, -para establecer una comparativa<sup>56</sup>-, acotamos la investigación a la prensa nacional. Una vez localizadas las fuentes, que nos proporcionasen la información necesaria, el paso siguiente fue la organización y procesamiento del material extraído y la selección de herramientas informáticas para el análisis. Todo ello siguiendo la metodología de la lingüística de corpus: listas de frecuencias, palabras clave, líneas de concordancia, colocaciones, patrones recurrentes y *clusters*.

Ya lo hemos dicho en la introducción y también al inicio de este capítulo. La génesis de la elaboración del CCF coincidió con un momento de efervescencia de publicaciones sobre información financiera a causa del avance de la crisis económica. Se presentaba una gran oportunidad para la creación de un corpus lo suficientemente potente que permitiese analizar

---

<sup>55</sup> *The Guardian*, como cabecera representativa del sector liberal y *The Daily Telegraph* del centro-derecha conservador.

<sup>56</sup> Analizar cómo la prensa anglosajona plasmaba acontecimientos de la crisis financiera y establecer conclusiones.

los rasgos principales de un lenguaje, en nuestro caso el lenguaje de la crisis financiera. Como ya vimos en el capítulo 1, el interés por la economía saltó de las páginas salmón a otras secciones. Valga el ejemplo que ya aportamos sobre el interés que ha despertado en la sociedad ya que fue el eje de debate del congreso celebrado en San Millán de la Cogolla en mayo de 2013 y en el que la Reina Letizia ya resaltó el surgimiento de un lenguaje propio de la crisis con vocablos específicos que se habían incorporado al lenguaje común. La investigación se ha desarrollado con la idea de que el periodista, en nuestro caso el de la prensa escrita, utiliza una serie de recursos léxico-gramaticales ya sea para expresar sus propios sentimientos o poner palabras en la boca de los principales actores sociales, es decir, los principales actores en la crisis financiera, que en este capítulo conoceremos. Vamos a descubrir cómo los periodistas llevan a cabo una construcción lingüística de la realidad social en sus artículos a través de los textos recopilados de prensa escrita en 2012.

Para la elaboración de nuestro CCF contamos con la disponibilidad de los textos periodístico en formato electrónico. Esta posibilidad ha hecho más sencilla y efectiva la construcción de un corpus especializado como éste (Barlow, 1996). En el capítulo dedicado al periodismo, y en concreto en el punto 2.4, vimos que el desarrollo científico y tecnológico han multiplicado las posibilidades de comunicación y también que internet ha ampliado el flujo de información entre las personas y ha hecho crecer la influencia de los medios de difusión más allá de la tradición de la original prensa escrita.

El corpus, y más específicamente el CCF, es, como bien asegura Francis (1992: 167), el resultado de un proceso empírico de recopilación de datos lingüísticos en combinación con la capacidad del autor de elicitación<sup>57</sup> de conocimiento previo y de introspección. Leech (1991: 74) sostiene que en los estudios lingüísticos es importante destacar la combinación del corpus y la intuición: *corpus plus intuition*. Precisamente, el análisis del lenguaje a partir de la intuición o introspección de la experiencia y conocimiento del investigador es algo a considerar como establece Wichmann et al. (1997: 3). Para estos autores es una rica fuente de datos auténticos

---

<sup>57</sup> Es una adaptación del verbo inglés *to elicit*. Aparece en textos de psicología con el sentido de provocar, suscitar u obtener, según el *Diccionario de la RAE*. En el ámbito de la lingüística se asocia a la obtención de manera provocada de información de una fuente.

que contiene estructuras, patrones y características predecibles que están esperando a ser desbloqueados por la inteligencia humana.

Con las aportaciones de los expertos en la materia desplegadas sobre la mesa, vamos a tratar de aplicarlas a nuestro trabajo, que es el análisis lingüístico de la crisis financiera. Es evidente el impacto social que la crisis financiera ha tenido y ha transmitido a las vidas de los ciudadanos. Insistimos en la idea de que engloba un momento histórico, de gran trascendencia y de interés en la sociedad española y que ha generado un aluvión de informaciones que hemos aprovechado para construir el CCF y tratar de demostrar cómo los periodistas realizan a través de sus escritos esa construcción lingüística de la realidad social.

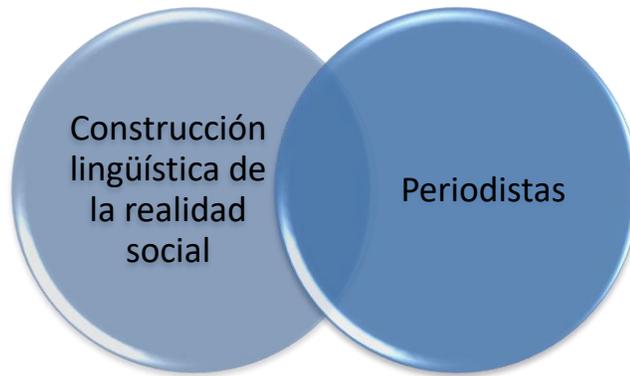


Figura 4.1: Diagrama de confluencia realidad social-relato periodístico

El estudio lingüístico de la crisis financiera, a partir de publicaciones periodísticas de España, respondió al importante impacto social por la situación económica en el ámbito internacional y la amplia demanda de información económica en que había derivado.

Pero, ¿qué es lo que convierte un tema determinado, -como puede ser la información financiera-, en algo noticioso? ¿Por qué la información salmón ha llegado a generar tanto interés? Recordamos, una vez más, lo que al respecto defiende uno de los exponentes de la denominada escuela norteamericana de periodismo. Warren (1979) propone los siguientes

ingredientes como indispensables para que un hecho acabe por incluirse en la agenda de los medios<sup>58</sup> de comunicación:

- Actualidad
- Proximidad
- Consecuencias
- Relevancia personal
- Suspense
- Rareza
- Conflicto
- Sexo
- Emoción
- Progreso

Como ya vimos en el capítulo dedicado a los géneros periodísticos (punto 2.2) estas diez características son readaptadas por el periodista Herraiz que añade el factor subjetivo en función de los lectores particulares de cada publicación y que varía a la hora de comparar el tratamiento concreto que una u otra publicación da a los acontecimientos de un mismo día. Los conceptos y las clasificaciones varían con el tiempo.

A la teoría expuesta por Warren, añadimos los nuevos valores de la noticia fijados por Martini (2000) referido a los criterios de noticiabilidad como los alineamientos que hacen posible que, a través de un proceso, el acontecimiento se convierta en noticia. Dice Martini que según los efectos que un acontecimiento puede tener sobre la sociedad, los valores-noticia más importantes son:

---

<sup>58</sup> La teoría del establecimiento periodístico de temas de discusión o *agenda-setting* postula que los medios de comunicación de masas tienen una gran influencia sobre el público al determinar qué historias poseen interés informativo y cuánto espacio e importancia se les da. El punto central de esta teoría es su capacidad para graduar la importancia de la información que se va a difundir, su orden de prioridad para obtener mayor audiencia, mayor impacto y una determinada conciencia sobre la noticia. Del mismo modo, deciden qué temas excluir de la agenda. Esta teoría, estudia el impacto de la prensa (impresa, electrónica) y la información que ésta maneja en el público, además del análisis de audiencias.

- Novedad. Implica la existencia de un hecho como ruptura de la cotidianeidad
- Originalidad, imprevisibilidad e ineditismo. Refuerzan el concepto de novedad. Un hecho original e inédito es más noticia porque es más novedad. Los hechos impredecibles circulan con mucha mayor rapidez entre los mismos medios y la sociedad, haciendo que ellos no desaparezcan fácilmente de la sociedad
- La evolución futura de los acontecimientos genera mucha expectativa por parte del público. Esto propicia un mayor consumo de información que detalle el desenlace final. Aquí enmarcamos las noticias que tienen que ver con secuestros, juicios públicos, catástrofes y en general, cualquier otro hecho que tenga que resolverse o que tenga un desarrollo secuencial
- El grado de importancia y de gravedad. Se mide en varios niveles y tiene que ver con la incidencia del hecho sobre la vida de la sociedad, a corto o largo plazo. Entre otras cosas, está ligado a otros valores como la cantidad y la jerarquía de las personas implicadas en el hecho
- La proximidad geográfica. Tiene que ver con la cercanía o lejanía del hecho con respecto al público
- La magnitud por la cantidad de personas y lugares implicados. Mientras más personas o ámbitos geográficos atañen al hecho, más valor tiene la noticia
- La jerarquía de los personajes implicados. Si son conocidos o no, despertará mayor o menor interés
- La inclusión de desplazamientos

La crisis financiera ha resultado un hecho de notable importancia y de gran calado social. Reúne algunas de las características citadas por Martini. Se trata de una serie de acontecimientos imprevisibles, importantes y graves. Son también cercanos al público, porque afectan a las economías de gran y pequeña escala, de gran magnitud y en que hay numerosos personajes públicos implicados. Esta investigación recoge el papel de los periodistas en la construcción lingüística de la realidad social mediante sus textos periodísticos ya sean géneros informativos o de opinión.

El concepto de evaluación en los textos ha sido esencial en nuestro estudio del CCF. El estudio de los recursos lingüísticos en diversas disciplinas ha sido objeto de estudio durante muchos

años. Los lingüistas han abordado la cuestión de la evaluación desde distintas perspectivas: semánticas o funcionales como en el caso de Martin y White (2005), discursiva (Hunston, 1994) o léxico-gramatical (Biber, 1995, 2006). Otros estudios relacionados son los de Hyland (1996), Thetela (1997) y Hunston (2010).

Pero antes de hablar de su análisis abordamos la construcción del corpus. Leech (1992) advierte que la investigación basada en un corpus es una metodología con peso científico suficiente ya que está abierta a la verificación objetiva de los resultados, además de otros aspectos como su simplicidad o su solidez. La primera fase en el desarrollo de un corpus implica definir cuestiones importantes como su finalidad, los límites temporales, geográficos y lingüísticos, el tamaño y tipo de textos y las proporciones temáticas. De todas estas cuestiones Villayandre (2006:38) destaca la finalidad como objetivo o motivo por el que se lleva a cabo la recopilación.

Puede ser una finalidad muy concreta (un corpus para un servicio automático de telefonía) o más general (un corpus para estudiar el funcionamiento de la lengua española). Mientras más amplio sea el objetivo o finalidad, más complicado es el diseño. También se consideran en este momento inicial cuestiones sobre la posible reutilización del corpus (lexicografía, procesamiento del lenguaje natural, fonética, etc.). Asimismo, hay que destacar que la finalidad del corpus condiciona muchas de las decisiones posteriores: no es lo mismo un corpus que va a servir como base para la elaboración de materiales de referencia (gramáticas, diccionarios) que otro cuya finalidad principal es el diseño de programas de procesamiento del lenguaje natural (correctores ortográficos, etiquetadores...).

Por otro lado, los límites temporales guardan relación con el período que han de contener los textos, los geográficos con las zonas y porcentajes de cada una de estas que deben integrar el corpus<sup>59</sup> y los lingüísticos a las lenguas o variedades lingüísticas a las han de pertenecer los textos si es multilingüe, bilingüe o monolingüe<sup>60</sup>. En cuanto al tamaño, éste está condicionado por su finalidad y tipo de textos lo que determinará de qué se nutrirá el corpus, tales como tipos

---

<sup>59</sup> Corpus de Referencia del Español Actual (50% de español de España-50% de español de América). Corpus del Español del Siglo XXI (30% de español de España-70% de español de América). *Corpus of Contemporary Spanish* (25% de español peninsular-25% de español de Argentina-50% de español de otras zonas de América del Sur). *The Bank of English* (proyecto COBUILD) (70% de inglés británico-20% de inglés americano-10% de inglés de otras zonas). Longman/Lancaster English Language Corpus (50% de inglés británico-40% de inglés americano- 10% de inglés de otras zonas).

<sup>60</sup> CRATER (*Corpus Resources and Terminology Extraction Project*) contiene textos en inglés, francés y español y el Corpus textual especializado plurilingüe del Instituto Universitario de Lingüística Aplicada de la Universidad Pompeu Fabra de Barcelona, está conformado por textos en cinco lenguas catalán, castellano, inglés, francés y alemán.

de textos (generales o específicos), distribución (textos concretos que integrarán el corpus: una lista que dé cuenta de los documentos que conformarán el corpus) y tamaño de las muestras (cantidad de texto que se tomará de cada documento para las muestras: textos íntegros o fragmentos; si son fragmentos de textos, la parte inicial, la central o la final de un texto).

Un corpus ha de ser representativo (Biber et al. 1998). Esto significa que un corpus debe proporcionar un modelo de las propiedades del lenguaje analizado de la manera más verosímil posible. Esta representatividad determina los tipos de hipótesis a formular y la obtención de conclusiones finales de la investigación. Para Sebba (1991: 23) cuanto más se acote la representatividad y mejor se defina el dominio del corpus más fiable serán sus afirmaciones. Al tratarse de un corpus especializado es necesario establecer criterios estrictos en la selección de textos. Sinclair (1991: 45) afirma que hay palabras y frases que resultan extraños en una muestra general del texto, pero pueden ser muy frecuentes en ciertos textos específicos.

Summers (1996: 186) considera que a menos que un corpus sea representativo será *ipso facto* poco fiable como medio de adquisición de conocimiento léxico. En este punto resulta crucial determinar la finalidad del corpus y la selección de textos que van a componerlo. Del tipo de discurso y conocimiento que se quiera representar dependerá el grado de especialización y la audiencia (Demarty-Warée, 1998: 100).

#### **4.2. Diseño del corpus**

El diseño y el tipo del corpus dependerán de su propósito y objetivo. Vargas (2006) establece una serie de parámetros que caracterizan su tipología y que veremos a continuación. El contenido del corpus marca si es general o especializado. Su tamaño nos indica si se trata de un corpus grande mediano o pequeño. A través de la codificación o nivel de procesamiento interno podemos clasificarlo como corpus simple (no acotado) o etiquetado. En el primer caso los textos recopilados no contienen anotación de ningún tipo frente a aquellos que contienen etiquetas descriptivas de los elementos constitutivos de cada texto o etiquetas que explicitan aspectos lingüísticos de cada ítem del corpus como las categorías gramaticales. El corpus también puede ser comparable o paralelo. En el primer caso está compuesto por textos que presentan características que comparten como es el caso del tema, el tipo de texto, período de tiempo en

que se redactaron, función comunicativa o el grado de especialización, por ejemplo, o corpus paralelo; en el caso de que los textos que contenga estén redactados en una lengua junto a traducciones a otros idiomas.

En el caso del CCF decidimos elaborar un corpus a partir de las publicaciones vinculadas a la crisis financiera en los principales periódicos de España. Ya hemos dicho que escogimos *El País* y *El Mundo* como representantes en la prensa escrita del bipartidismo gubernamental. *El País* está el relacionado con el sector progresista que representa el Partido Socialista Obrero Español (PSOE) y *El Mundo* en representación del ala conservadora que encarna en la época de la investigación el Partido Popular. Ambas formaciones políticas que han tenido responsabilidad de Gobierno en España durante el periodo de crisis económica analizada. En el punto 1.5 ya realizamos una descripción y evolución histórica de estos dos medios a partir de cuyas publicaciones hemos construido el CCF.

El objetivo es radiografiar las publicaciones del año 2012 en los medios mencionados y extraer el mensaje que ha quedado plasmado en publicaciones de la sección de economía y de opinión, analizarlo y organizarlo para poderlo representar. El corpus recopilado para este análisis es diacrónico y lo hemos trabajado por trimestres, lo cual nos va a permitir ver cómo evoluciona.

El contenido del corpus ha de estar estrechamente vinculado a la finalidad que se quiere alcanzar. Si echamos un vistazo a la tradición a la hora de elaborar un corpus comprobamos que han existido dos métodos básicos para determinar su contenido. Uno de ellos se basa en tomar extractos de entre 2.000 y 2.500 palabras de una variedad de textos predeterminados que proporcionen una muestra equilibrada. Stubbs (1996: 32) sostiene que algunas características lingüísticas de un texto se distribuyen uniformemente. Los primeros corpus, que tenían un tamaño limitado, seguían este método. La otra opción es utilizar textos completos algo que llegó a ser más común con el desarrollo de los corpus más grandes. En este punto destacamos el trabajo realizado por Scott (1997) con el *word-and-collocation-span model*. Esta aproximación comienza por el análisis de una palabra central o nodo y se basa en las colocaciones de palabras a una distancia predeterminada. La estadística de las palabras clave viene determinada al comparar el texto completo de un corpus con uno de referencia a través del software *Wordsmith Tools*.

Ya hemos visto en el apartado 3.12 cómo Vargas (2012: 73) divide las herramientas informáticas empleadas en el tratamiento y análisis de un corpus en etiquetadores morfosintácticos y programas de análisis textual o programas de concordancias. Estos últimos constituyen un conjunto de aplicaciones capaces de analizar un corpus y generar diferentes listas de resultados a través de la herramienta *Wordsmith Tools* y ofrecer datos mediante listas con estadísticas, listado monoléxico ordenado alfabéticamente, listado poliléxico de todas las palabras o de una selección a través de *clusters*, líneas de concordancias o listados de una palabra específica, de búsqueda o clave, listado de palabras clave, gráfico de distribución de la palabra de búsqueda o *plot* y lista de *collocates*. Asegura que “al crear listados de palabras mono y poliléxicos los programas de concordancias pueden ser utilizados como extractores de terminología”.

En este contexto, es importante no descuidar el trabajo que en materia de terminografía ha realizado Cabré (1993: 113), que da un paso respecto al concepto tradicional y establece que se trata de la identificación y recopilación de los términos que los especialistas utilizan en realidad. En terminografía, se utiliza el corpus especializado. Es un tipo de corpus especial diseñado con un propósito específico y una finalidad representativa de un tipo particular de lengua. Es el caso de un campo de especialidad o un grupo particular de hablantes. Durán (2011: 2) en su análisis de los criterios para la elaboración y diseño de los corpus especializados para la terminografía señala que:

Si un terminógrafo debe estudiar los términos que los especialistas utilizan en su trabajo diario, deberá consultar directamente con los especialistas del campo de especialidad en cuestión o realizar un estudio detallado de las producciones lingüísticas que estos especialistas crean para comunicarse entre ellos o con otros actores. De estas dos opciones, la primera no es siempre posible, ya que puede resultar complicado disponer de los especialistas adecuados y, cuando se puede acceder a ellos, a menudo encuentran dificultades a la hora de explicar el significado y el uso del lenguaje que emplean, al fin y al cabo, de forma intuitiva.

Durán (2011: 2-3) continúa explicando que los avances en el ámbito de la informática han posibilitado disponer de la información en formato electrónico y su procesamiento automático o semiautomático con lo que el uso de los corpus electrónicos se ha extendido en la actualidad, hasta convertirse en la herramienta esencial para la mayoría de los trabajos terminográficos y dando lugar a lo que Leech (1992: 106) considera una nueva forma de pensar sobre el lenguaje.

El análisis cuantitativo del CCF nos ha permitido conocer el índice de frecuencias de elementos de los textos periodísticos a través de herramientas informáticas (*Wordsmith*) lo cual nos ha proporcionado datos sobre las palabras utilizadas con un elevado número de apariciones sobre los doce actores principales o protagonistas de la crisis financiera en cada uno de los trimestres, y el entorno en el que aparecen, es decir, lo que conocemos como colocaciones y líneas de concordancia. Más adelante veremos cómo el método de análisis adaptado ha proporcionado evidencias sobre determinadas categorías semánticas que co-aparecen con estas entidades tales como adjetivos y sustantivos abstractos.

#### **4.2.1 Diseño de una base de datos**

La constitución del corpus es el proceso de establecimiento de los criterios que han de definir su diseño y recopilación de textos. Sinclair (1996) establece una serie de criterios internos y externos para la elaboración de un corpus. Mientras los criterios internos se basan en la distribución de palabras, los externos abordan cuestiones relacionadas con su entorno. Villayandre (2006) también aborda la descripción de los criterios internos y externos. De manera muy sucinta podemos señalar que la autora explica que los criterios internos se centran en la aparición de patrones o elementos diferenciadores de la variedad lingüística de un texto como la longitud de las oraciones y que dan lugar a los tipos de texto. Entre este tipo de criterios internos destaca el tema, dominio o ámbito al que pertenece el texto y el estilo. Al hablar del tema detalla una clasificación de seis hipercampos que se agrupan a su vez en un gran número de áreas temáticas, en el caso de los textos informativos, y aporta un séptimo hipercampo para textos de ficción. En lo relativo al estilo, señala la autora que es el tipo de modelo de lengua, si es informal o formal, y a su vez se enmarca en el subtipo privado o público, si se ha producido en un contexto natural, en una conversación telefónica o en los medios de comunicación. En este último caso pueden estar relacionados con las entrevistas, publicaciones científicas, reportajes, informaciones deportivas. En lo referente a los criterios externos destaca la cronología, el origen, estado, objetivo, género literario y medio de publicación.

Tras esta referencia a la literatura sobre el establecimiento de criterios para diseñar una base de datos retomamos la idea del objetivo primero fijado para esta investigación: la recopilación y diseño de una base de datos de publicaciones periodísticas sobre la crisis financiera en España del año 2012 para su posterior análisis. A la vista de lo establecido en el capítulo 3.5, en que

hemos determinado la tipología y clasificación de los corpus, recordamos ahora, y a modo de resumen, que el CCF se ha construido teniendo en cuenta los siguientes criterios:

- Es un corpus especializado de publicaciones relativas a la crisis financiera. (Criterio interno: tema)
- Es un corpus escrito elaborado con textos de medios de comunicación escritos. (Criterio externo: origen)
- Es un corpus monolingüe de textos en castellano. (Criterio interno: dominio)
- Es un corpus sincrónico ya que recoge textos de un tiempo determinado de un año natural (2012). (Criterio interno: ámbito)
- Es un corpus cerrado que se ciñe al momento de estudio, aunque no se descarta ir incorporando textos nuevos. (Criterio interno: ámbito)
- Es un corpus codificado o anotado al que se han añadido etiquetas de algunos elementos de los documentos recopilados. (Criterio externo: estado)
- Es un corpus finito. Abarca del 1 de enero al 31 de diciembre de 2012. (Criterio interno: ámbito)

Vistos ya los criterios de construcción ¿qué textos formarían el CCF? ¿Con tipo de informaciones crearíamos este corpus? ¿Cuáles han sido las metas fijadas inicialmente? El objetivo que marcamos inicialmente fue el estudio de la carga semántica de los textos, por lo que optamos en primera instancia por seleccionar, dentro de las publicaciones relacionadas con la crisis económica, los textos de las secciones de opinión de los mencionados periódicos.

Los dos periódicos seleccionados, tanto *El País* como *El Mundo*, disponían de edición digital. Esta circunstancia pone en nuestro favor el trabajo de recopilación. Como ya pusimos por escrito en el punto 3.9, en el que analizamos Corpus y la Web, internet ha provocado un cambio decisivo en los medios de comunicación y en la difusión de la información. Ya dijimos que los mayores proveedores de noticias han adoptado los medios digitales en la producción y transmisión de noticias. Lo explicamos al hablar del informe de la *Open Society Foundation, Mapping Digital Media: Spain 2* (2012: 7). El hecho de que los periódicos digitales emulen la función de la prensa tradicional, pero con el uso de nuevos canales, los hace mucho más accesibles (Salaverría, 2005).

Este avance de los medios de comunicación *online* convierte la web en una magnífica fuente para la elaboración de un corpus.

La adaptación de la prensa tradicional a la era digital aportó un notable impulso a la elaboración del CCF. Una vez más recordamos que un mes después del inicio de la recopilación de textos de las secciones de opinión de las versiones digitales de los mencionados diarios se produjo un importante cambio en la edición digital de *El País* que nos invitó a ampliar la perspectiva de nuestro corpus. Ante la coyuntura económica y una mayor demanda de información, en febrero de 2012, *El País* dio impulso a su oferta digital con el lanzamiento de una nueva sección de economía<sup>61</sup> que vino a potenciar la información y el análisis por parte de firmas destacadas, además de incorporar nuevos blogs, herramientas y un seguimiento permanente de los mercados. La nueva sección se presentaba con las ventajas informativas de la era digital y con la solvencia de la prensa tradicional. Información actualizada y publicada en formato web pero sobre fondo salmón, el color asociado internacionalmente a la prensa económica. La sección arrancaba el 31 de enero de 2012 con una serie de informaciones exclusivas y con accesos directos a los temas macroeconómicos, de empresas, mercados, bolsa, finanzas personales, vivienda y tecnología. En la web de *El País* se abría un apartado especial denominado “La crisis” con informaciones, reportajes y entrevistas sobre el impacto de la crisis económica en los ciudadanos, desde el drama del desempleo a los desahucios de viviendas pasando por los recortes sociales, el empobrecimiento de la clase media y la pérdida de poder adquisitivo de las familias. Vemos en este paso dado por *El País*, con la creación de esta nueva sección, aquello que Salaverría (2005) plantea sobre la función de los periódicos digitales al emular la prensa tradicional, pero con canales actualizados y con voluntad de mediación entre hechos y público y los criterios de uso y niveles de cualificación, multimedia, interactividad e hipertexto.

La decisión de incorporar los textos de esta nueva sección fue decisiva para obtener un corpus más potente para su posterior análisis. De este modo, los textos recopilados de la sección de Economía de *El País* quedarían identificados con las siglas *EP*, mientras que los de las secciones de Opinión de *El País*, relacionados con la crisis financiera llevarían las siglas *OEP* y los de la sección de Opinión de *El Mundo* sobre con la crisis financiera los identificaríamos como *OEM*.

---

<sup>61</sup> <http://economia.elpais.com/hemeroteca/economia/2012/01/31/n/portada.html> [26-XI- 2015]

Todos los textos recopilados se fueron clasificando por orden de publicación en un documento Excel que hemos denominado *Financial Crisis Corpus* o *Corpus de la Crisis Financiera* (CCF).

#### 4.2.2 Organización de los textos

La metodología aplicada comienza con la elaboración del CCF, pasando por su diseño, hasta el análisis del mismo utilizando los métodos y procesos característicos de la lingüística de corpus tales como listas de frecuencias, palabras clave, líneas de concordancia y colocaciones, patrones recurrentes y clústeres. Respecto al desarrollo del corpus, los textos del CCF se fueron recogieron y poco a poco analizando trimestralmente. De este modo las preguntas de la investigación podrían ser replicadas y reformuladas a medida que el estudio avanzaba. Más adelante veremos que la investigación realizada se ha apoyado además en los métodos relacionados con el análisis de redes sociales (Hanneman y Riddle, 2005).

El CCF elaborado para la presente investigación está construido con los siguientes metadatos clasificados en diversas hojas de un documento Excel que contienen:

- El nombre del periódico que hemos analizado. En este caso son dos *El Mundo* y *El País*. Como ya hemos avanzado los textos de la sección de Opinión los identificamos con los acrónimos *OEM* (Opinión *El Mundo*) y *OEP* (Opinión *El País*) y los de la sección de Economía de *El País* que identificamos como *EP*
- La fecha de publicación de los textos. Van por orden desde enero de 2012 hasta diciembre de ese mismo año
- El autor o editor de los textos. El periodista que firma la información o el lector que emite la carta al medio. Únicamente aparecen sin autor los editoriales ya que son textos que nunca van firmados en los medios porque representan la ideología general del periódico
- El género periodístico de las publicaciones. Si son noticias, crónicas, entrevista, reportaje (de los géneros periodísticos objetivos) o si son editoriales, columna, tribunas, artículos, cartas (de los géneros de opinión o subjetivos)

La siguiente tabla es un ejemplo de los metadatos del corpus y cómo los hemos organizado en el documento Excel.

| Periódico | Fecha      | Autor/Editor  | Artículo (género) | Texto id.    | Palabras |
|-----------|------------|---------------|-------------------|--------------|----------|
| El País   | 01/01/2012 | CARLOS E. CUE | Reportaje         | EP_01_01_12  | 2.300    |
| El Mundo  | 02/01/2012 | CARLOS CUESTA | Columna           | OEM_02_01_12 | 446      |

Tabla 4.1: Metadatos del CCF

En la primera columna (Periódico) figura el nombre de la publicación (*El País/El Mundo*), seguida de la columna de la Fecha, el nombre del Autor/Editor, en las columnas segunda y tercera, respectivamente. Cada texto, que denominamos Artículo (género), está asociado a un hipervínculo (Texto id.) que ocupan las columnas cuarta y quinta. El número de palabras de cada texto (Palabras) ocupa la sexta columna. Al final, sumadas nos dará el número total de palabras del CCF.

Según Meyer y Mackintosh (1996: 264) la consulta de documentación especializada en forma de corpus textual, es más accesible, rápida y directa y, por tanto, determinante en el trabajo terminológico:

While experts obviously know their domains, they do not all explain their knowledge clearly (whether orally or in writing), completely (it is up to the knowledge acquirer to make sure that all important areas in the field are covered), or consistently (experts often disagree with each other or change their minds).

Descritos los primeros objetivos, vemos en la tabla 4.2 un resumen de las tareas realizadas y metodología empleada en las distintas etapas de la investigación.

| <b>Análisis del Corpus de Crisis Financiera con herramientas informáticas + interpretación manual</b> |   |
|---|---|
| <b>OBJETIVO 1</b>   | Recopilar y diseñar una base de datos de artículos periodísticos sobre la crisis financiera en España del año 2012  |
| <b>OBJETIVO 2</b>   | Recopilar datos de cada trimestre por periódico/sección   |
| <b>OBJETIVO 3</b>   | Extraer listas de palabras de cada trimestre por periódico /sección   |
| <b>OBJETIVO 4</b>   | Diseñar tabla comparativa de palabras más frecuentes por trimestres   |
| <b>OBJETIVO 5</b>   | Determinar los 12 actores sociales más frecuentes por trimestres y extraer sus colocaciones y líneas de concordancia. Diseñar redes ilustrativas de las relaciones  |
| <b>OBJETIVO 6</b>   | Extraer los sustantivos abstractos del CCF  |
| <b>OBJETIVO 7</b>   | Seleccionar los adjetivos que acompañan a cada actor social   |
| <b>OBJETIVO 8</b>   | Establecer categorías semánticas de los adjetivos   |
| <b>OBJETIVO 9</b>   | Analizar manualmente las líneas de concordancia y asignar etiquetas de ejemplos de lenguaje evaluativo. Recopilar etiquetas más frecuentes. Transferir a un documento Excel las etiquetas propuestas con ejemplos |
| <b>OBJETIVO 10</b>  | Analizar textos completos (editoriales): macroestructura y lenguaje evaluativo  |

Tabla 4.2: Metodología: Etapas de la investigación

A partir de las diferentes etapas involucradas en el estudio, la metodología implementada combina el enfoque asistido por ordenador para el análisis de textos con la interpretación manual de los resultados obtenidos estadísticamente en etapas anteriores. Nuestra hipótesis inicial es que un corpus de textos de prensa sobre economía y su posterior procesamiento y análisis lingüístico es una metodología adecuada para:

- Identificar a los principales actores sociales
- Descubrir los principales acontecimientos

- Explorar la forma en que ambos son evaluados por los autores de los textos

La intención es descubrir cuáles son las entidades y escenarios de la crisis financiera, que comenzó en 2008 con la quiebra de Lehman Brothers y si existe alguna interrelación lingüística entre estas entidades. Las primeras hipótesis de la investigación se centraron en analizar el juego dialéctico que se libra entre los actores sociales y los editores en la elaboración de las noticias.

El uso de diferentes herramientas informáticas nos ha provisto de listas de palabras de nuestro CCF. Con ellos hemos identificado los principales actores sociales y sus colocaciones: qué personajes destacan y con qué términos se asocian. Hemos utilizado los métodos relacionados con el análisis de redes sociales (Hanneman y Riddle, 2005) para ver de manera gráfica esta relación. En este apartado hemos combinado el uso de las herramientas informáticas con el análisis manual. Veremos cómo los sustantivos abstractos han sido localizados y procesados debido a lo que consideramos términos con alta densidad semántica y de gran fuerza expresiva y que han aportado conclusiones interesantes en este estudio. También hemos explorado nuestros actores sociales y los adjetivos que los acompañan a fin de clasificarlos semánticamente y de manera manual. Las etapas finales (objetivos 9 y 10 de la tabla 4.2) en la investigación se extienden a las unidades de significado más largas (frases, líneas de concordancia, párrafos y textos). Veremos que la sección de resultados en este documento ofrecerá una visión general de las características más sobresalientes del CCF.

#### **4.2.2.1 Descripción del CCF**

La obtención del CCF constituye ya el primer resultado del trabajo de recopilación de textos periodísticos. Al finalizar la compilación de un año natural, el 2012, en los dos medios de comunicación analizados, hemos contabilizado 3.254.499 palabras que pertenecen a 5.480 textos periodísticos. Por la tipología textual (los géneros de opinión y por contenidos de específicos de la sección de economía) lo consideramos un corpus especializado. Como vimos en el punto 3.2.1 este tipo de corpus entra dentro de la clasificación de los que van desde 50.000 hasta 5 millones de palabras.

En el capítulo teórico de la lingüística de corpus, Sinclair (1991: 171) o Sampson (2001: 124) apuestan por este tipo de corpus, que proporciona una representación adecuada del lenguaje y suficientes y múltiples ocurrencias de los ítems objeto de la investigación y disponen de una muestra importante de uso del lenguaje de la vida real, para asegurar que existe evidencia suficiente para generar o probar hipótesis acerca del lenguaje.

Estos son los datos finales de la recopilación de artículos de los cuatro trimestres del año 2012 en los que hemos subdividido la investigación.

|                  | <b>Total palabras</b> | <b>Total artículos</b> |
|------------------|-----------------------|------------------------|
| <b>TOTAL CCF</b> | <b>3.254.499</b>      | <b>5.480</b>           |

Tabla 4.3: Datos del CCF

El desglose por trimestres de los datos del CCF, sobre el número total de palabras y de artículos o textos recopilados, nos muestra las primeras informaciones del estudio lingüístico de la crisis financiera en el período analizado. Si exploramos al detalle el CCF cada tres meses, vemos cómo se produce un claro despegue de las informaciones publicadas en los medios, según avanza el año. En concreto, a partir del segundo trimestre es cuando las publicaciones dedicadas a la información de economía van cobrando vuelo respecto al pasado.

En la tabla desglosamos estas cifras y vemos cómo evolucionan los dos diarios analizados en esta investigación:

| Textos / Procedencia                       | Número palabras | Número artículos |
|--|-----------------|------------------|
| Sección Economía El País Primer Trimestre  | 326.822         | 454              |
| Sección Economía El País Segundo Trimestre | 350.712         | 452              |
| Sección Economía El País Tercer Trimestre  | 533.545         | 746              |
| Sección Economía El País Cuarto Trimestre  | 467.542         | 704              |
| Sección Opinión El Mundo Primer Trimestre  | 172.301         | 413              |
| Sección Opinión El Mundo Segundo Trimestre | 202.445         | 437              |
| Sección Opinión El Mundo Tercer Trimestre  | 164.865         | 385              |
| Sección Opinión El Mundo Cuarto Trimestre  | 112.105         | 291              |
| Sección Opinión El País Primer Trimestre   | 170.821         | 294              |
| Sección Opinión El País Segundo Trimestre  | 262.792         | 484              |
| Sección Opinión El País Tercer Trimestre   | 259,803         | 421              |
| Sección Opinión El País Cuarto Trimestre   | 230.746         | 399              |

Tabla 4.4: Evolución por trimestres del CCF

Como paso previo al análisis semántico, hacia el que evolucionará esta investigación, nos fijamos primero de manera aséptica en los números y comprobamos que en los 2.356 textos recopilados en la sección de economía de *El País* relacionados con la crisis financiera hemos contabilizado 1.678.621 palabras. En el caso de la secciones de opinión vemos un primer dato de interés acerca de la diferencia del número de palabras que agrupan un casi número similar de artículos. Mientras que en la sección de opinión de *El Mundo* recopilamos 651.716 palabras contabilizadas en 1.526 artículos, en la sección de economía de *El País* en los 1.598 artículos analizados el número de palabras es muy superior y alcanza las 924.162 palabras.

Por último, en esta primera descripción esquemática del CCF, veamos en la figura 4.1 cómo afecta porcentualmente este reparto, según el número de palabras, al resultado final cada una de las tres secciones analizadas.

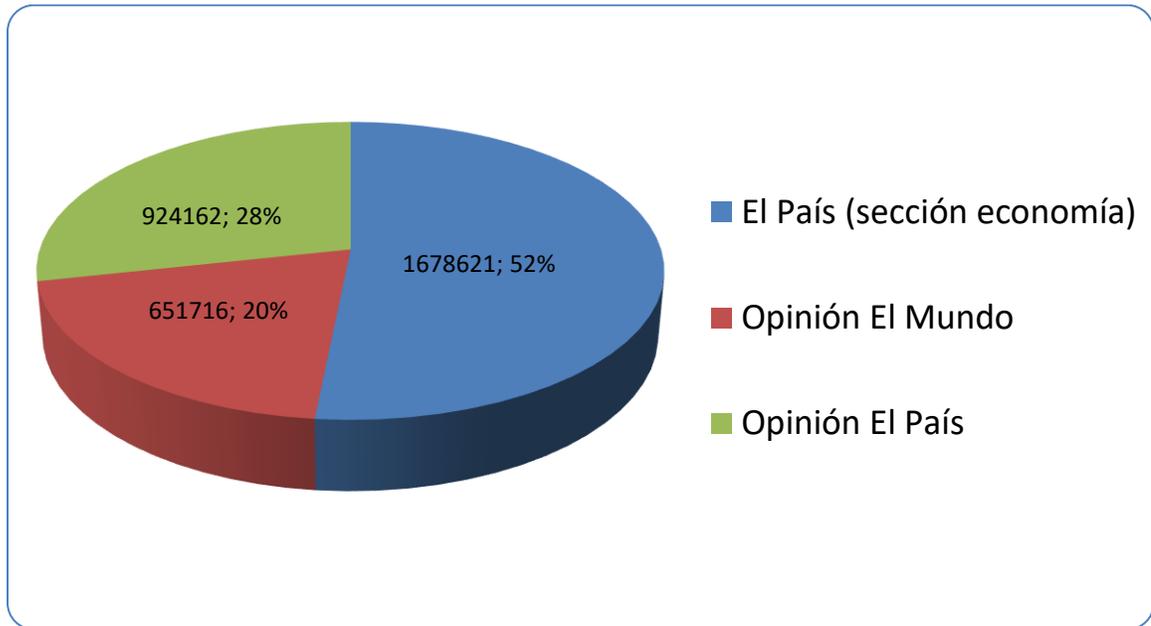


Figura 4.2: Descripción esquemática del CCF

Poco más de la mitad, el 52% del corpus lo conforman los textos recopilados de la sección de economía de *El País*, que como preveíamos, iba a aportar gran parte de la masa muscular del corpus a investigar. En concreto, 1.678.621 palabras. La otra mitad del CCF queda repartido entre un 28% proveniente de la sección de opinión de *El País*, que ha generado 924.162 palabras y el otro 20% restante es de la sección de opinión de *El Mundo*, con 651.716 palabras.

### 4.3 Análisis del corpus

Una vez establecidos los criterios de recopilación de textos y ya con el diseño definitivo de la base de datos sobre la crisis financiera en España a lo largo del año 2012, comenzó el proceso de compilación y agrupación vinculada. Los textos extraídos y guardados en formato plano se fueron asociando a un documento Excel, en el que los fuimos clasificando desde la publicación más antigua, situada al inicio de la tabla, a la más actual, que ocupa el final de la tabla. A continuación presentamos tres imágenes de la parte alta del CCF.

| 1  | Periodico | Fecha      | Autor / Editor                    | Artículo (género) | Texto id.                       | Palabras |
|----|-----------|------------|-----------------------------------|-------------------|---------------------------------|----------|
| 2  | El País   | 01/01/2012 | CARLOS E. CUÉ                     | Reportaje         | <a href="#">EP 01 01 12</a>     | 2.300    |
| 3  | El País   | 01/01/2012 | MIGUEL JIMÉNEZ                    | Noticia           | <a href="#">EP 01 01 12 [2]</a> | 454      |
| 4  | El País   | 01/01/2012 | AGENCIAS                          | Noticia           | <a href="#">EP 01 01 12 [3]</a> | 744      |
| 5  | El País   | 02/01/2012 | ANDREU MISSÉ                      | Reportaje         | <a href="#">EP 02 01 12</a>     | 1.552    |
| 6  | El País   | 02/01/2012 | ÁLVARO ROMERO                     | Noticia           | <a href="#">EP 02 01 12 [2]</a> | 822      |
| 7  | El País   | 02/01/2012 | EFE                               | Noticia           | <a href="#">EP 02 01 12 [3]</a> | 360      |
| 8  | El País   | 02/01/2012 | ALEJANDRO BOLAÑOS / ÍÑIGO DE BARI | Reportaje         | <a href="#">EP 02 01 12 [4]</a> | 804      |
| 9  | El País   | 03/01/2012 | EL PAIS                           | Noticia           | <a href="#">EP 03 01 12</a>     | 367      |
| 10 | El País   | 03/01/2012 | MIGUEL JIMÉNEZ                    | Noticia           | <a href="#">EP 03 01 12 [2]</a> | 864      |
| 11 | El País   | 03/01/2012 | ALICIA GONZALEZ                   | Noticia           | <a href="#">EP 03 01 12 [3]</a> | 551      |
| 12 | El País   | 03/01/2012 | MANUEL V. GÓMEZ                   | Reportaje         | <a href="#">EP 03 01 12 [4]</a> | 488      |

Figura 4.3: Inicio página EXCEL de textos El País (EP)

| 1  | Periodico | Fecha      | Autor / Editor            | Artículo (género) | Texto id.                        | Palabras |
|----|-----------|------------|---------------------------|-------------------|----------------------------------|----------|
| 2  | El País   | 01/01/2012 | JOSÉ IGNACIO TORREBLANCA  | Análisis          | <a href="#">OEP 01 01 12</a>     | 1.071    |
| 3  | El País   | 02/01/2012 | MIQUEL ALBEROLA           | Análisis          | <a href="#">OEP 02 01 12</a>     | 176      |
| 4  | El País   | 03/01/2012 | IGNACIO SOTELO            | Columna_opinión   | <a href="#">OEP 03 01 12</a>     | 721      |
| 5  | El País   | 03/01/2012 | PAUL KRUGMAN              | Columna_opinión   | <a href="#">OEP 03 01 12 [2]</a> | 966      |
| 6  | El País   | 03/01/2012 | JOSEP RAMONEDA            | Columna_opinión   | <a href="#">OEP 03 01 12 [3]</a> | 800      |
| 7  | El País   | 04/01/2012 | MANUEL RODRÍGUEZ RIVERO   | Columna_opinión   | <a href="#">OEP 04 01 12</a>     | 662      |
| 8  | El País   | 05/01/2012 | XAVIER VIDAL-FOLCH        | Columna_opinión   | <a href="#">OEP 05 01 12</a>     | 690      |
| 9  | El País   | 06/01/2012 | FRANCISCO TERRASA BESTARD | Carta_opinión     | <a href="#">OEP 06 01 12</a>     | 234      |
| 10 | El País   | 06/01/2012 | GERARDO SEISDEDOS         | Carta_opinión     | <a href="#">OEP 06 01 12 [2]</a> | 181      |
| 11 | El País   | 07/01/2012 | KATTYA CASCANTE           | Análisis          | <a href="#">OEP 07 01 12</a>     | 646      |
| 12 | El País   | 07/01/2012 | CARLOS ACOSTA             | Carta_opinión     | <a href="#">OEP 07 01 12 [2]</a> | 324      |

Figura 4.4: Inicio página EXCEL de textos de opinión de El País (OEP)

| 1  | Periodico | Fecha      | Autor / Editor           | Artículo (género) | Texto id.                        | Palabras |
|----|-----------|------------|--------------------------|-------------------|----------------------------------|----------|
| 2  | El Mundo  | 01/01/2012 |                          | Editorial         | <a href="#">OEM 01 01 12</a>     | 533      |
| 3  | El Mundo  | 02/01/2012 | CARLOS CUESTA            | Columna_opinión   | <a href="#">OEM 02 01 12</a>     | 446      |
| 4  | El Mundo  | 02/01/2012 |                          | Editorial         | <a href="#">OEM 02 01 12 [2]</a> | 146      |
| 5  | El Mundo  | 02/01/2012 | ANTONIO GALA             | Columna_opinión   | <a href="#">OEM 02 01 12 [3]</a> | 190      |
| 6  | El Mundo  | 02/01/2012 | CASIMIRO GARCÍA ABADILLO | Tribuna           | <a href="#">OEM 02 01 12 [4]</a> | 1.019    |
| 7  | El Mundo  | 02/01/2012 | ANTONIO LUCAS            | Columna_opinión   | <a href="#">OEM 02 01 12 [5]</a> | 475      |
| 8  | El Mundo  | 02/01/2012 | PEDRO J. PIQUERAS IBÁÑEZ | Carta_opinión     | <a href="#">OEM 02 01 12 [6]</a> | 129      |
| 9  | El Mundo  | 02/01/2012 | SERGIO RAMÓN RODRIGUEZ   | Carta_opinión     | <a href="#">OEM 02 01 12 [7]</a> | 287      |
| 10 | El Mundo  | 02/01/2012 | JUAN SÁNCHEZ MANRUBIA    | Carta_opinión     | <a href="#">OEM 02 01 12 [8]</a> | 82       |
| 11 | El Mundo  | 03/01/2012 | LUCÍA MÉNDEZ             | Columna_opinión   | <a href="#">OEM 03 01 12</a>     | 525      |
| 12 | El Mundo  | 03/01/2012 | MARISA MESTRES           | Carta opinión     | <a href="#">OEM 03 01 12 [2]</a> | 60       |

Figura 4.5: Inicio página EXCEL de textos de opinión de El Mundo (OEM)

Este formato nos facilita el estudio del CCF por trimestres, para ir analizando poco a poco los resultados y su evolución, según avanza el año. Como hemos visto, cada publicación tiene su hoja. Aquí matizamos que en el caso de *El País* se distribuyeron los textos en dos hojas: una para los textos de la información general y otra para sección de opinión. En el caso de *El Mundo*, le

corresponde una sola hoja del documento. Veamos, a continuación, la secuencia al detalle de los pasos dados para elaborar el CCF:

1- El procesamiento de los textos comienza en enero de 2012, con un rastreo de la información publicada en las ediciones digitales de *El País* y *El Mundo*. El primer objetivo eran las secciones de opinión, es decir localizar y recopilar los géneros de opinión que desglosamos en el apartado 2.2.4 como son editoriales, columnas, artículos de opinión, tribunas y cartas al director, de estos dos periódicos como representantes del bipartidismo que en ese año imperaba en el país. La intención era iniciar la confección de un corpus en el que se pudiera realizar el análisis del lenguaje evaluativo, por eso se optó por los textos de la sección de opinión.

2- Los textos seleccionados se han guardado en archivos en formato plano (a cada texto se le ha asignado un código de identificación). A través de un hipervínculo se relacionaban con el medio en que había sido publicado, la fecha, el autor, el género al que pertenecen y el número de palabras que aportan al CCF. Como hemos visto en las imágenes, toda esta información queda estructurada en diversas páginas en formato Excel.

3- Con la aparición de la nueva sección de economía, en febrero de 2012, ésta se incorpora al CCF, como hemos explicado anteriormente, para dotar de mayor potencia a este corpus. Por ello se amplía la clasificación de los textos y se incorpora una nueva página Excel con las referencias de los textos extraídos de la sección de Economía de *El País*. Se clasifican por orden de publicación, su autor, género periodístico y número de palabras que aportan al CCF. En el apartado de géneros de opinión se introduce la división de los géneros periodísticos objetivos cuya clasificación ya vimos en el apartado 2.2.1: noticia, crónica, reportaje y entrevista.

4- En la elaboración de las tablas para el CCF el patrón diseñado es el siguiente: el nombre del periódico que hemos analizado. En esta investigación son dos los medios de comunicación analizados, *El Mundo* y *El País*, ambos en la sección de opinión que identificamos con los acrónimos OEM y OEP, respectivamente, y en el caso de la sección de economía de *El País*, también analizada, sus textos los identificamos como EP. Como ya hemos avanzado, el siguiente metadato del CCF es la fecha de publicación de los textos. Están colocadas por orden de publicación desde enero de 2012.

5- Identificar a los autores o editor de los textos en la tercera columna del documento de clasificación de textos del CCF. Éste es el periodista que firma la información, el lector que emite la carta al medio, el columnista o el experto que plasma sus opiniones en el periódico en cuestión. Únicamente aparecen sin autor los editoriales ya que se trata de textos que nunca van firmados en los medios de comunicación por representar la ideología general del periódico, como ya vimos en las características de los géneros periodísticos.

6- Determinar las publicaciones por géneros periodísticos. Si son noticias, crónicas, entrevistas, reportajes, que se engloban en los géneros periodísticos objetivos o si, por otro lado, se trata de editoriales, columnas, tribunas, artículos de opinión, cartas al director, estas últimas enmarcadas en los géneros de opinión.

7- Asociar a cada texto un hipervínculo que conforma el documento-esqueleto del CCF y, por último, las tablas de este corpus reflejan el número de palabras de cada texto, que al final sumadas nos dará el número total de palabras contabilizadas.

#### **4.3.1 Extracción de listas de palabras**

Después del análisis cuantitativo del CCF, que hemos visto en el punto 4.2.2.1, y la observación de la estructura o esqueleto del corpus, el siguiente estudio va a ser semántico. Incidimos una vez más en la hipótesis de partida de esta investigación, que es la creación de un corpus de artículos de opinión y de información económica para su posterior tratamiento y análisis lingüístico con herramientas informáticas y también de manera manual, considerado todo ello como una metodología adecuada para extraer los principales actores, acontecimientos y cómo estos se evalúan. Nuestro interés final se centra en identificar en los textos extraídos de los periódicos las entidades y los escenarios de la crisis financiera en el periodo analizado y si hay establecida algún tipo de interrelación en el ámbito lingüístico entre dichas entidades.

El paso previo a este análisis terminológico es la toma de decisiones y la preparación de las herramientas informáticas que utilizaremos para la extracción de resultados y su valoración a la vista de los objetivos marcados. Este punto abarca la descripción del método utilizado para la extracción de esta información del CCF y la evaluación de otros métodos que la literatura sobre

este tema aborda. Doyle (2005) establece como punto conveniente y principio fundamental de toda investigación la obtención de la información siguiendo procedimientos diferentes que permitan comparar y contrastar los resultados obtenidos. Valga el siguiente ejemplo de ir por casa para ilustrar esta idea; a la hora de hacer la colada determinadas prendas las lanzamos directamente a la lavadora mientras que otras, por sus características, han de ser lavadas a mano para una mejor puesta a punto de nuestro vestuario. En el caso del análisis lingüístico del corpus la combinación del método a mano y a máquina nos permitirá aproximarnos a un corpus desde perspectivas distintas. De este modo, al replicar (Doyle 2005: 1) el proceso lo desarrollaremos de una manera óptima en cualquier investigación empírica:

Replication is a method of validating research in the sciences, and is generally considered to be a criterion for the acceptance of new theories and knowledge within those domains. In the context of testing hypothesis through empirical research, replication means to repeat an experiment in such a way that scientists can be certain that the results of two experiments are comparable and, therefore, that they are measuring the same phenomenon.

El uso de distintas herramientas informáticas nos ha provisto de listas de palabras extraídas de los textos que componen el CCF y que hemos distribuido por trimestres. El contenido de estas listas de palabras nos ha permitido la identificación de los actores sociales y sus *collocates*, es decir qué personajes salen más a escena y asociados con qué términos. Listas de palabras o *wordlists* y *collocates* son términos que ya identificamos y explicamos en el punto 3.4 cuando analizamos los conceptos clave de la lingüística de corpus.

¿Cómo obtuvimos las listas de palabras? La extracción de esta información lingüística se llevó a cabo de forma automática a partir del CCF en su totalidad. Dado que la recopilación de datos por medio de publicación y secciones había sido una premisa en la construcción del CCF, esta circunstancia iba a seguir siendo así en la extracción de listas de palabras. Recordamos que el análisis de los primeros datos empíricos del CCF, sobre número de palabras (3.254.499), se realizó a partir de la subdivisión en trimestres. De este modo, las listas de palabras se extrajeron como una combinatoria de ambas estructuras. Es decir, la extracción de listas de palabras del CCF se hizo por medio de publicación, por sección y trimestres. Esta tarea, con la que se perseguía la obtención de información útil de carácter estadístico, se realizó de forma manual combinada con el uso de herramientas informáticas (véase Tabla 4.3).

La idea inicial era el análisis de las estructuras típicas en los textos del corpus, de las estructuras léxico-gramaticales de dichos textos, el lenguaje evaluativo de los textos. La obtención de listas de palabras nos permitiría, en primera instancia, la realización de búsquedas de palabras y frases, la elaboración de listas de frecuencias y de concordancias, de información sobre la distribución de palabras y la evaluación de los resultados de los análisis de colocaciones mediante cálculos estadísticos.

La herramienta informática con la que iniciamos el proceso para la obtención de esta información es *Wordsmith Tools Wordlist*. Scott (2012: 5) sobre las listas de palabras precisa:

Este programa genera listas de palabras basadas en uno o más archivos de texto ANSI o ASCII. Las listas de palabras son una muestra tanto en orden alfabético y frecuencia. Se pueden guardar para su uso posterior, edición, impresión, copiado en el procesador de textos, o guardarse como archivos de texto.

Veamos a continuación en la siguiente figura una lista de las palabras más frecuentes obtenidas con la herramienta *Wordsmith Tools Wordlist*. Hay un gran desequilibrio de frecuencia de palabras en texto (Sinclair, 1991). De Beaugrande & Dressler (1981) distinguen entre palabras funcionales y palabras de contenido. Las palabras funcionales o gramaticales (artículos, preposiciones, verbos auxiliares, etcétera) serán las más frecuentes. En la lista comprobamos que las palabras más repetidas son los determinantes sin carga semántica. En primer lugar está la preposición “de”, le sigue el artículo femenino “la”, otros tantos artículos y preposiciones que son las palabras que con mayor frecuencia se repiten. La frecuencia de palabras en las lenguas y la notable presencia de algunas de ellas fue objeto de estudio que derivó en la *Ley de Zipf*. Esta ley establece que, dado un corpus de textos de lenguaje natural, la frecuencia de cualquier palabra es inversamente proporcional a su rango en la tabla de frecuencia. Así la palabra más frecuente ocurrirá aproximadamente el doble de veces que la segunda palabra más frecuente, tres veces más que la tercera palabra más frecuente, etcétera. La distribución rango-frecuencia es una relación inversa. Como vemos en la figura 4.6 la primera palabra “de” ocurre 17.565 veces y la segunda “la” ocurre 11.264 veces. No es exactamente la mitad de veces, pero sí aproximadamente. La tercera palabra “que” ocurre 9.241 veces y no es tres veces menos que la primera palabra. Sin embargo, la ley se cumple en líneas generales. En la lista de palabras extraída de nuestro corpus vemos como tras los artículos, preposiciones y determinantes, en la línea número 30 aparece la primera palabra con un significado propio: “España”. Con una

frecuencia de repetición de 730 en los textos recopilados en este grupo y que pertenecen al tercer trimestre de la sección de opinión de El País. Los siguientes escalones los ocupan las palabras *crisis* y *gobierno*, con una frecuencia de 702 y 659, respectivamente.

| N  | Word     | Freq.  | %    | Texts | % Lemmas | Set |
|----|----------|--------|------|-------|----------|-----|
| 1  | DE       | 17,565 | 6.68 | 419   | 100.00   |     |
| 2  | LA       | 11,264 | 4.28 | 418   | 99.76    |     |
| 3  | QUE      | 9,241  | 3.51 | 419   | 100.00   |     |
| 4  | EL       | 7,720  | 2.93 | 411   | 98.09    |     |
| 5  | EN       | 6,709  | 2.55 | 410   | 97.85    |     |
| 6  | Y        | 6,680  | 2.54 | 415   | 99.05    |     |
| 7  | A        | 5,522  | 2.10 | 412   | 98.33    |     |
| 8  | LOS      | 5,471  | 2.08 | 412   | 98.33    |     |
| 9  | DEL      | 3,463  | 1.32 | 382   | 91.17    |     |
| 10 | LAS      | 3,433  | 1.30 | 386   | 92.12    |     |
| 11 | SE       | 3,081  | 1.17 | 399   | 95.23    |     |
| 12 | NO       | 3,035  | 1.15 | 408   | 97.37    |     |
| 13 | UN       | 2,626  | 1.00 | 391   | 93.32    |     |
| 14 | #        | 2,617  | 0.99 | 312   | 74.46    |     |
| 15 | ES       | 2,473  | 0.94 | 381   | 90.93    |     |
| 16 | UNA      | 2,419  | 0.92 | 377   | 89.98    |     |
| 17 | POR      | 2,394  | 0.91 | 394   | 94.03    |     |
| 18 | PARA     | 2,194  | 0.83 | 386   | 92.12    |     |
| 19 | CON      | 2,067  | 0.79 | 379   | 90.45    |     |
| 20 | MÁS      | 1,655  | 0.63 | 357   | 85.20    |     |
| 21 | LO       | 1,540  | 0.59 | 348   | 83.05    |     |
| 22 | AL       | 1,535  | 0.58 | 349   | 83.29    |     |
| 23 | SU       | 1,256  | 0.48 | 320   | 76.37    |     |
| 24 | COMO     | 1,234  | 0.47 | 323   | 77.09    |     |
| 25 | PERO     | 1,015  | 0.39 | 303   | 72.32    |     |
| 26 | HA       | 1,013  | 0.38 | 296   | 70.64    |     |
| 27 | O        | 959    | 0.36 | 286   | 68.26    |     |
| 28 | SI       | 776    | 0.29 | 287   | 68.50    |     |
| 29 | SUS      | 759    | 0.29 | 278   | 66.35    |     |
| 30 | ESPAÑA   | 730    | 0.28 | 226   | 53.94    |     |
| 31 | CRISIS   | 702    | 0.27 | 234   | 55.85    |     |
| 32 | GOBIERNO | 659    | 0.25 | 240   | 57.28    |     |

Figura 4.6: Lista de palabras más frecuentes correspondiente a textos del tercer trimestre 2012

#### 4.3.2 Estudio comparativo de términos del CCF

El análisis cuantitativo del CCF ha consistido en conocer la tasa de frecuencia y otros aspectos estadísticos de los diferentes elementos en los textos periodísticos estudiados. Nos ha proporcionado datos sobre las palabras y las estructuras utilizadas con altas tasas de frecuencia. A continuación, en la Tabla 4.4, mostramos las listas de las 20 palabras de contenido más frecuentes en los textos de la sección de opinión y economía en el primer trimestre de 2012 de los dos periódicos analizados.

Como se puede observar, las palabras más frecuentes en el corpus son, como era de esperar, los términos relacionados con la economía, las instituciones, los países y las personas involucradas en este proceso: *gobierno, España, deuda, economía, crisis, bancos*.

Vale la pena mencionar el hecho de que no encontramos palabras como *ciudadanos, pueblo, población*, u otros términos relacionados con las personas más afectadas por la crisis. En los cuatro períodos analizados se repiten prácticamente las mismas palabras, aunque varían posiciones según avanza el año, lo mismo que la frecuencia. Se da la circunstancia además, por ejemplo, de que en el segundo trimestre hay palabras que vuelven a ocupar la misma posición que en los tres primeros meses del año.

Una parte destacada del corpus está compuesta por formas que aparecen una sola vez. Son los *hápx legomena*. En el CCF hemos localizado algunos de ellos. Vemos a continuación algunos ejemplos extraídos de las listas de palabras de textos publicados del tercer trimestre de 2012. La palabra *Alcoy* es un *hápx legomenon* localizado en los textos de *El Mundo (OEM)*, otro puede ser la palabra *corpus*, que sólo hemos hallado una sola vez en los textos agrupados en el tercer trimestre de *El País (OEP)*, o el caso de *doctor* que aparece también una vez en las publicaciones del tercer trimestre de la sección de economía de *El País*. En las siguientes figuras vemos los casos descritos.

| N      | Word      | Freq. | % | Texts | % Lemmas | Set |
|--------|-----------|-------|---|-------|----------|-----|
| 16,267 | DOBSON    | 1     |   | 1     | 0.13     |     |
| 16,268 | DÓCIL     | 1     |   | 1     | 0.13     |     |
| 16,269 | DOCILIDAD | 1     |   | 1     | 0.13     |     |
| 16,270 | DOCSIS    | 1     |   | 1     | 0.13     |     |
| 16,271 | DOCTOR    | 1     |   | 1     | 0.13     |     |

Figura 4.7: Hápx legomena hallados en El País 3T

|       |            |   |  |   |      |  |
|-------|------------|---|--|---|------|--|
| 8,316 | ALCANZARLO | 1 |  | 1 | 0.26 |  |
| 8,317 | ALCISTA    | 1 |  | 1 | 0.26 |  |
| 8,318 | ALCOHOL    | 1 |  | 1 | 0.26 |  |
| 8,319 | ALCOY      | 1 |  | 1 | 0.26 |  |

Figura 4.8: Hápx legomena hallados en El Mundo Opinión 3T

|        |                |   |  |   |      |  |
|--------|----------------|---|--|---|------|--|
| 12,975 | COREADOS       | 1 |  | 1 | 0.24 |  |
| 12,976 | CORPORACIÓN    | 1 |  | 1 | 0.24 |  |
| 12,977 | CORPORATIVAS   | 1 |  | 1 | 0.24 |  |
| 12,978 | CORPORATIVISTA | 1 |  | 1 | 0.24 |  |
| 12,979 | CORPUS         | 1 |  | 1 | 0.24 |  |

Figura 4.9: Hápx legomena hallados en El País Opinión 3T

A la izquierda vemos el número de palabra que ocupa en el grupo de textos localizado, a continuación, vemos la palabra seguida de la frecuencia, una vez, con que aparece en los textos. Algunas palabras no co-aparecen entre ellas en el corpus. Pero pueden co-ocurrir con aquellas palabras que se presentan en la parte superior de la lista de palabras por orden de frecuencias. Cabe esperar que se encuentre mayoritariamente en co-ocurrencia con ítems frecuentes como *El* o *es* y en este caso decidiremos si dicha co-aparición es casual. Pero puede ser que los ítems más frecuentes co-aparezcan unos con otros y se tendrá también que comprobar si cada uno de esos pares es inusual o no.

| <b>El País (1er. Trimestre)<br/>Posición/Frecuencia</b> | <b>Opinión El Mundo (1er. Trimestre)<br/>Posición/Frecuencia</b> | <b>Opinión El País (1er. Trimestre)<br/>Posición /Frecuencia</b> |
|---|--|--|
| 1. MILLONES (1.827)                                     | 4. MILLONES (377)  | 10. MILLONES (254)   |
| 2. GOBIERNO (1.002)                                     | 1. GOBIERNO (740)  | 1. GOBIERNO (432)  |
| 3. ESPAÑA (937)   | 2. ESPAÑA (414)  | 4. ESPAÑA (327)  |
| 4. EUROS (851)  | 16. EUROS (177)  | 22. EURO. (149)  |
| 5. DEUDA (786)  | 18. DEUDA (162)  | 11. DEUDA (227)  |
| 6. ECONOMÍA (624)                                       | 11. ECONOMÍA (217)   | 3. ECONOMÍA (372)  |
| 7. BANCO (594)  | 36. BANCOS (101)   | 39. BANCO (117)  |
| 8. REFORMA (524)  | 5. REFORMA (367)   | 5. REFORMA (296)   |
| 9. CRISIS (522)   | 9. CRISIS (247)  | 2. CRISIS (400)  |
| 10. DÉFICIT (492)                                       | 6. DÉFICIT (327)   | 6. DÉFICIT (293)   |
| 11. ENTIDADES (446)                                     | 84. ENTIDADES (65)   | 97. ENTIDADES (66)   |
| 12. GRECIA (434)  | 115. GRECIA (52)   | 36. GRECIA (125)   |
| 13. MERCADO (414)                                       | 48. MERCADOS (92)  | 20. MERCADO (158)  |
| 14. SECTOR (384)  | 33. SECTOR (108)   | 70. SECTOR (83)  |
| 15. LABORAL (341)                                       | 8. LABORAL (250)   | 12. LABORAL (220)  |
| 16. EMPLEO (326)  | 20. EMPLEO (148)   | 8. EMPLEO (290)  |
| 17. EMPRESAS (326)                                      | 23. EMPRESAS (146)   | 14. EMPRESAS (194)   |
| 18. MEDIDAS (323)                                       | 24. MEDIDAS (141)  | 31. MEDIDAS (136)  |
| 19. EUROPA (314)  | 29. EUROPA (115)   | 16. EUROPEA (251)  |
| 20. MINISTRO (311)                                      | 30. MINISTRO (115)   | 100. MINISTRO (65)   |

Tabla 4.5: Lista de palabras del primer trimestre.

### 4.3.3 Localización de los actores sociales en el CCF.

Hemos visto que la utilización de la herramienta informática *Wordsmith Tools Wordlist* nos ha provisto de manera automática de listas de palabras de los textos del CCF de los cuatro trimestres de 2012 en que hemos dividido los textos recopilados. A la información sobre frecuencias de palabras más repetidas, que hemos visto en el punto anterior, vamos a añadir la posibilidad de identificar los actores sociales más frecuentes en los textos y sus *collocates*, o lo que es lo mismo, qué personajes salen más a escena durante el período analizado de la crisis financiera, el año 2012, en la prensa española, y asociados a qué términos. Todo ello de nuevo por trimestres y por grupos de textos analizados -EP, OEP, OEM- para extraer sus colocaciones y líneas de concordancia.

¿Cómo localizar los términos que nos interesan para ser investigados? La literatura nos provee de distintos trabajos que nos orientan en la identificación de palabras que resultan específicas en un corpus al compararlo con otro corpus. Recurrimos a autores como Salton (1989) y su test sobre la frecuencia de una palabra y su distribución en diferentes partes del corpus. Church y Hanks (1990) plantean la fórmula de la *Información Mutua* (MI) para describir fenómenos de co-ocurrencia en corpus. Las aportaciones de estos teóricos marcan un punto de inflexión ya que a partir de entonces se ha utilizado también para precisar el vocabulario específico del corpus (Scott, 1997) y (Kilgariff, 2001). Este último (Kilgariff, 2001), en su interés por encontrar un método que hiciera posible la comparación de los corpus experimenta con otros como el test de rango de Mann-Whitney, el test-t, MI, *Log-likelihood*, el test exacto de Fisher y TF-IDF. Previamente Lafon (1980) ya describe la técnica que más adelante utilizarían Lebart y Salem (1994) basada en la distribución hipergeométrica y en las investigaciones más recientes del campo de la terminología computacional. Vemos investigadores que han utilizado la comparación de corpus para capturar términos simples (*uniterms*) y otros tipos de unidades léxicas. En esta línea Ahmad et al. (1994) y Chung (2003) han utilizado cálculos basados en frecuencia normalizada, mientras que Drouin (2003) ha recurrido a la distribución normal para aproximarse a la distribución binomial.

La herramienta informática *Wordsmith Tools Wordlist* de nuevo nos ayuda en la obtención de los listados de los actores sociales. Del repertorio de palabras y del repaso visual identificamos los nombres propios que con mayor frecuencia se repiten. El enfoque cuantitativo al corpus ha

proporcionado información sobre los principales actores o entidades relevantes de la crisis financiera y su contexto, es decir, el entorno lingüístico, en el que aparecen estos personajes (colocaciones y líneas de concordancia).

La tabla que vemos a continuación ilustra los principales actores sociales en los cuatro trimestres de 2012. De las listas de palabras, hemos extraído quiénes son los personajes o actores sociales más frecuentes. Seleccionamos los 12 que aparecían en la parte superior de los listados.

|     | ACTORES SOCIALES 1.T | ACTORES SOCIALES 2. T | ACTORES SOCIALES 3.T | ACTORES SOCIALES 4 T |
|-----|----------------------|-----------------------|----------------------|----------------------|
| 1.  | RAJOY                | RAJOY                 | RAJOY                | RAJOY                |
| 2.  | GUINDOS              | GUINDOS               | DRAGHI               | GUINDOS              |
| 3.  | MONTORO              | MERKEL                | MERKEL               | DRAGHI               |
| 4.  | ZAPATERO             | RATO                  | GUINDOS              | MERKEL               |
| 5.  | MERKEL               | HOLLANDE              | MONTORO              | MONTORO              |
| 6.  | RATO                 | ZAPATERO              | RATO                 | RATO                 |
| 7.  | DRAGHI               | DRAGHI                | ZAPATERO             | OBAMA                |
| 8.  | MONTI                | MONTORO               | HOLLANDE             | ZAPATERO             |
| 9.  | BÁÑEZ                | MONTI                 | WEIDMANN             | LAGARDE              |
| 10. | S. DE SANTAMARIA     | FDZ. ORDOÑEZ          | S. DE SANTAMARÍA     | REHN                 |
| 11. | CAMERON              | GOIRIGOLZARRI         | ALMUNIA              | BÁÑEZ                |
| 12. | OBAMA                | LAGARDE               | GOIRIGOLZARRI        | HOLLANDE             |

Tabla 4.6: Los doce actores sociales que más se repiten en los cuatro trimestres de 2012

Como se puede apreciar, los más repetidos son Rajoy, Guindos, Montoro, -del ámbito español- Dragui, y Merkel, -en el ámbito internacional-, por ejemplo. Son personajes con responsabilidad en el momento del análisis de los textos. Otros personajes aparecen en algún trimestre o no dependiendo de los acontecimientos que se narran. Goirigolzarri está presente, por ejemplo, en el 2º y 3er trimestre, por tratarse de los periodos de la crisis de Bankia.

También podemos ver que los personajes principales de la crisis son los mismos y el rango de posiciones son similares en todo el año. Sin embargo, en el segundo trimestre encontramos más actores europeos en el escenario con contenidos informativos dado que se centran en los acontecimientos que miran a Europa. Con estos doce nombres vamos a seguir trabajando para extraer nuevas conclusiones. Veamos a continuación con qué términos se relacionan.

#### 4.3.3.1 Actores sociales. *Collocates*.

Una vez identificados los actores de la crisis financiera en el CCF conoceremos cuáles son los términos con los que suelen venir asociados.

En el capítulo dedicado a la literatura sobre la lingüística de corpus ya vimos que el término *collocation* procede del inglés. Su origen los atribuimos a Palmer y Hornby (1933) y a Sinclair (1991). Este último autor centró la definición en la tendencia de dos o más palabras a aparecer en la misma línea de concordancia, a combinarse unas junto a otras. Así pues, en la lingüística de corpus conocemos por *collocations* o colocaciones la secuencia de palabras o términos que ocurren de manera coordinada, o co-ocurren, con mayor frecuencia de lo que puede ocurrir por azar. La propuesta de Church y Hanks (1990) sobre el uso de métodos estadísticos de coaparición para identificar colocaciones significativas, supuso una gran ventaja en este tipo de análisis. El método que se sigue es esencialmente el siguiente (el término ‘nodeword’, ‘palabra núcleo’, aquí se refiere a aquella palabra cuyo comportamiento combinatorio es objeto de la investigación):

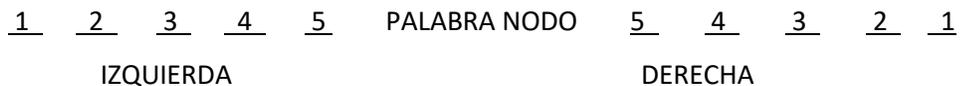
- para cada aparición en el corpus de la palabra núcleo, localiza todas las palabras que aparecen a  $n$  palabras de ésta; llevando cuenta de cada palabra que co-aparece
- para cada palabra que co-aparece, realiza un cálculo estadístico para cuantificar en qué medida la relación entre ellas es significativa
- clasifica las palabras según las estadísticas, mostrando solamente los ítems con los resultados más elevados

La estadística varía dependiendo de cómo se valore y se mida la significación. Esto se realiza averiguándose en qué grado la colocación es improbable, dadas las probabilidades de cada uno de sus componentes. Las probabilidades se estiman en base a frecuencias en el corpus. El reto para el matemático es calcular de forma precisa y comparar las probabilidades, dada la información sobre frecuencias. La fórmula de Información Mútua que Church & Hanks (1990) presentan es la siguiente:

$$MI(x,y) = \log_2(P(x,y)/(P(x).P(y)))$$

En la lingüística de corpus, desde los años 60 del siglo XX, se ha extendido la idea de establecer como horizonte hasta cuatro palabras más allá del nodo para una co-ocurrencia. La naturaleza de esta co-ocurrencia, por sí misma, puede decir mucho de la relación entre dos ítems. Debemos pues considerar qué tipo de relaciones de dependencia se podrían obtener entre ellas. Así, al analizar los resultados del estudio veremos cómo los fenómenos de relación semántica y gramatical, colocación y coligación, presentan direccionalidad, o lo que es lo mismo, la relación entre el nodo y el *collocate* no es simétrica, sino que suele ir hacia un sentido. ¿Qué representa todo esto? Pues, simplemente, que la co-aparición, que a primera vista puede parecer simple, puede, sin embargo, resultar compleja y abre la puerta a la investigación a través de métodos basados en un corpus.

El corpus lingüístico es una plataforma para la observación de patrones colocacionales, visible a través de la pantalla vertical de la palabra nodo y su *collocate*, que es la palabra o palabras con la que se produce co-ocurrencia. En general, se acepta que el comportamiento colocacional de naturaleza léxica se puede manifestar hasta una línea de cuatro o cinco palabras a la izquierda o a la derecha de la palabra nodo. El *Span* es el número de palabras situadas a la derecha y a la izquierda y cuya frecuencia de combinación determina el ordenador. Este sería el esquema:



En la figura 4.10 vemos un ejemplo de líneas de concordancia extraídas del CCF con Rajoy, como palabra nodo de uno de los 12 actores seleccionados. La herramienta *WordSmith Wordlist* nos permite ver a través de listados la presencia de este término y las palabras que lo acompañan inmediatamente antes y después. Este es un ejemplo, pero hemos realizado la misma operación con los 12 actores analizados, para conocer cuáles son los términos, palabras o frases que los acompañan en los textos del CCF a lo largo de los cuatro trimestres y cómo va evolucionando su presencia, según avanza el año analizado. Y esas son las palabras que en las fases posteriores analizaremos de manera manual para extraer el sentimiento que se expresa en los textos vinculado a los actores de la crisis.

| N  | Concordance   | Set | Word # |
|----|---|-----|--------|
| 1  | convirtió en el objetivo a batir. Mariano Rajoy y sobre todo Guindos          |     | 1,246  |
| 2  | del Estado en Bankia, dimitió ante Rajoy. Jugó fuerte a banquero y a          |     | 1,395  |
| 3  | Mario Monti, y el presidente Mariano Rajoy puede ser una buena ocasión        |     | 940    |
| 4  | al presidente del Gobierno, Mariano Rajoy, a "fajarse" y "pelear" con el      |     | 319    |
| 5  | dos euros y medio suplementarios. "Rajoy nos ha perdido perdón, pero me       |     | 176    |
| 6  | porque era la vuelta de las vacaciones. Rajoy extrema el control de la banca  |     | 1,112  |
| 7  | presión sobre los hombros de Mariano Rajoy, que deberá decidir si solicita    |     | 229    |
| 8  | hacia España. El Gobierno de Mariano Rajoy se ve así en la endiablada         |     | 206    |
| 9  | ya existentes, el Gobierno de Rajoy aprobará este viernes —con                |     | 491    |
| 10 | como mañana si el presidente Mariano Rajoy quiere evitarse un castigo         |     | 123    |
| 11 | a Grecia en las últimas semanas. Rajoy presentará mañana el                   |     | 532    |
| 12 | segundo salvavidas. En todo lo demás, Rajoy debería apresurarse en cumplir    |     | 733    |
| 13 | verá obligado el Gobierno de Mariano Rajoy para asegurar el cumplimiento      |     | 203    |
| 14 | respetado: las pensiones. "Parece que Rajoy ahora está dispuesto a hacer lo   |     | 822    |
| 15 | en cambio, reclamó a Báñez y Rajoy que no abandonen la senda                  |     | 755    |
| 16 | pocas diferencias entre las políticas de Rajoy y las de Zapatero. R. Creo que |     | 438    |
| 17 | la actual. Parecía que iba a llegar Rajoy y se iba a solucionar todo, y lo    |     | 471    |
| 18 | Zapatero como con el de Mariano Rajoy, duros ajustes en los últimos           |     | 46     |
| 19 | de cotización. Diciembre de 2011: Rajoy se estrena subiendo los               |     | 520    |
| 20 | el Gobierno presidido por Mariano Rajoy aprobaba un recorte                   |     | 548    |
| 21 | consejo de ministros presidido por Rajoy fueron las siguientes:               |     | 599    |
| 22 | En abril el Ejecutivo de Mariano Rajoy anunció un ajuste adicional de         |     | 1,234  |
| 23 | El 11 de julio de 2012 Mariano Rajoy anunciaba en el Congreso de los          |     | 1,292  |
| 24 | Rajoy va no se atreve a descartar que   |     | 0      |

Figura 4.10: Ejemplo líneas de concordancia de Rajoy en 3T

A fin de poder visualizar mejor la información de las líneas de concordancia de los textos hemos utilizado la función del programa *Concord* de *Wordsmith Tools*. Más adelante veremos cómo las colocaciones extraídas por este procedimiento se transfirieron a una base de datos para ir analizando la información tras el filtrado, revisión y organización manual.

#### 4.3.3.2 Actores sociales: líneas de concordancia

Las concordancias son, según Barnbrook (1996), las listas de palabras de un corpus, ordenadas dentro de un contexto. Son también conocidas como *Key Word In Context* (KWIC) ya que de estas últimas se selecciona la palabra para traer en concordancia. La palabra que se analiza

viene en una columna centrada. Una concordancia es una colección que recoge todas las apariciones de una palabra en un texto o conjunto de textos con un número determinado de caracteres de un co-texto y posterior. Es la lista alfabética de las palabras principales usadas en un corpus, con sus contextos inmediatos.

Se generaron matrices a partir de estas co-ocurrencias en las líneas de concordancia y se transfirieron a una hoja de cálculo. A continuación, cada una de estas matrices se transfirió al programa Ucinet y, por último, las redes se visualizaron con la aplicación Netdraw. A continuación, mostraremos tres ejemplos del análisis realizado con las listas de palabras extraídas del CCF. Pero primero explicamos el método utilizado y sus posibilidades.

A la vista de los resultados obtenidos vemos la viabilidad de extraer y representar la interacción semántica de los principales participantes de grandes eventos sociales en nuestra sociedad mediante el uso de métodos de lingüística de corpus en combinación con las formulaciones de la teoría de redes. Los actores que más se nombran en los textos de nuestro CCF son esos principales participantes y Ucinet es la herramienta que hemos utilizado para la representación de las redes sociales. Es la forma con la que vemos de manera gráfica la conexión entre los individuos u organizaciones y que define la posición que éstos ocupan en la red, los grupos y la estructura global de la red, el conocimiento y los flujos de información dentro de la red y de la red de relaciones que implican la influencia recíproca.

Ejemplos de este tipo de investigación se pueden encontrar en Newman (2001), Molina y Muñoz (2002), Sanz (2003), González et al. (2006). Todo esto es lo que podría llamarse redes sociales de la lengua en la que los individuos o actores no son sólo los miembros de un grupo, sino también los términos relacionados con la crisis financiera y los enlaces son las relaciones entre ellos. Metafóricamente hablando, de la misma manera que en las redes sociales son consideraciones en cuanto al tipo de interacciones entre los individuos: el número de veces que los participantes cumplan con ciertos usuarios en el sistema (otros actores sociales o sus colocaciones) implicará una mayor o menor relación significativa relevante (en nuestro caso, conceptual y densidad semántica).

La siguiente figura, la 4.11, muestra la red del principal actor social en el marco de cooperación, el presidente de España en el año analizado, Mariano Rajoy, y sus colocaciones. Se trata de una red simple, ya que sólo se trata de un solo actor.

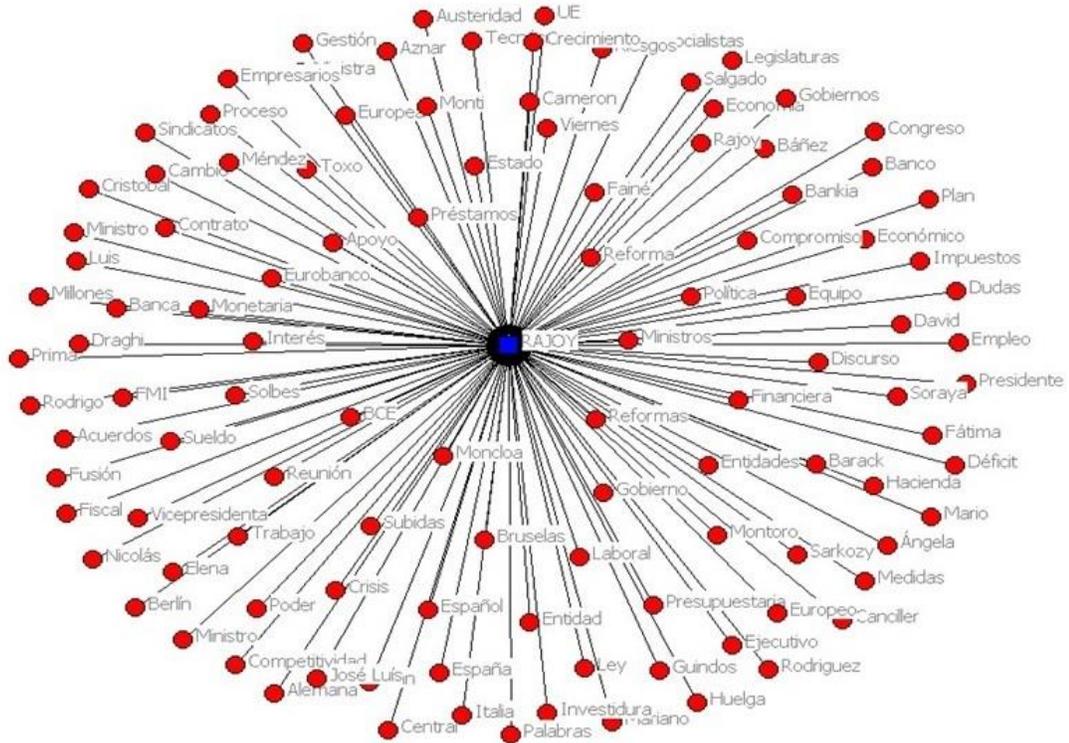


Figura 4.11: Rajoy y sus *collocates*

En los dos primeros trimestres analizamos con estas mismas herramientas informáticas la relación de los 12 actores más nombrados con sus respectivas colocaciones para comprobar qué personajes salían más a escena y asociados con qué términos. Los actores sociales más nombrados del primer trimestre son Rajoy, Guindos, Montoro, Zapatero, Merkel, Rato, Draghi, Monti, Báñez, Saénz de Santamaría, Cameron y Obama. La figura 4.12 muestra la densa relación entre los 12 actores sociales analizados con sus colocaciones.





### 4.3.3.3 Sustantivos abstractos del CCF

La recopilación de textos en el CCF nos ha permitido en primer lugar identificar a los actores sociales para posteriormente comprobar cuáles son los adjetivos que los califican. Esta relación nos ha permitido agruparlos bajo diversas categorías semánticas. Pero, primero veremos cómo se relacionan los actores con los sustantivos abstractos detectados en el CCF y que hemos seleccionado de manera manual. Hay términos que son *per se* evaluativos, que expresan subjetividad y opinión. Sin embargo, los textos a veces contienen un sentido implícito (por ejemplo, en forma de ironía) y enmascaran la naturaleza de evaluación de ciertas palabras o expresiones. En la figura 4.26 veremos un ejemplo de los análisis realizados de manera manual de las líneas de concordancia para poder asignarles etiquetas de ejemplos de lenguaje evaluativo. Hemos extraído también los sustantivos abstractos de las listas de palabras elaboradas a partir de CCF. Los sustantivos abstractos tras un filtrado han ido conformando un documento Excel para ser tratados de manera manual y poder proceder a la asignación de las etiquetas de evaluación. Hemos podido descubrir términos de carácter genérico (gobierno, mercado, empresas) específico (Europa, España, Grecia) cuantitativos (millones) y cualitativos (crisis). Vemos en la siguiente figura una de las listas de sustantivos abstractos del primer trimestre, tras el correspondiente filtrado.

| Word        | Freq. | %      | Texts | %      |
|-------------|-------|--------|-------|--------|
| REFORMA     | 1093  | 0,177  | 348   | 33,656 |
| ECONOMÍA    | 1066  | 0,1726 | 450   | 43,52  |
| DEUDA       | 1023  | 0,1656 | 355   | 34,333 |
| CRISIS      | 1007  | 0,1631 | 452   | 43,714 |
| DÉFICIT     | 973   | 0,1575 | 303   | 29,304 |
| EMPLEO      | 704   | 0,114  | 271   | 26,209 |
| MERCADO     | 594   | 0,0962 | 293   | 28,337 |
| ESTADO      | 589   | 0,0954 | 309   | 29,884 |
| CRECIMIENTO | 570   | 0,0923 | 233   | 22,534 |
| MEDIDAS     | 544   | 0,0881 | 314   | 30,368 |

Figura 4.14: Sustantivos abstractos 1T

El método analítico utilizado con la aplicación *WordSmith Tools* nos ha proporcionado listados de colocaciones con palabras de los que se han obtenido estas listas y que aportan evidencia acerca de categorías semánticas que co-ocurren con las entidades estudiadas, es el caso de:

- Categorías genéricas (Gobierno, Mercado, Empresas)
- Categorías específicas (Europa, España, Grecia)
- Categorías cuantitativas (Millones)
- Categorías cualitativas (de crisis)

Tras el análisis de las palabras que acompañan a los actores seleccionados detectamos sustantivos con una fuerte densidad semántica relacionada con el campo de la economía y, curiosamente, encontramos algunos ejemplos que en principio asociaríamos a un significado positivo, pero en los textos del CCF tienen una connotación negativa. Vemos algunos ejemplos extraídos de los textos recopilados:

- El término *rescate* puede darnos una idea de positividad, de que algo o alguien ha sido liberado, salido de un cautiverio, de un secuestro. Sin embargo, en el corpus periodístico hemos analizado la frase *rescate para España*, como sinónimo de una inversión económica. Dentro del entorno de la crisis financiera y económica tiene, por el contrario, connotaciones negativas ya que supone el sometimiento a unas duras condiciones de recuperación.
- Otro ejemplo es la palabra *reforma*. Si hablamos de la reforma, pensamos en una actualización para mejorar las cosas. Reformamos nuestra casa para darle nuevo aire, por ejemplo. Pero si hablamos en términos de economía y hablamos de *reforma laboral* nos da una idea de algo poco agradable y que va a suponer reducir derechos o poder adquisitivo.

Los métodos de análisis estadísticos han generado ejemplos que muestran que en los textos que el periodista o autor ha hecho uso de estructuras lingüísticas que van más allá de la selección de los elementos léxicos individuales, es decir, palabras individuales. Pero, hay ocasiones en las que el autor utiliza unidades semánticas, con gran riqueza expresiva, que requieren mayor análisis e interpretación independientemente de las herramientas de software utilizadas en las primeras etapas del estudio. Por lo tanto, en la siguiente etapa de la investigación abundaremos en el análisis cualitativo y la interpretación manual con el fin de proponer una categorización de los sentimientos.

#### 4.3.3.4 Actores sociales: categorización semántica de adjetivos

Para establecer la relación de los actores sociales y sus adjetivos hemos tomado listados extraídos de los cuatro trimestres del CCF. Este proceso de selección de los adjetivos que acompañan a cada actor social se ha realizado con una combinación de extracción de términos de las listas de palabras mediante la herramienta informática *Wordsmith Tools* y la posterior y distribución en documentos Excel como el que vemos a continuación, en la figura 4.24 y que pertenece a los adjetivos que acompañan a los actores sociales en los textos recopilados en el primer trimestre de 2012.

| 1  | Adjetivos       | RAJOY | GUINDOS | MONTORO | ZAPATERO | MERKEL | RATO | DRAGUI | MONTI | BÁÑEZ | S. DE SAN CAMERO | OBAMA |
|----|-----------------|-------|---------|---------|----------|--------|------|--------|-------|-------|------------------|-------|
| 2  | ABANDONADAS     |       |         | 1       |          |        |      |        |       |       |                  |       |
| 3  | ABIERTA         |       |         |         | 1        |        |      |        |       |       |                  |       |
| 4  | ABORDADO        |       |         |         |          |        |      |        |       |       |                  |       |
| 5  | ABSOLUTA        | 1     |         |         | 1        |        |      |        |       |       |                  |       |
| 6  | ABSURDO         |       |         |         | 1        |        |      |        |       |       |                  |       |
| 7  | ACABADA         |       |         |         |          |        |      |        |       |       | 1                |       |
| 8  | ACELERADO       | 1     |         |         |          |        |      |        |       |       |                  |       |
| 9  | ACEPTADO        | 1     |         |         |          |        |      |        |       |       |                  |       |
| 10 | ACORDADA        |       |         |         | 1        |        |      |        |       |       |                  |       |
| 11 | ACTIVOS         |       | 1       |         |          |        |      |        |       |       |                  |       |
| 12 | ACTUAL          |       |         |         |          |        | 1    |        |       | 1     | 1                |       |
| 13 | ACUÑADA         |       |         |         |          | 1      |      |        |       |       |                  |       |
| 14 | ADELANTADO      |       | 1       |         |          |        |      |        |       |       |                  |       |
| 15 | ADICIONAL       |       |         |         |          |        |      |        |       |       |                  |       |
| 16 | ADMITIDO        |       |         |         |          |        |      |        |       |       |                  | 1     |
| 17 | ADQUIRIDOS      | 1     |         |         |          |        |      |        |       |       |                  |       |
| 18 | ADVERTIDO       |       | 1       |         | 1        |        |      |        |       |       |                  |       |
| 19 | AGRADECIDO      | 1     |         |         |          |        |      |        |       |       |                  |       |
| 20 | AGRESIVA [*]    | 1     |         | 1       |          | 1      |      |        |       |       |                  |       |
| 21 | AHORRADORES     |       |         |         |          | 1      |      |        |       |       |                  |       |
| 22 | AJENOS          |       |         |         | 1        |        |      |        |       |       |                  |       |
| 23 | ALABADO         | 2     |         |         |          |        |      |        |       |       |                  |       |
| 24 | ALEMÁN [*]      | 1     |         |         | 1        |        | 15   |        |       | 1     |                  |       |
| 25 | ALENTADO [*]    |       |         |         |          | 1      | 2    |        |       |       |                  |       |
| 26 | ALTA [*]        |       | 1       |         | 1        | 2      |      |        |       | 1     |                  |       |
| 27 | AMBICIOSAS      | 1     |         |         |          |        |      |        |       |       |                  |       |
| 28 | AMENAZADORA     | 1     |         |         |          |        |      |        |       |       |                  |       |
| 29 | AMPLIADO [*]    |       |         |         |          | 1      |      |        |       |       |                  |       |
| 30 | ANTERIOR [*]    | 1     |         | 2       |          | 2      |      |        |       |       |                  |       |
| 31 | ANTICIPADO      |       | 1       |         |          |        |      |        |       |       |                  |       |
| 32 | ANTIDEMOCRÁTICO |       |         |         |          |        |      |        |       | 1     |                  |       |
| 33 | ANTIGUO         |       | 1       |         |          |        |      |        |       |       |                  |       |
| 34 | ANTROPOLÓGICO   |       |         |         |          | 1      |      |        |       |       |                  |       |
| 35 | ANUNCIADA [*]   | 1     |         | 1       | 3        | 1      |      | 1      |       |       |                  |       |
| 36 | AÑADIDO [*]     |       | 2       |         | 1        |        |      |        |       |       | 1                | 1     |
| 37 | APADRINADA      | 1     |         |         |          |        |      |        |       |       |                  |       |
| 38 | APLICADA [*]    | 2     |         |         |          |        |      |        |       |       |                  |       |
| 39 | APOSTADO        |       |         |         |          |        |      |        |       |       |                  |       |
| 40 | APROBADA [*]    |       |         | 1       | 1        |        |      |        |       |       |                  | 1     |
| 41 | APUNTADO        |       |         | 1       |          |        |      |        |       |       |                  |       |
| 42 | AQUILATADOS     |       |         |         |          |        |      |        |       |       | 1                |       |

Figura 4.15: Documento sobre la relación de los actores sociales y sus adjetivos en el 1er trimestre

Los resultados obtenidos en este análisis en el primer y segundo trimestre es el siguiente. La lista del primer trimestre es de 473 adjetivos ordenados por orden alfabético. En la figura 4.11 visualizamos las primeras 42 posiciones con el número de veces que se repite con cada uno de los actores. Cuando adjetivo aparece en masculino o femenino se advierte con un asterisco. Lo

vemos en esta figura en los ejemplos de *agresiva, alemán, alentado, anunciada, añadido o aplicada*.

En el segundo trimestre los 12 actores sociales analizados están relacionados con 614 adjetivos según los textos recopilados en nuestro CCF. En el tercer trimestre el número de adjetivos relacionados con los actores sociales es de 771 y en el cuarto trimestre son 495 el número de adjetivos los que se relacionan.

Los 2.353 adjetivos que hemos contabilizado junto a los actores sociales en los cuatro trimestres del CCF los hemos agrupado, a continuación, en diversas categorías semánticas adaptadas de la Teoría de la Valoración. Es el modelo que conocemos como *Appraisal Theory* (Martin & White, 2005). De acuerdo con los principios expuestos por estos autores hemos clasificado sentimientos diferentes, o lo que es lo mismo, los juicios, las emociones y actitudes, tanto positivas (*notable, efectiva, competitivas*) como negativas (*preocupante, fracasado*).

¿Cuál ha sido el método utilizado para la obtención de la categorización semántica de los adjetivos que acompañan a los actores sociales? Una vez más los adjetivos se han filtrado de forma manual mediante los siguientes pasos:

- Colocación de cada actor social, ya sea adjetivos con connotación positiva (*notable, efectiva, competitivas*) o de sentido negativo (*preocupante, fracasado*).
- Agrupación en tres categorías principales: actitud, compromiso y graduación (Martin, 2000).

Cada uno de estos sistemas incluye un subconjunto de valores por el cual los autores de los textos expresan y negocian su posición ideológica. Destacamos los resultados encontrados para el término juicio, que evalúa los individuos o grupos humanos de acuerdo con las convenciones sociales, es decir, lo que está moral o legalmente aceptado por las normas y reglamentos en instituciones (la normalidad, la capacidad, la tenacidad, la propiedad, veracidad). Resaltamos adjetivos que acompañan a nuestros actores sociales expresan la forma en que se evalúan en términos de honestidad, integridad y comportamiento ético: *falso, hipócrita, laxo, rencoroso, incapaz, inútil, mediocre, sincero, sabio, solvente*.

Según esta teoría, en los textos del CCF hemos detectado emociones y actitudes, tanto positivas (*notable, efectiva, competitivas*) como negativas (*preocupante, fracasado*). Veamos a continuación ejemplos de los adjetivos extraídos en el CCF para estas categorías y sus subcategorías. Se trata de una aproximación semántica a nuestros textos que nos ha proporcionado un marco con el que se han podido contemplar las opiniones, actitudes o emociones de los periodistas como categorías o *systems of meanings* (Martin y White, 2005: 42).

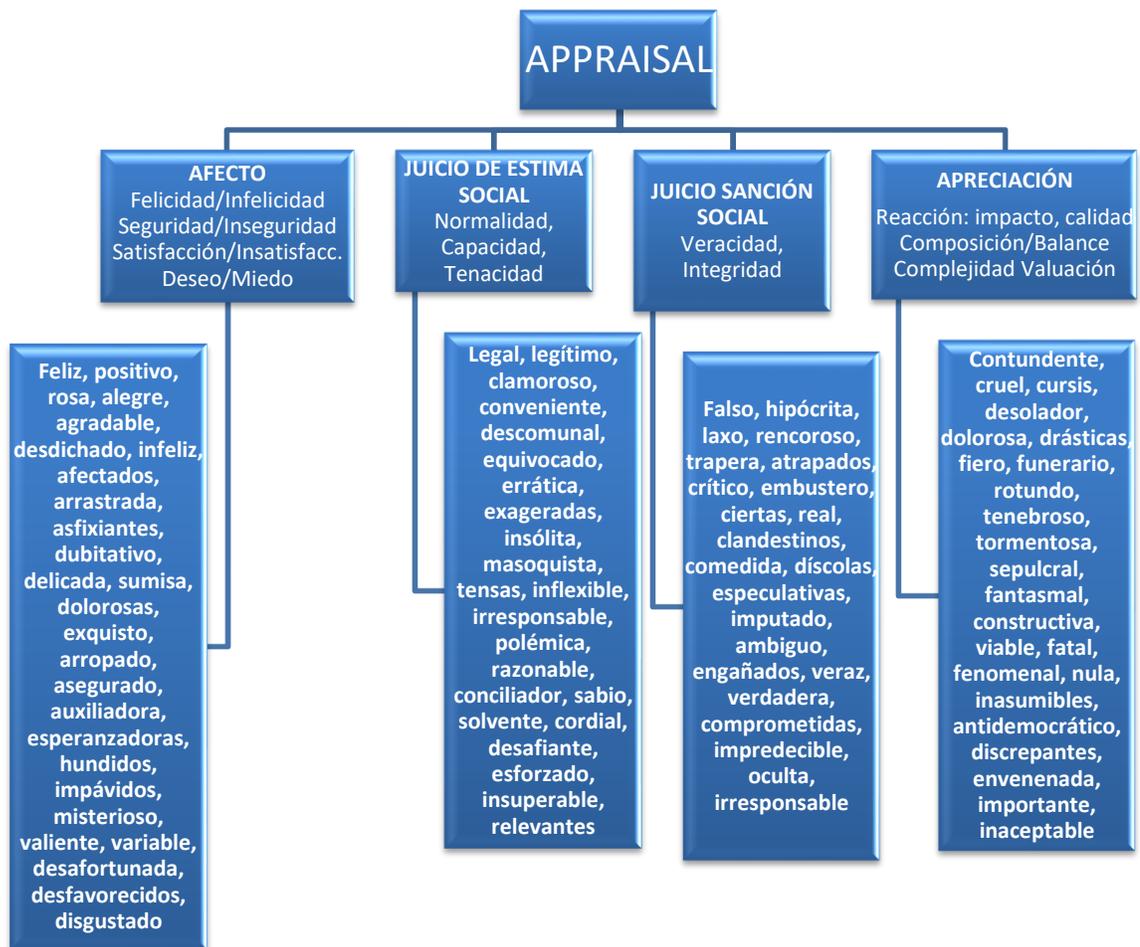


Figura 4.16: Esquema categorización semántica de los adjetivos, según *Appraisal Theory*

Hunston (2014:12), en este sentido, expresa: “collocation, concordance lines and wider context explicate the meaning, cultural or ideological significance and function of the words and phrases so identified”.

En esta etapa de la investigación nos dedicamos a profundizar en la interpretación cualitativa de los sentimientos que los autores de los artículos expresan mediante diferentes recursos. Para este estudio se tomó una muestra de las líneas de concordancias obtenidas a partir de los doce actores principales. De cada una de estas líneas se ha explorado el mensaje que el autor quiere transmitir acerca de la actuación, en cada momento, del personaje en cuestión. El procedimiento que se siguió fue anotar, recopilar y clasificar los valores o emociones expresados con el objetivo de establecer una categorización de sentimientos. Los ejemplos seleccionados para cada actor contienen información que podría pasar desapercibida cuando se analizan los términos corpus de forma individual. Por el contrario, el estudio de los ejemplos de mayor longitud de texto revela una serie de ideas que se expresan de manera explícita o implícitamente y que el lector podrá percibir de manera inconsciente. El seguimiento de esta pauta motivó el análisis de las líneas de concordancia de los 12 actores sociales por trimestres del 2012. Centramos el trabajo en descubrir:

- ¿Quién realiza la evaluación?
- ¿Qué está realmente expresando el periodista o autor del texto en su mensaje mediante sus palabras o a través de los testimonios de los protagonistas del CCF o de las instituciones involucradas?

La finalidad es desggranar quién opina, qué opina, cómo valora el autor los acontecimientos que se presentan o las actuaciones o palabras de los personajes de la noticia. Sirvan los siguientes ejemplos extraídos del CCF para ilustrar esta idea:

- *Guindos avisa de que la recesión económica en España se agrava (¿Qué dice Guindos?)*
- *Ali Babá y los 659 evasores españoles puso la lista en manos de Cristóbal Montoro; si no trinca a los evasores es porque no le dejan (¿Quién dice esto? ¿Cómo se comporta Montoro?)*
- *Las altísimas expectativas que Mario Draghi había generado la semana pasada se han disuelto como un azucarillo (¿Qué piensa el autor acerca de Draghi?)*

Para llegar a este punto primero hemos asignado etiquetas. Frases como *Rajoy aprobaba un recorte presupuestario histórico y subía los impuestos*, contiene algunas de las etiquetas con las que hemos trabajado el CCF de manera manual. En este ejemplo vemos la evaluación negativa

del autor, que utiliza antítesis como *recorte presupuestario y subir los impuestos* con condena social.

En la figura 4.25 mostramos un ejemplo de las extraídas a uno de los 12 actores sociales del CCF, Montoro, en textos del primer trimestre de 2012.

The image shows a concordance tool interface. On the left, a text snippet is displayed with a yellow highlight. On the right, a list of 12 extracted labels is shown, each with a corresponding evaluation. The labels are:

- Comentario [C1]: Metáfora "siembra dudas". Evaluación negativa
- Comentario [C2]: Evaluación positiva
- Comentario [C3]: Evaluación positiva/negativa. Metáfora. Síndrome entra en el campo semántico de enfermedad, lo mismo que contagio. Así que también es una ironía, se piensa lo contrario de lo que se expresa
- Comentario [C4]: Evaluación negativa
- Comentario [C5]: Evaluación
- Comentario [C6]: Evaluación
- Comentario [C7]: Figura retórica, hipérbole
- Comentario [C8]: Evaluación negativa y metáfora
- Comentario [C9]: Evaluación positiva, metáfora
- Comentario [C10]: Evaluación positiva, metáfora
- Comentario [C11]: Evaluación
- Comentario [C12]: Evaluación positiva y metáfora

Figura 4.17: Extracción de etiquetas a través de las líneas de concordancia de Montoro

Vamos a analizar algunos ejemplos que hemos localizado en esta figura de las líneas de concordancia de Montoro:

- *Montoro siembra dudas sobre la capacidad para cumplir el déficit* (Metáfora: *siembra dudas*. Evaluación negativa del autor)
- *El síndrome Merkel parece haber contagiado a Montoro* (Metáfora: Merkel entendida como una enfermedad. Ironía. Evaluación negativa del autor)
- *Montoro tiene ahora un Stringer Bell aporreando calculadoras* (Figura retórica: hipérbole. Evaluación negativa)
- *Ahora Montoro anuncia poco menos que un revival de la Transición para cuadrar números* (Ironía, evaluación negativa)

- *Montoro lo hará bien. De las Cortes de Cádiz viene ese galgo* (Figura retórica, hipérbole, ironía, evaluación negativa)
- *Montoro ya no es una servilleta arrugada donde alguien ha pintarrajeado un 8* (Metáfora, evaluación negativa)
- *Para Cristóbal Montoro probablemente interpreten el Nessum dorma* (nadie duerma) (Perífrasis, intertextualidad, evaluación negativa)
- *El ministro de Hacienda, Cristóbal Montoro, ha evitado presionar más a las anémicas comunidades* (Metáfora, evaluación positiva)
- *Montoro está intentando escribir derecho con los renglones torcidos y eso no tiene mérito.* (Intertextualidad, metáfora, evaluación negativa)
- *Como ha expuesto el dibujante Montoro, en una viñeta de Mingote* (Ironía, intertextualidad y evaluación negativa)
- *Al acabar el telediario debería salir Montoro con un palo frente a un mapa señalando los déficits* (Ironía, Evaluación negativa)

El siguiente paso será concretar las etiquetas de lenguaje evaluativo que se han obtenido como más frecuentes tras el análisis manual de los textos.

#### **4.4 Análisis manual: etiquetas de lenguaje evaluativo**

La localización de los sustantivos abstractos inicial y la posterior categorización semántica de los adjetivos para su análisis manual, ha dado paso a la asignación de etiquetas de evaluación.

Las líneas de concordancia analizadas han permitido extraer el mensaje del autor sobre lo que dijeron los actores de la crisis, cómo se comportaban o, incluso, el concepto o la opinión que el autor de los textos tiene de ellos en el momento de escribir. El método adoptado fue identificar, recoger y ordenar los valores o las emociones contenidas en los textos.

Hemos identificado distintos matices en el mensaje, que se traduce en una serie de etiquetas que reflejan significados que pueden estar presentes en el texto, ya sea en forma de varias realizaciones lingüísticas o que el lector debe interpretar o descodificar utilizando su conocimiento previo. Las etiquetas semánticas que proponemos son las siguientes:

- Evaluación: En esta categoría se incluye la opinión positiva / negativa que el autor quiere expresar.
- Los recursos retóricos: A través de esta categoría se identifican los recursos estilísticos utilizados por los periodistas. Los más comunes son: la ironía, la antítesis, metáfora, el contraste, la comparación, la perífrasis, juego de palabras, la paradoja, pregunta retórica, la hipérbole, la alegoría y la metonimia.
- Citas textuales: El periodista reproduce en ellos las palabras exactas pronunciadas por los personajes u otros agentes de prensa con el objetivo de distanciarse de las opiniones expresadas por ellos.
- Sentimiento: En esta categoría se estableció si el autor o aplaude condena las palabras o comportamiento de un actor social.
- Referencia: El periodista interactúa con el lector usando expresiones culturalmente conocidos (conocimiento cultural compartido). Estos pueden ser dichos o referencias familiares para el lector que provienen del cine, y los campos literarios o deportivos, por ejemplo.
- Neologismo: Es una técnica muy común en el que el autor introduce términos de su propia creación, una adaptación de otros idiomas, o combina elementos comúnmente utilizados con otro contenido semántico, para que el texto sea más elocuente y así atraer la atención del interlocutor.
- Adjetivos calificativos: Se recogieron bajo esta categoría adjetivos con una alta densidad semántica con los que el autor quiere representar una acción o comportamiento.

A medida que se fueron analizando los diferentes actores y su entorno lingüístico, se fueron descubriendo diversas realizaciones lingüísticas contenidas en el mensaje, que fuimos concretando con las etiquetas descritas. Se recopilaron aquellos recursos que aparecían de forma repetida y los categorizamos. Los siguientes ejemplos extraídos del CCF ilustran esta idea:

- *Después de regresar triunfante de Bruselas, en donde habló por señas, Rajoy recibe una puñalada trapera; tras resignarse al 5,3%. (Ironía)*
- *(...) pueden resultar más nocivas que las ocurrencias zapatéticas. Es verdad que Rajoy en solitario lleva unas semanas sentado en la silla curul y que es necesario (...).*

(Neologismo: inventado la palabra - Zapatero + patético; intertextualidad: referencia al Imperio Romano)

- *Las previsiones de Fátima Báñez y Cristóbal Montoro han saltado por los aires.*  
(Metáfora)
- *Que cada palo aguante su vela, señaló el ministro Cristóbal Montoro a finales de mayo.*  
(Referencia: refrán)
- *Montoro tendrá sus hombres de negro.* (Intertextualidad: referencia a películas)
- *Valeriano Gómez una Montoro: "Usted es el ministro Más irresponsable de España"*  
(Entrecomillado)

En este punto recurrimos al trabajo de Tognini-Bonelli (2001: 84-85), quien introdujo la distinción entre *corpus-based and corpus-driven language study*. Los estudios basados en corpus (*corpus-based*) suelen utilizar datos del corpus para explorar una teoría o hipótesis, con el objetivo de validarla, refutarla o refinarla. *Corpus-driven language study* sostiene que el propio corpus debería ser la única fuente de nuestras hipótesis sobre el lenguaje. El enfoque *corpus-driven* es una metodología mediante la cual el corpus sirve como una base empírica a partir de la cual los lingüistas extraen sus datos y detectan fenómenos lingüísticos sin premisas y expectativas previas (Tognini-Bonelli, 2001).

Los ejemplos recogidos por este método se transfirieron entonces a una hoja de cálculo con las diferentes categorías semánticas de evaluación que vemos a continuación. Como se puede apreciar en los ejemplos de cada actor, cada una de las instancias contiene información que resultaría inapreciable al analizar términos del corpus de forma individual. Por el contrario, el estudio de líneas de concordancia nos revela una serie de ideas que el autor quiere transmitir en su mensaje, cargadas de una elevada densidad semántica y que el lector va captando de manera inconsciente.

| Ejemplo   | Evaluación | Figura retórica | Entrecomillado  | Sentimiento | Referencia | Neologismo   | Calificativo              |
|---|------------|-----------------|-----------------|-------------|------------|--------------|---------------------------|
| Báñez, Primera ministra de Trabajo del rajoyato                             | Negat.     |                 |                 | Condena     | Histórica  | Parasíntesis | primera ministra          |
| “El plan de Cameron da confianza; yo haría algo similar en España”          | Posit.     | Comparación     | Distanciamiento | Aplauso     |            |              |                           |
| Draghi conectó la manguera de préstamos a tres años                         | Negat.     | Metáfora        |                 | Condena     |            |              |                           |
| Guindos presenta su plan económico. 40 días y 40 noches ha tardado          | Negat.     | Hipérbole       |                 | Condena     | Bíblica    |              |                           |
| actitud beligerante de la canciller alemana Merkel                          | Posit.     |                 |                 | Aplauso     | Bélica     |              | beligerante               |
| el experimento de Monti va a seguirse con gran expectación                  | Negat.     | Metáfora        |                 | Condena     | Histórica  |              |                           |
| Para Montoro probablemente interpreten el Nessun dorma                      | Negat.     | Ironía          |                 | Condena     | Musical    |              |                           |
| Obama ve las reformas en España “un paso positivo”                          | Posit.     |                 | Cita textual    | Aplauso     | Histórica  |              | positivo                  |
| Ahora Rajoy se enfrenta a un momento decisivo                               | Posit.     |                 |                 | Aplauso     |            |              | decisivo                  |
| Rato ha demostrado ser un buen gestor y ahora le toca lidiar                | Posit.     | Metáfora        |                 | Aplauso     | Taurina    |              | buen                      |
| Santamaría es rigurosa, aplicada y seria                                    | Posit.     |                 |                 | Aplauso     |            |              | rigurosa, aplicada, seria |
| Lo malo de Zapatero no es malo porque lo haya hecho él sino porque era malo | Negat.     | Repetición      |                 | Condena     |            |              | malo                      |

Tabla 4.7: Etiquetas con ejemplos

Mediante esta metodología de análisis de las opiniones contenidas en el contexto de los protagonistas del CCF hemos podido agrupar los diferentes recursos empleados con sus correspondientes ejemplos de uso (por actor/trimestre). Ello nos ha proporcionado la base para las siguientes fases de la investigación, en que se ha contemplado el concepto de evaluación tomando como objeto unidades de mayor o menor complejidad y diferentes niveles textuales.

#### 4.4.1 Evaluación explícita a nivel de palabra y frase

Los ejemplos de términos y expresiones recopilados del CCF, tanto adjetivos, como verbos y sustantivos tienen un valor positivo. Retomamos el análisis de algunas líneas de concordancia de nuestro corpus para ver algunos ejemplos:

- *Los mercados han respaldado la gestión de Monti mucho más que la de Rajoy (positivo)*
- *El sabio Cristóbal Montoro no ha meditado sobre el punto más vidrioso de su propuesta (positivo)*
- *Guindos tiene hora con Mario Draghi como quien la tiene con un médico (neutral)*
- *Buena jugadora, Ángela Merkel consigue vendernos las migajas (positivo)*
- *Cristobal Montoro le dio una lección acerca de cómo se miente en público (condena)*
- *Merkel ha tenido la gentileza de aderezar un pastel (aplauzo)*

Los términos y expresiones subrayados (adjetivos, verbos, sustantivos): *respaldado, sabio, buena, consigue, le dio una lección, gentileza, aderezar*, tienen un valor positivo. Sin embargo, como se puede observar, el mensaje en todos ellos es negativo y está camuflado mediante la ironía. Del mismo modo, lo que a priori podríamos considerar una afirmación neutral, como es el caso del tercer ejemplo y a la vista de que, estadísticamente, ninguno de los términos de la oración entraría en la polaridad positivo/negativo, la interpretación manual nos revela que las connotaciones son negativas al hacer referencia a la reunión que se va a celebrar entre los dos dirigentes para tratar el grave problema financiero de España. Draghi es considerado metafóricamente como el especialista que va a diagnosticar y tratar de proponer tratamiento a la enfermedad de los españoles. Vemos en el CCF ejemplos similares con la figura de Draghi como “sanado”, pero con matices negativos por su contexto lingüístico. Es el caso de *dottore, píldora, operación de drenaje de liquidez*.

- *el dottore Draghi, actuando de la peor manera posible*
- *una píldora que Draghi ve absolutamente necesaria*
- *la institución presidida por Mario Draghi realizará mañana una operación de drenaje de liquidez*

Del mismo modo que hemos visto con los adjetivos, los sustantivos abstractos más frecuentes en el CCF, términos con una elevada densidad semántica, adquieren en los ejemplos que vemos a continuación connotaciones opuestas a su valor positivo. Es el caso de *medidas*, *liquidez*, *crecimiento* y *rescate*.

- *Las medidas de Hollande -sobre todo, las fiscales- están dificultando el débil crecimiento*
- *Draghi de momento descarta reabrir la barra libre a tres años de liquidez*
- *Hasta Draghi es consciente de la miopía fundamentalista y habla de introducir crecimiento*
- *Draghi debe actuar con otro manguerazo de liquidez*
- *Guindos advierte: “Si se pide el rescate, no hay marcha atrás”*
- *Rajoy no logra frenar un lunes negro en los mercados tras el rescate de Bankia*
- *Almunia advierte de que la incertidumbre sobre el rescate aumenta los riesgos*

En el punto 4.3.3 cuando comenzamos a abordar la selección de los sustantivos abstractos ya avanzamos lo que ahora vemos con ejemplos de manera más clara; que existen términos que son *per se* evaluativos y pueden aparecer en el texto con un valor totalmente contrario al que cabría esperar. Expresiones como *barra libre* o *manguerazo*, de naturaleza positiva, esconden, sin embargo, un mensaje negativo del autor: España está al borde de la quiebra. De este modo podemos afirmar que para poder apreciar el sentido real de determinadas palabras de naturaleza evaluativa cabe analizar unidades mayores de significación para captar el verdadero mensaje real.

Existen otros casos en que los autores de los textos que hemos recopilado en el CCF recurren a coloquialismos y neologismos, es decir, palabras y expresiones cercanas al lector y

comprensibles para el hombre de la calle, procedentes de otras lenguas e, incluso, de nueva creación por parte del autor. Mediante estos recursos se capta la atención del público al encontrarse con términos que utiliza en su conversación cotidiana. Se sitúa de este modo el escritor-lector en un mismo nivel (*dicho en más suave, rojería, racaneo, cabezas de chorlito, puñalada trapera, adefesio garlero, requetehistórica cumbre*). Este mismo efecto se consigue con palabras como *bashing, mon dieu, tempus fugit*, que quizá desconoce por tratarse de extranjerismos. De manera similar en el CCF encontramos “nuevos” términos inventados por el escritor, normalmente tomado como base palabras relacionadas con los políticos a los que se añaden prefijos y/o sufijos (*zapatético, zapateresco, peperero, rajoyato, Merkiavelli*). El discurso cobra un ritmo desenfrenado y dinámico, cargado de gran expresividad. Además de la evaluación explícita a través de las categorías y recursos anteriormente mencionados, determinadas estructuras gramaticales constituyen otra forma de manifestar la valoración personal del autor sobre el mensaje del texto, tales como que conocemos como *attitude markers*. En el CCF ejemplos de modales deónticos, adverbios actitudinales y la estructura ser+ adjetivo valorativo /comparativo, siguiendo el análisis de Dafouz (2006, 2008).

- debe acometer sin demora los deberes que ya tenía impuestos
- *una incógnita que debe despejarse*
- *tendríamos que volver a la sencillez aforística*
- *deberían dar ejemplo renunciando a sus emolumentos y prebendas*
- *se han aprovechado de nuestro dinero impunemente*
- *el sistema está perversamente diseñado*
- *La cumbre ha sido especialmente positiva para España*
- *Las comunidades son las responsables del mayor desvío en el déficit*
- *Las dificultades de liquidez son una auténtica oportunidad para imponer duras condiciones*
- *La responsabilidad de las autonomías en el desvío del déficit es dramática e incontestable*
- *Pero el problema es mucho más profundo que el incumplimiento de 2011*
- *el modelo está sesgado a favor de la Administración central*
- *Es una herencia compartida y no producto de la obsesión gastadora del rojería*
- *el problema es mucho más profundo*

#### 4.4.2 Evaluación implícita en el CCF

El carácter persuasivo del discurso periodístico lleva al escritor a camuflar tras su discurso el mensaje real. El autor acude a diferentes recursos estilístico para embellecer su lenguaje y fuerza al lector a descodificar e interpretar el contenido de su artículo (a valorar un acontecimiento o institución o las palabras pronunciadas por un personaje público) y establecer sus propias conclusiones. El objetivo prioritario será, en todos los casos, convencer a su interlocutor de la verdad de sus palabras.

Un recurso frecuente en el CCF es la ironía, mediante el cual el autor trata de ridiculizar un personaje o una situación. Con la ironía dice todo lo contrario a lo que realmente quiere expresar. Vemos algunos ejemplos:

- *merece comentario aparte Fátima Báñez, primera ministra de Trabajo del rajoyato*
- *Draghi hombre prudentísimo al que hemos visto conduciendo sin cinturón*
- *Para Cristóbal Montoro probablemente interpreten el Nessun dorma*
- *Ayer Rajoy gobernó. Bien, mal o regular, pero tomó una decisión importante*
- *el optimismo panglossiano de Zapatero*
- *se escandalizaba porque Zapatero se atrevió a subir el IVA de «los chuches»*
- *Hollande, que tiene pinta de profesor de guardería venido a menos*
- *Buena jugadora, Angela Merkel consigue vendernos las miqajas*
- *En manos de la muchacha, apodo con el que Kohl se refería por aquel entonces a Merkel, su antigua protegida*

Otra de la figura muy repetida es la metáfora. Muchas veces la ironía se presenta en el texto mediante este recurso (*diabólico tripartito, el clan de “los mejores”, el síndrome Merkel parece haber contagiado a Montoro*). El análisis del CCF ha demostrado que el periodista se vale con mucha frecuencia de la metáfora para crear imágenes sobre la realidad con las que expresa un mayor número de significados con menos palabras. El lenguaje figurado ayuda al lector a interpretar el discurso y, a partir de las valoraciones emitidas implícitamente, sumarse a ellas o cuestionarlas. En este sentido, son una herramienta potente para la creación y propagación de ideologías.

Las metáforas que hemos detectado suelen hacer referencia a deportes, anuncios publicitarios, la religión, el cine, hechos históricos y bélicos, la ciencia o la literatura, por ejemplo. A menudo, son referencias que nos resultan familiares por ser expresiones del léxico común. Vemos algunos ejemplos:

- *Guindos echó balones fuera: “Tenemos que ver cuál es la inflación de noviembre”*
- *Guindos asegura que “las reformas empiezan a dar sus frutos”*
- *Zapatero esparció las siembras de Caín sobre los surcos de Cataluña*
- *casi todos los acuerdos concedidos por Angela Merkel a François Hollande en la requetehistórica cumbre de junio, los cañones están apuntados de nuevo a París*
- *Esto sí que era un algoritmo de altos vuelos. Guindos se ha convertido así en alguien como el doctor Hoffmann*
- *Merkel, a la que llama Merkiavelli y compara con el Príncipe*
- *a Rato lo arrojaron por la borda*

El juego dialéctico escritor-lector, en palabras de Thompson y Thetela (1995) “*the reader-in-the-text*”, mediante el cual de forma inconsciente ambos intercambian opiniones y se evalúan los contenidos se consigue también con la inclusión de la audiencia en el propio texto, bien dirigiendo su discurso hacia ella con imperativos o vocativos (*miren, imaginen, ya ven, comparen, cabe preguntarse, no debe olvidarse que*), plurales sociativos, con los que se sitúa afectivamente próximo al lector (*nos ha hecho esclavos, nos causan, que le hemos dado, nos lleva, nos ha conducido*) y con el uso de preguntas retóricas. En este caso, se plantea un interrogante para que reflexione sobre la cuestión y siga leyendo el texto, donde, frecuentemente, encontrará la respuesta o él mismo tendrá que contestar.

- *¿Logrará Merkel domesticar la animalidad de los mercados?*
- *¿a qué precio valorará el Mecanismo de Estabilidad el capital de Bankia y del resto de las entidades nacionalizadas?*
- *¿Cómo va a recuperar lo que no ha perdido? O mentía en la primera parte de la proposición o mentía en la segunda, aunque, conociéndole, quizá mintiera en las dos.*
- *(...) ¿Es razonable? En parte, pues es el Estado quien dispensa los grandes estímulos y el que pecha con las dos grandes partidas en tiempo de vacas flacas (...)*

- (...) *¿Ha habido despilfarro? ¿Por supuesto. Pero es poco relevante en términos macro — otra cosa es la escandalera (...)*

El uso de citas textuales es otro mecanismo con el que el escritor se mantiene a cierta distancia de las palabras que otros han pronunciado. En el CCF, que el autor reproduce las palabras exactas de uno de los actores sociales o un personaje referido en su publicación. También podemos interpretar este recurso como una forma de enfatizar el contenido, evaluándolo implícitamente.

- *"Las dificultades de liquidez son una auténtica oportunidad para imponer duras condiciones" (...) como se sinceró Luis de Guindos*
- *"le hace un flaco favor al Estado de bienestar", comenta Vilalta*
- *"ninguna prestación sin contrapartidas" dice Merkel*
- *Obama considera las reformas en España "un paso positivo"*
- *Guindos ha hablado de "apoyo financiero" (+ eufemismo)*
- *Rajoy prepara "medidas contundentes" (+ eufemismo)*

Hemos podido comprobar con todos estos ejemplos cómo la evaluación está presente en los artículos periodísticos a nivel de palabra, frase y oración tanto de forma explícita, como oculta detrás de diferentes recursos retóricos. Igualmente, el análisis ha confirmado que la evaluación cumple las tres funciones definidas por Hunston y Thompson (2000) expresar opinión, mantener relaciones entre el escritor y el lector y organizar el discurso. El siguiente paso va a consistir en comprobar cómo la evaluación se distribuye por todo un texto y actúa como herramienta que contribuye a la organización textual confiriendo cohesión.

#### **4.5 Análisis de textos completos: macro estructura y lenguaje evaluativo**

La observación de las muestras de texto obtenidas mediante el procedimiento descrito nos ha permitido reflexionar y establecer conclusiones sobre la naturaleza de las expresiones evaluativas: la evaluación puede manifestarse de forma explícita, mediante la selección de términos y expresiones propiamente valorativos, pero también se encuentra en el texto de forma implícita, escondida detrás de figuras retóricas. La evaluación puede apreciarse de forma local (en palabras, frases, oraciones) o estar presente de manera global, continuada en todo un

texto (en nuestro caso, en el artículo periodístico) mediante el uso de diferentes estrategias que describiremos y explicaremos mediante ejemplos.



Figura 4.18: Metodología para el análisis de la crisis financiera en España

Nuestro análisis proporciona muestras de la selección lingüística por parte de los medios de comunicación para atraer a la opinión pública en un evento social importante como es la crisis financiera en España. Los recursos identificados en esta etapa incluyen técnicas utilizadas por los periodistas para establecer las relaciones con los lectores (repetición, intertextualidad, metáforas interpersonales) y poder dar forma su propia realidad social. En la figura 4.27 hemos comprobado la descripción de la metodología desarrollada e implementada hacia un perfil semántico del discurso de los medios. Partimos de la idea de que el periodismo es una actividad social y los medios de comunicación, en nuestro caso, la prensa escrita, reflejan a través de sus artículos el contexto socioeconómico. La creación de un corpus periodístico y su análisis informático y lingüístico nos permite detectar y analizar el lenguaje evaluativo presente en los textos de forma explícita o implícita y desenmascarar evidencias sobre el pensamiento político e ideológico en los textos.

De este modo, después de analizar las líneas de concordancia hemos dado un paso más en la investigación con la exploración de textos completos: su macroestructura y sus características principales. En la última parte de nuestro estudio nos propusimos explorar cómo se expresa y se distribuye la opinión en una selección de artículos de opinión del CCF, mayoritariamente editoriales. El editorial es el género periodístico que siguen aquellos que quieren conocer el estado de la opinión pública de un país. Los gobiernos ven en los editoriales una fuente muy importante de información como una interpretación de los medios de comunicación en sus decisiones. El editorial es el género periodístico que, según Martínez Albertos (1983: 379) sirve para juzgar a una noticia de especial importancia y establece la postura ideológica del periódico.

Parafraseando a Joseph Pulitzer, célebre director de *The New York World*, Martínez Albertos (1983) establece que el lector debe conocer el punto de vista del periódico y considera inmoral que se oculte tras la aparente neutralidad de la noticia y con ello oculte al lector de sus propios planteamientos ideológicos. *The Committee on Modern Journalism* (1967) sostiene que, si bien el periodista ayuda a los lectores a aprender, el editorialista les ayuda a entender.

La metodología de análisis siguió el modelo de las fases anteriores: se analizó cómo el periodista expresa su opinión e intenta atraer a su audiencia a través de unidades de menor a mayor extensión (palabra, frase, líneas de concordancia), para finalmente observar cómo se distribuyen y entrelazan todos éstos en la unidad mayor, el texto completo.

Nos centramos principalmente en dos aspectos: por una parte, la detección de funciones comunicativas en diferentes partes o secciones del artículo y, por otra, la organización y selección léxico-gramatical para presentar y valorar acontecimientos de gran calado social. En cuanto el primer objetivo, nos referimos a la macro-estructura, que reveló una tendencia a seguir un patrón común en las muestras analizadas:

|   |
|---|
| Macro-estructura de los editoriales   |
| <b>Título: evaluativo / hipótesis / llamada implícita al lector</b>   |
| <b>Cuerpo: Declaración y evaluación de las noticias</b>   |
| <b>Cuerpo: Declaración y evaluación de las noticias / información / argumentación y evaluación / análisis / perspectivas de futuro y predicciones</b> |
| <b>Conclusión: Predicción de eventos futuros / llamada a la acción / llamada a la autoridad y reformulación de la hipótesis inicial (repetición)</b>  |

Tabla 4.8: Macroestructura de los editoriales

El título suele ya evaluar los eventos que se van a describir, con fórmulas breves que atrapan desde el inicio al lector, a quien lleva de la mano y con quien interactúa hasta la parte de cierre. En la sección final solemos ver una proyección hacia el futuro, a modo de advertencia de posibles consecuencias y la apelación a las instituciones, autoridades o agentes sociales llamándoles a la acción. Suele concluir con la reformulación de título, creando con esta repetición una especie de círculo cerrado.

En el editorial publicado por el diario *El Mundo* el 3 de abril de 2012 leemos el siguiente titular: *Rajoy sabe al menos cuál es el buen camino*. El título ya contiene una evaluación y la hipótesis inicial. El periodista comienza manifestando su opinión: *sabe al menos cuál es el buen camino*. Seguimos leyendo en el texto:

*Consciente de las dudas que ha suscitado en diversos ámbitos el proyecto de Presupuestos, que ayer era criticado por el diario británico *Financial Times*, Mariano Rajoy realizó ante el Comité Ejecutivo Nacional del PP una convinciente defensa de su política y explicó por qué se ha visto obligado a tomar medidas impopulares, que han generado un cierto descontento social. En la primera parte de su intervención, Rajoy efectuó un diagnóstico de los tres graves problemas que lastran la economía y que ha heredado del pasado: un déficit público que el año anterior ascendió a 90.000 millones de euros, el estrangulamiento del crédito y la falta de competitividad empresarial.*

Vemos en el cuerpo del texto la exposición de una noticia sobre los Presupuestos y la opinión que en el autor del texto suscita a través de términos como *dudas, criticado, medidas impopulares, descontento social, graves problemas, déficit, estrangulamiento*. Comprobamos en el cuerpo el desarrollo de la noticia y la opinión del autor. Ésta se intercala con los eventos. Al final vemos la argumentación y evaluación de la noticia con análisis y apelación a la autoridad y llamada a la acción.

El siguiente objetivo ha sido la observación de forma global de la concentración y distribución de las unidades estudiadas en las diferentes etapas de la investigación, que hemos descrito anteriormente, para poder extraer conclusiones sobre las estrategias lingüísticas que se utilizan

en el mensaje periodístico para influir en la audiencia y convencerla de la validez de sus argumentos. Destacamos el hecho de que en muestras de textos individuales, apreciamos en una lista de palabras de contenido obtenida de uno de los artículos analizados, de breve extensión, que los términos tienden a aglutinar significados similares que se pueden agrupar bajo categorías semánticas: palabras relacionadas con la banca, (*Banca, Bankia, institución*) sustantivos concretos y abstractos de elevada carga semántica y adjetivos que en ocasiones guardan estrecha relación (*esclavos- víctimas, abundancia-exceso- libertades, poderes, sabiduría- conocimiento- extrañeza, , libertad-sencillez, bueno-malo, obvio, alto- altura, aforismos, aforísticas*)

| <b>PALABRA</b> | <b>FRECUENCIA</b> | <b>PALABRA</b> | <b>FRECUENCIA</b> |
|----------------|-------------------|----------------|-------------------|
| BANCO          | 5                 | ALTURAS        | 1                 |
| BORDA          | 3                 | ANCHO          | 1                 |
| BUENO          | 3                 | ANTAÑO         | 1                 |
| MALO           | 3                 | ARROJABAN      | 1                 |
| ARROJARON      | 2                 | CAUSAN         | 1                 |
| BANKIA         | 2                 | COMPAREN       | 1                 |
| CUÁNTA         | 2                 | CONDUCE        | 1                 |
| ESPAÑA         | 2                 | CONDUCIDO      | 1                 |
| EXCESO         | 2                 | CONOCIÉNDOLE   | 1                 |
| INFORMACIÓN    | 2                 | CONOCIMIENTO   | 1                 |
| LIBERTADES     | 2                 | CONTRARIO      | 1                 |
| LÓGICA         | 2                 | ESCLAVOS       | 1                 |
| MENTÍA         | 2                 | EXTRAÑEZA      | 1                 |
| OBVIO          | 2                 | INSTITUCIÓN    | 1                 |
| RECUPERAR      | 2                 | LIBERTAD       | 1                 |
| SABIDURÍA      | 2                 | PODERES        | 1                 |
| ABUNDANCIA     | 1                 | RUBIA          | 1                 |
| ACCIONES       | 1                 | SENCILLEZ      | 1                 |
| ACTUAL         | 1                 | TENDRÍAMOS     | 1                 |
| AFORISMOS      | 1                 | VÍCTIMAS       | 1                 |
| AFORÍSTICA     | 1                 | VOLUNTARIA     | 1                 |

Tabla 4.9: Lista de palabras y frecuencias

Podemos afirmar que los sentidos expresados tienden a concentrarse, por lo que, a pesar de encontrar un índice bajo de frecuencias, suelen repetirse significados o sentidos similares. Todo ello crea el hilo conductor del texto.

Después de comparar los términos individuales, el siguiente paso consistió en analizar las expresiones de mayor longitud, lo que de nuevo nos ofreció claros ejemplos de ese flujo de valoraciones, opiniones y sentimientos que el lector va descubriendo a medida que lee el

artículo. Para poder mostrar ejemplos reales de uso, vemos una carta al director, de corta extensión, tomada del CCF y cómo el análisis se llevó a cabo y los hallazgos más significativos. En un solo párrafo se condensan términos con connotaciones negativas (*despilfarro, megalomanía, excesos, disparates, descaro, causas más peregrinas, pufos y mentira del déficit*), expresiones que agrupamos bajo etiquetas evaluativas que cumplen las funciones comunicativas que describimos anteriormente y figuras retóricas que con frecuencia coexisten en una misma oración o párrafo (metáforas continuadas, el lenguaje cargado de términos e ideas contradictorias que dan lugar a la antítesis y al oxímoron con un tono irónico, o los eufemismos, con los que se suaviza una idea que pudiera resultar un tanto dura).

Esta repetición o reformulación de ideas produce una prosodia semántica. Se puede apreciar también con esas llamadas al lector mediante expresiones inclusivas o el uso recurrente de preguntas retóricas.

*La crisis tiene su parte **positiva***

***Gracias a la crisis** están saliendo a relucir el despilfarro, la megalomanía y los excesos de nuestros gobernantes. A nivel nacional **conocimos** los disparates del Plan E, los 400 euros de ZP para comprar el voto de 2008 con el mayor descaro, y las donaciones y subvenciones a las causas más peregrinas del mundo a cargo de las ministras de cuota. Luego han salido más pufos y la mentira del déficit.*

***Pero** con el anuncio de la aportación del Estado para pago a los proveedores de ayuntamientos se ha destapado la caja de pandora: las facturas de hace décadas y los cientos de miles de deuda en poblaciones de 200 habitantes. **Bienvenida** sea la ruina si sirve para que se sonroje la caterva de políticos ineptos, pésimos administradores, presumidos electoralistas, que **nos han llenado** el país y la vida de rotondas, polideportivos vacíos, farolas y setos al más puro estilo hortera y nuevo rico. Los ciudadanos con algo de sentido común hace tiempo que **temíamos** que esto iba a explotar, que España tuviera más aeropuertos por kilómetro cuadrado que el resto de Europa y las fachadas más rimbombantes y subvencionadas. ¿En manos de quién **estamos**? ¿Qué cabezas de chorlito se han aprovechado de **nuestro dinero** impunemente para satisfacer su ego y comprar votos (esto sin hablar de la corrupción)? **Hay que desmantelar este régimen de desaprensivos, levantar las alfombras y sanear la casa de una vez. **Ahora o nunca****.*

Oxímoron + repetición:

- *La crisis tiene su parte positiva*
- *Gracias a la crisis*
- *Bienvenida sea la ruina (ironía)*

Metáforas:

- *se ha destapado la caja de pandora*
- *levantar las alfombras y sanear la casa de una vez*
- *nos han llenado el país y la vida de rotondas*

Expresiones plurales (inclusivas):

- *conocimos los disparates*
- *nuestros gobernantes*
- *Los ciudadanos con algo de sentido común hace tiempo que temíamos que esto iba a explotar*
- *nos han llenado el país*
- *temíamos*
- *estamos*

Modalidad deóntica:

- *Hay que dismantelar este régimen de desaprensivos, levantar las alfombras y sanear la casa de una vez. Ahora o nunca*
- *no podía ser que*

Preguntas retóricas (estilo directo) + respuesta:

- *¿En manos de quién estamos?*
- *¿Qué cabezas de chorlito se han aprovechado de nuestro dinero impunemente para satisfacer su ego y comprar votos (esto sin hablar de la corrupción)?* Respuesta por medio de una pregunta a una pregunta anterior
- La regla de tres, para crear contenido atractivo:
  - *el despilfarro, la megalomanía y los excesos*
  - *políticos ineptos, pésimos administradores, presumidos electoralistas*

- *Hay que dismantelar este régimen de desaprensivos, levantar las alfombras y sanear la casa de una vez*

Coloquialismos:

- *nos han llenado el país y la vida de rotondas, polideportivos vacíos, farolas y setos al más puro estilo hortera y nuevo rico*
- *fachadas más rimbombantes y subvencionadas*
- *cabezas de chorlito*

|  |
|--|
| Macro-estructura del texto   |
| <b>Titular: evaluativo (ironía + antítesis) /llamada implícita al lector</b>                                   |
| <b>Cuerpo: El autor desarrolla la información / Expresa su opinión acerca de la actuación del gobierno /</b>   |
| <b>Cuerpo: El autor evalúa la actuación del gobierno (Gobierno Zapatero) evalúa su pérdida de credibilidad</b> |
| <b>Conclusión: Llamada a la acción / llamada a la autoridad</b>  |

Tabla 4.10: Macroestructura del texto

A continuación, destacamos el uso continuado de recursos estilísticos, que a menudo consiste en la reformulación un mismo recurso, pero también en la repetición exacta, lo que dota al texto de textura y coherencia, además de aportarle un toque humorístico de gran expresividad. Veamos, por ejemplo, la repetición en otro de los textos del CCF en el que se hace una referencia literaria además de utilizar un extranjerismo con connotaciones negativas.

|   |
|---|
| (OEP_01_03_12_[2]) <b>Delenda est autonomia</b>   |
| <b>Subtítulo</b> <i>Las comunidades son las responsables del mayor desvío en el déficit, pero el grueso es del Estado</i>   |
| <b>Cuerpo: Delenda est Carthago</b> —Cartago debe ser destruida—, proclamaba el senador romano Catón, con ocasión y sin ella, un siglo y medio antes de Cristo.<br><b>Delenda est Monarchia</b> , escribió José Ortega el 15 de noviembre de 1930 en <i>El Sol</i> , sembrando el camino a la República.<br><b>Delenda est autonomia</b> , sugieren algunos desde que la FAES publicó en 2010 su pionero estudio <i>Por un Estado autonómico racional y viable</i> , que abrió paso a la gran campaña contra el "despilfarro" regional. |

Tabla 4.11: Repetición de figuras retóricas

Ejemplos como este muestran la relación dialéctica entre los textos escritos en dos medios de comunicación y su evaluación donde las noticias son un tipo textual y evaluación es explorada en cuanto a la forma, el contenido y la estructura de este tipo de texto. La evaluación es una herramienta utilizada para la expresión de la opinión del escritor, para construir relaciones entre el escritor y el lector (con fines persuasivos, en el caso de los géneros periodísticos) y para organizar el texto (Thompson y Hunston, 2000).

Las categorías semánticas sugeridas anteriormente para el análisis de la evaluación en los textos escritos son el resultado de los datos obtenidos del corpus (*data-driven semantic categories*). El uso de métodos informáticos de análisis (listas de palabras, palabras clave, colocaciones, líneas de concordancia, clústeres) permiten el procesamiento de grandes cantidades de datos y, por lo tanto, la recogida de muestras representativas del uso del lenguaje, así como de ejemplos de patrones lingüísticos recurrentes. Sin embargo, la interpretación manual de los resultados estadísticos obtenidos revela ejemplos de evaluación explícitos e implícitos a medida que el análisis se extiende de las características más sobresalientes de elementos individuales a patrones de elementos del discurso de mayor extensión.

#### 4.6 Análisis detallado de textos completos

En este punto del estudio, quisimos mostrar al lector de manera detallada cómo la evaluación se articula en todo un texto. Por lo tanto, a continuación, se ofrece un análisis de veinte artículos de opinión del CCF que se estudiaron con el objetivo de detectar las diferentes técnicas que utiliza el periodista a la hora de evaluar actuaciones/declaraciones de actores sociales, entidades, procesos en dichos textos. Los textos se presentan uno a uno con los comentarios pertinentes.

| Periódico | Fecha      | Autor / Editor     | Artículo (género) | Texto id.                        | Palabras |
|-----------|------------|--------------------|-------------------|----------------------------------|----------|
| El Mundo  | 02/01/2012 | Institucional      | Editorial         | <a href="#">OEM_02_01_12_[2]</a> | 146      |
| El Mundo  | 03/02/2012 | Institucional      | Editorial         | <a href="#">OEM_03_02_12</a>     | 679      |
| El País   | 04/02/2012 | Institucional      | Editorial         | <a href="#">OEP_04_02_12</a>     | 482      |
| El Mundo  | 01/03/2012 | LUIS MARÍA ANSON   | Columna           | <a href="#">OEM_01_03_12_[2]</a> | 558      |
| El Mundo  | 01/03/2012 | ANTONIO GALA       | Columna           | <a href="#">OEM_01_03_12_[3]</a> | 192      |
| El País   | 01/03/2012 | XAVIER VIDAL-FOLCH | Análisis          | <a href="#">OEP_01_03_12_[2]</a> | 750      |
| El País   | 01/03/2012 | Institucional      | Editorial         | <a href="#">OEP_01_03_12</a>     | 503      |
| El Mundo  | 19/03/2012 | PILAR BLASCO       | Carta             | <a href="#">OEM_19_03_12_[2]</a> | 249      |

|          |            |                           |           |                                  |     |
|----------|------------|---------------------------|-----------|----------------------------------|-----|
| El País  | 20/03/2012 | Institucional             | Editorial | <a href="#">OEP_20_03_12</a>     | 428 |
| El País  | 01/05/2012 | Institucional             | Editorial | <a href="#">OEP_01_05_12</a>     | 486 |
| El País  | 18/05/2012 | JUAN JOSÉ MILLÁS          | Columna   | <a href="#">OEP_18_05_12 [3]</a> | 348 |
| El Mundo | 30/06/2012 | Institucional             | Editorial | <a href="#">OEM_30_06_12</a>     | 556 |
| El País  | 30/06/2012 | Institucional             | Editorial | <a href="#">OEP_30_06_12</a>     | 604 |
| El País  | 03/07/2012 | PEDRO PABLO PELLÓN PULIDO | Carta     | <a href="#">OEP_03_07_12 [2]</a> | 198 |
| El País  | 04/07/2012 | Institucional             | Editorial | <a href="#">OEP_04_07_12</a>     | 473 |
| El Mundo | 08/07/2012 | TOMÁS CAMPOS IBÁÑEZ       | Carta     | <a href="#">OEM_08_07_12 [3]</a> | 140 |
| El Mundo | 01/08/2012 | Institucional             | Editorial | <a href="#">OEM_01_08_12 [2]</a> | 526 |
| El País  | 01/10/2012 | Institucional             | Editorial | <a href="#">OEP_01_10_12</a>     | 490 |
| El País  | 02/11/2012 | Institucional             | Editorial | <a href="#">OEP_02_11_12</a>     | 483 |
| El País  | 02/12/2012 | Institucional             | Editorial | <a href="#">OEP_02_12_12 [2]</a> | 418 |

Tabla 4.12: Textos analizados de manera detallada

### Texto 1 [EDITORIAL OEM\_02\_01\_12\_[2]]

Se trata de un texto muy breve con un mensaje concreto que utiliza un estilo similar al del texto 19 (que analizaremos más adelante) al tomar un tema en Madrid y extenderlo a todo el país. En el texto 19, la cuestión es el pago de los pensionistas por sus medicamentos, que se extendió a la cuestión de la privatización de los servicios sanitarios. En el texto 1, la fusión de la Cámara de Comercio de Madrid con la federación empresarial local se utiliza como argumento para que el gobierno fusione todas las Cámaras de Comercio de toda España con sus federaciones de negocios locales y así crear sinergias importantes y un notable ahorro de dinero. El escritor termina el texto incitando al gobierno a tomar la acción legislativa necesaria.

- (1) El Gobierno debería realizar los cambios legislativos necesarios para promover esta integración.

La evaluación está en el uso de la modalidad deóntica (*debería, necesarios*). La modalidad deóntica indica la obligatoriedad de que se dé el contenido enunciado. La evaluación no es explícita (en el sentido de que haya una crítica del gobierno), pero la intención persuasiva del escritor es que debe haber un cambio legislativo y que el escritor lo considera como algo bueno (evaluación orientada a objetivos).

### Texto 2 [EDITORIAL OEM\_03\_02\_12]

El texto 2 comienza con una cita que contiene una fuerte evaluación negativa. Esta cita tiene como fuente a otra fuente especificada (Fundación de Estudios de Economía Aplicada).

- (1) «TRAS CUATRO años de crisis, no sólo se sigue destruyendo empleo, sino lo que es peor, la intensidad de esta destrucción aumenta».

De hecho, la primera parte del artículo está llena de referencias negativas al desempleo en España.

- (2) ...son las cifras de una auténtica emergencia nacional.
- (3) Es muy grave...
- (4) ...si se mantiene en el tiempo hace inviable el sistema.
- (5) ...incapaz de frenar la sangría del paro
- (6) ...la sociedad le pide ya con urgencia soluciones.

Toda esta evaluación negativa tiene como fuente el redactor/escritor del artículo. La segunda parte del artículo está dedicada casi exclusivamente al antídoto contra el problema del desempleo en España. El escritor nos da primero las malas noticias y luego la buena noticia de que el gobierno tiene un plan que reforma el sistema financiero. El título del artículo es:

- (7) Que a Guindos no le tiemble el pulso al aplicar su plan

Esta misma frase se repite en el último párrafo (en la penúltima frase) del editorial. Es un mensaje claro para el lector (y para el gobierno sobre esta cuestión) que el Ministro de Economía debe ser decisivo y seguir adelante con su plan de limpieza ("saneamiento", "limpieza", "sanearse") del sistema financiero en España para que el crédito vuelva a fluir hacia las empresas y las familias, estimulando así el crecimiento económico. El hecho de que la línea editorial esté en consonancia con el gobierno se realiza con la metáfora de un viaje:

- (8) Un plan contundente y bien encaminado...
- (9) ¿Será ésta la reforma definitiva del sector o se quedará a medio camino como las del Gobierno socialista? El tiempo lo dirá, pero el plan Guindos es mucho más realista.
- (10) El Gobierno dio ayer un buen paso adelante con esta reforma ...

El artículo termina con un claro mensaje tanto para los lectores como para el gobierno "... que se vuelva a crear empleo en España". Una de las fuerzas motrices detrás de la evaluación en el texto es el carácter del discurso, orientado al logro de objetivos. El escritor no sólo está tratando

de persuadir al lector de la necesidad de las reformas del Ministro de Economía, sino también de incitar y empujar al gobierno a tomar medidas firmes y decisivas para crear empleo. La evaluación en el texto es a menudo una cuestión de lograr metas y el logro o no de objetivos puede ser positivo o negativo. Por lo tanto, en un artículo de investigación, el logro de resultados significativos es una evaluación positiva.

**Texto 3 [EDITORIAL OEP\_04\_02\_12]**

El principal objeto de evaluación en este texto es: la nueva reforma financiera. Se evalúa de la siguiente manera:

- (1) añade intensidad para restablecer el funcionamiento del Sistema

La complejidad inherente es que la evaluación positiva en esta afirmación está implícita. Algo que añade intensidad no es *per se* “bueno”. Sin embargo, el lector sabe que el editor está diciendo que la reforma debe aplicarse rigurosamente con determinación para que el sistema financiero pueda funcionar. Dos reformas financieras anteriores (dos reformas financieras impulsadas por los Gobiernos de Rodríguez Zapatero) que fallaron se evalúan como:

- (2) que no han surtido el efecto deseado

Aunque esta es claramente una afirmación negativa, el problema de analizar su significado evaluativo es que está contenido dentro de una cláusula relativa, una expresión idiomática, una cadena fija donde la evaluación negativa es el resultado de la colocación. La evaluación negativa se dispersa a través de la cláusula relativa (y no se localiza fácilmente).

Esta reforma se evalúa además como:

- (3) la vocación de ser la llave definitiva que restablecerá la confianza en las entidades crediticias españolas

La pregunta que podría formularse aquí es cómo el lector sabe que vocación conlleva algún tipo de ironía, la duda sobre si la reforma financiera realmente restablecerá la confianza en las entidades financieras españolas.

El artículo en general siembra continuamente la duda en el lector acerca de si la reforma financiera será eficaz y logrará sus objetivos. Una de las maneras en que el escritor del texto logra esto es a través de señales lingüísticas. A continuación, se presentan algunas de estas:

- (4) Esos principios son, de entrada, correctos.
- (5) Poco se puede oponer a ...
- (6) ... es un ejercicio comprensible de voluntarismo.
- (7) Otra cosa es que pueda organizarse un control efectivo de esa exigencia.
- (8) Las dudas comienzan ...
- (9) Nada garantiza que ...
- (10) Y, naturalmente, también puede suscitar dudas ...
- (11) ... la pretensión de que ...
- (12) El ministro y su equipo supone ...
- (13) ... parece razonable pensarlo así ...
- (14) ... pero habrá que esperar a ver si se cumple.

Todas estas afirmaciones representan problemas al analizar el significado evaluativo. En (4), *correctos* señala evaluación positiva explícita, pero *de entrada* prepara al lector para una inversión en la opinión del escritor. En (5), *poco se puede oponer* implica que puede haber oposición. En (6), el *voluntarismo* es culturalmente negativo. En (7), *otra cosa es que* significa algo como que es otra cosa (con duda implícita). En (8), *las dudas* es la duda explícita como es en (9) *nada garantiza* y en (10) *puede suscitar dudas*. En (11), *la pretensión* expresa dudas implícitas. El escritor continúa en la misma línea en (12), (13) y (14). El problema para el analista (no para el lector que puede captar inmediatamente las señales negativas) es que la evaluación negativa a veces es explícita y otras veces está oculta o enterrada en una frase idiomática (*a ver si se cumple*). Sin embargo, en términos generales, el escritor parece apoyar la reforma financiera del gobierno y espolvorea todo el texto con aspectos positivos de la reforma:

- (15) ... la reforma restaurará el crédito...
- (16) ... acelerar su saneamiento inmobiliario...
- (17) ... saneamiento contra resultados...
- (18) ... saneamiento contra capital...
- (19) ... salvar el destino de las entidades intervenidas.

Sin embargo, al final del texto, el escritor advierte al gobierno y es implícitamente negativo sobre la recién introducida reforma financiera.

(20) Por todo ello cabe pedir al Ejecutivo que ... se muestre extremadamente vigilante sobre los posibles efectos secundarios de la reforma aprobada ayer.

La principal señal de evaluación negativa es la expresión *posibles efectos secundarios*. Es una metáfora médica, el gobierno puede estar aplicando un diagnóstico y tratamiento con efectos secundarios no deseados.

Una distinción sutil pero importante al analizar la evaluación en el texto es la diferencia entre la voz y el punto de vista. Esto es particularmente cierto cuando se evalúa lo que dijo alguien que a su vez puede estar evaluando en su discurso a otro actor social o algún proceso social. En el siguiente ejemplo, podemos ilustrar esta idea:

(21) El ministro y su equipo suponen que de esa manera se incentiva la venta de pisos a precios bajos.

El uso del verbo *suponen* en la subordinada indirecta es la opinión del editor/escritor, pero la información que se reproduce es lo que el ministro del gobierno piensa (en este contexto, *precios bajos* son algo positivo pero el lector tiene que tener conocimiento cultural/del mundo para saber esto en la situación socioeconómica de España). Esto se puede mostrar con los siguientes ejemplos:

- ...se incentiva la venta de pisos a precios bajos. *valor positivo* (en el contexto del artículo)
- ...hemos tenido que vender nuestro piso a un precio bajo. *valor negativo* (= hemos malvendido nuestra vivienda)

#### **Texto 4 [COLUMNA\_01\_03\_12 [2]\_OEM]**

El periodista Luis María Ansón en su columna de opinión del 1 de marzo de 2012 realiza una crítica a la gestión del gobierno de Zapatero en temas económicos. En el título se atribuyen los problemas políticos del momento a los problemas económicos.

Nos encontramos ante un artículo de una gran riqueza expresiva cargada de matices que se presentan en forma de diferentes recursos. Los objetos que se evalúan son actores sociales

como *Rodríguez Zapatero, la ministra Salgado, Montoro* e instituciones como el *Partido Popular, las Comunidades Autónomas, Bildu, los partidos catalanes y Amaiur*. Se evalúan, además, los procesos en torno a estos agentes sociales tales como el *déficit público, la gestión socialista, la gestión del PP* o el *Estatuto secesionista catalán*.

El autor está presente en todo su artículo a título individual y como parte del colectivo autor/lector.

- (1) No sé por qué
- (2) Ignoro
- (3) Dirigí
- (4) añadía
- (5) A mi manera de ver
- (6) Hemos creado
- (7) Conozcamos
- (8) No podremos
- (9) Todos sabíamos que

Mediante el marcador de atribución de la autoridad se transmite certeza o seguridad al mensaje, nadie va a poner en duda la veracidad de lo que expone.

- (10) Todos sabíamos que
- (11) La verdad es que

También vemos su voz al citarse a él mismo, haciendo referencia a un artículo anterior. Para ello, hace uso del entrecomillado:

- (12) «Salvo milagro zapateresco o maquillajes tórpidos, la cifra en que se mueve en estos momentos el déficit español oscila entre el 7 y el 8'5%»
- (13) «¿No sería mejor que dijeras la verdad? -le preguntaba yo a Elena Salgado- ¿Qué ventaja tiene mentir a la opinión española y europea?»
- (14) «Si esa fuera la herencia envenenada que dejáis a Rajoy, el líder popular no podrá cumplir con la parte sustancial de su programa y se fragilizarán sus propuestas sobre impuestos, actuación en favor de la pequeña y mediana empresa, inversiones públicas... Donde ha dicho digo, el líder popular tendrá que decir Diego»

Encontramos términos con marcado carácter evaluativo: *dulcificar, mentir, cumplir, fragilizarán, derroche, contener, despilfarro, reducir, déficit, mintieron, endosado, abrumadora,*

*especialmente preocupante, enseñorea, totalitario, amedrentar, dadivoso, deuda galopante, tragedia, hipócrita, brujulea, manipulan.*

Junto a este léxico con claros valores positivos y negativos detectamos neologismos, con los que el escritor acuña nuevas palabras: *zapateresco* (analogía con el estilo arquitectónico), *zapatética* (Zapatero+patético). Entabla igualmente un juego dialéctico con el interlocutor mediante la ironía en (15) a (19) y mediante el uso continuado de metáforas, como podemos apreciar en (20) a (28).

- (15) en un intento final por dulcificar los resultados
- (16) Salvo milagro zapateresco o maquillajes tórpidos
- (17) cataclismo desconocido
- (18) El expresidente dadivoso
- (19) los casi 6.000.000 de parados son una tragedia en sí mismos pero tal vez solo una anécdota en comparación con la descarga política que se prepara
- (20) estados de pitiminí, pero manirroto
- (21) brujulea (andar sin rumbo fijo)
- (22) herencia envenenada
- (23) el déficit público como un milagro zapateresco
- (24) juegan abiertamente a la secesión
- (25) dos bombas de relojería sobre la mesa política española que estallarán
- (26) descarga política
- (27) se rasgan las vestiduras como plañideras con vocación de histeria
- (28) rasgarse ahora las vestiduras parece un tanto hipócrita

De los ejemplos anteriores cabe mencionar la acumulación de diferentes recursos en una misma oración: metáfora, ironía referencias a frases populares y religiosas, como en (27) y (28), en los que aluden a los antiguos judíos y que viene de una tradición religiosa que consistía en que una persona ofendida rompía su ropa para rechazar así el agravio recibido. En la actualidad la usamos para criticar a quienes se escandalizan con excesiva ira ante una débil ofensa. Esta especie de ira da lugar al siguiente recurso estilístico en la misma oración, el símil, que se utiliza para evocar la antigua tradición en la que se contrataba a mujeres para que llorasen en los funerales (esa *histeria* a que hace referencia).

Otros casos de intertextualidad en el mismo texto con tono irónico y crítico son los siguientes:

- Frase hecha popular en español, reformulada por el escritor (para rectificar una afirmación o desdecirse): *Donde ha dicho digo, el líder popular tendrá que decir Diego*
- Referencia religiosa (reformulada por el escritor): *Amaiur, por obra y gracia de José Luis Rodríguez Zapatero\_brujulea en el Congreso de los Diputados*
- Repetición (aparece la misma frase tres veces en el mismo párrafo): *También por obra y gracia de José Luis Rodríguez Zapatero*

Estructuras prototípicas evaluativas que se detectan con frecuencia:

1. X (objeto de evaluación) **es** Y (valor de la evaluación).

**El 8'5% de déficit, la deuda galopante y los casi 6.000.000 de parados son una tragedia en sí mismos**

2. Y (valor de la evaluación) **es** X (objeto de evaluación).

**Lo grave de la herencia zapatética no es la economía; es la política**

3. Y (valor de la evaluación) of X (objeto de evaluación)

**el derroche de esos 17 Estados de pitiminií ...**

### Texto 5 [COLUMNA\_01\_03\_12 [3]\_OEM]

Antonio Gala en su columna de opinión del 1 de marzo manifiesta abiertamente cuál es su opinión sobre la situación de crisis que vive España a la vez que expresa su angustia por la incertidumbre sobre el futuro del país (*La situación es pésima y el euro, insostenible*). Una vez más, vemos al autor presente en *Me gustaría, no veré, Compadezco*. También vemos su voz presente mediante el uso del marcador de autoridad, que el propio escritor resalta con las mayúsculas, junto con la modalidad deóntica:

- A NADIE puede extrañar que

Imprime un tono diferente a su artículo ya que lanza diferentes preguntas al tiempo que las responde, convirtiéndose él mismo en un hipotético interlocutor. El diálogo que se articula dota de ritmo al texto, captando de esta forma al lector, quien encuentra en el propio texto las respuestas a las cuestiones que se plantean, con argumentos convincentes.

- (1) ¿Pesimismo? ¡Qué va!
- (2) Da igual cambiar de Gobierno: ¿cómo no va a dudarse de las estadísticas?
- (3) Si los grandes timos provienen de los grandes proyectos, gubernamentales o no, ¿en quién confiar?
- (4) ¿En quién confiar? En nadie, ni dentro ni fuera
- (5) ¿Cómo colaborar en un proyecto que, si existe, no es fiable?

En (1), que es el título del artículo, el interrogante que se plantea y su respuesta, un giro idiomático que significa *no hay pesimismo*, darían como resultado un sentimiento positivo en un análisis computacional. Sin embargo, la ironía enmascara el valor real de la oración, marcadamente negativo. El mismo tono irónico se percibe en (2) y (3), que requieren de la interpretación por parte del lector de su sentido implícito. En (4) se plantea la misma cuestión anterior, esta vez proporcionando una respuesta clara y tajante, que no deja cabida a interpretaciones erróneas y que, de nuevo se reformula en (5) con la repetición del objeto de evaluación (*proyecto/s*) y también del uso continuado de términos con los que se realiza una evaluación negativa. Podrían enmarcarse dentro de la categoría “Actitud” según la Teoría de la Valoración (Martin y White, 2005), que comprende aquellos casos en los que se emite una valoración subjetiva de los procesos que se describen y de sus participantes. En este sistema se incluye la categoría “Juicio”, que hace referencia a aquellos significados en los que se evalúan conductas y comportamientos sociales. Se consideran como recursos de “sanción social” aquellos que miden cuestiones relacionadas con la integridad y la moralidad. En nuestro caso, encontramos ejemplos como: *dudarse, timos, confiar, fiable*, todos ellos para cuestionar el cumplimiento de lo que todos consideramos como reglas o regulaciones sociales.

La interacción con el lector se logra también en el artículo mediante la intertextualidad, es decir, acudiendo a otras fuentes que resultan familiares para ambas partes. Encontramos referencias al mundo del deporte, concretamente al tenis en (6) y (7), mientras que en (8) se recurre a una frase que se originó en los naufragios de la Antigüedad y que en la actualidad se utiliza en situaciones no tan apremiantes, pero sí delicadas.

- (6) el peloteo entre Gobierno y comunidades
- (7) Lo económico es un juego de tenis entrampado
- (8) sálvese quien pueda

Vemos ejemplos de lenguaje figurado, además, mediante el uso de metáforas con connotaciones negativas, de fácil comprensión para el lector, a quien incluye en (10).

- (9) algún diario -más cuanto más polémico- se hunda
- (10) este lodo que nos ahoga
- (11) Todas las naciones están hechas polvo

Otro caso en los que se evidencia la dificultad que puede entrañar en ocasiones determinar la polaridad de una cadena simplemente por la suma de sus componentes es la paradoja, la unión de dos ideas opuestas que resultan contradictorias. Los siguientes ejemplos podrían considerarse como de signo neutro al comparar términos positivos y negativos en la misma expresión. De nuevo, su connotación negativa es algo que el lector deberá desenmascarar.

- (12) ganó para perder
- (13) Los pocos que trabajan son los que ganan menos
- (14) Lo único duradero es la caducidad
- (15) Si los grandes timos provienen de los grandes proyectos (paralelismo)

El uso metafórico del lenguaje alberga matices implícitos, que en algunos casos nos resultan familiares, por lo que es otra de las técnicas empleadas para situarse en el mismo nivel de la audiencia. En (16) y (17) se presupone un conocimiento por parte del lector sobre el BCE como la entidad que emite la moneda europea, o el cierre de un periódico nacional.

- (16) El Banco Europeo es la editorial que más imprime
- (17) algún diario -más cuanto más polémico- se hunda
- (18) este lodo que nos ahoga
- (19) Todas las naciones están hechas polvo

El texto concluye con una especie de sentencia, a modo de doctrina, en la que el columnista expresa su angustia por la incertidumbre sobre el futuro de España, realizando predicciones de futuro con una visión fatalista a través de contradicciones.

- (20) El final de esto será quizá el de todo
- (21) Lo único duradero es la caducidad

**Texto 6 [ANALISIS\_01\_03\_12 EP [2]\_OEP]**

En este texto, el autor inicia su reflexión sobre el déficit español y el gasto que supone para el gobierno central los gobiernos autonómicos tomando una frase literaria, *Delenda est Carthago*, adaptada por José Ortega *Delenda est Monarchia* y reformulándola *Delenda est autonomía*. El título ya constituye en sí mismo un claro ejemplo de cómo se presupone un conocimiento del mundo real, del mundo literario y de lo que se ha escrito en otros textos. Se presupone además que el lector va a ser capaz de interpretar o traducir la frase literaria escrita originalmente en latín. A partir de este momento, va a introducir a su lector en un entramado de recursos estilísticos y mecanismos lingüísticos con el único fin de convencerle de que los argumentos que se presentan son totalmente ciertos.

Como podríamos imaginar, si miramos de forma individual las palabras que se utilizan con una carga semántica positiva o negativa para plantear y describir la situación que se describe en el texto, encontramos términos de diferentes categorías gramaticales -sustantivos, adjetivos, verbos, adverbios- y que podríamos agrupar bajo diferentes categoría que hacen referencia a valores de justicia o moralidad tales como: *despilfarro, déficit, desfase, desempleo, deslealtad, gasto, bienestar, desvío, recortes, problema, incumplimiento, obsesión, despilfarro, bienestar, expectativas, austeridad, racaneo; sesgado, inquietantes, dramática e incontestable, incumplido, imposible, relevante, responsables, auténtica; perversamente. Los verbos expresan fluctuaciones económicas: ahorrarían, reducirían, suprimiesen, aumenta, baja, decrece, desbordado, incrementa*. El tono positivo o negativo en estos casos no plantearía dudas a la hora de formalizarlos mediante procedimientos algorítmicos.

Ahora bien, el autor hace uso de combinaciones más sofisticadas cuya interpretación requiere la capacidad del lector de descodificar el mensaje. Son metáforas que aluden a la naturaleza en (1) y (2) o, como en los ejemplos (3) y (4), que consideran al déficit como un ser animado. En (5) recurre a frases hechas para calificar a la dura crisis económica que atraviesa España en ese momento y en (6) califica al sistema político, imponiendo duras medidas económicas, como un laboratorio donde se hacen ensayos y experimentos, quizá si saber cuál va a ser el resultado. Ironiza al referirse a ciertos políticos en (7) y (8), contrastando *diabólico* con *los mejores* para

referirse a los miembros de dos formaciones políticas y compara el modelo español con la austeridad que vivió Grecia tras su quiebra, algo bien conocido y temido por los españoles.

- (1) arrecia la tormenta antiautonómica
- (2) desbordado su propia cifra
- (3) tripas del déficit
- (4) león del déficit
- (5) en tiempo de vacas flacas
- (6) laboratorio de la política de recortes
- (7) diabólico tripartito
- (8) el clan de “los mejores”
- (9) austeridad casi griega

Otro recurso recurrente y que se distribuye por todo el texto es el uso de marcadores expresivos formados por la combinación *ser + forma adjetival* que reflejan claramente la voz del autor. En todos los casos, vemos como el autor realiza una serie de afirmaciones atribuyéndose a sí mismo autoridad o conocimiento de la realidad. La voz en estos casos es del autor, que utiliza un tono persuasivo, aunque aparentemente se mantiene a distancia. En todos los casos nos podemos preguntar: ¿quién dice esto? Términos y expresiones como *responsables, grueso, condiciones, responsabilidad, problema, incumplimiento, herencia compartida, obsesión, modelo, sistema* (con connotaciones negativas o neutras) vienen calificadas por expresiones con un marcado sentimiento negativo: *mayor desvío en el déficit, imponer duras condiciones, es dramática e incontestable, es mucho más profundo, está sesgado, no producto de la obsesión gastadora, perversamente diseñado, poco relevante, gasto muy poco elástico, muy expansivo.*

- (10) Las comunidades son las responsables del mayor desvío en el déficit, pero el grueso es del Estado
- (11) Las dificultades de liquidez son una auténtica oportunidad para imponer duras condiciones
- (12) La responsabilidad de las autonomías en el desvío del déficit es dramática e incontestable
- (13) Pero el problema es mucho más profundo que el incumplimiento de 2011
- (14) el modelo está sesgado a favor de la Administración central (x2: repetición)
- (15) Es una herencia compartida y no producto de la obsesión gastadora del rojerío
- (16) es que el sistema está perversamente diseñado
- (17) Pero es poco relevante

(18) Pero no en tanta medida

(19) Es un gasto muy poco elástico (más difícil de recortar que el militar), y muy expansivo

En (11) nos llama la atención la combinación de sentimientos con signo contrario:

- Las dificultades de liquidez [sentimiento negativo] + son una auténtica oportunidad [sentimiento positivo] para imponer duras condiciones [sentimiento negativo]

Esta distribución y acumulación de valoraciones negativas, en ocasiones con sentidos muy similares, van calando en la mente del lector creando una especie de prosodia o ritmo negativo. En este sentido, se percibe como una sensación de evaluación de forma global en el texto en su conjunto.

Escuchamos diferentes voces (Thompson, 1996) que en ocasiones se materializan en el texto mediante el empleo de recursos como el entrecomillado. En (20) y (21), el periodista pretende mantenerse al margen de ciertas valoraciones que han pronunciado otros (*unspecified others*), y que revisten cierta dureza. En (22) y (23) recurre al estilo directo para reproducir las palabras exactas de actores sociales (*specified others*) impregnadas un tono irónico, cargado de contradicciones que pretenden manipular al lector y provocarle, hacerle reaccionar.

(20) sugieren algunos desde que la FAES publicó en 2010 su pionero estudio Por un Estado autonómico racional y viable, que abrió paso a la gran campaña contra el "despilfarro" regional

(21) o de los gastos causados por la "deslealtad" institucional

(22) "Las dificultades de liquidez son una auténtica oportunidad para imponer duras condiciones" (...) como se sinceró Luis de Guindos

(23) "le hace un flaco favor al Estado de bienestar", comenta Vilalta

Vemos la voz del periodista en el mensaje, que interactúa abiertamente con su audiencia, incluyéndose como uno de ellos al utilizar la primera persona del plural y creando un sentido de comunidad, con conocimientos y sentimientos compartidos (*community*). Esa voz del articulista se materializa cuando se dirige abiertamente a sus interlocutores e, incluso, formula preguntas repetidas cuyas respuestas él mismo proporciona (*unspecifiable others*). Intenta además situarse a su mismo nivel recurriendo a coloquialismos, expresiones que los ciudadanos utilizamos en nuestra conversación cotidiana para expresar nuestros sentimientos o actitud hacia los dirigentes.

- Plurales sociativos:

- Operemos
- Pongamos
  
- Apelación directa al lector:
  - Imaginen
  - Miren
  
- Pregunta retórica (directa) + respuesta + repetición:
  - ¿Ha habido despilfarro? Por supuesto.
  - ¿Ha habido despilfarro? Por supuesto. Pero es poco relevante (...)
  - ¿Es razonable? En parte, (...)
  
- Coloquialismo:
  - dicho en más suave
  - rojerío
  - racaneo

Hemos podido comprobar cómo las voces en el texto forman parte del propio mensaje, que a su vez contiene de forma explícita o implícita las actitudes que el periodista quiere expresar.

#### **Texto 7 [EDITORIAL OEP\_20\_03\_12]**

Los actores sociales en torno a los cuales se articula el texto 7 son:

- Gobierno de Mariano Rajoy (x2)
- Ejecutivo de José Luis Rodríguez Zapatero; anterior Ejecutivo
- representantes
- Gobiernos de uno y otro color.
- sociedad democrática
- ciudadano/s (x3); contribuyentes; ciudadanía; los españoles

Cabe mencionar la presencia de términos e ideas contradictorias que, como hemos ido viendo en el análisis de textos completos, es un rasgo recurrente en el lenguaje periodístico.

- ley de transparencia que rompa la cultura del ocultismo
- Opacidad / se les niegue el derecho a conocer
- Utilizarlo iría en contra de la habitual inercia sectaria de la política española, pero sería una magnífica señal de buen gobierno

Percibimos la presencia o voz del escritor emitiendo sus propias valoraciones, que se pone de manifiesto mediante el uso de la modalidad deóntica (*lo que la futura ley debería promover*) así como con la construcción formada por *ser/resulta* + adjetivo valorativo.

- sería un importante paso adelante
- No es una cuestión menor
- es la mejor herramienta
- no es suficiente para colmar las expectativas
- sería una magnífica señal
- Resulta intolerable que
- resultan sin embargo imposibles de ocultar a una

Otro rasgo que merece mención especial en este artículo es la reformulación del titular y la repetición de ideas que giran alrededor de éste y que se distribuyen por todo el artículo. Se trata de mecanismos de cohesión textual, que consisten en la repetición de ideas similares.

- Romper el ocultismo
- España necesita una ley de transparencia que obligue a rendir cuentas de la gestión pública
- España necesita una ley de transparencia (x2) tener cumplida información de la actividad de sus representantes y de cómo estos administran el dinero público que se les ha confiado
- La transparencia en la gestión pública es la mejor herramienta
- El secretismo hurta a los contribuyentes este derecho esencial
- favorecer la corrupción y el mal gobierno
- La arraigada cultura española del secretismo
- resultan sin embargo imposibles de ocultar a una ciudadanía cada vez más crispada
- sacar adelante una ley que obligara a las Administraciones a aportar a los ciudadanos la información requerida
- que rompa la cultura del ocultismo
- han extendido la misma opacidad que imperaba en la política nacional
- a los españoles se les niegue el derecho a conocer
- justo lo que la futura ley debería promover

#### Texto 8 [CARTA OPINIÓN\_19\_03\_12\_[2] OEM]

El texto se articula en torno a la irónica idea de que la crisis financiera tiene como contrapartida su lado positivo pues ha sacado a relucir los abusos y tramas financieras que se estaban cometiendo en España. Mediante el oxímoron, es decir, la contradicción o incoherencia de

ideas, el autor está apelando al lector para que vea la gravedad de los hechos. Este argumento se reformula en el artículo, siendo esta valoración la que le confiere ritmo y coherencia.

- (1) La crisis tiene su parte positiva (ironía + oxímoron)
- (2) Gracias a la crisis (ironía + oxímoron)
- (3) Bienvenida sea la ruina (ironía + oxímoron)

El autor atribuye a nuestros gobernantes las causas de la crisis y utiliza diversas fórmulas para calificar sus actuaciones. Como objetos de evaluación encontramos a *nuestros gobernantes, ZP, las ministras, los políticos*. Ahora bien, llama la atención los calificativos que aparecen a su alrededor tanto para denominarlos (*caterva, hortera, nuevo rico, cabezas de chorlito*) y para calificarlos (*electoralistas, ineptos, pésimos, presumidos, desaprensivos*) como para sancionar su gestión (*despilfarro, megalomanía, excesos, disparates, descaro, pufos, mentira del déficit, deuda, ruina, impunemente*).

Es evidente que estas valoraciones vienen del editor, quien apela a sus lectores mediante el plural sociativo:

- (4) conocimos los disparates
- (5) nuestros gobernantes
- (6) Los ciudadanos con algo de sentido común hace tiempo que temíamos que esto iba a explotar
- (7) nos han llenado el país
- (8) estamos

Incluso en (6) el autor se “atreve” atribuirse a sí mismo *ese sentido común* o capacidad de vislumbrar la mala gestión, mitigado por el plural, por lo que está también pronosticando ese conocimiento previo por parte del lector. Incluso, propone formas de remediar la situación mediante la modalidad deóntica, con la que concluye el artículo (*Hay que dismantelar este régimen de desaprensivos. Ahora o nunca*).

Al igual que hemos visto en los artículos anteriores, la metáfora es un recurso estilístico frecuente, haciendo alusión a elementos bien conocidos por todos. La metáfora está presente en (10) y (11) en forma de coloquialismos, mezcla de registros en los que vemos incluso referencias a nuestra vida cotidiana para llamar la atención y sumar audiencias:

- (9) se ha destapado la caja de pandora
- (10) levantar las alfombras y sanear la casa de una vez
- (11) nos han llenado el país y la vida de rotondas

La interacción con sus interlocutores se logra, además, mediante la interrogación retórica, en este caso proporcionando la respuesta mediante otra pregunta:

- (12) ¿En manos de quién estamos?
- (13) ¿Qué cabezas de chorlito se han aprovechado de nuestro dinero impunemente para satisfacer su ego y comprar votos (esto sin hablar de la corrupción)? [Respuesta (mediante pregunta) a la interrogación anterior]

Nos ha llamado la atención el uso de la estructura en forma de tríadas para captar la atención de la audiencia. Se trata de un recurso con el que se estructuran las ideas (en series de tres) y se dota de ritmo a la oración:

- (14) el despilfarro, la megalomanía y los excesos
- (15) políticos ineptos, pésimos administradores, presumidos electoralistas
- (16) Hay que dismantelar este régimen de desaprensivos, levantar las alfombras y sanear la casa de una vez

La ironía es un recurso que está presente en todo el texto; ideas que podrían verse como positivas tienen una marcada connotación negativa, como vimos anteriormente en (1) a (3) o en el siguiente ejemplo:

- (17) que España tuviera más aeropuertos por kilómetro cuadrado que el resto de Europa

#### **Texto 9 [EDITORIAL OEP\_01\_02\_12]**

El texto 9 es interesante porque es un ejemplo clásico de evaluación explícitamente negativa. La evaluación en texto tiende a ser sutil, implícita e incluso subrepticia, pero este texto pide al gobierno que actúe ahora para limpiar el sistema bancario español. El título del artículo es una llamada a la acción: *Que se haga pronto*. El artículo también termina con la misma llamada a la acción: *mejor que lo haga pronto*.

Entre el principio y el final del artículo, somos bombardeados por una serie de adjetivos, sustantivos y verbos negativos: *ineficaz, frustrante, desconfiando, preocupante, depreciación, interminable, fracaso, parcheo, encubrimiento, pérdidas ocultas, se deprecia, recesión, morosidad, inútil, parálisis, bancos zombies*, etc. El escritor describe las medidas del gobierno para llevar a cabo las reformas financieras como *ineficiente, frustrante, inútil, preocupante, un fracaso* entre muchos otros modos de calificar al gobierno de una manera abiertamente negativa. De hecho, la evaluación negativa estructura el primer párrafo.

La primera oración del texto nos dice que las reformas financieras del gobierno son ineficientes y frustrantes.

- 1) La reforma financiera aprobada por el Gobierno está resultando hasta ahora ineficaz y frustrante.

A continuación, el escritor explica por qué, utilizando la siguiente estructura lingüística (que es una especie de paralelismo, pero en realidad es más una tematización de los dos adjetivos de la primera oración):

- Ineficaz, porque...
- Frustrante, porque...

La puesta en relieve, es decir, la transformación en tema de la evaluación negativa (los adjetivos *ineficaz* y *frustrante*) resalta la opinión del autor, que no puede ser más escéptico y pesimista sobre el manejo que hace el gobierno de la crisis económica, algo que es de esperar de El País. La estructura lingüística utilizada por el autor en este caso es la tematización de los adjetivos, que es aquel mecanismo sintáctico en virtud del cual el tema aparece en la posición inicial.

#### **Texto 10 [EDITORIAL OEP\_01\_03\_12]**

El texto 10, si bien es temáticamente similar en el sentido de que trata de la economía española y, más concretamente, del sistema bancario español, la forma en que el escritor evalúa las acciones del Banco Central Europeo (BCE) en su intento de inyectar liquidez en el sistema bancario europeo y español es más argumentativa. La argumentación es una variedad discursiva con la cual se pretende defender una opinión y persuadir de ella a un receptor mediante

razonamientos. En otras palabras, la evaluación en este texto viene en forma de declaraciones y contra-afirmaciones. La primera oración del texto ilustra esta estrategia:

- La subasta del BCE aliviará los vencimientos de la banca, pero no elevará el crédito a las empresas

La primera parte de la oración habla positivamente sobre el alivio de los pagos bancarios, pero contrasta con la segunda parte (*pero* marca el cambio en el sentimiento), donde es evidente que el crédito no fluirá a las empresas como resultado. En la segunda oración del texto, se mantiene este tono y estructuración del argumento. El escritor habla positivamente de la inyección de liquidez del BCE como un éxito técnico. Sin embargo, el periodista afirma que hay dudas y utiliza una extraña metáfora para referirse a esta liquidez (*barra libre financiera*), lo que sugiere una falta de control. De hecho, el periodista destaca la naturaleza fluida del dinero que emana del BCE (*inundación de liquidez, último riego de liquidez*).

Esta forma de evaluar a través de contra-afirmaciones continúa en la cuarta oración:

- El BCE aplica el principio sensato..., ejercicio de sensatez que *contrasta abruptamente con los corsés presupuestarios procedentes de Alemania*.

La oración con que comienza el segundo párrafo es, de nuevo, optimista al empezar con la expresión: *El éxito del último riego de liquidez...* . Esto contrasta con el comienzo del tercer párrafo que se abre de la siguiente manera: *Pero el problema financiero europeo, y el español en particular, no es solo cuestión de liquidez*. Una vez más, el periodista utiliza, *pero* para destacar un cambio en sentimiento, pero también señala que algo negativo viene a continuación con la palabra *el problema*. El argumento del periodista es que, a pesar de una enorme inyección de liquidez en los bancos europeos y españoles, esto no se traducirá en préstamos privados y de empresa, ni en inversiones y en la creación de puestos de trabajo, sino más bien *se conforman con refinar viejas deudas, cobertura de activos dudosos* etc. Esta actitud de los bancos *es un obstáculo formidable para la recuperación de la economía*. El periodista utiliza un argumento razonado para contrastar el estado real del sistema bancario con un estado deseado para realizar su evaluación del estado de la economía española.

**Texto 11 [COLUMNNA OEP 18\_05\_12 [3]]**

De nuevo, en este texto el periodista entabla un juego dialéctico con su lector mediante el uso de recursos estilísticos que se distribuyen de forma continuada por todo el artículo.

Encontramos un antetítulo, “Aforismos”, que a priori nos anticipa la idea de doctrina o regla y que consiste en una declaración breve que pretende expresar un principio de una manera concisa, coherente y en apariencia cerrada. El título nos proporciona la máxima: *¡Cuánta sabiduría hay en lo obvio! El banco malo no es bueno*. La ironía constituye un modo de valorar de manera implícita a los personajes, sus acciones y declaraciones.

El lenguaje de las contradicciones constituye el hilo conductor del artículo, como hemos podido apreciar ya en el título. Los ejemplos de oxímoron son numerosos, como lo es también la reformulación e, incluso, la repetición de oraciones, todas ellas con un elevado tono irónico.

- (1) El exceso de información nos ha convertido en una sociedad ignorante

Y (valor de la evaluación) **de** X (objeto de evaluación)

- (2) la abundancia de libertades económicas nos ha hecho esclavos de los poderes financieros Y (valor de la evaluación) de X (objeto de evaluación)

- (3) Bankia fue intervenida porque funcionaba bien

- (4) a Rato lo arrojaron por la borda para premiar una gestión brillante

- (5) El banco malo no es bueno; el hombre alto no es bajo; la mujer rubia no es morena; el ascensor estrecho no es ancho; el caballo no es perro, etc.

X (objeto de evaluación) **es** Y (valor de la evaluación).

- (6) el Banco de España, dijo, es una institución muy prestigiosa que pronto recuperará su prestigio.

X (objeto de evaluación) **es** Y (valor de la evaluación).

Los objetos de evaluación en este caso son el banco malo, Bankia, Rato, Botín, Luis de Guindos y el Banco de España. Como se puede apreciar en (1) a 6), los términos valorativos abundan en el

artículo y podrían agruparse, bajo la categoría de “Actitud” (Martin y White, 2005), y las subcategorías de afecto (*libertades-esclavos*), estima social (*ignorante, prestigiosa, brillante*), sanción social (*forrado*) y apreciación (*malo-bueno*). Alrededor de los agentes sociales se crea como una red de aparentes silogismos cuya finalidad es evaluar, o mejor dicho, condenar la mala gestión económica del país.

No obstante, en ocasiones, detrás de estos juegos de palabras se está encubriendo el mensaje real. Sirva como ejemplo la siguiente oración:

- En realidad, no lo arrojaron por la borda, se fue porque quiso, pero se fue porque quiso al mismo tiempo que lo arrojaban por la borda.

En el ejemplo anterior comprobamos cómo la metáfora es, de nuevo, otro recurso estilístico recurrente en el periodismo que se combina con otras figuras.

- lo arrojaron por la borda
- lo que pilló tacita a tacita mientras hundía Bankia - Intertextualidad (referencia a un anuncio publicitario) + metáfora
- lo que le hemos dado de indemnización por dejar de hundirla, se va con el riñón forrado - Ironía + metáfora (x2)

Igualmente, la voz del lector está presente en el texto como interlocutor a quien apela directamente el autor (*unspecifiable others*): *Ya ven, cabe preguntarse, comparen*. También como parte del mismo colectivo: *nos ha hecho esclavos, nos causan, que le hemos dado, indemnizamos, nos lleva, nos ha conducido, tendríamos que* (modalidad deóntica). El diálogo con la audiencia se **logra** además con la formulación de preguntas retóricas, que el propio autor responde.

- Cabe preguntarse por qué indemnizamos a quien nos lleva a la ruina en lugar de a sus víctimas. Pues por la misma lógica que el exceso de información nos ha conducido a la ignorancia y el de libertades a la esclavitud - Pregunta retórica (indirecta) + respuesta

- ¿Cómo va a recuperar lo que no ha perdido? O mentía en la primera parte de la proposición o mentía en la segunda, aunque, conociéndole, quizá mintiera en las dos. Pregunta retórica (directa) + respuesta

Otras marcas lingüísticas en el texto de la presencia autorial se perciben mediante expresiones como: *En realidad, Ni el mismo Rato, Además le da igual, Pues por la misma lógica*. De forma implícita, el autor está emitiendo sus propias valoraciones. Tenemos ejemplos de otras voces en el texto (*specified others*):

- que es *lo que hizo Botín el otro día al proclamar* que el banco malo no es bueno
- Comparen ese discurso diáfano con el de *Luis de Guindos* acerca del Banco de España: el Banco de España, *dijo*, es una institución muy prestigiosa que pronto recuperará su prestigio.

Insistimos en la idea de que la reformulación de ideas y las repeticiones del título o de oraciones idénticas son una constante en el texto, lo cual puede entenderse como un recurso para la evaluación confiriendo coherencia al discurso periodístico.

- mentía en la primera parte de la proposición o mentía en la segunda,  **aunque**,  **conociéndole**, quizá  mintiera en las dos.

#### Texto 12 [EDITORIAL\_30\_06\_12\_OEM]

En el editorial de El Mundo del 30 de junio de 2012 se recogen las impresiones tras una cumbre europea. El artículo pone una nota de optimismo: *Rajoy podrá hacer ajustes sin que le hunda la deuda*. El título se reformula en el texto:

- consigue una segunda oportunidad para  recortar el déficit sin ser ahogado por la deuda
- pero  es mejor que Rajoy pueda emprender el  nuevo ajuste sin la angustia de la rentabilidad del bono situada en niveles de intervención.

Los objetos de evaluación en el artículo son el consejo Europeo, la cumbre, Rajoy, España, Italia, el Gobierno, las bolsas europeas, las primas de riesgo, el Ibex. Nos llama la atención el marcado tono positivo que se percibe por todo el artículo para evaluar las decisiones y resultados tras las negociaciones. Se materializan de forma explícita (adjetivos, adverbios y verbos de signo

positivo) e implícita mediante diferentes recursos estilísticos como se puede apreciar en los ejemplos:

- (1) Los resultados conseguidos parecen invitar al optimismo (metáfora+modalidad epistémica)
- (2) Hay que alegrarse (modalidad interpersonal)
- (3) la moneda única ha salido fortalecida
- (4) Si nadie duda de que una cumbre sin acuerdos hubiera supuesto iniciar el proceso de defunción del euro, el esfuerzo conjunto por conseguirlos demuestra que todos los países mantienen su confianza en este proyecto común (marcador de atribución de autoridad)
- (5) seis meses para que el BCE pueda estar en condiciones de supervisar a la banca europea es un plazo muy corto para los tiempos comunitarios y, sin embargo, así se ha acordado en el Consejo
- (6) La cumbre ha sido especialmente positiva para España
- (7) Rajoy -gracias al frente común con Monti y al apoyo de Hollande- ha logrado lo esencial
- (8) eliminando así una de las incertidumbres que habían encarecido hasta niveles críticos nuestra prima de riesgo
- (9) España e Italia han logrado también que, con una celeridad impropia en la UE, los fondos de rescate se adecuen legalmente
- (10) Se trata de una especie de red de seguridad
- (11) pero también plenamente consciente de que debe acometer sin demora los deberes que ya tenía impuestos
- (12) el resumen de este trascendental Consejo Europeo es que Rajoy, que llegó a la cumbre en una situación límite, consigue una segunda oportunidad para recortar el déficit
- (13) las bolsas europeas ganaron
- (14) la mayor subida en dos años
- (15) el euro se reforzará aún más
- (16) cambio de tono en los mercados (metáfora)
- (17) el Ibex registró la mayor subida en dos años
- (18) las primas de riesgo se relajaron considerablemente: la española bajó (metáfora)
- (19) pero es mejor que Rajoy pueda emprender (marcador expresivo (adjetivo evaluativo) + modalidad epistémica)

Merece mención especial cómo en este mismo sentido positivo, que resulta paradójico, se pronuncia el editor al afirmar:

- el Gobierno se ha quedado corto en las reformas que necesitaba el país

El mensaje contrasta totalmente con las valoraciones que veíamos en los artículos anteriores sobre la cumbre europea donde las decisiones iban, de nuevo, a repercutir de forma negativa en el bienestar de los españoles, que están viendo disminuidos sus derechos por el aumento de las presiones fiscales (*más esfuerzos, más recortes, Rajoy ha vuelto a dar una nota disonante*).

En este repertorio de sentimientos positivos encontramos términos y expresiones metafóricas que hacen referencia a sentimientos personales o estados de ánimo: ser *ahogado, angustia, optimismo, esfuerzo, confianza, incertidumbres, seguridad, consigue, relajaron*.

Como hemos ido observando, la voz del articulista está presente en los ejemplos anteriores bien al exponer las noticias de forma “objetiva” (cifras y estadísticas, medidas anunciadas) o al especificar quién dijo qué:

- (1) Rajoy declarara ayer que
- (2) los analistas hablaban ayer
- (3) decía Montoro

En ocasiones, su voz está infiltrada, entrando a realizar sus propias valoraciones conforme a su punto de vista. Sirvan las siguientes muestras para ver cómo esa voz se manifiesta de diferentes modos:

- (4) Hay que alegrarse
- (5) La cumbre ha sido especialmente positiva para España
- (6) De ahí que Rajoy declarara ayer que
- (7) Si se consigue que España e Italia dejen de ser un problema, el euro se reforzará aún más

De nuevo, apela al lector, mediante el plural sociativo (*si hemos llegado, tenemos, todo lo que hagamos*) y al dirigirse a su audiencia con expresiones metafóricas e idiomáticas que le van a hacer comprender perfectamente de qué está hablando (sentido de comunidad con conocimientos compartidos):

- (8) proceso de defunción del euro
- (9) con un colchón de seguridad
- (10) ataque de los mercados
- (11) La Troika no aterrizará en España

(12) los hombres de negro que decía Montoro se van a situar en lo alto de los Pirineos para escrutar todo lo que hagamos

Los ejemplos (11) y (12) requerirán que el lector reactive sus conocimientos previos sobre la situación del país y los riesgos y amenazas que sufrimos si no se reúnen determinados criterios económicos. Se trata de mensajes que a todos nos eran familiares en su momento por haber aparecido de manera reiterada en los diferentes medios de comunicación.

### **Texto 13 [EDITORIAL\_30\_06\_12\_OEP]**

En el texto 13, el periodista trata de informar al lector sobre los acuerdos adoptados tras la reciente cumbre europea. En su descripción sobre las medidas propuestas y las decisiones adoptadas en el proceso de la unión bancaria y fiscal encontramos, tal y como cabría esperar, partes/porciones de texto en las que se presentan quiénes han sido los actores sociales o participantes en el proceso y cómo han sido sus actuaciones, decisiones o declaraciones. Es decir, encontramos una serie de objetos que se evalúan conforme a diferentes patrones o, en este caso, aspectos sociales y financieros.

En primer lugar, en el artículo vemos ejemplos claros en los que el periodista relata sobre lo sucedido siguiendo un estilo informativo, propio de este género, manteniéndose a una cierta distancia para simplemente exponer los hechos.

- (1) La cumbre europea ha conseguido desbloquear varias decisiones cruciales para la estabilidad de la zona euro.
- (2) Es decir, el Mecanismo podrá comprar en el mercado
- (3) La secuencia de acontecimientos, a grandes rasgos, es esta: (...)
- (4) La cumbre decidió además que la deuda europea para sanear la banca no tendrá prelación sobre el resto de activos.
- (5) En este marco de acuerdos favorables y detalles por negociar, Hollande, Monti y Rajoy han conseguido lo que querían

En los ejemplos anteriores la cumbre europea, el Mecanismo de Estabilidad Europea (MEDE), o Hollande, Monti y Rajoy son objetos de evaluación a los que el editor simplemente se refiere como mero agente informador. Detectamos, no obstante, claras señales de lenguaje evaluativo

para calificar a estos objetos (instituciones, políticos, gobiernos) tales como: ha *conseguido*, *decisiones cruciales*, *sanear*, *acuerdos favorables* y *han conseguido*. Se trata de expresiones de signo positivo que transmiten al lector el mensaje de que las negociaciones han resultado propicias para nuestro país.

Sin embargo, en ocasiones, la voz del autor está presente con el fin de emitir sus propias opiniones, juicios o valoraciones bien mediante el uso de una serie de señales, pero también de forma un tanto infiltrada o encubierta a lo largo del texto, aunque a simple vista no exista ningún signo de su presencia. El uso de términos antitéticos en expresiones como *un deterioro de la solvencia del país* y *encadenar la solución al grave problema* constituyen un ejemplo de este mecanismo.

En los ejemplos 6) a 11) esa presencia o voz se concretan en el texto mediante expresiones, que destacamos mediante el subrayado, que nos llevan a preguntarnos ¿quién dice/opina que esto es así?

- (6) Estas son las piezas capitales de un acuerdo al que le queda un largo periodo de negociación para concretar detalles decisivos.
- (7) La importancia del pacto radica (...)
- (8) Pocas cumbres europeas han sido tan resolutivas y apegadas a la realidad como la última.
- (9) Anclar la recapitalización directa a la exigencia de avances hacia la unión bancaria es un factor de incertidumbre.
- (10) Tales avances no se van a conseguir con facilidad
- (11) Es algo que contradice la evidencia (“ninguna prestación sin contrapartidas” dice Merkel) y daña la credibilidad exterior del Gobierno.

Nos encontramos ante valoraciones positivas (*importancia*, *resolutivas*, *apegadas a la realidad*) y negativas (*le queda un largo periodo de negociación*, *es un factor de incertidumbre*, *es algo que contradice*, *daña la credibilidad*) de los acontecimientos cuya autoría no viene explícitamente indicada pero que tienen como finalidad calar en la mente del lector. Nos ofrecen ejemplos claros de la intención persuasiva y manipulativa del género periodístico de opinión.

En los extractos (12) a (14), estos fines persuasivos se materializan mediante el uso de la modalidad deóntica (*debe despejarse*) y la modalidad epistémica (*parece obligado crear*, *parece*

*oportuno por tanto negociar*). Se está incitando a los españoles a que cuestionen determinadas acciones, o planteen modelos de actuación para sus gobernantes. La interrogación retórica en (13) plantea una pregunta de difícil respuesta que quizá causará inquietud a quienes la lean.

(12) (...) una incógnita que debe despejarse.

(13) Pero ¿a qué precio valorará el Mecanismo de Estabilidad el capital de Bankia y del resto de las entidades nacionalizadas? No debe olvidarse que (...)

(14) Parece oportuno por tanto negociar

En el siguiente ejemplo, el editor apela al lector con ese tono de reiteración *ha vuelto a dar un nota disonante*, creando una voz de comunidad (Thompson, 1996) en la que se presupone un mundo o conocimientos compartidos entre el autor y su audiencia. El uso de la metáfora *una nota disonante* dota al texto de expresividad al recrear la imagen de una orquesta en la que todos sus miembros deben tocar al unísono para que la representación sea considerada como correcta.

(15) Pero Rajoy ha vuelto a dar una nota disonante

En el ejemplo (16) se acumulan varios recursos estilísticos en una misma oración: pleonismo (redundancia: *silencio discreto*) + marcador expresivo (verbo ser +adjetivo evaluativo: *hubiera sido lo mejor*). El mensaje que se transmite tiene un tono moralizador que podría incluso interpretarse como una manifestación de postura ideológica y de llamada a la acción.

(16) Un silencio discreto hubiera sido lo mejor.

Como hemos podido observar, las muestras de texto encierran matices y sutilezas de cierta complejidad que deben tomarse en cuenta al realizar el análisis lingüístico de los sentimientos que un texto recoge. En este sentido, nos planteamos la siguiente cuestión: ¿pueden cuantificarse todos estos sentimientos (valores, emociones, juicios) o clasificarse conforme a determinadas etiquetas? Igualmente, nos preguntamos si es posible modelizarlos mediante el uso de herramientas computacionales para el análisis de sentimientos que miden la intensidad de polaridad (positiva o negativa) y asignan puntuaciones conforme se hayan registrado en su diccionario o bases de datos.

Vemos, pues, que el lenguaje humano presupone un conocimiento de situaciones del mundo real y acude a múltiples recursos y mecanismos lingüísticos difíciles de cuantificar y representar mediante métodos matemáticos.

**Texto 14 [CARTA OEP\_03\_07\_12]**

El texto 14 es una carta al director, también considerado como un género de opinión. A pesar de su corta extensión (alrededor de 250 palabras), aglutina un gran número de las estrategias que hemos ido explorando en este recorrido por algunas de las muestras del CCF. Cubre la misma noticia que la del primer texto de este estudio (El País, 30 de junio de 2012): las reacciones tras la cumbre europea. En este caso, la fuente es también El País (3 de julio de 2012). No obstante, ante el título en ambos textos nos planteamos si la postura del editor va a ser diferente: *Por el buen camino versus Más esfuerzos*. A medida que avanzamos en la lectura de ambos artículos comprobamos cómo la ideología política se plasma y se puede detectar con un análisis lingüístico de estas características.

En el artículo se teje como un tapiz formado por un entramado de diferentes voces que se entrelazan. Por una parte, la voz del editor en (1), por otra, la del editor que se suma a sus lectores en (2) y el entrecomillado en (3), que en realidad no es lo que “otro” dice textualmente, sino lo que el autor quisiera que el ministro dijera, es decir, una apelación, en este caso al ministro.

(1) Veo; digo yo; me temo; le pediría

(2) Nuestro ministro; los españolitos de abajo (repetición); somos los que haremos más esfuerzos, nosotros sufriremos; nuestras famélicas nóminas; aceptaremos y sufriremos; nuestra casta superior; no se nos estresen; somos el único país; todos nos quedaremos

(3) Por lo tanto, le pediría al señor de Guindos, que matice un poco y mejor diga: "Algunos españoles harán más esfuerzos para contener el déficit"

La voz colectiva viene reforzada por un marcado tono irónico gracias a las metáforas y coloquialismos, que resaltamos mediante el subrayado en (4) y (5). Se trata de que el “español de a pie” (metáfora que usa de forma repetida) interprete sin dificultades cuáles son las

maniobras del gobierno y cómo se nos está pidiendo que hagamos grandes esfuerzos económicos (cualquier lector va a entender qué es la subida del IRPF, de la luz, etc.). De igual manera, queda muy claro su mensaje al decir “aceptaremos y sufriremos”.

- (4) Al menos los españolitos de abajo\* somos los que haremos más esfuerzos, nosotros sufriremos más que nadie las subidas del IVA, de la luz, el agua, la gasolina... todas las subidas que hagan falta\*\*; nuestras famélicas nóminas asumirán el IRPF que el Gobierno quiera; aceptaremos y sufriremos la Sanidad y Educación sin recursos que las reducciones presupuestarias dejen. (\*metáfora; \*\* coloquialismos: uso de términos conocidos por el lector – de su vida cotidiana- para que sea consciente del alcance de los hechos)
- (5) Todo ello para que nuestra casta superior cuyos ingresos sean difíciles de controlar por Hacienda, no se nos estresen

Este espíritu irónico se materializa en esta carta también a través de la paradoja (unión de dos ideas opuestas que resultan contradictorias) en (6) o el uso del sarcasmo en (7) y (8).

- (6) Por algo somos el único país del mundo que indemniza a los directivos que desfalcán y/o llevan a la quiebra a bancos y cajas de ahorros
- (7) Bien, eso está muy bien
- (8) así será más exacto y todos nos quedaremos más contentos

Ante estos ejemplos, de nuevo, nos preguntamos: ¿Es posible determinar la polaridad positiva/negativa de estas oraciones mediante métodos estadísticos?

#### **Texto 15 [EDITORIAL OEP\_04\_07\_12]**

En el texto 15, podemos ver lo complicado que es el punto de vista y la voz.

- (1) Las estadísticas de junio muestran además una **incoherencia significativa**.

Las estadísticas no muestran inconsistencias significativas. Las estadísticas son simplemente números que son analizados e interpretados por alguien. El editor/redactor del texto 15 ha decidido que hay inconsistencias significativas (nadie más). Ésta es la opinión y el punto de vista

del escritor, pero él/ella está intentando camuflar esto tratando las estadísticas como la fuente/voz de esta opinión. Todos los ejemplos que ofrecemos a continuación siguen un patrón similar:

- (2) El descenso del paro... ha sorprendido al Gobierno y a los analistas.
- (3) ..., la prudencia exigiría confirmar en periodos sucesivos el buen dato de junio
- (4) ... la prudencia debe extremarse todavía más, ...
- (5) ... no hay razones para suponer una mejora en la tendencia depresiva del mercado laboral.
- (6) Todo hace suponer que ...
- (7) Ninguna proyección laboral pronostica atisbos de creación neta de puestos de trabajo ...
- (8) Las hipótesis de trabajo para explicar esta discordancia sugieren un agravamiento del efecto desánimo ...
- (9) El problema social más grave que provoca ...
- (10) La Comisión Europea acaba de dar la voz de alarma al recordar que ...
- (11) ... pero se ha olvidado de legislar para que la contratación no castigue a la mitad de los profesionales jóvenes del país.

¿Quién dice que la caída en el desempleo *ha sorprendido* al gobierno? Seguramente, es el periodista/editor del periódico. La *prudencia* no puede ser un actor. Otra vez este es el punto de vista del periodista. El uso de estructuras impersonales tales como *no hay razones para suponer ...* y *todo hace suponer que ...* o el uso de la reflexiva *se en se ha olvidado de ...*. *Olvidarse* es un verbo pronominal en español son otras técnicas para esconder opiniones del articulista. Estos verbos expresan significados relacionados con el afecto, la memoria y el conocimiento. Típicamente, son procesos afectivos y cognitivos.

- (12) ME temo que sea ya tarde. (ésta es claramente una opinión personal)

En los ejemplos anteriores, hay otras maneras de ocultar la opinión de alguien o pretender que otro está diciendo lo que en realidad éste está afirmando. El escritor emite su propia opinión refiriendo como fuente una entidad abstracta: *Ninguna proyección laboral pronostica...*, *Las hipótesis de trabajo para explicar...*, *El problema social más grave que provoca...*. O se puede simplemente atribuir una evaluación a una fuente, aunque nunca la realizara: *La Comisión Europea acaba de dar la voz de alarma al recordar que...*

Hay múltiples opciones lingüísticas disponibles para que un escritor disimule una opinión atribuyendo la opinión a otra fuente, ya sea a través de estructuras impersonales (*unspecified others*), el uso de entidades abstractas (*specified others*) o de actores/procesos sociales (*Comisión Europea, problema social*) que podemos etiquetar como comunidad (*community*). Los escritores también tienen la opción de declarar explícitamente “este soy YO (mismo)”: *ME temo que sea ya tarde*.

#### Texto 16 [CARTA OEM\_08\_07\_12\_[3]]

En el texto 16 el autor acude a los recursos que hemos destacado como frecuentes en los artículos anteriores para evaluar personas, instituciones y situaciones. El autor manifiesta en primera persona su decepción y malestar: *Me sonrojo, Me preocupa*. Utiliza la forma del plural para atraer al lector: *nos encontramos*, a quien apela directamente mediante el imperativo *recordemos*.

El objeto de evaluación del título, *Los políticos*, se reformula utilizando expresiones como *responsables, ellos* (que se repite a lo largo del texto), *a sí mismos, todos*.

La ironía vuelve a ser un recurso destacado que se logra, como vemos, mediante la selección de términos muy expresivos (*servir, vocación, hasta el último concejal, emolumentos y prebendas*), recursos como la modalidad deóntica, el coloquialismo o la apelación al lector y las repeticiones (para referirse a los gobernantes), que se concentran en unas pocas líneas.

- Si están para servir a España con vocación, entonces todos, hasta el último concejal, deberían dar ejemplo renunciando a sus emolumentos y prebendas. Un buen pico sí que aportarían. Recordemos que en España hay en torno a 450.000 políticos.

#### Texto 17 [EDITORIAL OEM\_01\_08\_12\_[2]]

El texto 17 tiene un objetivo similar para persuadir al lector, pero esta vez para persuadir al lector de que es necesaria la intervención en el funcionamiento de los gobiernos regionales. Esta es una llamada a la acción mucho más fuerte. El artículo contiene varias formas de evaluación y vamos a ilustrar cómo un escritor puede utilizar para evaluar palabras individuales, frases, construcciones gramaticales y usos metafóricos del lenguaje. Las palabras evaluativas pueden

ser de cualquier categoría gramatical (sustantivos, verbos, adjetivos, adverbios). Los siguientes son algunos ejemplos de palabras (adjetivos y sustantivos) que son explícita e inherentemente negativas.

- adjetivos: *irreparable, irresponsable, ineficaces, incumplimiento, nefastos, ilógico*
- sustantivos: *daño, desafío, desconfianza, plante, rehén, deslealtad*

Aunque los sustantivos y los adjetivos son el tipo más común de evaluación léxica, verbos como *dinamita, ha fracasado* tienen un peso evaluativo negativo. Si examinamos las frases en el texto, encontramos ejemplos como: *mano dura, romper la baraja, el desbarajuste autonómico*. Se debe observar la naturaleza idiomática y metafórica de estas frases.

Ejemplos de construcciones gramaticales evaluativas típicas son:

- (1) X (objeto de evaluación) **es** Y (valor de la evaluación).
  - ... **los instrumentos del Gobierno** para alcanzar ese objetivo **son ineficaces**.  
(adjetivo)
  - La cumbre** *ha sido* **especialmente positiva** ... (adverbio + adjetivo)
- (2) Y (valor de la evaluación) de X (objeto de evaluación)
  - ... **el enfoque blando y conciliador de Rajoy**
- (3) X (objeto de evaluación) **hace algo a** Y (valor de la evaluación).
  - El daño irreparable a España** que ayer **causaron** **los gobiernos de Cataluña y Andalucía** ...

#### Texto 18 [EDITORIAL OEP\_01\_10\_12]

En el texto 18, tenemos una extensión de la construcción gramatical:

X (objeto de evaluación) **es** Y (valor de la evaluación).

Adopta la siguiente forma:

X = oración de infinitivo (objeto de evaluación) **es** Y (valor de la evaluación).

La oración de infinitivo es normalmente una acción que se considera que es positiva o negativa.

A continuación se muestran algunos ejemplos:

- **Revalorizar las prestaciones en función del elevado IPC sería una equivocación**
- **es una equivocación autorizar una subida de las pensiones**

- sería **contraproducente** aplicarla
- **No** es **coherente**, ... , presentar un programa de estabilidad en Europa
- sería **una decisión errónea** revalorizar las pensiones en función del IPC

En este caso, todos los ejemplos son negativos ...

#### **Texto 19 [EDITORIAL OEP\_02\_11\_12]**

En el texto 19, nos gustaría destacar el uso del argumento como parte del proceso evaluativo en el texto. Este texto inicia su argumento utilizando contraposición o contraste para establecer una evaluación del gobierno español (y, más concretamente, del Partido Popular). Antes de entrar en el análisis de este Texto, quisiéramos explicar por qué creemos que el argumento es una parte importante de la forma en que los escritores evalúan en el texto. La argumentación textual estructura las justificaciones de las reivindicaciones y críticas en un texto. Estas pueden ser implícitas o explícitas dentro de un texto. Por lo tanto, tomemos una afirmación como la siguiente:

- (1) El PP, que criticó la tasa cuando se instauró en Cataluña, la aplica ahora en Madrid

Contiene una crítica implícita porque contrapone dos acciones del Partido Popular (PP) y por lo tanto destaca una contradicción en el comportamiento de este partido político. De hecho, todo el texto se estructura inicialmente en torno a la idea de que el cobro a los ancianos/pensionistas por sus medicamentos fue una vez criticado por el PP en Cataluña, pero ahora se aplica en Madrid. El editor utiliza este argumento como base para atacar cualquier tipo de privatización de los servicios médicos en Madrid.

La argumentación puede utilizarse para guiar al lector hacia el punto de vista del escritor o editor. El periodista es la voz principal en el texto a pesar de que hay otras voces como: Esperanza Aguirre, Ignacio González, Javier Fernández-Lasquetty, el gobierno de Rajoy, etc. Todos ellos son políticos pertenecientes al PP, conservador. Con algunos ejemplos más, podemos ver cómo el escritor de este texto establece la contraposición en sus argumentos:

- (2) la lanza una vez que el gobierno de Rajoy ya ha puesto en marcha el medicamentazo
- (3) los madrileños acumulan en sus casas 45 millones de medicamentos

- (4) no tiene afán recaudatorio, sino disuasorio, para reducir el consumo
- (5) sí calcula que se recaudarán 83 millones de euros adicionales

A medida que el texto avanza, la crítica se vuelve más abierta:

- (6) Es una medida que va a afectar fundamentalmente a los jubilados y a los enfermos crónicos...
- (7) ... hay que hacerlo extremando las precauciones para no penalizar a los sectores más débiles.
- (8) González presenta un ambicioso plan privatizador de la gestión de la sanidad pública madrileña.
- (9) Hay razones para la inquietud.
- (10) La privatización de servicios públicos esenciales requiere en todo caso una estricta y transparente supervisión pública que hasta ahora no se ha acabado de establecer.
- (11) ... con este tipo de medidas se hace recaer el precio de la crisis en el ciudadano de a pie y en los más vulnerables —jubilados y enfermos— a los que además se les presenta como responsables de abusar del sistema.

#### **Texto 20 [EDITORIAL OEP\_02\_12\_12\_[2]]**

El tema del texto 20 son los bancos malos. Un banco malo, o en términos económicos, fondo para la reestructuración de los activos financieros, es una entidad financiera que compra los llamados activos tóxicos (activos contabilizados en el balance de un banco por un valor sobrevalorado) a fin de salvar el sistema financiero. Un banco malo es donde un banco puede aparcar sus activos tóxicos para que el banco limpie sus balances. El título del texto es tortuoso: *Dudoso 'banco malo'* y anuncia el tono del artículo que un banco malo siembra dudas. Hay varias cadenas léxicas desde el principio hasta el final del texto en este sentido.

- *Dudoso-suscita incertidumbre-suscita dudas-dificultades logísticas importantes-poco hay organizado-superficialmente optimista-está hilvanado a medias-eventuales pérdidas-déficit-deuda-adolesce de otra inconsecuencia importante-tampoco es muy consecuente-mal podría hacerlo-créditos dudosos-no contribuyen a infundir confianza.*

A estas cadenas léxicas, obviamente hay que añadir otros aspectos lingüísticos del texto como estas dos cláusulas subordinadas (*una sociedad que se haga cargo de los activos depreciados del*

*ladrillo que lastran los balances bancarios*) o la oración (*En teoría, la idea es correcta*) que señala que el redactor va a criticar la idea de dicha sociedad (el banco malo). Sin embargo, vamos a enfocar nuestra atención exclusivamente en estas cadenas léxicas ilustrando su efecto acumulativo en el texto.



Figura 4.19: El 'banco malo'

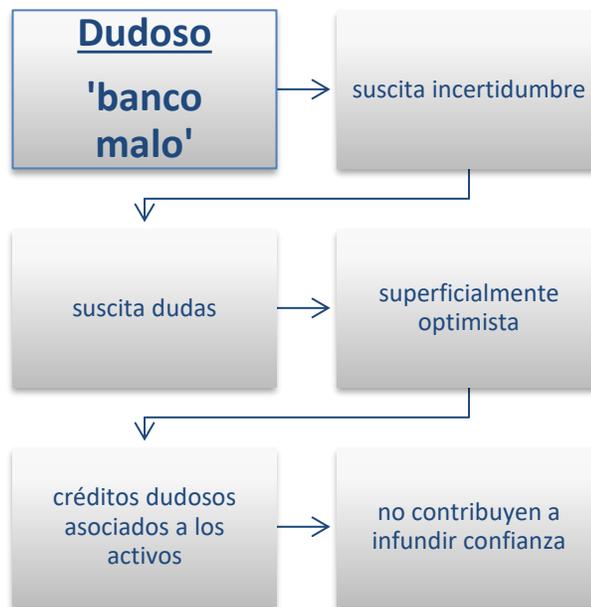


Figura 4.20: Dudoso



Figura 4.21: El 'banco malo' es dudoso porque produce pérdidas, déficit, deuda



Figura 4.22: El 'banco malo' es dudoso porque es difícil gestionar

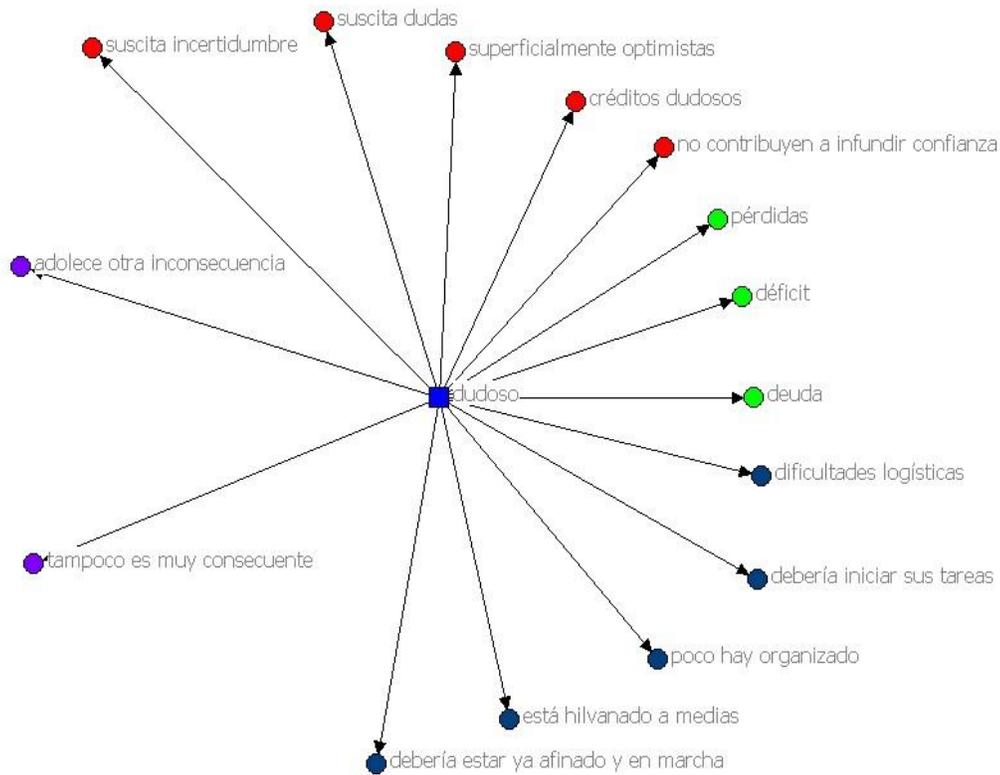


Figura 4.23: Las dudas que suscita el 'banco malo'

#### 4.7 Conclusión

La evaluación expresa una actitud hacia una persona (un actor social), una situación, un proceso, un producto y alguna otra entidad que es subjetiva y que se sitúa dentro de un sistema de valores de la sociedad. El uso del lenguaje evaluativo para indicar una actitud hacia algo es importante en los actos de habla socialmente significativos como la persuasión y la argumentación (que hemos encontrado en los textos de nuestro corpus, CCF). En estos textos, la evaluación se expresa frecuentemente de manera acumulativa e implícita. Estas características de acumulación e implícidad significan que el lenguaje evaluativo es más adecuado para un enfoque basado en texto en el que un análisis detallado y una lectura profunda pueden ayudar a explicar el lenguaje evaluativo en el texto, como hemos visto en la última parte de este capítulo a través de nuestro análisis minucioso de 20 textos. Los métodos cualitativos basados en texto complementan los métodos de lingüística de corpus, que nos proporcionan detalles

cuantitativos y facilitan la observación de patrones del lenguaje mediante el uso de programas de concordancias. En este capítulo, hemos visto algunos de estos patrones del lenguaje:

X = oración de infinitivo (objeto de evaluación) **es** Y (valor de la evaluación).

**Revalorizar las prestaciones en función del elevado IPC sería una equivocación**

**es una equivocación autorizar una subida de las pensiones**

Nos acercamos al final de esta tesis y me gustaría hacer algunas observaciones sobre el análisis de líneas de concordancia del lenguaje evaluativo. Una concordancia nos abre una ventana sobre el lenguaje en el corpus para que podamos ver patrones léxicos y gramaticales comunes de co-ocurrencia. Cuando analizamos una línea de concordancia, necesitamos ser claros sobre la naturaleza de los elementos lingüísticos que intervienen (palabras/lemas, unidades múltipalabra, unidades gramaticales o unidades de significado) y que implica decidir el número de elementos lingüísticos que se van a analizar (por defecto en Wordsmith [Scott, 2012] es 5 a la izquierda y 5 a la derecha de la palabra nodo, pero se puede aumentar hasta 25 en cada parte). También tenemos que tener en cuenta el número de veces que un patrón lingüístico se observa antes **de que cuente**. En otras palabras, necesitamos saber **lo que cuenta**. En principio, una expresión lingüística debe observarse con más frecuencia de lo esperado.

Una de las críticas que se pueden hacer al uso de líneas de concordancia para analizar texto es que se toman esas líneas fuera de contexto. Como la evaluación es altamente dependiente del contexto, esto puede ser una debilidad. Sin embargo, preferimos pensar en la evaluación en diferentes niveles de texto y contexto. En el nivel de oración, las líneas de concordancia son útiles. De hecho, esperamos haber establecido ya en esta tesis y, especialmente en este capítulo, que el lenguaje evaluativo funciona a diferentes niveles de organización del discurso:

- Léxico-gramatical (palabra, frase, cláusula, oración): operadores lingüísticos tales como adjetivos, sustantivos, expresiones idiomáticas, negación, modalidad e intensificadores.
- Nivel textual: conectores discursivos, estructura del discurso, figuras retóricas (lenguaje figurado que abarca una variedad de fenómenos, que incluye la metáfora, el oxímoron, la ironía y el sarcasmo) y relaciones retóricas tales como el contraste y la condición.
- Nivel contextual (contexto socioeconómico): valores culturales y sociales, conocimiento de sentido común, dependencia del dominio, género y limitaciones generales de espacio/tiempo.

Hemos visto cómo la evaluación está presente en todos los textos, incluso donde no hay signos explícitos de esa evaluación. Ejemplos de evaluación se pueden encontrar no sólo dentro de tramos cortos de texto, sino sobre tramos más largos del discurso.

Como hemos comprobado, hay una prosodia semántica de los actos de evaluación que tiene una función discursiva en unidades extendidas de significado en la escritura periodística. Los ejemplos de evaluación contribuyen a dar a los artículos periodísticos textura y estructura. Por otra parte, los ejemplos de evaluación identificados tienden a aparecer en los textos de una manera acumulativa. En los textos, hay hilos y ondas de evaluación que impregnan el texto con juicios de valor y de carácter personal.

Hemos visto cómo los medios de comunicación juegan un papel clave como instituciones sociales que moldean la cultura, la política y la vida social de un país y reflejan cómo éstas se forman y expresan. Nos encontramos con que los hechos y los juicios de los artículos son inseparables de los eventos a los que se refieren (Núñez Ladevèze, 1995).

Se ha orientado este estudio al análisis de la evaluación en textos periodísticos aplicando técnicas a partir de ejemplos de uso del CCF (corpus-based + corpus-driven).

El uso de métodos analíticos informáticos (listas de palabras, colocaciones, líneas de concordancias) facilita el procesamiento de grandes cantidades de datos y, por tanto, la recopilación de muestras representativas de uso del lenguaje. Sin embargo, la interpretación manual de los resultados obtenidos por métodos estadísticos revela casos de evaluación tanto de forma explícita como implícita cuando el análisis se extiende de los rasgos individuales más notables a patrones de elementos discursivos de mayor extensión.

Descubrimos una serie de ideas que el autor quiere transmitir en su mensaje, cargadas de una elevada densidad semántica y que el lector captará de manera inconsciente. Es así como a través de la negociación escritor/lector se forja un sentimiento colectivo que va filtrando en diferentes grupos sociales.

## CAPÍTULO 5: CONCLUSIONES Y REFLEXIONES DE CIERRE

En este capítulo realizaremos una reflexión a modo de sumario sobre las premisas teórico-prácticas en que se ha basado el estudio que conforma la presente Tesis Doctoral. Dichas reflexiones versarán en torno a tres aspectos:

- Hipótesis de partida
- Objetivos: su consecución
- Limitaciones y perspectivas de futuro

### 5.1. Hipótesis

Como vimos en la Introducción a nuestro trabajo, seis hipótesis constituyen el punto de partida e hilo conductor del mismo. A continuación, se exponen cómo hemos ido dando respuesta a cada una de ellas, y por ende, en qué medida se ha logrado alcanzar los objetivos – generales y específicos - que nos marcamos desde los comienzos de la investigación.

En primer lugar, hemos trabajado en el análisis del Corpus de la Crisis Financiera siguiendo las diferentes etapas de la metodología propia de la lingüística de corpus, pudiendo comprobar lo complejo que resulta cuantificar y enumerar todos los casos de lenguaje evaluativo presentes en el discurso periodístico ya que el lenguaje evaluativo se distribuye por todo el texto, es acumulativo y se manifiesta de formas muy diversas.

En cuanto a la segunda hipótesis, hemos podido descubrir una cantidad ingente de recursos retóricos (hipérbole, ironía, metáfora, oxímoron, paralelismo, repetición, símil) a los que recurre el periodista para opinar sobre los acontecimientos sociales. En ocasiones, se trata de un lenguaje coloquial, desenfadado y, en otras ocasiones, sofisticado, casi literario (uso del latín y referencias al mundo clásico, la biblia, por ejemplo).

Hemos comprobado que, efectivamente, nuestra tercera hipótesis se ha confirmado: el lenguaje evaluativo en los artículos periodísticos está presente en diferentes niveles (palabra, frase, cláusula, oración, e incluso como hilo conductor en todo un texto). Hasta la intertextualidad (*delenda est autonomia*) puede tener una función evaluativa como hemos visto en la segunda hipótesis.

La cuarta hipótesis refiere a nuestra metodología de trabajo, nuestro sistema siguió la propuesta de Tognini-Bonelli (2001: 85), conocida como ‘corpus-driven linguistics’, en la que el propio corpus de artículos de prensa económica actuó como informante y nos proporcionó a primera vista evidencias que, por su carácter especial, nos llevaron a profundizar en su análisis, lo cual nos permitiría establecer conclusiones sobre determinadas particularidades del lenguaje evaluativo de la prensa española (como hemos visto antes, la abundancia de figuras retóricas).

La metodología nos ha permitido detectar cuáles y quiénes son las instituciones y los actores sociales objetos de la evaluación por parte de los periodistas. También, hemos podido descubrir aspectos sobre los procesos más relevantes (saneamiento de la banca, reformas financieras y laborales, los recortes, el paro) de la crisis financiera en España.

En líneas generales, el análisis del Corpus de la Crisis Financiera en sus diferentes etapas ha proporcionado ejemplos claros sobre la complejidad del lenguaje evaluativo, cargado de recursos estilísticos y matices, y nos ha permitido profundizar en la cobertura periodística sobre esta crisis social que tanto ha afectado a la sociedad española en general.

## 5.2 Consecución de Objetivos

1. Objetivo general del estudio: Realizar una radiografía lingüística de la crisis financiera. A continuación ofrecemos algunos ejemplos obtenidos del CCF de cada uno de los trimestres del año 2012. Como se puede apreciar, el análisis manual de las líneas de concordancia nos ha proporcionado claras muestras de los recursos retóricos utilizados por los periodistas para filtrar de forma clara, directa y, en ocasiones, coloquial su opinión entre los lectores. Podemos observar de forma esquemática algunos de los mecanismos empleados para forjar ideologías.

| EJEMPLO DE USO-1erT   | FIGURA RETÓRICA |
|---|-----------------|
| merece comentario aparte Fátima Báñez, primera ministra de Trabajo del rajoyato no se parece a Arenas | Ironía          |
| Báñez se está fingiendo adormidera  | Metáfora        |
| Draghi advierte del peligro de dar marcha atrás en las metas de déficit                               | Perífrasis      |
| Draghi hombre prudentísimo al que hemos visto conduciendo sin cinturón                                | Ironía          |
| Mario Draghi conectó la manguera de préstamos a tres años   | Metáfora        |
| el manguerazo de liquidez que el organismo que preside Mario Draghi está dando a las entidades        | Metáfora        |

|  |                                     |
|--|-------------------------------------|
| Ni hay Vicepresidencia, ni hay De Guindos, ni el Gobierno tiene un plan, ni Montoro es   | Repetición                          |
| Guindos presenta su plan económico Cuarenta días y cuarenta noches ha tardado  | Hipérbole/<br>Bíblica               |
| De las Cortes de Cádiz viene ese galgo. De Guindos no se caerá de su apellido  | Refrán                              |
| ¿Logrará Merkel domesticar la animalidad de los mercados?  | Pregunta<br>retórica/Alegoría       |
| si quienes vencen en el pulso son los halcones de Merkel, el sacrificio será de más de 40.000 millones   | Metáfora                            |
| el actual matrimonio de conveniencia entre Merkel y Sarkozy también podría empezar a hacer agua  | Metáfora                            |
| el síndrome Merkel parece haber contagiado a Montoro   | Ironía/Metáfora                     |
| Ahora Montoro anuncia poco menos que un revival de la Transición para cuadrar números  | Comparación                         |
| ha esbozado Cristóbal Montoro y que podríamos traducir como «uno para todos y todos para uno»  | Intertextualidad<br>(Literaria)     |
| Para Cristobal Montoro probablemente interpreten el Nessun dorma   | Intertextualidad/<br>Ironía/Musical |
| Montoro está intentando escribir derecho con renglones torcidos  | Refrán                              |
| le pone en bandeja a Rajoy una oportunidad única para aplicar la guadaña del recorte   | Metáfora                            |
| la estrategia política de Rajoy es, de esta manera, una suerte de ducha escocesa   | Comparación                         |
| Me resultó a la vez divertido y útil comparar a Zapatero con el doctor Dulcamara que vende su falso curativo   | Comparación                         |
| Compadezco a Rajoy. Ganó para perder   | Antítesis                           |
| Ayer Rajoy gobernó. Bien, mal o regular, pero tomó una decisión importante   | Ironía                              |
| Rajoy decidió nombrarse a sí mismo vicepresidente económico. «Roma locuta, causa finita»   | Refrán                              |
| la excavadora de Rajoy no puede dragar las cloacas de la corrupción  | Metáfora/ironía                     |
| Ningún cangrejo ha caminado nunca con tanta rapidez hacia atrás. Zapatero y sus ministros deberían ser incorporados al Manual de zoología fantástica | Metáfora/Ironía                     |
| Lo malo de Zapatero no es malo porque lo haya hecho él sino porque era malo  | Repetición                          |
| Lo políticamente correcto es denostar la gestión de José Luis Rodríguez Zapatero al frente del Gobierno  | Ironía/contradicción                |
| como quedó bien demostrado en la etapa de Zapatero. Los parados no son hoy ejército de reserva o carne de cañón                                      | Comparación                         |

Tabla 5.1: Figuras retóricas 1T del CCF

| EJEMPLO DE USO-2ºT  | FIGURA RETÓRICA |
|---|-----------------|
| Y este antipático de Draghi que no quiere aflojar y encima nos falta  | Metáfora        |
| seguir golpeando las puertas del castillo del BCE, donde Draghi tiene peroles de aceite hirviendo preparados para lanzarlos | Hipérbole       |

|   |                                |
|---|--------------------------------|
| La institución presidida por Mario Draghi realizará mañana una operación de drenaje de liquidez   | Metáfora                       |
| Draghi debe actuar con otro manguerazo de liquidez  | Metáfora                       |
| Draghi de momento descarta reabrir la barra libre a tres años de liquidez   | Metáfora                       |
| la CAM a ser calificada como "lo peor de lo peor", por el exgobernador Miguel Ángel Fernández Ordóñez   | Comparación                    |
| y el menos malo, que era MAFO (Miguel Ángel Fernández Ordóñez)  | Comparación                    |
| Se trata de meterle un cuerno al gobernador Miguel Ángel Fernández Ordóñez, al que se atribuye una actitud pusilánime en la crisis financiera               | Metáfora                       |
| Un gobernador entre dos fuegos Fernández Ordóñez, se ha convertido ahora en el pimpampum del PP   | Contraposición/<br>Onomatopeya |
| El equipo de Goirigolzarri ha llegado a la cifra después de analizar las tripas de la entidad   | Metáfora                       |
| Goirigolzarri ha tenido un aterrizaje forzoso en Bankia   | Metáfora                       |
| A su sucesor, José Ignacio Goirigolzarri, se le supone un mejor conocimiento del oficio de banquero   | Ironía/comparación             |
| Guindos tiene hora con Mario Draghi como quien la tiene con un médico   | Comparación                    |
| Guindos atribuye la escalada de la prima de riesgo (el sobreprecio...)  | Perifrasis/Retruecano          |
| Hollande, que tiene pinta de profesor de guardería venido a menos,  | Ironía                         |
| En solo unas semanas, Hollande ha pasado de ser Flanby (flan de sobre) a ser un peligro para Europa   | Comparación                    |
| puñetero martirio económico, le mandaría el dossier a personas como Christine Lagarde, directora del FMI, que debe de observar desde muy lejos el desamparo | Ironía                         |
| Buena jugadora, Angela Merkel consigue vendernos las migajas  | Ironía                         |
| En manos de la muchacha, apodo con el que Kohl se refería por aquel entonces a Merkel, su antigua protegida   | Ironía                         |
| Lo de Monti fue como el codazo de Tassoti a Luis Enrique en el Mundial de 1994  | Comparación                    |
| Los tres caballeros (Monti, Rajoy y Hollande)   | Metáfora                       |
| Lo cierto es que, según el propio Mario Monti, la lucha diaria consiste en driblar al abismo  | Metáfora                       |
| oye que pedalee Monti, ya que decís que es tan listo  | Metáfora                       |
| Los hombres de negro», como calificó Montoro a los enviados de la troika  | Metáfora                       |
| La tijera de Montoro  | Metáfora                       |
| Montoro no es autoritario, parece un liberal de tensión baja. Habló ayer, con cierta resonancia nasal, en el Congreso                                       | Contradicción                  |
| Como Sinuhé el egipcio La actitud de Merkel con Rajoy me recuerda a la relación de unos personajes de la novela de Mika Waltari,                            | Comparación                    |
| Rajoy intenta hacer virtud de la necesidad  | Ironía                         |
| Rajoy vendió la piel del oso antes de cazarlo   | Metáfora                       |

Tabla 5.2: Figuras retóricas 2T del CCF

| EJEMPLO DE USO-3ºT  | FIGURA RETÓRICA       |
|---|-----------------------|
| las condiciones creo que no van a ser, en el caso de que se llegue a ese memorando, diferentes a las que ya se conocen", señaló Almunia | Perífrasis            |
| Almunia advierte de que la incertidumbre sobre el rescate aumenta los riesgos   | Redundancia           |
| la ya famosa frase que Draghi pronunció la semana pasada. "Haré todo lo necesario para sostener el euro"                                | Ironía                |
| Berlín envía señales tibias de apoyo a Draghi   | Metáfora/metonimia    |
| Pero si Draghi no despliega este jueves toda la artillería  | Metáfora              |
| Las altísimas expectativas que Mario Draghi había generado la semana pasada se han disuelto como un azucarillo                          | Metáfora              |
| Draghi tiró la piedra la semana pasada en Londres al asegurar: "El BCE hará todo"   | Metáfora/refrán       |
| Draghi nadó en la ambigüedad  | Metáfora              |
| Con ustedes Mario Draghi, el hombre más poderoso de Europa  | Ironía                |
| todo el poder para Draghi, una suerte de nuevo Richelieu de la política   | Ironía                |
| Goirigolzarri: "Lo atípico es formular cuentas sin borrador de auditoría"   | Ironía                |
| Goirigolzarri, pregunta a ver si hay suerte a un espejo mágico  | Metáfora              |
| El desfile para solicitar los rescates lo iniciaron las entidades financieras, con Goirigolzarri en el papel de gran majorette          | Ironía                |
| la expectativa de que el Gobierno de François Hollande no pueda evitar recortes presupuestarios   | Eufemismo             |
| la tríada católica de Monti, Rajoy y Hollande contra la campeona  | Ironía                |
| Hollande ha querido aparecer como un presidente "normal"  | Ironía                |
| Merkel hace honor a su viejo apodo —la canciller de teflón  | Ironía/metáfora       |
| Angela Merkel llega a Madrid con más fuerza de la que tenía Hitler  | Comparación           |
| Angela Merkel se sigue moviendo entre la bipolaridad  | Ironía                |
| ...que cada palo aguante su vela", señaló el ministro Cristóbal Montoro a finales de mayo   | Refrán                |
| Montoro tendrá sus 'hombres de negro'   | Metáfora              |
| Rajoy hizo un extraordinario esfuerzo por hablar como si la realidad no existiera   | Metáfora              |
| En la partida de póquer entre Mariano Rajoy y Mario Draghi habrá que esperar a ver quién desvela antes sus cartas                       | Metáfora              |
| Hasta que se muestren los naipes, la pelea de Rajoy es demostrar que basta con las condiciones ya impuestas                             | Metáfora              |
| "Necesito que se me ayude en la batalla europea", pidió Rajoy a la oposición"   | Metáfora              |
| En 'Esperando a Rajoy' continúa con la comparación médica   | Ironía                |
| Consideran a Rato "el problema dentro del problema"   | Reiteración           |
| Lo que pareció su salvación, fue su tumba. Rodrigo Rato, expresidente de Bankia, celebró con cava la salida a Bolsa de la entidad       | Antítesis/comparación |
| Otra pregunta impagable: "Señor Rato, ¿cuántas preferentes de Bankia tiene usted?"  | Pregunta retórica     |

Tabla 5.3: Figuras retóricas 3T del CCF

| EJEMPLO DE USO-4ºT  | FIGURA RETÓRICA             |
|---|-----------------------------|
| Fátima Báñez esa mujer piadosa que se acoge en su trabajo a la autoridad del cielo ("De la Virgen del Rocío siempre llega un capote") | Metáfora                    |
| Draghi se mostró irritado con el alemán: "¿A qué vienen esas sospechas porque soy italiano?"  | Pregunta retórica           |
| la ponencia de Draghi fue "muy convincente". "Se ha revelado como un prusiano del Europa del sur"                                     | Comparación                 |
| Guindos advierte: "Si se pide el rescate, ya no hay marcha atrás"   | Paradoja                    |
| Guindos, el peor ministro de Economía de Europa en la lista de 'Financial Times   | Comparación                 |
| Esto sí que era un algoritmo de altos vuelos. Guindos se ha convertido así en alguien como el doctor Hoffmann                         | Ironía/Comparación          |
| Lagarde, por su parte, bromeó con el apelativo "hombres de negro"   | Metáfora                    |
| Angela Merkel está dispuesta a poner una pica más en Flandes:   | Metáfora                    |
| Aunque Angela Merkel intentó poner buena cara el miércoles en su visita a Downing Street  | Ironía/metáfora             |
| Angela Merkel, inquieta porque una gripe francesa acabe contagiando a Berlín  | Ironía/metáfora             |
| Merkel, a la que llama Merkiavelli y compara con el Príncipe  | Comparación                 |
| ¿Cómo va a creer en Rajoy y en sus sinsorgos crucigramas, o en Merkel, a la que sólo ella se importa?                                 | Pregunta retórica           |
| El equipo de Montoro ha aprovechado esa debilidad para jugar al palo y la zanahoria   | Metáfora                    |
| ¿Es Montoro el ministro de Hacienda de IU?  | Pregunta retórica<br>Ironía |
| Alí Babá y los 659 evasores españoles puso la lista en manos de Cristóbal Montoro; si no trinca a los evasores es porque no le dejan  | Metáfora                    |
| El objetivo de Obama era suturar las heridas de un país dividido por las disputas   | Metáfora                    |
| Las cosas no han salido como Obama soñó. Han sido cuatro años de infarto  | Ironía/<br>Comparación      |
| Obama se juega una baza decisiva. Debilitar y dividir a los montaraces republicanos   | Metáfora                    |
| Esperando a Rajoy, la versión española de 'Esperando a Godot', según el NYT 'The New Y  | Intertextualidad            |
| Zapatero esparció las siembras de Caín sobre los surcos de Cataluña   | Metáfora                    |

Tabla 5.4: Figuras retóricas 4T del CCF

A continuación resumimos algunos de los recursos de la lengua española utilizados para plasmar la evaluación en diferentes niveles del texto e incluso los aquellos utilizados para expresar sentimientos y valores, para lo cual el periodista y lector deberán acudir a su conocimiento cultural y del mundo.

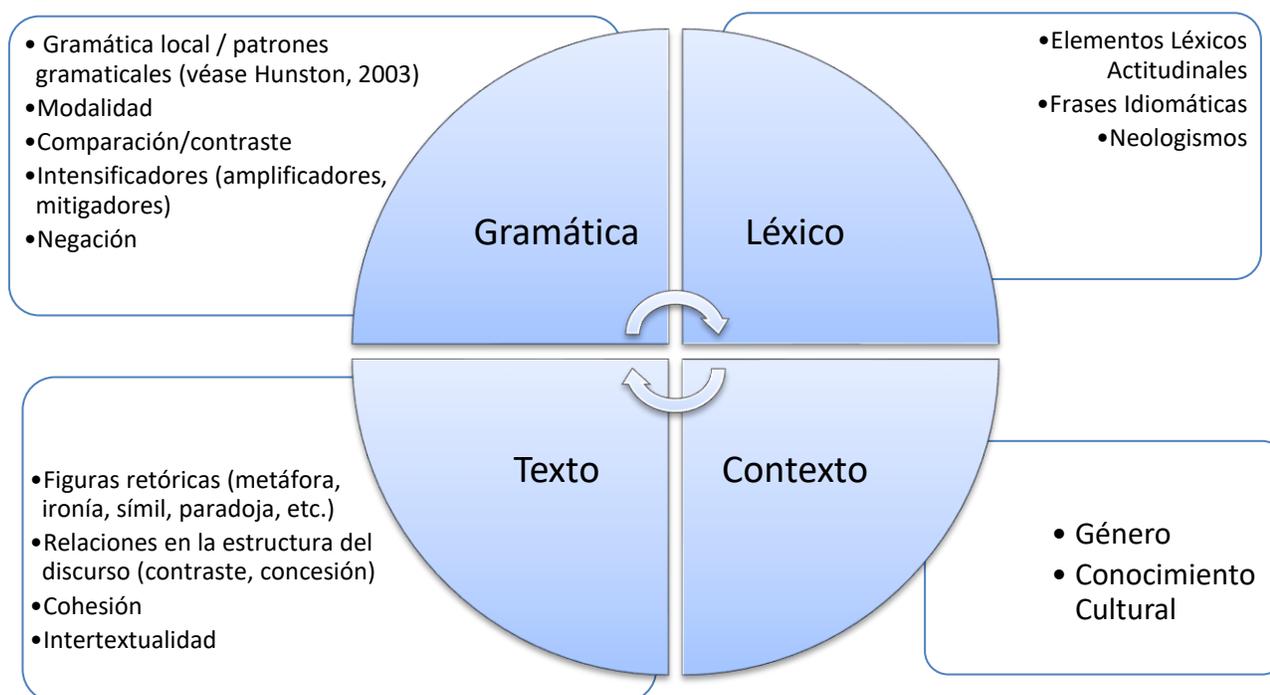


Figura 5.1: Modelización del lenguaje evaluativo

## 2. Objetivos específicos:

Podemos afirmar que los objetivos específicos que nos marcamos para el análisis del corpus y para el diseño de la metodología se han alcanzado, tal y como hemos podido comprobar en el capítulo 4 capítulo y como resumimos a continuación:

- Se ha recopilado y diseñado una base de datos de artículos periodísticos sobre la crisis financiera en España del año 2012 (véase Anexo 1\_CCF)
- Disponemos de datos de cada trimestre por periódico (*El País* y *El Mundo*) y sección (Economía y Opinión) (véase páginas 138-139)
- Se han extraído listas de palabras de cada trimestre por periódico y sección (véase Anexo II\_Lista de palabras)
- Se han diseñado tablas comparativas de las palabras más frecuentes por trimestres (véase Anexo II\_Lista de palabras)

- Hemos detectado cuáles son los 12 actores sociales más frecuentes en el CCF por trimestres y extraído sus colocaciones y líneas de concordancia. Hemos representado sus relaciones de forma reticular (Anexo III\_Actores Sociales\_Colocaciones\_Concordancias)
- Los adjetivos que acompañan a cada actor social figuran en el Anexo IV\_Actores Sociales\_Adjetivos
- Se han establecido las categorías semánticas de los adjetivos. Se recogen en el documento del Anexo IV\_Adjetivos\_1T\_2T\_3T\_4T\_Categorías semánticas
- Se ha llevado a cabo un análisis manual de las líneas de concordancia asignando etiquetas de ejemplos de lenguaje evaluativo. Hemos recopilado las etiquetas más frecuentes y las hemos transferido a un documento Excel con ejemplos (Anexo V\_Actores Sociales\_Evaluación y Anexo VI\_Etiquetas con ejemplos)
- En la última etapa del estudio analizamos textos completos (editoriales): su macroestructura y lenguaje evaluativo (Anexo VII\_Análisis manual del lenguaje evaluativo en textos completos)

### **5.3 Limitaciones y Perspectivas de Futuro**

#### **5.3.1 Limitaciones**

En este estudio hemos encontrado dos limitaciones principales: 1) la naturaleza del lenguaje evaluativo; 2) la naturaleza de nuestro corpus. El análisis de la evaluación en el texto plantea desafíos metodológicos significativos, que surgen principalmente del hecho de que:

- a) los significados evaluativos pueden expresarse a través de una amplia gama de recursos léxico-gramaticales;
- b) pueden abarcar varias palabras;
- c) el contexto y el co-texto desempeñan un papel clave en la determinación del significado evaluativo de palabras o frases;
- d) la interpretación de la evaluación en el texto depende de la posición del lector / analista y, por tanto, es necesariamente subjetiva.

Estas complejidades hasta ahora han limitado seriamente el desarrollo de la descripción de este fenómeno a partir de corpus (*corpus-driven*). La observación de Hunston (2004: 157) de que “much evaluative meaning is not obviously identifiable, as it appears to depend on immediate context and on reader assumptions about value” es particularmente pertinente aquí.

Al diseñar un corpus, es importante que el corpus se desarrolle tomando en cuenta ciertos criterios. En nuestro caso, los criterios fueron los siguientes. Los artículos de prensa debían ser de:

1. *El País* o *El Mundo* como representantes de la política bipartidista española
2. el 1 de enero de 2012 al 31 de diciembre de 2012
3. artículos de opinión y de la sección de Economía (sobre la crisis financiera en España)

Nuestro corpus es un corpus especializado. Un corpus especializado contiene textos limitados a una o más áreas temáticas, dominios y temas. Un corpus de estas características es útil para estudiar cómo se usa el lenguaje especializado. En nuestro caso, se ha estudiado el uso del lenguaje evaluativo en el corpus especializado. La ventaja de este corpus especializado es que es *representativo* del género de artículos de opinión de periódicos españoles y de un período específico en la historia socioeconómica de España. En la lingüística de corpus, como en otras ciencias de datos, la *representatividad* es una relación que se mantiene entre una muestra y una población. Una muestra *representa* una población mayor si la muestra se obtuvo de manera rigurosa y sistemática en relación con una población bien definida. Si la muestra no es representativa de esta manera, es una *muestra arbitraria*. En otras palabras, no se obtuvo de manera rigurosa y sistemática en relación con una población bien definida. La representatividad de nuestro corpus nos permite examinar la muestra y sacar conclusiones sobre la población. Sin embargo, debido a que es un corpus bien definido, resulta difícil extrapolar dichas conclusiones a la lengua española en su conjunto. Los resultados representan una población lingüística específica. Cuanto más rigurosamente se defina un corpus, más difícil resulta extrapolar los resultados.

### 5.3.2 Perspectivas de Futuro

A pesar de sus limitaciones, este estudio plantea un problema con el lenguaje evaluativo que necesita un planteamiento diferente. Requiere un pensamiento creativo y crítico. A continuación se presentan seis posibles áreas de estudio futuro que podrían resultar interesantes al tratar con la naturaleza del lenguaje evaluativo:

1. *Markedness* (Marcado/no marcado)
2. Flujo del sentimiento
3. Redes de sentimiento dinámicas
4. Argumentación
5. Detección de la intención
6. Actitudes evaluativas y retóricas

En primer lugar, comenzaremos con el concepto de *markedness*. Este concepto se refiere a la relación entre un par de términos antónimos o complementarios. Los primeros relatos formales sobre las relaciones de marcado sugerían que un elemento de un par está marcado físicamente mientras que el otro no está marcado; el marcado semántico sugiere que las palabras dentro del par se diferencian por la presencia o ausencia de una determinada propiedad (Jakobson, 1962; Levinson, 1983). El marcado morfológico puede ocurrir cuando una de las dos palabras se deriva de la otra a través, por ejemplo, de la afijación (*acceptable/inacceptable; cierto/incierto; feliz/infeliz; posible/imposible*). La “marca” adicional (en nuestros ejemplos, un prefijo) sobre un ítem de un par de opuestos estableció la noción de *markedness*. Los elementos no marcados tienen una amplia aplicabilidad contextual y aparecen con más frecuencia que sus equivalentes marcados (por ejemplo, “acceptable” = 22366 versus “inacceptable” = 10835; “cierto” = 559301 versus “incierto” = 11488; “feliz” = 209647 versus “infeliz” = 9089; “posible” = 678591 versus “imposible” = 194073) – Corpus del Español (<http://www.corpusdelespanol.org/>). Podemos concluir de esta discusión que una palabra marcada será menos frecuente que su equivalente no marcada y, en nuestros ejemplos, más negativa en su evaluación. De hecho, Halliday y James (1993) encontraron que las cláusulas positivas ocurren abrumadoramente con más frecuencia que las negativas, en una proporción de aproximadamente 9: 1. Parece, al menos en inglés, que las cláusulas negativas están marcadas y las cláusulas positivas son la opción no marcada. Por lo tanto, podría ser una hipótesis interesante comprobar la relación entre el marcado, la

frecuencia y la evaluación. La hipótesis básica sería que los ítems marcados y evaluativos son menos frecuentes y que el marcado está relacionado con la evaluación en que las opciones altamente marcadas son altamente evaluativas (sean evaluadas positivamente o negativamente).

Otra forma de modelar la evaluación en el texto sería el concepto de flujo de sentimiento. Dado un texto T que incluye oraciones  $\{s_1, s_2, \dots, s_n\}$ ; cada oración  $s_i$  tiene un conjunto de elementos lingüísticos  $X_i = \{x_1, x_2, \dots, x_n\}$ . Seleccionamos un clasificador que coge los elementos lingüísticos  $\{x_1, x_2, \dots, x_n\}$  como input y asigna etiquetas de polaridad de sentimiento como output. Hay oraciones ambiguas que el clasificador no es capaz de clasificar correctamente. Para mejorar esto, asumimos que debe haber un flujo de sentimiento dentro de un texto, que se rompa sólo bajo condiciones específicas. Así, oración tras oración, el editor de un artículo de opinión dice cosas positivas (flujo) pero luego una oración es negativa (flujo-violación), marcada bien por un término de contraste como *pero*, *sin embargo*, etc. o por un término negativo fuerte, un recurso retórico negativo, o los tres.

Un seguimiento lógico es la idea de tratar un texto como una red dinámica de sentimiento, donde los valores se calculan sobre la base de pesos dados a palabras individuales o diferentes elementos lingüísticos. Varias combinaciones de pesos producen varios tipos de dinámica, que simulan una amplia gama de efectos semánticos evaluativos. Esta variedad asegura la expresividad semántica de la dinámica. En particular, puede ser posible representar las sutiles propiedades evaluativas de adverbios y conectores con respecto al flujo de sentimientos. La dinámica también nos permite modelizar los efectos semánticos que surgen al comparar con un estándar.

El flujo de un texto depende en gran medida de la argumentación. El seguimiento de argumentos en el texto consiste en identificar su estructura argumentativa, incluyendo las conexiones entre el argumento y las relaciones de contra-argumento. La minería de la argumentación es un área relativamente nueva en el Procesamiento del Lenguaje Natural (Mochales y Moens, 2011; Peldszus y Stede, 2013; Hasan y Ng, 2014). Lawrence y Reed (2015) afirman que la argumentación se basa en una lista de marcadores discursivos que conectan premisas adyacentes divididas en dos grupos según sea su efecto sobre la argumentación:

marcadores de soporte y marcadores de ataque. Por ejemplo, las conectivas adversativas tales como *pero* y *aunque* conectan argumentos opuestos, mientras que las conjuntivas como *y*, *o*, y *entonces/pues* enlazan argumentos que persiguen el mismo objetivo. La argumentación se podría analizar utilizando la Teoría de la Estructura Retórica (Mann y Thompson, 1988) o la noción de proyección de Halliday (Halliday, 1994).

Otra manera de contemplar la evaluación en el texto puede ser el reconocimiento de la intención. La detección de la intención está en sus primeras etapas en el Procesamiento del Lenguaje Natural (Wang et al., 2015), pero la teoría de los actos de habla de Searle (1969, 1975) o el Principio de Cooperación de Grice (1975) pueden resultar útiles para estos propósitos.

Por último, nos gustaría hablar de actitudes evaluativas y retóricas que han surgido con fuerza de este estudio. Parece que, en ciertos géneros en particular, la actitud evaluativa y retórica del hablante / escritor impactará fuertemente en cómo se realiza la evaluación en el texto. En otras palabras, podemos detectar una cierta actitud retórica a través del lenguaje evaluativo, pero la actitud retórica también nos llevará a detectar qué lenguaje se está utilizando para expresar esta actitud. Las actitudes evaluativas suelen estar relacionadas con mecanismos retóricos. La indignación se expresa a través de exclamaciones y preguntas retóricas, la deferencia a través de marcadores específicos de enunciación, el escepticismo o el desprecio a través de la ironía. El tratamiento de estos recursos retóricos es uno de los problemas más difíciles a que se enfrenta el procesamiento del lenguaje natural.

Goebbels (1897-1945) fue un político alemán, ministro de la Propaganda de la Alemania nazi. A él se atribuye la frase *“una mentira repetida mil veces se convierte en una verdad”*. La máxima aún se pronuncia en las aulas de Periodismo y sirve de ejemplo de la influencia que puede ejercer un determinado mensaje sobre el receptor. Parece que los métodos democráticos tienen serias limitaciones. Están sesgados bajo el denominador común de la información superficial, en detrimento del conocimiento especializado, que no se comparte de manera extensiva.

Me licencié en Periodismo y llevo ejerciendo como periodista desde hace algo más de 30 años. En este tiempo he ido acumulando conocimientos y experiencias en los distintos medios de

comunicación en los que he trabajado. Esta tesis es el resultado del estudio lingüístico de las publicaciones periodísticas de la crisis financiera con el que poder aportar una explicación científica a los diferentes estilos utilizados por los periodistas para construir la realidad en los textos. Al escribir sobre lo que ha pasado en una rueda de prensa o en un acontecimiento el periodista se deja llevar por una serie de condicionantes que van conformar un estilo. De manera consciente o inconsciente alguno de estos elementos se van a repetir en el texto escrito de manera tan asidua que van a calar en el lector como un mantra. Aparecen con tanta frecuencia que llegan a convertirse en la única verdad para el lector.

Con la crisis económica ha sucedido lo mismo. Los periodistas, que han seguido de cerca a los actores de la crisis, han transmitido una realidad e interpretación de los hechos a través de ideas y conceptos que han creado un pensamiento circular sobre la crisis financiera. A la vista de las publicaciones en dos medios de referencia en España en el año 2012, *El País* y *El Mundo*, ambos con una línea editorial muy distante, y como representantes del bipartidismo de la época, la primera conclusión es que el mensaje que, *a priori*, debería ser diferente. Sin embargo, en términos lingüísticos lo que cuenta la prensa conservadora no difiere demasiado de lo que narra la considerada progresista.

La influencia de los medios de comunicación de masas en la sociedad es un asunto que ya abordó Eco (1965) en su obra *Apocalípticos e Integrados*. El autor lo analiza desde su estructura, su influencia y su desarrollo hasta complejos conceptos de semiótica, tratados siempre desde un punto de vista de una comunicación práctica. Para Eco (1965), son integrados aquellos que hacen una interpretación benévola la cultura de masas, como hizo en su día Marshall McLuhan con su teoría de la Aldea Global. Eco (1965) considera los medios de comunicación de masas como generadores de modelos y situaciones humanas que aunque no tienen conexión con la audiencia pueden llegar a entretenerla. Para los apocalípticos esta cultura de masas y sus medios de difusión son los que destruyen las características de cada grupo ya que en esta cultura el lector no se muestra conforme a lo que se le ofrece a través de los medios de comunicación. Por otro lado, los integrados tienen una visión más optimista de la situación, que denominan cultura de masas en donde los ciudadanos participan y se les tiene en cuenta. Sin embargo, esta cultura de masas la producen grupos con poder económico que obtienen beneficios y persiguen un mercado con ganancias y contenidos simples, entretenidos y que lleguen a la mayoría de los hogares del país.

La construcción y análisis del CCF nos ha permitido obtener conclusiones sobre las decisiones tomadas por los autores en la presentación de ideas y eventos a través de los textos.

El CCF ha demostrado el sentido evaluativo de los textos ya sea abierta o implícitamente expresados en los textos transmitidos por medio de figuras retóricas. Los recursos lingüísticos (repetición, la intertextualidad, metáfora) y el metadiscurso interpersonal son recursos intelectuales clave para convencer a los lectores individuales e incluso dar forma a las opiniones colectivas e ideologías de los medios de comunicación. Tras el análisis de los textos se puede decir que existe una 'semántica' de la crisis financiera en España en el discurso periodístico analizado con el que conducir a la sociedad.

**ANEXOS**

## Anexo 0\_Manuales de Estilo

- Manual\_de\_estilo\_El Mundo
- Manual\_de\_estilo\_El País

## Anexo I\_CCF

- Corpus\_CCF
  - EP
    - EP 1 TRIMESTRE
    - EP 2 TRIMESTRE
    - EP 3 TRIMESTRE
    - EP 4 TRIMESTRE
  - OEM
    - OEM 1 TRIMESTRE
    - OEM 2 TRIMESTRE
    - OEM 3 TRIMESTRE
    - OEM 4 TRIMESTRE
  - OEP
    - OEP 1 TRIMESTRE
    - OEP 2 TRIMESTRE
    - OEP 3 TRIMESTRE
    - OEP 4 TRIMESTRE
- Corpus\_de\_la\_crisis\_financiera. Hoja de cálculo Microsoft Office Excel

## Anexo II\_Listas de palabras

- Wordlist\_1T
- Wordlist\_2T
- Wordlist\_3T
- Wordlist\_4T

## Anexo III\_Actores Sociales\_Colocaciones\_Concordancias

- 1T\_Anexo III\_Actores Sociales\_Colocaciones\_Concordancias
  - Bñez
  - Cameron

- Draghi
  - Guindos
  - Merkel
  - Monti
  - Montoro
  - Obama
  - Rajoy
  - Rato
  - Santamaría
  - Zapatero
- 2T\_ Anexo III\_Actores Sociales\_Colocaciones\_Concordancias
    - Draghi
    - Fdez Ordóñez
    - Goirigolzarri
    - Guindos
    - Hollande
    - Lagarde
    - Merkel
    - Monti
    - Montoro
    - Rajoy
    - Rato
    - Zapatero
- 3T\_ Anexo III\_Actores Sociales\_Colocaciones\_Concordancias
    - Almunia
    - Draghi
    - Goirigolzarri
    - Guindos
    - Hollande
    - Méndez
    - Merkel
    - Montoro
    - Rajoy
    - Rato
    - Santamaría
    - Weidman
    - Zapatero
- 4T\_ Anexo III\_Actores Sociales\_Colocaciones\_Concordancias
    - Báñez

- Draghi
- Guindos
- Hollande
- Lagarde
- Merkel
- Montoro
- Obama
- Rajoy
- Rato
- Rehn
- Zapatero

Anexo VI\_Actores Sociales\_Adjetivos

- 1T\_Rajoy\_et al\_Adj
- 2T\_Rajoy\_et al\_Adj
- 3T\_Rajoy\_et al\_Adj
- 4T\_Rajoy\_et al\_Adj
- Adjetivos\_1T\_2T\_3T\_4T\_Categorías\_semánticas

Anexo V\_Actores sociales\_Colocaciones\_Concordancias

- 1T
  - Báñez
  - Cámeron
  - Draghi
  - Guindos
  - Merkel
  - Monti
  - Montoro
  - Obama
  - Rajoy
  - Rato
  - Santamaría
  - Zapatero
- 2T
  - Draghi
  - Fdez. Ordoñez
  - Goirigolzarri
  - Guindos
  - Hollande
  - Lagarde

- Merkel
  - Monti
  - Montoro
  - Rajoy
  - Rato
  - Zapatero
- 3T
  - Almunia
  - Draghi
  - Goirigolzarri
  - Guindos
  - Hollande
  - Mendez
  - Merkel
  - Montoro
  - Rajoy
  - Rato
  - Santamaría
  - Weidman
- 4T
  - Báñez
  - Draghi
  - Guindos
  - Hollande
  - Lagarde
  - Merkel
  - Montoro
  - Obama
  - Rajoy
  - Rato
  - Rehn
  - Zapatero

Anexo VI\_Etiquetas con ejemplos

- Etiquetas con ejemplos

Anexo VII\_Análisis manual del lenguaje evaluativo en textos completos

## BIBLIOGRAFÍA

- Ahmad, K., Davies, A., Fulford, H. y Rogers, M. (1994). What's in a Term? The semi-automatic Extraction of Terms from Text. *Translation Studies. An Interdiscipline*. John Benjamins, (pp. 267-278).
- Arrese, A. (2002). Prensa económica. De la Lloyd's list al wsj.com. Pamplona: Eunsa, (pp. 440). Recuperado de [http://www.unav.es/fcom/communication-society/es/resena.php?art\\_id=111](http://www.unav.es/fcom/communication-society/es/resena.php?art_id=111)
- Barlow, M. (1996). Corpora for Theory and Practice. En C. Chapella [ed.] *Computers and Applied Linguistics*. Rice University.
- Barnbrook, G. (1996). *Language and computers*. Edinburgh: Edinburgh University Press.
- Baroni, M., Bernardini, S., Ferraresi, A. y Zanchetta, E. (2009). The WaCky Wide Web: A collection of very large linguistically processed Web-crawled corpora. *Journal of Language Resources and Evaluation*. 43 (3) (pp. 209-226). DOI: 10.1007/s10579-009-9081-4.
- Beaugrande, R. de y Dressler, W. (1981). *Introduction to text linguistics*. Austin, TX: University of Texas Press.
- Beißwenger, M. y Storrer, A. (2008). Corpora of computer-mediated communication. En Lüdeling, A. y Kytö M. [eds.]. *Corpus linguistics. An international handbook, vol. 1*, Berlin/New York: Mouton de Gruyter, (pp. 292-308). Recuperado de <http://www.michael-beisswenger.de/pub/hsk-corpora.pdf>.
- Biber, D. (1995). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D., Conrad, S. y Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press (Cambridge Approaches to Linguistics).
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, (pp. 5-2, 97-116). Recuperado de: <http://dx.doi.org/10.1016/j.jeap.2006.05.001>.
- Bowker, L. y Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora*. London/New York: Routledge.
- Borrat, H. (1993). Hacia una teoría de la especialización periodística. En *Revista Anàlisi*. Facultad de Ciencias de la Información de la Universidad Autónoma de Barcelona, 15 (pp. 79-84). Recuperado de: <http://www.raco.cat/index.php/Analisi/article/view/41185>.
- Borrat, H. (2002) Paradigmas alternativos y redefiniciones conceptuales en comunicación periodística. En *Revista Anàlisi*. Facultad de Ciencias de la Información de la Universidad Autónoma de Barcelona, 28 (pp. 55-77). Recuperado de: <http://www.raco.cat/index.php/analisi/article/viewFile/15102/14943>

- Cabré, M. T. (1993). *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Editorial Empúries.
- Casasús, J.M., Núñez, L. (1991). *Estilo y géneros periodísticos*. Barcelona: Ariel.
- Cebrián, M. (2003). Globalidad comunicativa de la solidaridad y defensa del ciudadano frente a los poderes mediáticos. *Revista latina de Comunicación Social*, 53. Recuperado de: <http://www.ull.es/publicaciones/latina/200353cebrian.htm> [31-05-2015]
- Claridge, C. (2007). Constructing a Corpus from the Web: Message Boards. En M. Hundt, N. Nesselhauf y C. Biewer [eds.] *Corpus Linguistics and the Web*, Amsterdam: Rodopi. (pp. 87-108).
- Committee on Modern Journalism. VV. AA. (1967). *Periodismo Moderno*. México: Editorial Letras.
- Chomsky, N. (2002). *On Nature and Language*. Cambridge: Cambridge University Press.
- Chung, T. M. (2003). A corpus comparison approach for terminology extraction. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 9 (2) (pp. 221-246). DOI: 10.1075/term.9.2.05chu. Recuperado de: <http://www.benjamins.com/jbp/series/TERM/9-2/art/0004a.pdf>. [03-02-16]
- Church, K. y Hanks, P. (1990). Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, (pp. 16-1: 22-29) Recuperado de: <https://pdfs.semanticscholar.org/b413/0e315df099e981342542536edfe36f77a20f.pdf>
- Demarty-Warée, J. (1998). Le travail sur Corpus: une entrée par les paradigmes designationnels. *I jornades Catalanes per a necessitat específiques*. Barcelona: Universitat de Barcelona.
- Díaz Noci, J. y Salaverría Aliaga, R. (2003). *Manual de redacción Ciberperiodística*. Barcelona: Grupo Planeta.
- Diezhandino, P., Aguado, G., Carrera, P., Fernández, J., Martín, O. y Muro, I. (2012). *El periodista en la encrucijada*. Ariel. Colección Fundación Telefónica.
- Doyle, P. (2005). Replicating Corpus-Based Linguistics: Investigating Lexical Networks in Text. *Proceedings from the Corpus Linguistics Conference Series*, 1 (1). Recuperado de: <http://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2005-journal/Lexiconodf/COLING2005paper.pdf> [09/03/2016]
- Dovifat, E. (1959 y 1960). *Periodismo*, 2 vols. México: UTEHA.
- Drouin, P. (2003). Term Extraction Using Non-technical Corpora as a Point of Leverage. *Terminology*. (pp 9 -1: 99 -115). Recuperado de: [http://olst.ling.umontreal.ca/pdf/Terminology\\_2003.pdf](http://olst.ling.umontreal.ca/pdf/Terminology_2003.pdf)

- Druker, P. F. (1994). The Age of Social Transformation. En *The Atlantic Monthly*, 273 (11). Boston. Recuperado de:  
<http://www.theatlantic.com/past/docs/issues/95dec/chilearn/drucker.htm>
- Durán, I. (2011). Criterios específicos para la elaboración y diseño de los corpus especializados para la terminografía. En Carrió, M. L. y Candel, M. A. *Las tecnologías de la información y las comunicaciones: Presente y futuro en el análisis de corpora*. Actas del III Congreso Internacional de Lingüística de Corpus. Valencia: Universitat Politècnica de València, (pp. 43-50). Recuperado de: [http://www.lexytrad.es/assets/duran\\_2011d\\_0.pdf](http://www.lexytrad.es/assets/duran_2011d_0.pdf)
- Eco, U. (1965). *Apocalípticos e integrados*. Barcelona: Lumen.
- Eco, U. (1977). *Come si fa una tesi di laurea: le materie umanistiche*. Toscabili Bompiani, Milano. (2001) *Cómo se hace una tesis: técnicas y procedimientos de investigación, estudio y escritura*. Edición traducida por Lucía Baranda y Alberto Clavería Ibáñez. Barcelona: Gedisa.
- Fattorello, F. (1969). *Introducción a la técnica social de la información*. Caracas: Universidad Central de Venezuela.
- Faúndez, H. (2004). *Los límites a la libertad de expresión*. México: Universidad Nacional Autónoma de México. Recuperado de: <https://es.scribd.com/doc/308722836/Los-Limites-de-La-Libertad-de-Expresion-Hector-Faundez-Ledesma>.
- Fernández del Moral, J. (1993). *Fundamentos de la información periodística especializada*. Madrid: Síntesis.
- Fernández del Moral, J. (2004). El periodismo especializado: un modelo sistémico para la difusión del conocimiento. En: Fernández del Moral, J. (coord.). *Periodismo Especializado*. Barcelona: Ariel.
- Fletcher, W. H. (2004). Facilitating the compilation and dissemination of ad-hoc web corpora. En G. Aston, S. Bernardini y D. Stewart [eds.], *Corpora and Language Learners*. Amsterdam: John Benjamins. *Studies in Corpus Linguistics* 17 (pp 271-300). Recuperado de:  
<http://kwicfinder.com/FletcherTaLC5.pdf>
- Fontcuberta, M. y Borrat, H. (2006). *Periódicos: sistemas complejos, narradores en interacción*. Buenos Aires: La Crujía Ediciones.
- Fontcuberta, M. (1993). *La noticia. Pistas para percibir el mundo* Editorial. Barcelona: Editorial Paidós.
- Fontcuberta, M. (1997). Propuestas sistémicas para el análisis y producción de información periodística especializada. En Esteve, F. (coord.). *Estudios sobre información periodística especializada*. Valencia: Fundación Universitaria San Pablo.
- Fondevila Gascón, J. F. (2009). Relación entre multimedialidad, hipertextualidad e interactividad en la prensa digital española: análisis empírico. En *I Congreso Internacional "Sociedad Digital"*, Madrid, 26-28 de octubre.

- Francis, W. N. (1992). Language Corpora British Corpus. En J. Svartvik [ed.]: *Directions in Linguistics: Proceedings of Nobel Symposium 82* (Stockholm, 4-8 August 1991). Berlin/New York: Mouton de Gruyter, (pp. 17-32).
- González, G., Valderrama, J.C., Aleixandre, R., Alonso, A., De Granda, J.I. y Villanueva, S. (2006). Redes de coautoría y colaboración de las instituciones españolas en la producción científica sobre drogodependencias en biomedicina 1999-2004. *Trastornos Adictivos* 8 (pp.78-114). Recuperado de: [file:///E:/Users/Luc%C3%ADa/Downloads/13090606\\_S300\\_es.pdf](file:///E:/Users/Luc%C3%ADa/Downloads/13090606_S300_es.pdf)
- Gomis, L. (1974). *El medio media: la función política de la prensa*. Madrid: Hora h, Seminarios y Ediciones, S.A.
- Grice, P. (1975). Logic and conversation. En Peter Cole y Jerry L. Morgan [eds.], *Speech Acts. Syntax and Semantics, Volume 3*. Academic Press, New York, (pp. 41–58).
- Grishman, R. (1986). *Computational linguistics: an introduction*. Cambridge: Cambridge University Press.
- Halliday, M. A. K. (2004). *An Introduction to Functional Grammar*. [3ª Edición] London: Arnold.
- Halliday, Michael A. K., y James, Z. L. (1993). A quantitative study of polarity and primary tense in the English finite clause. En John Sinclair, Michael Hoey y Gwyneth Fox (Eds.), *Techniques of description: Spoken and written discourse*. New York: Routledge, (pp. 32-66).
- Hanneman, R. A. y Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California, Riverside. Recuperado de: <http://faculty.ucr.edu/~hanneman/> [11-04-2016]
- Hasan, K. y Ng, V. (2014). Why are you taking this stance? Identifying and classifying reasons in ideological debates. En *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Doha (pp. 751–762).
- Heidenreich, M (2003): Die Debate um die Wissensgesellschaft. En Böschen, Sthephan & Schulz-Schaeffer, Ingo [ed.]. *Wissenschaft in der Wissensgesellschaft*. Opladen: Westdeutscher Verlag. Recuperado de: [http://www.sozialstruktur.uni-oldenburg.de/dokumente/wissensgesellschaft\\_2002.pdf](http://www.sozialstruktur.uni-oldenburg.de/dokumente/wissensgesellschaft_2002.pdf) [11-04-2016]
- Herrero, C. (2004). Géneros para la divulgación periodística. En Fernández del Moral, J. (coord.). *Periodismo Especializado*. Barcelona: Ariel.
- Hoey, M. (1998). Introducing Applied Linguistics: 25 Years On. *Plenary paper in the 31st BAAL Annual Meeting: Language and Literacies*. University of Manchester.
- Hulteng, J.L. (1979). *The news media: What makes them tick?* New Jersey: Prentice-Hall.

- Hunston, S. (1994). Evaluation and organization in academic discourse. En M. Coulthard [ed.] *Advances in written text analysis*. London: Routledge, (pp. 191-218).
- Hunston, S. y Thompson, G. (2000). *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.  
Recuperado de: <http://assets.cambridge.org/052180/1710/sample/0521801710ws.pdf>
- Hunston, S. (2004). Counting the uncountable: problems of identifying evaluation in a text and in a corpus. En Partington, A., Morley, J. y Haarman, L. [eds.] *Corpora and Discourse*. Peter Lang, (pp. 157-188).
- Hunston, S. y Thompson, G. (2006). *System and Corpus: Exploring connections*. Equinox: University of Liverpool.
- Hunston, S. (2010). *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London: Routledge.
- Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, 17 (pp. 433-454).
- Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text* 18 (3) (pp. 349-382).
- Jakobson, R. (1962). *Phonological studies*. S-Gravenhage: Mouton.
- Janowitz, M. y Schultz, R. (1969). Tendencias de la investigación en el campo de la comunicación de masas. En Eco U et al, *Los efectos de Comunicación de masas*. Buenos Aires: Ed. Jorge Álvarez, (pp. 47).
- Johansson, S. y Hofland, K. (1987). The Tagged LOB Corpus: Description and Analysis. En W., MEIJS [ed.] *Corpus Linguistics and Beyond*. Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora. Amsterdam: Rodopi.
- Jurafsky, D. y Martin, J.H. (2000). *Speech & Language Processing*. Library Binding. Prentice Hall.  
Recuperado de:  
[http://www.deepsky.com/~merovech/voynich/voynich\\_manchu\\_reference\\_materials/PDFs/jurafsky\\_martin.pdf](http://www.deepsky.com/~merovech/voynich/voynich_manchu_reference_materials/PDFs/jurafsky_martin.pdf)
- Kennedy, G.D. (1998). *An introduction to Corpus Linguistics*. Londres: Longman (Studies in Language and Linguistics).
- Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, 6 (1) (pp. 1-37).
- Kilgarriff, A. y Grefenstette, G. (2003). Introduction to the special issue on the web as corpus. *Computational Linguistics*, 29 (3) (pp. 333-348).

- Knorr-Cetina, K. (1998): Sozialität mit Objekten. Soziale Beziehungen in post-traditionalen Wissensgesellschaften. En Werner Rammert [ed.]: *Technik und Sozialtheorie*. Frankfurt am Main/New York: Campus, (pp. 83-120).
- Krohn, W. (1997). Rekursive Lernprozesse: Experimentelle Praktiken in der Gesellschaft. En Rammert, Werner y Bechmann, Gotthard (Ed.) *Technik und Gesellschaft*. Jahrbuch 9: Innovation – Prozesse, produkte, Politik. Frankfurt/ New York: A.M, (pp. 65-89).
- Lafon, P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *MOTS* 1, (pp. 128-165).
- Landow, G.P. (2009) *Hipertexto 3.0: La teoría crítica y los nuevos medios en una época de globalización*. Barcelona: Paidós Comunicación.
- Lawrence, J. y Reed, C. (2015). Combining argument mining techniques. En *Working Notes of the 2nd Argumentation Mining Workshop*, Denver, CO, (pp. 127–136).
- Lebart, L. y Salem, A. (1994). *Statistique textuelle*. Dunod.
- Leech, G. (1991). Corpora en Malmkjaer [ed.] *The Linguistics Encyclopedia*. Londres: Routledge, (pp. 73-80).
- Leech, G. (1992). Corpus Linguistics and Theories of Linguistic Performance. En J. Svartvik [ed.] *New Directions in Corpus Linguistics*. Berlin: Mouton De Gruyter, (pp. 105 -134).
- Levinson, S. (1983). *Pragmatics*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- López de Zuazo, A. (1980). *Catálogo de periodistas españoles del siglo XX*. Editorial Fundación Universidades E.
- Levinson, P. (2013). *New New Media*. 2<sup>nd</sup> Edition. Levinson. Pearson Paper.
- Lewin, K. (1947). Frontiers in group dynamics II: Channels of group life; social planning and action research. *Human Relations*, 1 (pp. 143-153). Recuperado de: <http://journals.sagepub.com/doi/pdf/10.1177/001872674700100201>.
- Llorens, C., Luzón, V. y Grau, H.P. (2012). *Los medios digitales: España. Un informe de Open Society Foundations*. Open Society Foundations. Recuperado de: <https://www.opensocietyfoundations.org/sites/default/files/mapping-digital-media-spain-esp-20131023.pdf>
- Lundvall, B. y Johnson, B. (1994). The Learning Economy. *Journal of Industry Studies* 1, (pp. 23-42). Recuperado de: <http://dx.doi.org/10.1080/13662719400000002>.
- Mann, W. y Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8 (3), (pp. 243-281).

- Martini, S. (2000). *Periodismo, noticia y noticiabilidad*. Buenos Aires: Norma, (pp. 84-99).  
Recuperado de: file:///E:/Users/Luc%C3%ADa/Downloads/886765714.noticiabilidad.pdf
- Marro, M. y Dellamea, A. (1993). *Producción de textos: Estrategias del escritor y recursos del idioma*. Buenos Aires: Editorial Docencia.
- Martin, J. R. (2000). Beyond Exchange: Appraisal systems in English. En Hunston, S. and Thompson, G. [eds.] *Evaluation in Text*. Oxford: Oxford University Press, (pp. 142–175).
- Martin, J.R. y White, P.R.R. (2005). *The Language of Evaluation, Appraisal in English*. London: Palgrave Macmillan.
- Martín, C. (2009). Aproximación a la Lingüística de Corpus y su contribución en la elaboración de diccionarios. *Cuadernos del Instituto de Historia de la Lengua*, ISSN 1889-0709, (pp. 11-24). Recuperado de: file:///E:/Users/Luc%C3%ADa/Downloads/Dialnet-AproximacionALaLinguisticaDeCorpusYSuContribucionE-3186529.pdf.
- Martín Vivaldi, G. (1964). Del pensamiento a la palabra. Curso de redacción. Teoría y práctica de la composición y del estilo. Madrid: Editorial Paraninfo.
- Martín Vivaldi, G. (1972). *Gran enciclopedia Rialp*, VIII. Madrid: Rialp.
- Martín Vivaldi, G. (1973). Géneros periodísticos: reportaje, crónica, artículo (análisis diferencial). Madrid: Editorial Paraninfo.
- Martínez Albertos, J. L. (1974). Redacción periodística. Los estilos y los géneros en la prensa escrita. Barcelona: ATE.
- Martínez Albertos, J.L. (1983). *Curso general de redacción periodística*. Barcelona: Editorial Mitre.
- Martínez Albertos, J. L. (1997). *El ocaso del periodismo*. Barcelona: CIMS.
- Martínez Santiago, F., Ureña López, L.A., García Vega, M. (2001) WWW como Fuente de Recursos Lingüísticos para su Uso en PLN. *Procesamiento del lenguaje natural*. Universidad de Jaén, 27 (pp. 1-7). Recuperado de:  
<http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/3353/1841>
- McCarthy M. (2001). *Issues in Applied Linguistics*. Cambridge: Cambridge University Press.  
Recuperado de: file:///E:/Users/Luc%C3%ADa/Downloads/issues-in-applied-linguistics-paperback-frontmatter%20(1).pdf.
- McCombs, M. (1996). Influencia de las noticias sobre nuestras imágenes del mundo. En Bryant, J. y Zillmann, D. *Los efectos de los medios de comunicación. Investigaciones y teorías*. Barcelona: Paidós, (pp. 13-34).
- McEnery, T. y Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press (Edinburgh Textbooks in Empirical Linguistics), [2ª Edición].

- McLuhan, M. (1962). *The Gutenberg Galaxy: The Making of Typographic Man*. University of Toronto Press.
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. London, New York.  
Recuperado de: [http://robynbacken.com/text/nw\\_research.pdf](http://robynbacken.com/text/nw_research.pdf)
- McLuhan, M. (1968). *Guerra y paz en la aldea global*. Barcelona: Planeta Agostini.
- McLuhan, M. (1996). *Comprender los Medios de Comunicación*. Barcelona: Paidós.
- McQuail, D. (1992). *Introducción a la teoría de comunicación de masas*. Barcelona: Paidós Comunicación.
- Meyer, I., Mackintosh, K. (1996). The Corpus from a Terminographer's Viewpoint. *International Journal of Corpus Linguistics*, 1 (2) (pp. 257-285).
- Meyer C. F. (2002). *English Corpus Linguistics. An Introduction*. Cambridge University Press.  
Recuperado de:  
[http://assets.cambridge.org/97805218/08798/frontmatter/9780521808798\\_frontmatter.pdf](http://assets.cambridge.org/97805218/08798/frontmatter/9780521808798_frontmatter.pdf).
- Miralles Castellano, A. M. (1998). La construcción de lo público desde el periodismo cívico. *Grupo de Trabajo Comunicación y Ciudad*. Asociación Latinoamericana de Investigación de la Comunicación.
- Mochales, R. y Moens, M. (2011). Argumentation mining. *Artificial Intelligence and Law*, 19 (1), (pp. 1-22).
- Molina, L. y Muñoz, J.L. (2002). Redes de publicaciones científicas: un análisis de la estructura de coautorías. *REDES-Revista hispana para el análisis de redes sociales*, 1 (3). Recuperado de: [http://revista-redes.rediris.es/html-vol1/vol1\\_3.htm](http://revista-redes.rediris.es/html-vol1/vol1_3.htm) [12-09-2007]
- Moltó, D. (2015). Comunicación personal. Periodista de *elmundo.es*.
- Mostaza, B. et al. (1966). *Enciclopedia del periodismo*. Cuarta edición. Barcelona: Editorial Noguer.
- Muñoz Torres, J.R. (1997). Aproximación al concepto de información periodística especializada. En Esteve, F. (coord.). *Estudios sobre información periodística especializada*. Valencia: Fundación Universitaria San Pablo CEU.
- Newman, M. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review*. Recuperado de: <http://www-personal.umich.edu/~mejnpapers/016131.pdf> [15-05-2007].
- Núñez Ladevéze, L. (1995). *Introducción al periodismo escrito*. Barcelona: Ariel Comunicación.

- OECD. (1996). *The Knowledge-based Economy*. Arbeitspapier Nr. OECD/GD(96)102. París.  
Recuperado de:  
[http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD\(9\)102&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=OCDE/GD(9)102&docLanguage=En).
- Orive, P. y Fagoaga, C. (1974). *La especialización en el periodismo*. Madrid: Dossat.
- Palmer, H. E. y Hornby, S. (1933). *Second Interim Report on English Collocations*. Tokyo: IRET.
- Peldszus, A. y Stede, M. (2013). From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence*, 7(1), (pp. 1–31).
- Pérez, M.C. (2002). Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento. En *Estudios de Lingüística Española (ELIES)*, (16). Universidad de Málaga. Recuperado de:  
<http://elies.rediris.es/elies18/>.
- Pérez, M. C. y Moreno, A. (2009). Lingüística computacional y Lingüística de Corpus. Potencialidades para la investigación textual. *Teoría y Literatura artística en la sociedad digital: Construcción y aplicabilidad*. Universidad de Málaga: Departamento Filología Inglesa, Alemana y Francesa.
- Portela, P. (1999). La información económica en España. Una revisión, y a la hora de abordar la situación de la prensa económica. El profesional de la información. *Revista internacional científica y profesional*. Recuperado de:  
[http://www.elprofesionaldelainformacion.com/contenidos/1999/diciembre/la\\_informacion\\_economica\\_en\\_espaa\\_una\\_revision.html](http://www.elprofesionaldelainformacion.com/contenidos/1999/diciembre/la_informacion_economica_en_espaa_una_revision.html).
- Quesada, M. (1998). *Periodismo Especializado*. EUINSA; Ediciones Internacionales Universitarias.
- Quirk R. et al. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Real Academia Española. (2001). Diccionario de la lengua española (22.a ed.). Consultado en <http://www.rae.es/rae.html>
- Reich, R. B. (1992). *The Work of Nations. Preparing Ourselves for the 21st Century*. New York: Vintage Books.
- Resnik, P. y Smith, N. A. (2003). The Web as a parallel corpus, *Computational Linguistics*, 29 (3) (pp. 349-380). DOI: 10.1162/089120103322711578.
- Rundell, M. (2000). The biggest corpus of all. *Humanising Language Teaching*, 2 (3). Recuperado de: <http://www.hltmag.co.uk/may00/idea.htm> [21/09/2013]
- Salaverría, R. (2005). *Cibermedios. El impacto de internet en los medios de comunicación en España*. Sevilla: Comunicación Social Ediciones y Publicaciones. Recuperado de:  
<http://hdl.handle.net/10171/34332>.

- Salton, G. (1989). *Automatic text processing: the transformation, analysis and retrieval of information by computer*, Addison Wesley.
- Sampson, G. (1992). Probabilistic parsing. En Svartvik, J [ed. ]. *Directions in corpus linguistics*. Berlin: Mouton de Gruyter.
- Sampson, G. (2001). *Empirical Linguistics*. London, New York: Continuum.
- Santamaría, L. (1990). *El comentario periodístico. Los géneros periodísticos*. Madrid: Editorial Paraninfo.
- Sanz, L. (2003). Análisis de redes sociales: o cómo representar las estructuras sociales subyacentes. *Apuntes de Ciencia y Tecnología*, 7 (pp. 21-9).
- Saperas, E. (1987.) *Los efectos cognitivos de la comunicación de masas*. Barcelona: Ariel Comunicación.
- Scott, M. (1997). PC analysis of key words - and key key words. *System*, 25 (1) (pp.233-345).
- Scott, M. y Tribble, C. (2006). *Textual patterns: Keyword and corpus analysis in language education*. Amsterdam: Benjamins.
- Scott, M. (2012). *WordSmith Tools version 6.0*. Stroud: Lexical Analysis Software.
- Searle, J. (1969). *Speech Acts*. Cambridge: Cambridge University Press.
- Searle, J. (1975). Indirect speech acts. En Peter Cole y Jerry L. Morgan [eds.], *Speech Acts. Syntax and Semantics, Volume 3*. Academic Press, New York, (pp. 59–82).
- Sebba, M. (1991). The Adequacy of Corpora in Machine Translation. En A. McEnery [ed.] *Applied Computer Translation*. 1 (1) (pp. 15-28).
- Sinclair, J.M. [ed.] (1987a). *Collins Cobuild English Language Dictionary*. Londres: Harper Collins.
- Sinclair, J.M. [ed.] (1987b). *Looking Up: an Account of the COBUILD Project in Lexical Computing*. Londres: Collins.
- Sinclair J. (1991). *Corpus, Concordance, Collocation*. Oxford University Press.
- Sinclair, J. (1996). Preliminary Recommendations on Corpus Typology. EAGLES Document EAG-TCWG-CTYP/P, May 1996. Recuperado de: <http://www.ilc.cnr.it/EAGLES96/corpusyp/corpusyp.html>
- Sinclair, J. (2004). *Trust the text: language, corpus and discourse*. London: Routledge.
- Sinclair, J. (2005). Corpus and Text: Basic Principles. *Developing Linguistic Corpora: a Guide to Good Practice*, ed. M. Wynne. Oxford: Oxbow Books: 1-16. Recuperado de: <http://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm> [2016-03-25]

- Stehr, N. (1994). *Arbeit, Eigentum und Wissen. Zur Theorie von Wissensgesellschaften*. Frankfurt: Suhrkamp.
- Stuart K.D. (2006). Towards an analysis of academic weblogs. *Revista Alicantina de Estudios Ingleses*. Universidad de Alicante, 19 (pp. 387-404).
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell.
- Summers, D. (1996). Computer Lexicography. En J. Thomas and M. Short [eds.] *Using Corpora for Language Research*. Londres: Longman
- Summers, D. (1997). The Role of Dictionaries in Language Learning. En R. Carter y M. McCarthy [eds.] *Vocabulary and Language Teaching*. Londres: Longman.
- Svartvik, J. [ed.] (1992). *Directions in Corpus Linguistics*. Berlin: Mouton de Gruyter.
- Tascón, M. y Cabrera, M. (2012). *Escribir en Internet: Guía para nuevos medios y las redes sociales*. Fundación del español urgente. Fundeu BBVA. Barcelona: Galaxia Gutenberg.
- Teuber, W., Cermáková, A. (2007). *Corpus Linguistics: a short introduction*. New York: Bloomsbury Publishing.
- Thetela, P. (1997). Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes*, 16 (2) (pp. 101-118).
- Thompson, G. y Thetela, P. (1995). The sound of one hand clapping: The management of interaction in written discourse. *Text*, 15 (pp. 103-27).
- Thompson, G. (1996). Voices in the text: Discourse perspectives on language reports. *Applied Linguistics*, 17/4, (pp. 501-530).
- Tognini E. Bonelli (2001) *Corpus Linguistics at work*. Amsterdam: John Benjamins Publishing.
- Torrero, A. (2008). La crisis financiera internacional. *Documentos de trabajo 8/2008. Instituto Universitario de Análisis Económico y Social*. Universidad de Alcalá. Recuperado de: <http://hdl.handle.net/10017/6525>.
- Torruella, J. y Llisterri, J. (1999). Diseño de corpus textuales y orales. En J.M. Blecua, G. Clavería, C. Sánchez, J. Torruella [eds.] *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Barcelona: Seminario de Filología e Informática, Departamento de Filología Española, Universidad Autónoma de Barcelona-Editorial Milenio, (pp. 45-47). Recuperado de: [http://liceu.uab.es/~joaquim/publicacions/Torruella\\_Llisterri\\_99.pdf](http://liceu.uab.es/~joaquim/publicacions/Torruella_Llisterri_99.pdf).
- van Dijk, T. A. (1990). *La noticia como discurso. Comprensión, estructura y producción de la información*. Barcelona: Paidós Comunicación.
- van Dijk, T. A. (2002). El conocimiento y las noticias. *Quaderns de Filologia. Estudis de Comunicació*, Vol. 1., (pp. 249-270).

- Vargas Sierra, C. (2006). Diseño de un corpus especializado con fines terminográficos: el Corpus de la Piedra Natural, Debate Terminológico 2 ISSN 1813-1867. París: RITERM (Red Iberoamericana de Terminología).
- Vargas Sierra, C. (2012). La tecnología de un corpus en el contexto profesional y académico de la traducción y la terminología: panorama actual, recursos y perspectivas. En Candel-Mora, M.A. y Ortega-Arjonilla, E. [eds.]. *Tecnología, Traducción y Cultura*. Valencia: Editorial Tirant lo Blanc, (pp.67-99).
- Villayandre, M. (2006). *Lingüística computacional II. Lingüística de corpus*. León: Universidad de León. Recuperado de: <http://www3.unileon.es/dp/dfh/Milka/Milka.htm> [09/03/2016]
- Wang, J., Cong, G., Xin Zhao, W. y Li, X. (2015). Mining user intents in Twitter: A semi-supervised approach to inferring intent categories for tweets. In *Proceedings of the Twenty-Ninth Conference on Artificial Intelligence*, Austin, TX, (pp. 318–324).
- Warren, C. N. (1951). *Modern news reporting*. Nueva York: Harper.
- Warren, C.N. (1975). *Géneros periodísticos informativos*. Barcelona: ATE.
- Warren, C. N. (1979). Géneros periodísticos informativos: nueva enciclopedia de la noticia. Barcelona: ATE.
- Weingart, P. (2001). Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft. *Weilerswist: Velbrück*. Recuperado de: <http://hsozkult.geschichte.hu-berlin.de/rezensionen/768.pdf>.
- Wichmann, A., Fligelstone, S., McEnery, T. y Knowles, G. [eds.] (1997). *Teaching and Language Corpora*. Harlow: Longman.
- Willke, H. (1998). *Systemisches Wissensmanagement*. Stuttgart: UTB/Lucius & Lucius.